*Revised manuscript with changes marked

GC-MS BASED METABOLOMICS OF COLON CANCER CELLS USING DIFFERENT EXTRACTION SOLVENTS

Clara Ibáñez[a], Carolina Simó[b], Mine Palazoglu[c], Alejandro Cifuentes[a*]

(a) Laboratory of Foodomics, CIAL (CSIC), Nicolas Cabrera 9, 28049 Madrid, Spain.

(b) Molecular Nutrition and Metabolism, CIAL (CSIC), Nicolas Cabrera 9, 28049 Madrid, Spain.

(c) University of California Davis, Genome Center, 451 E. Health Sci. Dr., Davis, CA 95616, USA.

Author for the correspondence*: Alejandro Cifuentes (a.cifuentes@csic.es). Laboratory of Foodomics, CIAL (CSIC), Nicolas Cabrera 9, 28049 Madrid, Spain. Phone: +34-91-0017955

ABSTRACT

The increasing incidence of colorectal cancer enforces the development of novel methodologies and protocols to deepen in the molecular mechanisms that govern disease pathophysiological events. The aim of this work is to deepen in the optimum metabolite extraction protocol from adherent mammalian cells of colon cancer for high throughput metabolomics using gas chromatography coupled to mass spectrometry (GC-MS). GC-MS results showed that metabolic information obtained from colon cancer cells was highly dependent on metabolite extraction selection, which at the same time is extremely influenced by the analytical platform. A further purpose of this investigation is to uncover an unexplored portion of HT-29 colon cancer cells metabolome, complementary to other already explored by CE-MS and LC-MS methods. At this respect, a total of ~~a large number of~~150 metabolites were identified in HT-29 colon cancer cells by GC-MS. ~~In addition, metabolite extraction protocol was observed to be crucial for the determination of potentially interesting clusters of metabolites. In summary, t~~The extraction protocol with acetonitrile~~involving ACN i~~Isopropanol-water was ~~shown to be~~ the most appropriate ~~extraction procedure~~ for fatty acids and~~/or~~ related pathways analysis~~ among the four metabolic extraction procedures~~. Most of the metabolites involved in pathways of amino acids, glutathione, amino sugars and other polar metabolites ~~present higher performance of extraction in the~~were better extracted with acidified water~~ extract~~, although ~~metabolic species in~~ water extraction showed the best overall reproducibility. Although pathways involving nitrogenous bases could be investigated using organic or aqueous extracts, a higher number of ~~more~~ metabolites involved in th~~e~~ose pathways ~~can be~~were identified in the aqueous extracts. In addition, metabolite extraction protocol was observed to be crucial for the determination of potentially interesting clusters of metabolites.

Keywords: Colon cancer, Sample treatment, Metabolomics, gas chromatography-mass spectrometry, GC-MS.

Abbreviations: CRC, colorectal cancer; MSEA, metabolite set enrichment analysis; MPA, metabolic pathway analysis; MSTFA, N-methyl-N-(trimethylsilyl)-trifluoroacetamide; ORA, overrepresentation analysis; PBS, phosphate buffered saline; PCA, principal component analysis; RIM, internal retention index markers; TIC, total ion chromatogram

1. INTRODUCTION

In the past few years, cancer biology research has increasingly been focused on metabolism in cancer cells [1]. The complexity and diversity of biological alterations inherently linked to cancer metabolism is in accordance with the increasing unresolved issues to determine specific causes in cancer development, to assess the progression and to unravel molecules or pathways to target in cancer therapy and/or prevention. It is known that cancer cells show different metabolism from healthy cells. One of the well-known metabolic alterations in cancer cells is the glycolysis utilization over the oxidative phosphorylation~~. Normal cells rely on glycolysis followed by oxidative phosphorylation to generate ATP molecules while cancer cells use glycolytic pathway for rapid proliferation and biosynthesis. This dependence upon aerobic glycolysis, also known as Warburg effect, is the best-accepted alteration of tumour cells~~ [2, 3]. In the past years, other perturbations in specific metabolic pathways have been addressed~~. Widespread alterations in the metabolism of cancer cell are~~ linked to energy metabolism [4] such as pyruvate production [5] and tricarboxylic acid (TCA) cycle [6]. ~~Normal healthy cells suffer from complex biochemical changes to be transformed into tumour cells, not well elucidated yet. Widespread cancer cells characteristics include special needs such as high energy status or increased biomolecules generation among others, suggesting that cancer disease encompasses heterogeneous and complex parts of the metabolism [1].~~ The progress of Metabolomics offers valuable opportunity to better understand biochemical changes produced in cancer metabolism for improving early detection, progression and therapy monitoring of cancer disease [7]. Metabolomics is able to meet the challenge of detecting hundreds of metabolites in short time analysis thanks to the evolution of analytical technologies and software tools for data

processing and statistics [8]. With the evolution of metabolomics platforms, rapid and comprehensive analyses of the metabolome complexity can be performed in a high-throughput manner with improved sensitivity, accuracy and resolution [1].Typical metabolomic studies are based on mass spectrometry (MS) or nuclear magnetic resonance (NMR) whose advantages and limitations have been discussed elsewhere [9, 10]. It is now assumed that the coverage of the human metabolome is impossible to achieve with a single analytical methodology and an increasingly common practice is the combination of analytical techniques to achieve complementary information [11]. Metabolomics potential has already been shown with the detection of novel biomarkers involved in different metabolic pathways related to breast [12], liver [13], prostate [14], colon [15] or lung cancer [16] among the vast amount of investigations. From those publications it can be assumed that different cancer subtypes exhibit different phenotypes and therefore show different metabolic alterations and biomarkers. This diversity enforces the investigation of cancer with special attention to most worldwide harmful cancer types.

Colorectal cancer is the second leading cause of cancer mortality in the USA [17] and Europe [18] with a continuous increasing incidence. Identifying soon the colon cancer apparition and/or progression will increase our knowledge on efficient drug discovery research and prevention. With this aim, in the last decade, Metabolomics has emerged to understand pathophysiological processes related to colon cancer. From the about 30 works related to colon cancer Metabolomics published so far, a minor part have been accomplished using NMR [19-23] with a clear trend toward the use of MS-based analytical platforms either alone [24-26] or in combination with separation techniques such as

capillary electrophoresis (CE) [27-30], liquid chromatography (LC) [24, 27-29, 31-34] and gas chromatography (GC) [15, 31-33, 35-44].

Metabolomics investigations of colorectal cancer by GC-MS are summarized in Table S1. As can be seen in Table S1 a variety of samples including biofluids, tissues and cell culture models have been under scrutiny. ~~The minimum invasive nature with relative ease of sampling and less limited volumes result in a predominance of biofluids over the other samples types for metabolomics studies. From those sample types, cell culture presents unique advantages such as sample supply of live human cells, easy control of experimental factors and much higher reproducibility in studies inter-laboratories.~~ The use of cell cultures is included in common reference model systems and is considered an invaluable biomedical research tool. However, the application of cell culture in metabolomics requires further development and standardization of study design steps, metabolism quenching method selection and optimization of metabolite extraction protocols, among others [45, 46]. Namely, quenching step at the time of harvesting aims to prevent metabolic content to be altered before the analysis so that the metabolic state of the cell is preserved. There is a great controversy and continuous research in the determination of the optimum quenching method. Until date the most widely spread techniques to quench the metabolism in cells is a shock maintaining the cells below -20 ºC [47, 48], the addition of cold solvents such as cold methanol [49], use of cold isotonic PBS [49-51] or the combination in a single step of quenching and extraction procedures in mammalian adherent cells [45, 52].

As can be observed in Table S1, three cell lines (namely SW, HT and Caco-2) have been used to investigate colon cancer following metabolomics approaches using GC-MS. All these cell lines are very well established for in-vitro studies of colorectal cancer and vary in

phenotype, growth rate, differentiation and morphology among other characteristics [53, 54]. Metabolic fingerprinting by GC- MS from two SW cell lines (SW-1116 and SW-480) was first reported by Zimmermann et al., [35]. More recently LC-MS and CE-MS based metabolomics have been used to investigate HT-29 cell line revealing significant metabolic information in colon cancer metabolism after treatment with different polyphenol-rich extracts [28, 29, 55].

Considering that no single method is appropriate for the determination of all intracellular metabolites, metabolic information obtained from cell cultures will depend on multiple factors and metabolite extraction selection is highly influenced by the analytical platform [56]. Optimization and selection of the protocol for metabolite extraction from HT-29 cells [57] and evaluation of the subsequent cytosol sample treatment [30] for CE-MS metabolomics have been previously carried out. CE-MS is particularly suited for the rapid separation of ionic, weakly ionic, and/or highly polar metabolites and aqueous solvent was the preferred for metabolite extraction of HT-29 cells [57]. On the contrary, GC-MS is preferred for the analysis of less polar, volatile (or amenable to chemical derivatization) compounds and requires thermal stability of the analytes. Considering that no single analytical platform is able to determine all intracellular metabolites, in this work, a GC-MS method has been developed due to the high complementarity nature of GC-MS used in this work and the CE-MS [57] and LC-MS [28, 29] methods already developed by our group. Further, taking into account all the above considerations, in this work GC-MS has been used to evaluate the coverage and reproducibility of four metabolic extraction solvents selected according to the physicochemical diversity of metabolites including the large variation on solubility (in aqueous or organic solvents) and based on previous

published works [30, 57-63]. To our knowledge this is the first time that HT-29 colon cancer cells are examined using GC-MS Metabolomics which could be of special importance in future studies to detect biochemical alterations due to colon cancer or to explore new preventive interventions.

2. MATERIALS AND METHODS

2.1. Chemicals

All reagents were of analytical grade. Phosphate buffered saline (PBS) was purchased from Lonza (Barcelona, Spain). For cell counting, trypan blue was purchased from Sigma Aldrich (St. Louis, MO, USA). Metabolite extraction solvents were of MS grade: formic acid and 2-propanol were from Riedel-de Haën (Seelze, Germany) while water and acetonitrile were from Labscan (Gliwice, Poland). Reagents for derivatization included methoxyamine hydrochloride from Sigma Aldrich, pyridine (silylation grade) and N-methyl-N-(trimethylsilyl)-trifluoroacetamide (MSTFA) purchased in 1 mL bottles from Pierce (Rockford IL USA).

For GC-MS analysis, a mixture of internal retention index markers (RIM) was prepared using fatty acid methyl esters of C8, C9, C10, C12, C14, C16, C18, C20, C22, C24, C26, C28 and C30 linear chain length, dissolved in chloroform (from Mallinckrodt Baker Inc., Phillipsburg. NJ, USA) at a concentration of 0.8 mg/mL (C8-C16) and 0.4 mg/mL (C18-C30). Fatty acid methyl esters were from Sigma Aldrich.

FC43 (Perfluorotributylamine) was ~~purchased~~acquired from Leco (Stockport, Cheshire, UK) for MS mass calibration.

2.2. Cell culture

The human HT-29 colorectal carcinoma cell line was purchased from the American Type Culture Collection (ATCC- LGC, Promochem, UK). HT-29 cells were cultured in McCoy's 5A medium supplemented with 10% heat-inactivated fetal calf serum, 50 U/mL penicillin G, and 50 U/mL streptomycin, at 37 ºC in humidified atmosphere with 5% CO_2. Once cells reached about 50% confluency, they were trypsinized, neutralized with culture medium and plated at 10000 cells/cm^2 density in different P150 culture dishes. Cell cultures were allowed to adhere overnight at 37 ºC and then plates were pooled. Next, cell count was performed using the trypan blue exclusion test to examine the viability and to estimate the number of cells. For that purpose, 5 µL cell suspension pool were further diluted 1:1 in trypan blue stock solution (0.4%, w/v in sterile PBS) and counted in a Neubauer counting chamber using a light microscope (ID3, Carl Zeiss, Jena, Germany). The estimated volume to have 10 million cells from the culture suspension was added to 20 aliquots. After aliquots were centrifuged (300 xg, 10 min) and supernatant was discarded, pellets containing 10 x 10^6 cells were subjected to metabolite extraction.

2.3. Metabolite extraction and derivatization

The 20 aliquots were divided in four solvent extraction groups: A, ACN; B, ACN-IsopOH-Water (3:3:1, v/v/v); C, Water; and D, 5% formic acid in water. Five independent metabolite extractions (with10 million cells each) were performed for each extraction solvent group. Metabolic content extraction from HT-29 cells have been previously optimized in our laboratory [57] and applied in this work with slight modifications. Briefly, 300 µL of extraction solvent and 0.3 g glass beads (212-300 µm) from Sigma-Aldrich were

added to 10 x 10^6 cell pellets. Then metabolism quenching and cell disruption was accomplished by using three cycles of snap-freeze (liquid N2, 1 min), thawing in ultrasound bath (50 Hz, 3 min) (ultrasonic cleaning bath Ultrasons from JP Selecta, Barcelona, Spain) and grinding using a mixer mill for 3 min at 30/s (MM 400, Retsch GmbH, Germany). After this, cell insoluble material and glass beads were removed by centrifugation (24000 xg for 10 min at 4ºC) and the supernatant was collected, and 50 µL metabolic extracts were vacuum-dried and kept at -80ºC until derivatization. The supernatants obtained in A and B extraction groups are protein-free considering organic extraction solvents lead to protein precipitation. On the contrary, supernatants obtained from aqueous-based extractions (C and D extraction groups) were submitted to ultrafiltration with 3 kDa centrifugal filters (Amicon Ultra 0.5 mL, Millipore,Billerica, MA, USA) to remove proteins according to the manufacturer's protocol. 50 µL of both metabolic extracts were vacuum-dried and kept at -80ºC until derivatization.

The best sensitivity was achieved with the following optimized derivatization protocol. First, the dried extracts were dissolved in 10 µL of a solution consisting of 40 mg/mL methoxyamine in pyridine and shaken for 15 min at 60 ºC to protect aldehyde and ketone groups. Then 50 µL of MSTFA with 1% RIM was added for trimethylsilylation of acidic protons and shaken at 37 °C for 30 min. Derivatized samples were immediately transferred to 2 mL clear glass autosampler vials with microinserts (Agilent, Santa Clara CA), closed by 11 mm T/S/T crimp caps (MicroLiter, Suwanee GA) and analyzed by GC-MS.

Formatted: Superscript

2.4. GC-TOF MS analysis

An Agilent 6890 gas chromatograph (Santa Clara CA) and a Leco Pegasus IV time of flight mass (TOF) spectrometer were controlled by the Leco ChromaTOF software vs. 2.32 (St. Joseph, MI). A 30 m long, 0.25 mm ID Rtx-5Sil MS column with 0.25 μm 95% dimethyl 5% diphenyl polysiloxane film and additional 10 m integrated guard column was used (Restek, Bellefonte PA). High purity helium with built-in purifier (Airgas, Radnor PA) was set at constant flow of 1 mL/min. The oven temperature was held constant at 50°C for 1 min and then ramped at 20°C/min to 330°C at which it was held constant for 5 min. The transfer line temperature between gas chromatograph and mass spectrometer was set to 280°C. A Gerstel automatic liner exchange system with multipurpose sample MPS2 dual rail and two derivatization stations was used in conjunction with a Gerstel CIS cold injection system (Gerstel, Muehlheim, Germany). The injector operated in splitless mode, opening the split vent after 25 s. Electron impact ionization at 70V was applied maintaining ion source temperature at 250°C. For every 10 samples, a fresh multibaffled liner was inserted (Gerstel) using the Maestro1 Gerstel software vs. 1.1.4.18 (Gerstel). Before and after each injection, the 10 μL injection syringe was washed three times with 10 μL ethyl acetate. 1 μL sample was filled using 39 mm vial penetration at 1 μL/s fill speed, injecting 0.5 μl at 10 μL/s injection speed. GC oven temperature was maintained at 50°C during one minute followed by a temperature ramp at 20°C/min until 330°C at which it was held constant for 5 min.

TOF-MS analysis in the positive ion mode was accomplished applying 1775 V detector voltage. Data was recorded in the 85-500 m/z range for 1200 s after 290 s solvent delay. Mass resolving power was 600 and acquisition rate was set at 17 spectra/s. Daily quality

controls were used to assure quality of data. The instrument performed autotuning for mass calibration using FC43 before each sequence analysis using the following m/z values: 93, 95,100, 101, 112, 113, 114, 119, 131, 132, 145, 150, 164, 169, 176, 181, 214, 219, 220, 226, 264, 265, 314, 326, 352, 376, 414, 415, 426, 448, 464, 502 and 503.

2.5. Data mining

GC-TOF MS data were acquired as ChromaTOF specific .peg files, and stored also in .txt and .cdf generic formats. Data mining was performed as already described [60]. Briefly, absolute spectra intensities from all sample files were automatically processed and filtered by BinBase algorithm using previously optimized parameters [64]. Then post-processing module to replace missing values was carried out. The resulting report contained row identifier, mass spectra ions, quantification by means of unique ion height, retention index and identification using defined databases [65]. Additional mass signal filtering was manually performed to remove not identified and low confident ions from the table. Then MS signals belonging to the same ion were grouped and data was normalized based on total ion height for each sample. The table with high-confident metabolite information from all extraction replicates obtained with the four solvents (five replicates per extraction solvent) was then examined in detail. Principal component analysis (PCA) and hierarchical clustering was performed to explore possible clustering, similarities and differences in metabolic coverage obtained among the four extraction protocols. Then comparison of metabolome coverage among the four extracts was performed by metabolite set enrichment analysis (MSEA) and metabolic pathway analysis (MPA). MSEA was performed by means of overrepresentation analysis (ORA). ORA assess whether a set of metabolites is enriched in a particular metabolic pathway compared to the total of annotated metabolites for that

pathway ~~[66]~~ (i.e. if the metabolites identified in the study represent a high percentage of the metabolites theoretically annotated in that pathway) [66]. Statistical analysis, MSEA and MPA were performed using metaboanalyst [67]. MBRole [68] was used to assist in the biological interpretation.

3. RESULTS AND DISCUSSION

3.1. Metabolic extraction assessment

Typical total ion chromatograms (TIC) for the cell extracts obtained using the four different extraction solvents tested in this work are represented in Fig. 1. As can be observed, the four metabolite profiles were very different in terms of intensity and distribution of the peaks along the chromatogram. Although aqueous extracts (Fig. 1C and 1D) presented greater overall intensity, peaks were mainly placed at the beginning of the chromatogram.

After data processing, 364 metabolic signals were detected. Signal grouping and filtering processes (i.e. unidentified metabolites and metabolites not found in at least 3 out of 5 replicates of each extraction solvent) rendered a total of 290 metabolites from which a group of 150 metabolites could be identified (see Table ~~2~~1 and Table S2). Then, each of the metabolic extracts were investigated in detail in terms of detected metabolites and identified metabolites found in each extract as well as other parameters to evaluate metabolite extraction performance and reproducibility. An average variation of the internal standards of 15% RSD was observed. Given that value the number of identified metabolites with RSD < 15% was also determined for the four extraction solvents.

Among the two aqueous based extractions C (water) and D (acidified water), C showed higher overall metabolome coverage in terms of number of detected (172 vs. 168) and

identified metabolites (100 vs. 94). However, deepening in the identified metabolites in both aqueous extracts, D showed a vast gain over C with respect to metabolite recovery performance (76 vs. 25 metabolites found at maximum intensity in D and C respectively) but the inverse was observed in terms of reproducibility (RSD <15% in 62 vs. 20 metabolites in C and D, respectively) (see Table ~~2~~1). As can be observed in Table ~~2~~1, a more evident enhancement was observed in B metabolic coverage when compared with A extract among the two organic extracts. Namely, 96 metabolites were identified out of the 182 detected in B, while for A only 82 metabolites could be identified out of the 141 detected. Moreover, triple metabolite recovery performance (i.e. metabolites at maximum intensity) from cell culture was observed for B with respect to A. In addition, metabolic content in A was extracted with less reproducibility among the 5 extraction replicates analysed, including only 17 identified metabolites with RSD values lower than 15%. Thus, among the organic solvents, B extract clearly offered the best overall performance when compared to A.

3.2. Colon cancer cell metabolome identified by GC-MS

Metabolic information obtained from the four extraction protocols was then statistically investigated in order to find patterns and possible clustering in metabolic coverage obtained applying the four extraction solvents. Fig. 2 represents the unsupervised hierarchical clustering of the samples (i.e. metabolic extracts) ~~attending to~~according to the relative content of each metabolite in each sample (represented by the colour of the cell). As can be observed from the clustering information of Fig. 2, the four extracts were mainly differentiated ~~attending to~~according to the polarity of the solvent used in the extraction. Thus, as expected two clear clusters were observed corresponding to A with B (i.e. organic

based solvents), and C with D (i.e. aqueous based solvents). This information was also observed and corroborated by means of PCA (see Fig. S1) and Venn diagram (see Fig. 3). In addition, PCA also showed that the metabolic information obtained from both organic solvents was more similar (presented less variation in their distribution in the plane defined by the main principal components from PC1 to PC5) than the one more heterogeneous obtained from both aqueous based extracts (more dispersion of aqueous samples in the plane defined by the PCs) (see Fig. S1).

Venn representation (Fig. 3) corroborated that the highest similarity of metabolic coverage was ~~attending to~~according to the extraction solvent polarity. Namely, organic based extracts (A and B) shared 72 metabolites, among which, 18 metabolites were exclusively detected in A and B. Similarly 80 metabolic species were common in both aqueous based extracts (C and D) from which 14 were only found in C and D. In addition, it could be observed that ACN and water extraction solvents (A and C, respectively) lead to the most different metabolic extracts with only 57 metabolites in common (from which 49 were found in all the four extracts) showing high complementarity.

To deepen in the coverage of the different extracts ~~attending to~~according to the nature of the extracted metabolites and pathways involved, two web servers were used. Information about the biological role and nature of the identified metabolites was derived from MBRole web server analysis. Thus, the total set of metabolites with available KEGG identifiers in the four extracts (107 out of 150 identified metabolites) were submitted to analysis. Most important groups of metabolites identified were amino acids and peptides followed by fatty acids and lipids, nucleic acids and carbohydrates. Other groups of metabolites were represented to a lesser extent such as vitamins and cofactors (data not shown). Main

metabolic pathways involving metabolites preferentially extracted (i.e. showing maximum levels) using the organic-based (A and B) and aqueous-based (C and D) protocols were analyzed using Metaboanalyst (Table S31 and S24 respectively). Fatty acid and nitrogenous bases (i.e. pyrimidine and purine) related pathways are well represented in organic based extracts. Among the two organic extracts, contribution of B extract is clearly superior to A whose metabolites are almost exclusively involved in nucleotide sugars metabolism. The main reason for this inequality could be explained for the lowest number of metabolites at highest intensity in A (see Table 21). On the other hand, pathways involving biosynthesis and metabolism of amino acids, glutathione, amino sugars and nitrogenous bases among others, are highly represented in aqueous extracts (C and D). Most of the metabolites involved in those pathways present higher performance of extraction using acidified water extraction (D) when compared with C (water extraction). Although in lower relative concentration, metabolic species in C showed the best reproducibility.

Finally, identified metabolites were submitted to enrichment analysis (MSEA) by means of overrepresentation analysis (ORA) to determine the more represented metabolic pathways in the defined set of metabolites. As can be observed in Fig. S2, “protein biosynthesis” is the metabolic pathway in which a higher number of the metabolites are involved. However, more useful information can be obtained if solvents are subdivided regarding polarity. When each protocol for metabolic extraction is investigated in detail, different trends are observed. MSEA results of metabolites preferentially extracted (i.e. showing maximum levels) in organic and aqueous based extracts are shown in Fig. 4 and Fig. 5 respectively. As can be observed depending on the nature and polarity of the solvents used in the extraction process diverse metabolic pathways are more suitable to study due to the very

different metabolic information. Some of the most relevant metabolic pathways that can be studied with more efficiency in organic based solvents are “RNA transcription” and “beta oxidation of very long fatty acids” (Fig. 4), while pathways most suitable to be evaluated in aqueous based extracts include protein, sugars and amino acids biosynthesis and metabolism (Fig. 5). Interestingly, pathways involving metabolism of glutathione and nucleotides can be approached using both organic and/or aqueous extracts.

CONCLUDING REMARKS

To date it is not possible to study the whole metabolome of a biological system and this is partially due to the selection of a certain sample treatment and analytical platform. In this work, four reliable sample preparation protocols have been evaluated to extract metabolites from HT-29 colon cancer cells and to uncover a portion of this cell line metabolome by GC-MS.

The unsupervised statistical analysis showed that the four extracts evaluated were mainly differentiated ~~attending to~~according to the polarity of the solvent used in the extraction process and that the overall GC-MS variability of aqueous extracts was higher compared to the organic ones.

In summary, the protocol involving ACN-Isopropanol-water (3:3:1, v/v/v) was shown to be the most appropriate extraction procedure for fatty acid and/or related pathways analysis among the four metabolic extraction procedures. Metabolic extraction with ACN yielded the lowest reproducibility of extraction together with the smallest metabolite coverage. On the other hand, metabolic coverage was similar in both aqueous extracts. However, overall performance of extraction of individual metabolites seemed to be higher using acidified

water (5% formic acid) compared to just water, which on the contrary showed better GC-MS reproducibility. Most of the metabolites involved in pathways of amino acids, glutathione, amino sugars and other polar metabolites present higher performance of extraction in the acidified water extract, although metabolic species in water extract show the best overall reproducibility. Although pathways involving nitrogenous bases could be investigated using organic or aqueous extracts, more metabolites involved in those pathways can be identified in the aqueous extracts.

In light of the results obtained, qualitative information about the metabolic coverage and related pathways exhibited high complementarity among the extracts. This information is useful for future hypothesis-driven studies for the evaluation of a certain metabolic pathway or a group of metabolites (i.e. if they are thought to be altered in HT29 colorectal cells by the action of a drug or a natural bioactive compound).

Conflict of interest statement

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

This work was supported the projects AGL2011-29857-C03-01 (Ministerio de Educación y Ciencia, Spain) and AGL2014-53609- P (Ministerio de Economía y Competitividad, Spain). CI thanks to Ministerio de Economía y Competitividad for her stay abroad grant (EEBB-I-12-03995) and for her Juan de la Cierva postdoctoral research contract (FJCI-2014-19601). Authors thank Prof. Oliver Fiehn (UC Davis, CA, USA) for his help with this work. Current address of CI: IMDEA Food Institute, Ctra. Cantoblanco 4, 28049 Madrid, Spain.

Formatted: Font: Not Bold

Formatted: Font: Not Bold

References

[1] R.A. Cairns, I.S. Harris, T.W. Mak, Regulation of cancer cell metabolism, Nat. Rev. Cancer 11 (2011) 85-95.

[2] O. Warburg, On the origin of cancer cells, Science 123 (1956) 309-314.

[3] M.D. Williams, R. Reeves, L.S. Resar, H.H. Hill Jr, Metabolomics of colorectal cancer: past and current analytical platforms, Anal. Bioanal. Chem. 405 (2013) 5013-5030.

[4] D. Hanahan, R.A. Weinberg, Hallmarks of cancer: the next generation, Cell 144 (2011) 646-674.

[5] M.G. Vander Heiden, J.W. Locasale, K.D. Swanson, H. Sharfi, G.J. Heffron, D. Amador-Noguez, H.R. Christofk, G. Wagner, J.D. Rabinowitz, J.M. Asara, L.C. Cantley, Evidence for an alternative glycolytic pathway in rapidly proliferating cells, Science 329 (2010) 1492-1499.

[6] H. Kondoh, M.E. Lleonart, J. Gil, J. Wang, P. Degan, G. Peters, D. Martinez, A. Carnero, D. Beach, Glycolytic enzymes can modulate cellular life span, Cancer Res. 65 (2005) 177-185.

[7] N.J. Serkova, K. Glunde, Metabolomics of cancer, Methods Mol. Biol. 520 (2009) 273-295.

[8] R.D. Beger, A review of applications of metabolomics in cancer, Metabolites 3 (2013) 552-574.

[9] K. Dettmer, P.A. Aronov, B.D. Hammock, Mass spectrometry-based metabolomics, Mass Spectrom. Rev. 26 (2007) 51-78.

[10] A. Smolinska, L. Blanchet, L.M. Buydens, S.S. Wijmenga, NMR and pattern recognition methods in metabolomics: from data acquisition to biomarker discovery: a review, Anal. Chim. Acta 750 (2012) 82-97.

[11] W.B. Dunn, D.I. Ellis, Metabolomics: current analytical platforms and methodologies, TrAC-Trend. Anal. Chem. 24 (2005) 285-294.

[12] C. Denkert, E. Bucher, M. Hilvo, R. Salek, M. Orešič, J. Griffin, S. Brockmöller, F. Klauschen, S. Loibl, D.K. Barupal, Metabolomics of human breast cancer: new approaches for tumor typing and biomarker discovery, Genome Med. 4 (2012) 37.

[13] Q. Huang, Y. Tan, P. Yin, G. Ye, P. Gao, X. Lu, H. Wang, G. Xu, Metabolic characterization of hepatocellular carcinoma using nontargeted tissue metabolomics, Cancer Res. 73 (2013) 4992-5002.

[14] T. Zhang, D.G. Watson, L. Wang, M. Abbas, L. Murdoch, L. Bashford, I. Ahmad, N. Lam, A.C. Ng, H.Y. Leung, Application of holistic liquid chromatography-high resolution mass spectrometry based urinary metabolomics for prostate cancer detection and biomarker discovery, PloS one 8 (2013) e65880.

[15] S. Nishiumi, T. Kobayashi, A. Ikeda, T. Yoshie, M. Kibi, Y. Izumi, T. Okuno, N. Hayashi, S. Kawano, T. Takenawa, A novel serum metabolomics-based diagnostic approach for colorectal cancer, PloS one 7 (2012) e40459.

[16] S. Hori, S. Nishiumi, K. Kobayashi, M. Shinohara, Y. Hatakeyama, Y. Kotani, N. Hatano, Y. Maniwa, W. Nishio, T. Bamba, A metabolomic approach to lung cancer, Lung Cancer 74 (2011) 284-292.

[17] American Cancer Society, Colorectal Cancer Facts & Figures 2011-2013, Atlanta: American Cancer Society (2011).

[18] F. Bray, The burden of cancer in Europe, Responding to the challenge of cancer in Europe.. in: M.P. Coleman, D.M. Alexe, T. Albreht, M. McKee (Eds.) Responding to the challenge of cancer in Europe. Institute of Public Health of the Republic of Slovenia, Ljubljana, 2008, pp. 7-40.

[19] D. Monleon, J.M. Morales, A. Barrasa, J.A. Lopez, C. Vazquez, B. Celda, Metabolite profiling of fecal water extracts from human colorectal cancer, NMR Biomed. 22 (2009) 342-348.

[20] M. Tessem, K.M. Selnæs, W. Sjursen, G. Tranø, G.F. Giskeødegård, T.F. Bathen, I.S. Gribbestad, E. Hofsli, Discrimination of patients with microsatellite instability colon cancer using 1H HR MAS MR spectroscopy and chemometric analysis, J. Proteome Res. 9 (2010) 3664-3670.

[21] H. Wang, D.E. Schiller, V. Tso, C. Slupsky, C.K. Wong, R.N. Fedorak, A novel highly sensitive test for detecting colon cancer using spot urine metabolomics, Gastroenterology 140 (2011) S-40.

[22] I. Bertini, S. Cacciatore, B.V. Jensen, J.V. Schou, J.S. Johansen, M. Kruhoffer, C. Luchinat, D.L. Nielsen, P. Turano, Metabolomic NMR fingerprinting to identify and predict survival of patients with metastatic colorectal cancer, Cancer Res. 72 (2012) 356-364.

[23] B. Jiménez, R. Mirnezami, J. Kinross, O. Cloarec, H.C. Keun, E. Holmes, R.D. Goldin, P. Ziprin, A. Darzi, J.K. Nicholson, 1H HR-MAS NMR spectroscopy of tumor-induced local metabolic “field-effects” enables colorectal cancer staging and prognostication, J. Proteome Res. 12 (2013) 959-968.

[24] S.A. Ritchie, P.W. Ahiahonu, D. Jayasinghe, D. Heath, J. Liu, Y. Lu, W. Jin, A. Kavianpour, Y. Yamazaki, A.M. Khan, Reduced levels of hydroxylated, polyunsaturated ultra long-chain fatty acids in the serum of colorectal cancer patients: implications for early screening and detection, BMC Med. 8 (2010) 1.

[25] F. Li, X. Qin, H. Chen, L. Qiu, Y. Guo, H. Liu, G. Chen, G. Song, X. Wang, F. Li, Lipid profiling for early diagnosis and progression of colorectal cancer using direct-infusion electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry, Rapid Commun. Mass Sp. 27 (2013) 24-34.

[26] E.G. Armitage, H.L. Kotze, J.S. Fletcher, A. Henderson, K.J. Williams, N.P. Lockyer, J.C. Vickerman, Time-of-flight SIMS as a novel approach to unlocking the hypoxic properties of cancer, Surf. Interface Anal. 45 (2013) 282-285.

[27] A. Hirayama, K. Kami, M. Sugimoto, M. Sugawara, N. Toki, H. Onozuka, T. Kinoshita, N. Saito, A. Ochiai, M. Tomita, H. Esumi, T. Soga, Quantitative metabolome profiling of colon and stomach cancer microenvironment by capillary electrophoresis time-of-flight mass spectrometry, Cancer Res. 69 (2009) 4918-4925.

[28] C. Ibáñez, C. Simó, V. García-Cañas, Á Gómez-Martínez, J.A. Ferragut, A. Cifuentes, CE/LC-MS multiplatform for broad metabolomic analysis of dietary polyphenols effect on colon cancer cells proliferation, Electrophoresis 33 (2012) 2328-2336.

[29] C. Ibáñez, A. Valdés, V. García-Cañas, C. Simó, M. Celebier, L. Rocamora-Reverte, Á Gómez-Martínez, M. Herrero, M. Castro-Puyana, A. Segura-Carretero, Global Foodomics strategy to investigate the health benefits of dietary constituents, J. Chromatogr A 1248 (2012) 139-153.

[30] C. Simó, C. Ibáñez, Á Gómez-Martínez, J.A. Ferragut, A. Cifuentes, Is metabolomics reachable? Different purification strategies of human colon cancer cells provide different CE-MS metabolite profiles, Electrophoresis 32 (2011) 1765-1777.

[31] B. Tan, Y. Qiu, X. Zou, T. Chen, G. Xie, Y. Cheng, T. Dong, L. Zhao, B. Feng, X. Hu, Metabonomics identifies serum metabolite markers of colorectal cancer, J. Proteome Res. 12 (2013) 3000-3009.

[32] D.C. Montrose, X.K. Zhou, L. Kopelovich, R.K. Yantiss, E.D. Karoly, K. Subbaramaiah, A.J. Dannenberg, Metabolic profiling, a noninvasive approach for the detection of experimental colorectal neoplasia, Cancer. Prev. Res. (Phila) 5 (2012) 1358-1367.

[33] Y. Cheng, G. Xie, T. Chen, Y. Qiu, X. Zou, M. Zheng, B. Tan, B. Feng, T. Dong, P. He, Distinct urinary metabolic profile of human colorectal cancer, J. Proteome Res. 11 (2011) 1354-1363.

[34] Y. Ma, H. Qin, W. Liu, J. Peng, L. Huang, X. Zhao, Y. Cheng, Ultra-high performance liquid chromatography–mass spectrometry for the metabolomic analysis of urine in colorectal cancer, Dig. Dis. Sci. 54 (2009) 2655-2662.

[35] D. Zimmermann, M. Hartmann, M.P. Moyer, J. Nolte, J.I. Baumbach, Determination of volatile products of human colon cell line metabolism by GC/MS analysis, Metabolomics 3 (2007) 13-17.

[36] E.C.Y. Chan, P.K. Koh, M. Mal, P.Y. Cheah, K.W. Eu, A. Backshall, R. Cavill, J.K. Nicholson, H.C. Keun, Metabolic profiling of human colorectal cancer using high-resolution magic angle spinning nuclear magnetic resonance (HR-MAS NMR) spectroscopy and gas chromatography mass spectrometry (GC/MS), J. Proteome Res. 8 (2008) 352-361.

[37] C. Denkert, J. Budczies, W. Weichert, G. Wohlgemuth, M. Scholz, T. Kind, S. Niesporek, A. Noske, A. Buckendahl, M. Dietel, Metabolite profiling of human colon carcinoma–deregulation of TCA cycle and amino acid turnover, Mol. Cancer 7 (2008) 72.

[38] Y.L. Ma, W.J. Liu, J.Y. Peng, P. Zhang, H.Q. Chen, H.L. Qin, Study on specific metabonomic profiling of serum from colorectal cancer patients by gas chromatography-mass spectrometry, Zhonghua Wei Chang Wai Ke Za Zhi 12 (2009) 386-390.

[39] Y. Ma, W. Liu, J. Peng, L. Huang, P. Zhang, X. Zhao, Y. Cheng, H. Qin, A pilot study of gas chromatograph/mass spectrometry-based serum metabolic profiling of colorectal cancer after operation, Mol. Biol. Rep. 37 (2010) 1403-1411.

[40] Y. Qiu, G. Cai, M. Su, T. Chen, Y. Liu, Y. Xu, Y. Ni, A. Zhao, S. Cai, L.X. Xu, Urinary metabonomic study on colorectal cancer, J. Proteome Res. 9 (2010) 1627-1634.

[41] Y. Kondo, S. Nishiumi, M. Shinohara, N. Hatano, A. Ikeda, T. Yoshie, T. Kobayashi, Y. Shiomi, Y. Irino, T. Takenawa, Serum fatty acid profiling of colorectal cancer by gas chromatography/mass spectrometry, Biomark. Med. 5 (2011) 451-460.

[42] M. Mal, P.K. Koh, P.Y. Cheah, E.C.Y. Chan, Metabotyping of human colorectal cancer using two-dimensional gas chromatography mass spectrometry, Anal. Bioanal. Chem. 403 (2012) 483-493.

[43] T. Yoshie, S. Nishiumi, Y. Izumi, A. Sakai, J. Inoue, T. Azuma, M. Yoshida, Regulation of the metabolite profile by an APC gene mutation in colorectal cancer, Cancer Sci. 103 (2012) 1010-1021.

[44] L.C. Phua, M. Mal, P.K. Koh, P.Y. Cheah, E.C.Y. Chan, H.K. Ho, Investigating the role of nucleoside transporters in the resistance of colorectal cancer to 5-fluorouracil therapy, Cancer Chemother. Pharmacol. 71 (2013) 817-823.

[45] M.A. Lorenz, C.F. Burant, R.T. Kennedy, Reducing time and increasing sensitivity in sample preparation for adherent mammalian cell metabolomics, Anal. Chem. 83 (2011) 3406-3414.

[46] C.A. Sellick, D. Knight, A.S. Croxford, A.R. Maqsood, G.M. Stephens, R. Goodacre, A.J. Dickson, Evaluation of extraction processes for intracellular metabolite profiling of mammalian cells: matching extraction approaches to cell type and metabolite targets, Metabolomics 6 (2010) 427-438.

[47] B. Álvarez-Sánchez, F. Priego-Capote, M.L. de Castro, Metabolomics analysis II. Preparation of biological samples prior to detection, TrAC-Trend. Anal. Chem. 29 (2010) 120-127.

[48] B. Gonzalez, J. François, M. Renaud, A rapid and reliable method for metabolite extraction in yeast using boiling buffered ethanol, Yeast 13 (1997) 1347-1355.

[49] S. Dietmair, N.E. Timmins, P.P. Gray, L.K. Nielsen, J.O. Krömer, Towards quantitative metabolomics of mammalian cells: Development of a metabolite extraction protocol, Anal. Biochem. 404 (2010) 155-164.

[50] S. Dietmair, M.P. Hodson, L. Quek, N.E. Timmins, P. Chrysanthopoulos, S.S. Jacob, P. Gray, L.K. Nielsen, Metabolite profiling of CHO cells with different growth characteristics, Biotechnol. Bioeng. 109 (2012) 1404-1414.

[51] J. Kronthaler, G. Gstraunthaler, C. Heel, Optimizing high-throughput metabolomic biomarker screening: a study of quenching solutions to freeze intracellular metabolism in CHO cells, Omics 16 (2012) 90-97.

[52] Q. Teng, W. Huang, T.W. Collette, D.R. Ekman, C. Tan, A direct cell quenching method for cell-culture based metabolomics, Metabolomics 5 (2009) 199-208.

[53] I. Chantret, A. Barbat, E. Dussaulx, M.G. Brattain, A. Zweibaum, Epithelial polarity, villin expression, and enterocytic differentiation of cultured human colon carcinoma cells: a survey of twenty cell lines, Cancer Res. 48 (1988) 1936-1942.

[54] D. Ahmed, P. Eide, I. Eilertsen, S. Danielsen, M. Eknæs, M. Hektoen, G. Lind, R. Lothe, Epigenetic and genetic features of 24 colon cancer cell lines, Oncogenesis 2 (2013) e71.

[55] S. Fernández-Arroyo, A. Gómez-Martínez, L. Rocamora-Reverte, R. Quirantes-Piné, A. Segura-Carretero, A. Fernández-Gutiérrez, J. Ferragut, Application of nanoLC-ESI-TOF-MS for the metabolomic analysis of phenolic compounds from extra-virgin olive oil in treated colon-cancer cells, J. Pharm. Biomed. Anal. 63 (2012) 128-134.

[56] M. Čuperlović-Culf, D.A. Barnett, A.S. Culf, I. Chute, Cell culture metabolomics: applications and future directions, Drug Discov. Today 15 (2010) 610-621.

[57] C. Ibáñez, C. Simó, A. Valdés, L. Campone, A.L. Piccinelli, V. García-Cañas, A. Cifuentes, Metabolomics of adherent mammalian cells by capillary electrophoresis-mass spectrometry: HT-29 cells as case study, J. Pharm. Biomed. Anal. 110 (2015) 83-92.

[58] H. Meyer, H. Weidmann, M. Lalk, Methodological approaches to help unravel the intracellular metabolome of Bacillus subtilis, Microb. Cell. Fact. 12 (2013) 69.

[59] A.D. Patterson, H. Li, G.S. Eichler, K.W. Krausz, J.N. Weinstein, A.J. Fornace, F.J. Gonzalez, J.R. Idle, UPLC-ESI-TOFMS-based metabolomics and gene expression dynamics inspector self-organizing metabolomic maps as tools for understanding the cellular response to ionizing radiation, Anal. Chem. 80 (2008) 665-674.

[60] D.Y. Lee, O. Fiehn, High quality metabolomic data for Chlamydomonas reinhardtii, Plant Methods 4 (2008) 7.

[61] J.L. Au, M.H. Su, M.G. Wientjes, Extraction of intracellular nucleosides and nucleotides with acetonitrile, Clin. Chem. 35 (1989) 48-51.

[62] L. von Stechow, A. Ruiz-Aracama, B. van de Water, A. Peijnenburg, E. Danen, A. Lommen, Identification of cisplatin-regulated metabolic pathways in pluripotent stem cells, PloS one 8 (2013) e76476.

[63] O. Fiehn, T. Kind, Metabolite profiling in blood plasma, Method. Mol. Biol. 358 (2007) 3-17.

[64] O. Fiehn, G. Wohlgemuth, M. Scholz, Setup and Annotation of Metabolomic Experiments by Integrating Biological and Mass Spectrometric Metadata. In: Ludäscher B., Raschid L. (eds) Data Integration in the Life Sciences. DILS 2005. Lecture Notes in Computer Science, vol 3615. Springer, Berlin, Heidelberg.

[65] M. Scholz, O. Fiehn, SetupX--a public study design database for metabolomic projects, Pac. Symp. Biocomput. (2007) 169-180.

[66] M. Chagoyen, F. Pazos, Tools for the functional interpretation of metabolomic experiments, Brief Bioinform. 14 (2013) 737-744.

[67] J. Xia, N. Psychogios, N. Young, D.S. Wishart, MetaboAnalyst: a web server for metabolomic data analysis and interpretation, Nucleic Acids Res. 37 (2009) W652-60.

[68] M. Chagoyen, F. Pazos, MBRole: enrichment analysis of metabolomic data, Bioinformatics 27 (2011) 730-731.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

565 **FIGURE CAPTIONS**

566 **Figure 1**. GC-MS total ion chromatograms (TIC) of metabolite extracts obtained using
567 ACN (A), ACN-IsopOH-water (3:3:1, v/v/v) (B), water (C), and 5% formic acid in water
568 (D). GC-MS conditions are described in the text.

569 **Figure 2**. Hierarchical clustering and heat map of total identified metabolites (rows) in all
570 samples (columns). A (ACN), B (ACN-IsopOH-water (3:3:1, v/v/v)), C (water) and D (5%
571 formic acid) extracts are represented in red, green, dark blue and light blue respectively.
572 Cell color represents metabolite relative content.

573 **Figure 3**. Venn diagram representation of number of HT-29 metabolites identified in A, B,
574 C and D extracts.

575 **Figure 4.** Summary plot for Metabolite Set Enrichment Analysis (MSEA) of metabolites
576 identified in the organic extracts (A and B) from HT-29 cells.

577 **Figure 5.** Summary plot for Metabolite Set Enrichment Analysis (MSEA) of metabolites
578 identified in the aqueous extracts (C and D) from HT-29 cells.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

~~**Table 1.** Summary of colorectal cancer (CRC) studies via GC-MS metabolomics published at Web of Science (December 2016) sorted by year of publication.~~

Matrix	Samples	Metabolomic approach	Analytical platform	Reference
Cell culture	Two colon cancer cell lines (SW1116 y SW480) versus a control cell line (NCM460)	Metabolic fingerprinting	GC-TOF MS	[35]
Tissue	Tissue from patients with primary CRC (n=27) and control tissue (n=18)	Metabolic fingerprinting	GC-TOF MS	[37]
Tissue	Tissue from tumour and control area from CRC patients (n=31)	Metabolic fingerprinting	HR-MAS NMR GC-TOF MS	[36]
Serum	CRC patients (n=31) and healthy controls (n=8)	Metabolic fingerprinting	GC-TOF MS	[38]
Serum	CRC patients (n=30) before and after surgery	Metabolic fingerprinting	GC-TOF MS	[39]
Human and rat urine	Human: CRC patients (n=60) and healthy controls (n=63) Rats before and after surgery treated with DMH (n=8) and control rats (n=8)	Metabolic fingerprinting DMH[b] metabonomic analysis	GC-TOF MS	[40]
Serum	CRC patients (n=42) and healthy controls (n=8)	Fatty acids profiling	GC-TOF MS	[41]
Urine	CRC patients (n=61) and healthy controls (n=62) Validation group: CRC patients (n=40) and healthy controls (n=41)	Metabolic fingerprinting	GC-TOF MS UHPLC-Q/TOF MS	[33]
Tissue	Tissue from tumour and control area from CRC patients (n=31)	Metabolic fingerprinting	GC x GC-TOF MS	[42]
Mice tissue, plasma and feces	Samples collected at 3, 5 and 7 weeks after azoxymethane treatment (n=40) and placebo (n=35)	Metabolic fingerprinting	UHPLC-MS/MS GC-TOF MS	[32]
Serum	CRC patients (n= 60) divided in four cáncer stages (12 patient/stage)) and healthy controls (n=60) Validation group: CRC patients (n=59) and healthy controls (n=63)	Metabolic profiling	GC-TOF MS	[15]

Mice tissue and serum Cell culture	Mice tissue (n=131) with and without gen APC mutation Mice serum (n=126) with and without gen APC mutation Cell line SW480 (n=75) with partial/total expression of gen APC	Metabolic fingerprinting	GC-TOF MS	[43]
Tissue	Patients with recurrent (n=4) and non-recurrent (n=3) CRC after treatment with 5-fluorouracil	Nucleotides and derivatives profiling	GC x GC-TOF MS	[44]
Serum	CRC patients (n=101) and healthy controls (n=102)	Metabolic fingerprinting	GC-TOF MS UHPLC-Q/TOF MS	[31]
Cell culture	Caco-2 cells versus other cancer cells (breast, neuroblastoma) and non cancer cell line	Metabolic profiling	GC-TOF MS	[61]
Tissue and feces	Feces and tissue from tumour and control area from CRC patients (n=17)	Metabolic fingerprinting	UHPLC-MS/MS GC-TOF MS	[62]
Plasma	Plasma from CRC patients (n=15) and healthy controls (n=15)	Metabolic fingerprinting	GC-TOF MS	[63]
Serum	Serum from patients suffering adenoma (n=31), various stages of CRC (320) and healthy controls (n=254)	Metabolic fingerprinting	GC-TOF MS	[64]

TABLES

Table ~~2~~1. Summary of metabolite extraction performance of the four extraction solvents used in this work.

Metabolic extract	Extraction solvent	Detected metabolites	Identified metabolites	Metabolites at maximum intensity [a]	Metabolites RSD<15% [a]
A	ACN	141	82	13	17
B	ACN-IsopOH-H_2O (3:3:1, v/v/v)	182	96	36	51
C	H_2O	172	100	25	62
D	5% HCOOH	168	94	76	20

[a]Metabolites at maximum intensity (5th column) and with RSD <15% (6th column) out of the total identified metabolites.

SUPPORTING INFORMATION

Supporting Figures

Figure S1. Principal component analysis (PCA) of the samples represented in the plane defined by the main principal components (PC) from PC1 to PC5. A, B, C and D extracts are represented in red (Δ), green (+), dark blue (X) and light blue (◊) respectively.

Figure S2. Summary plot for metabolite set enrichment analysis (MSEA) of metabolites identified in HT-29 cells. Top 50 metabolic pathways are shown.

Supporting Tables

Table S1. Summary of colorectal cancer (CRC) studies via GC-MS metabolomics published at Web of Science (December 2016) sorted by year of publication.

Matrix	Samples	Metabolomic approach	Analytical platform	Reference
Cell culture	Two colon cancer cell lines (SW1116 y SW480) versus a control cell line (NCM460)	Metabolic fingerprinting	GC-TOF MS	[35]
Tissue	Tissue from patients with primary CRC (n=27) and control tissue (n=18)	Metabolic fingerprinting	GC-TOF MS	[37]
Tissue	Tissue from tumour and control area from CRC patients (n=31)	Metabolic fingerprinting	HR-MAS NMR GC-TOF MS	[36]
Serum	CRC patients (n=31) and healthy controls (n=8)	Metabolic fingerprinting	GC-TOF MS	[38]
Serum	CRC patients (n=30) before and after surgery	Metabolic fingerprinting	GC-TOF MS	[39]
Human and rat urine	Human: CRC patients (n=60) and healthy controls (n=63) Rats before and after surgery treated with DMH (n=8) and control rats (n=8)	Metabolic fingerprinting DMH[b] metabonomic analysis	GC-TOF MS	[40]

Serum	CRC patients (n=42) and healthy controls (n=8)	Fatty acids profiling	GC-TOF MS	[41]
Urine	CRC patients (n=61) and healthy controls (n=62) Validation group: CRC patients (n=40) and healthy controls (n=41)	Metabolic fingerprinting	GC-TOF MS UHPLC-Q/TOF MS	[33]
Tissue	Tissue from tumour and control area from CRC patients (n=31)	Metabolic fingerprinting	GC x GC-TOF MS	[42]
Mice tissue, plasma and feces	Samples collected at 3, 5 and 7 weeks after azoxymethane treatment (n=40) and placebo (n=35)	Metabolic fingerprinting	UHPLC-MS/MS GC-TOF MS	[32]
Serum	CRC patients (n= 60) divided in four cáncer stages (12 patient/stage)) and healthy controls (n=60) Validation group: CRC patients (n=59) and healthy controls (n=63)	Metabolic profiling	GC-TOF MS	[15]
Mice tissue and serum Cell culture	Mice tissue (n=131) with and without gen APC mutation Mice serum (n=126) with and without gen APC mutation Cell line SW480 (n=75) with partial/total expression of gen APC	Metabolic fingerprinting	GC-TOF MS	[43]
Tissue	Patients with recurrent (n=4) and non-recurrent (n=3) CRC after treatment with 5-fluorouracil	Nucleotides and derivatives profiling	GC x GC-TOF MS	[44]
Serum	CRC patients (n=101) and healthy controls (n=102)	Metabolic fingerprinting	GC-TOF MS UHPLC-Q/TOF MS	[31]
Cell culture	Caco-2 cells versus other cancer cells (breast, neuroblastoma) and non cancer cell line	Metabolic profiling	GC-TOF MS	[61]
Tissue and feces	Feces and tissue from tumour and control area from CRC patients (n=17)	Metabolic fingerprinting	UHPLC-MS/MS GC-TOF MS	[62]
Plasma	Plasma from CRC patients (n=15) and healthy controls (n=15)	Metabolic fingerprinting	GC-TOF MS	[63]

Serum	Serum from patients suffering adenoma (n=31), various stages of CRC (320) and healthy controls (n=254)	Metabolic fingerprinting	GC-TOF MS	[64]

Table S2. Identification of metabolites including KEGG and PubChem identifiers sorted by retention index.

Identification	Retention Index	Quantmass	KEGG ID	PubChem ID
lactic acid	217657	191	C01432	19789253
glycolic acid	227636	177	C03547	3698251
alanine	244189	116	C00041	5950
maleimide	245118	154	C07272	10935
butylamine	249493	174	-	8007
hydroxylamine	253158	146	C00192	787
butyrolactam	277199	142	-	12025
methanolphosphate	289520	241	-	13130
methylmalonic acid	311617	147	C02170	487
valine	313502	144	C00183	6287
hydroxycarbamate	325948	278	-	-
pyrophosphate	327517	110	-	-
urea	332913	189	C00086	1176
benzoic acid	339067	179	C00539	20144841
caprylic acid	343457	201	C06423	379
ethanolamine	343800	174	C00189	700
glycerol	344466	205	C00116	753
phosphate	345365	314	-	1061
leucine	346101	158	C00123	6106
2-deoxyerythritol	355045	117	-	-
1-deoxyerythritol	357129	117	-	-
isoleucine	361332	158	C00407	6306
proline	366784	142	C00148	145742
glycine	368707	248	C00037	5257127
succinic acid	370608	247	C00042	1110
glyceric acid	379720	189	-	752
uracil	385735	241	C00106	1174
fumaric acid	390016	245	C00122	21883788

serine	395020	218	C00065	5951
pelargonic acid	399229	117	C01601	8158
threonine	409568	218	C00188	6288
thymine	420133	255	C00178	1135
glutaric acid	421260	261	C00489	23322899
beta-alanine	435564	248	C00099	239
homoserine	443878	218	C00263	12647
capric acid	451790	229	C01571	2969
3-aminoisobutyric acid	452655	248	C05145	25201103
aminomalonate	455754	218	C00872	100714
malic acid	463180	233	C00711	20130941
parabanic acid	464991	100	-	67126
threitol	467595	217	C16884	169019
nicotinamide	469827	179	C00153	936
erythritol	471922	217	C00503	222285
n-acetylglutamate	477776	158	C00624	70914
1,5-anhydroglucitol	478275	101	C07326	64960
aspartic acid	480387	232	C00049	44367445
methionine	483560	176	C00073	6137
trans-4-hydroxyproline	484934	140	C01157	5810
oxoproline	485935	156	C01879	7405
4-aminobutyric acid	488730	304	C00334	119
isothreonic acid	489385	292	-	-
cysteine	500158	220	C00097	5862
creatinine	502599	115	C00791	588
2-hydroxyglutaric acid	506306	247	C02630	43
dodecanol	507619	243	C02277	8193
2-(4-hydroxyphenyl)ethanol	509853	179	C06044	10393
erythronic acid	512029	217	-	2781043
digitoxose	521798	117	-	-
ornithine	527113	142	C00077	6262
glutamic acid	529100	246	C00025	33032
5-aminovaleric acid	536657	174	C00431	138
phenylalanine	537804	218	C00079	6140
4-hydroxyphenylacetic acid	542795	179	C00642	127
xylose	546699	307	-	-
lauric acid	547906	117	C02679	3893
N-acetylaspartic acid	548028	158	C01042	65065
asparagine	553078	188	C00152	6267
ribose	553135	217	C00121	5779

taurine	558857	326	C00245	1123
xylitol	567437	217		6912
levoglucosan	569637	204		2724705
lyxitol	573587	217	C01904	94154
beta-glycerolphosphate	574470	243		126740
fucose	578299	160		
UDP-glucuronic acid	587601	217	C00167	17473
putrescine	588872	174	C00134	1045
glycerol-alpha-phosphate	590747	357	C03189	754
diglycerol	591074	103		42953
glucose-1-phosphate	594647	217		439165
ribonic acid	599680	292		
glutamine	600315	156	C00064	5961
phosphoethanolamine	603912	299	C00346	1015
3-phosphoglycerate	610734	227	C00597	724
citric acid	617342	273	C00158	19782904
hypoxanthine	619128	265	C00262	790
citrulline	621404	157	C00327	6992098
pinitol	622466	260		164619
UDP-N-acetylglucosamine	623789	226	C00043	445675
UDP GlcNAc	627437	226		
dehydroascorbic acid	632888	173		
myristic acid	634414	117	C06424	11005
fructose	639442	307	C02336	439709
mannose	645856	205	C00159	18950
adenine	646534	264	C00147	190
galactose	648756	319	C00984	439357
glucose	650867	319	C00221	64689
lysine	663483	156	C00047	5962
histidine	663790	154	C00135	6274
tyramine	664737	174	C00483	5610
hexuronic acid	667373	160		18845
sorbitol	667922	217	C00794	5780
tyrosine	671252	218	C00082	6057
pentadecanoic acid	674647	117	C16537	13849
1-hexadecanol	679596	299	C00823	2682
pantothenic acid	690887	291	C12276	11306073
xanthine	701688	353	C00385	1188
hexitol	704741	217	C00392	453
palmitoleic acid	706508	117	C08362	445638

palmitic acid	710061	313	C00249	985
cysteine-glycine	715335	220		174707
N-acetylmannosamine	722897	129	C00645	439281
N-acetyl-D-mannosamine	726375	319		
myo-inositol	730022	305		892
heptadecanoic acid	751309	117		10465
octadecanol	755409	327		8221
glutathione	761296	213	C02471	124886
tryptophan	780482	202	C00078	6305
oleic acid	781527	129		
stearic acid	787622	117	C01530	5281
spermidine	792924	144	C00315	1102
fructose-6-phosphate 1	804279	315		
cystine	804619	218	C01420	24798687
glycerol-3-galactoside	805227	204		
hexose-6-phosphate 1	806282	387		
pseudo uridine	813899	217	C02067	15047
glucose-6-phosphate	818275	387	C00092	5958
nonadecanoic acid	822782	117	C16535	12591
inositol-4-monophosphate	846510	315		161368
phosphogluconic acid	847013	299		
oleamide	849710	144	C19670	5283387
arachidic acid	856421	117	C06425	10467
uridine	861508	217	C00299	6029
5-methoxytryptamine	864466	174	C05659	1833
inosine	897184	230	C00294	6021
1-monopalmitin	901749	129		14900
sucrose	915139	271	C00089	5988
adenosine	918039	236	C00212	60961
behenic acid	920648	117	C08281	8215
cellobiose	932179	204	C01971	6255
lactose	935640	191		
1-monoolein	952623	129		5283468
guanosine	954962	324	C00387	6802
1-monostearin	959214	129		24699
5'-deoxy-5'-methylthioadenosine	967036	236	C00170	439176
cytidine-5-monophosphate	974825	169		
uridine-5-monophosphate	977896	169		
inosine 5'-monophosphate	1016811	169	C00130	8582
adenosine-5-monophosphate	1041375	315	C00020	6083

cholesterol	1078536	129	C00187	5997
cholestan-3-ol	1082070	215	_	3240

Table S3. Main metabolic pathways involving metabolites preferentially extracted using the organic based solvents (A and B).

Pathway	Total	Expected	Hits	Raw p	-log(p)	Holm p	FDR	Impact
Fatty acid biosynthesis	49	0.7939	8	6.24E-07	14.287	4.99E-05	4.99E-05	0
Fatty acid metabolism	50	0.8101	3	0.0456	3.0869	1	1	0.0296
Pyrimidine metabolism	60	0.9722	3	0.0714	2.6396	1	1	0.1104
Purine metabolism	92	1.4907	3	0.1851	1.6867	1	1	0.1193
Galactose metabolism	41	0.6643	3	0.0274	3.5964	1	1	0.1174
Glycosylphosphatidylinositol(GPI)-anchor biosynthesis	14	0.2268	1	0.2049	1.5851	1	1	0
Citrate cycle (TCA cycle)	20	0.3241	1	0.2797	1.2742	1	1	0.0633
Sphingolipid metabolism	25	0.4051	1	0.3367	1.0887	1	1	0.0129
Arginine and proline metabolism	77	1.2476	2	0.3566	1.0312	1	1	0.0264
Fatty acid elongation in mitochondria	27	0.4375	1	0.3582	1.0267	1	1	0
Glutathione metabolism	38	0.6157	1	0.4651	0.7656	1	1	0.2374
Nitrogen metabolism	39	0.6319	1	0.4739	0.7468	1	1	0.0606
Glycerophospholipid metabolism	39	0.6319	1	0.4739	0.7468	1	1	0.0713
Inositol phosphate metabolism	39	0.6319	1	0.4739	0.7468	1	1	0.137
Ascorbate and aldarate metabolism	45	0.7291	1	0.5238	0.6466	1	1	0
Primary bile acid biosynthesis	47	0.7615	1	0.5394	0.6173	1	1	0.0552
Lysine degradation	47	0.7615	1	0.5394	0.6173	1	1	0.0651
Glyoxylate and dicarboxylate metabolism	50	0.8101	1	0.5619	0.5765	1	1	0.0033
Cysteine and methionine metabolism	56	0.9074	1	0.6037	0.5047	1	1	0.0074
Amino sugar and nucleotide sugar metabolism	88	1.4258	1	0.7688	0.263	1	1	0.0259
Steroid hormone biosynthesis	99	1.6041	1	0.8082	0.2129	1	1	0.0039

Table S4. Main metabolic pathways involving metabolites preferentially extracted using the aqueous-based solvents (C and D).

Pathway	Total	Expected	Hits	Raw p	-log(p)	Holm p	FDR	Impact

Aminoacyl-tRNA biosynthesis	75	2.7732	14	2.85E-07	15.07	2.28E-05	2.28E-05	0.169
Alanine aspartate and glutamate metabolism	24	0.8874	8	1.16E-06	13.671	9.12E-05	4.62E-05	0.5575
Arginine and proline metabolism	77	2.8471	13	2.65E-06	12.839	0.000207	7.08E-05	0.4744
Nitrogen metabolism	39	1.442	9	7.33E-06	11.824	0.0005641	0.0001465	0.0076
Starch and sucrose metabolism	50	1.8488	8	0.0003822	7.8695	0.02905	0.0056231	0.3363
beta-Alanine metabolism	28	1.0353	6	0.0004217	7.7711	0.03163	0.0056231	0.3232
Pyrimidine metabolism	60	2.2185	8	0.0013446	6.6117	0.099501	0.015367	0.135
Cyanoamino acid metabolism	16	0.5916	4	0.0022609	6.092	0.16505	0.02039	0
Glutathione metabolism	38	1.4051	6	0.0022938	6.0775	0.16516	0.02039	0.0642
Galactose metabolism	41	1.516	6	0.0034173	5.6789	0.24263	0.027338	0.0485
Amino sugar and nucleotide sugar metabolism	88	3.2538	9	0.004433	5.4187	0.31031	0.03224	0.2073
Phenylalanine metabolism	45	1.6639	6	0.0054974	5.2035	0.37932	0.03665	0.1665
Propanoate metabolism	35	1.2941	5	0.0083368	4.7871	0.56691	0.051304	0.0866
Pentose and glucuronate interconversions	53	1.9597	6	0.012229	4.4039	0.81935	0.06988	0.0685
Pantothenate and CoA biosynthesis	27	0.9983	4	0.015981	4.1364	1	0.085232	0.253
Purine metabolism	92	3.4017	8	0.018586	3.9853	1	0.092931	0.1436
Pentose phosphate pathway	32	1.1832	4	0.028469	3.5589	1	0.13397	0.1574
Glycine serine and threonine metabolism	48	1.7748	5	0.030202	3.4998	1	0.13423	0.4927
Cysteine and methionine metabolism	56	2.0706	5	0.053576	2.9267	1	0.21767	0.1344
Butanoate metabolism	40	1.479	4	0.057939	2.8484	1	0.21767	0.0461
Tyrosine metabolism	76	2.8101	6	0.059324	2.8247	1	0.21767	0.1258
D-Glutamine and D-glutamate metabolism	11	0.4067	2	0.05986	2.8157	1	0.21767	0.139
Valine leucine and isoleucine biosynthesis	27	0.9983	3	0.075419	2.5847	1	0.26233	0.0265
Glycolysis or Gluconeogenesis	31	1.1462	3	0.10445	2.2591	1	0.34815	0
Methane metabolism	34	1.2572	3	0.12868	2.0505	1	0.41176	0.0175
Sulfur metabolism	18	0.6656	2	0.14121	1.9575	1	0.43448	0.0378
Citrate cycle (TCA cycle)	20	0.7395	2	0.1674	1.7874	1	0.49599	0.0313
Valine leucine and isoleucine degradation	40	1.479	3	0.18221	1.7026	1	0.5206	0.0223
Thiamine metabolism	24	0.8874	2	0.2218	1.506	1	0.61187	0
Lysine degradation	47	1.7378	3	0.25063	1.3838	1	0.63876	0.2262
D-Arginine and D-ornithine metabolism	8	0.2958	1	0.26056	1.3449	1	0.63876	0
Fructose and mannose metabolism	48	1.7748	3	0.26074	1.3442	1	0.63876	0.07
Phenylalanine tyrosine and tryptophan biosynthesis	27	0.9983	2	0.26349	1.3337	1	0.63876	0.008
Glyoxylate and dicarboxylate metabolism	50	1.8488	3	0.28111	1.269	1	0.66144	0.0398
Lysine biosynthesis	32	1.1832	2	0.33288	1.1	1	0.73486	0.0999
Glycerolipid metabolism	32	1.1832	2	0.33288	1.1	1	0.73486	0.2091
Biotin metabolism	11	0.4067	1	0.33987	1.0792	1	0.73486	0
Inositol phosphate metabolism	39	1.442	2	0.42658	0.85196	1	0.89806	0.012
Nicotinate and nicotinamide	44	1.6269	2	0.48927	0.71483	1	0.97855	0.0383

metabolism								
Histidine metabolism	44	1.6269	2	0.48927	0.71483	1	0.97855	0.1404
Primary bile acid biosynthesis	47	1.7378	2	0.52482	0.64469	1	1	0.0164
Taurine and hypotaurine metabolism	20	0.7395	1	0.53073	0.6335	1	1	0.3309
Caffeine metabolism	21	0.7765	1	0.54823	0.60107	1	1	0.0305
Sphingolipid metabolism	25	0.9244	1	0.61199	0.49104	1	1	0
Pyruvate metabolism	32	1.1832	1	0.70288	0.35257	1	1	0.1376
Ubiquinone and other terpenoid-quinone biosynthesis	36	1.3311	1	0.745	0.29437	1	1	0
Porphyrin and chlorophyll metabolism	104	3.8455	3	0.75021	0.2874	1	1	0
Glycerophospholipid metabolism	39	1.442	1	0.77266	0.25791	1	1	0.0556
Ascorbate and aldarate metabolism	45	1.6639	1	0.81939	0.19919	1	1	0.008
Tryptophan metabolism	79	2.9211	1	0.95152	0.04969	1	1	0

Figure

Click here to download Figure: Figures_GC-MS_HT29_2016-12-16_BN.pptx

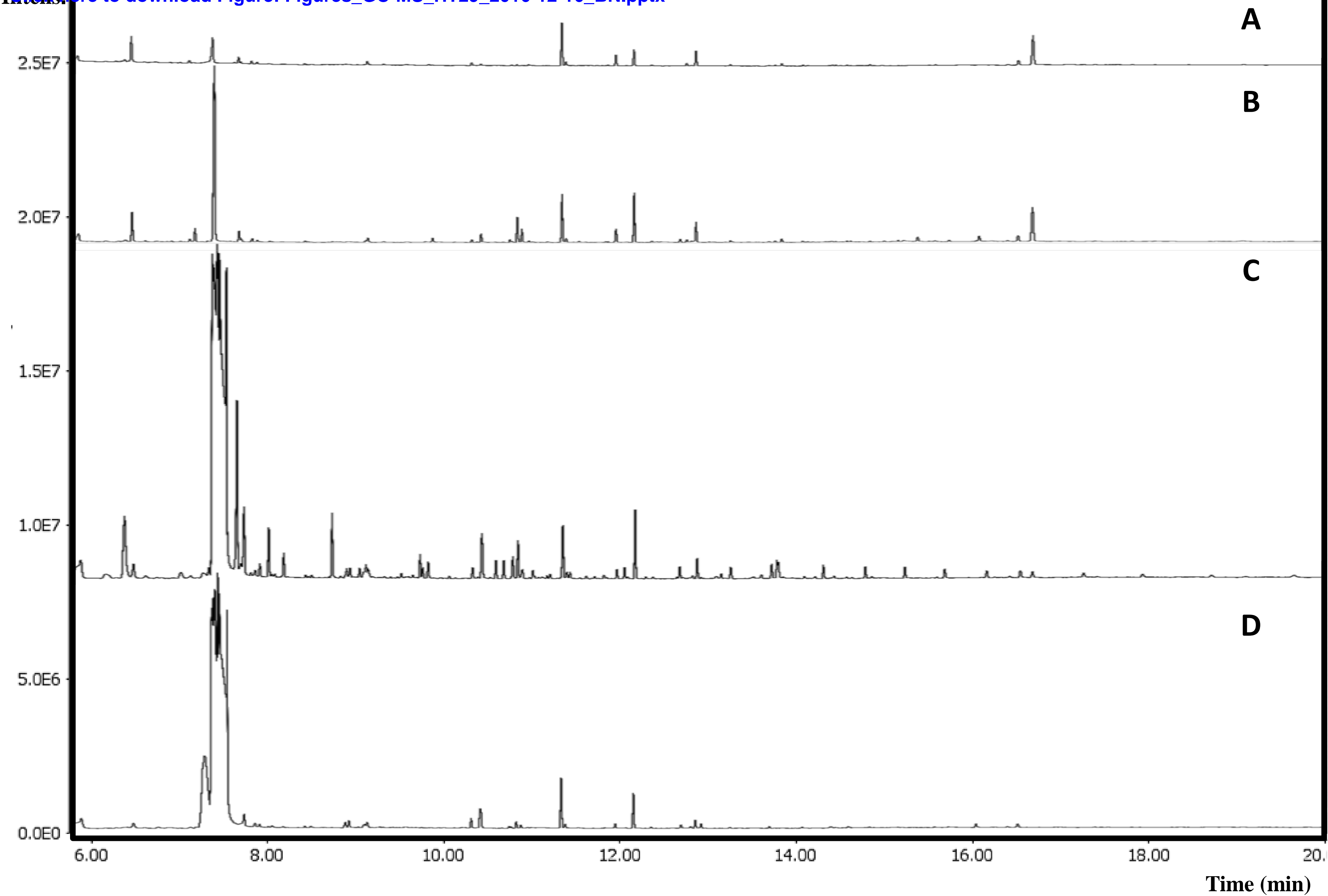

Figure 1.

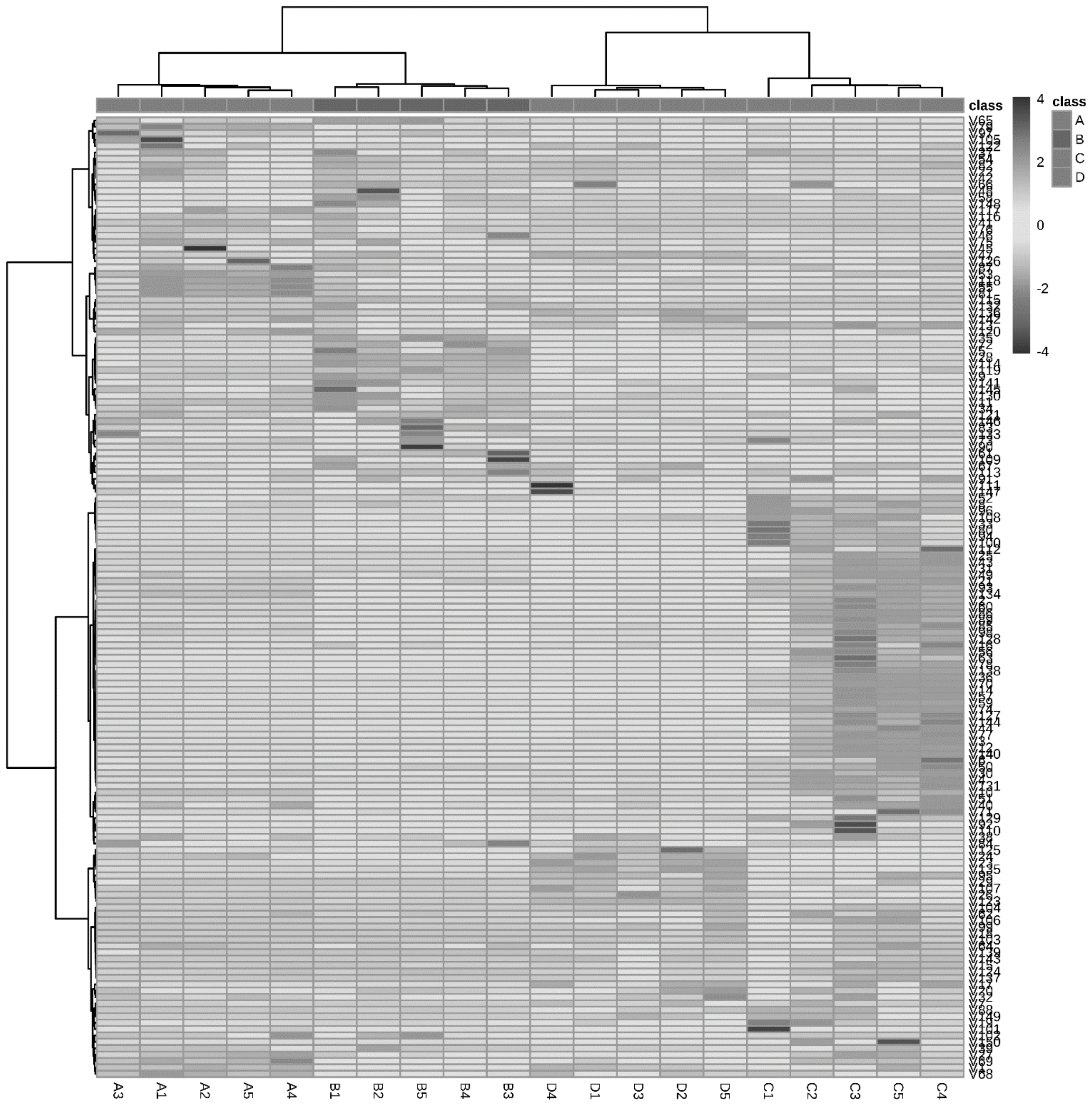

Figure 2.

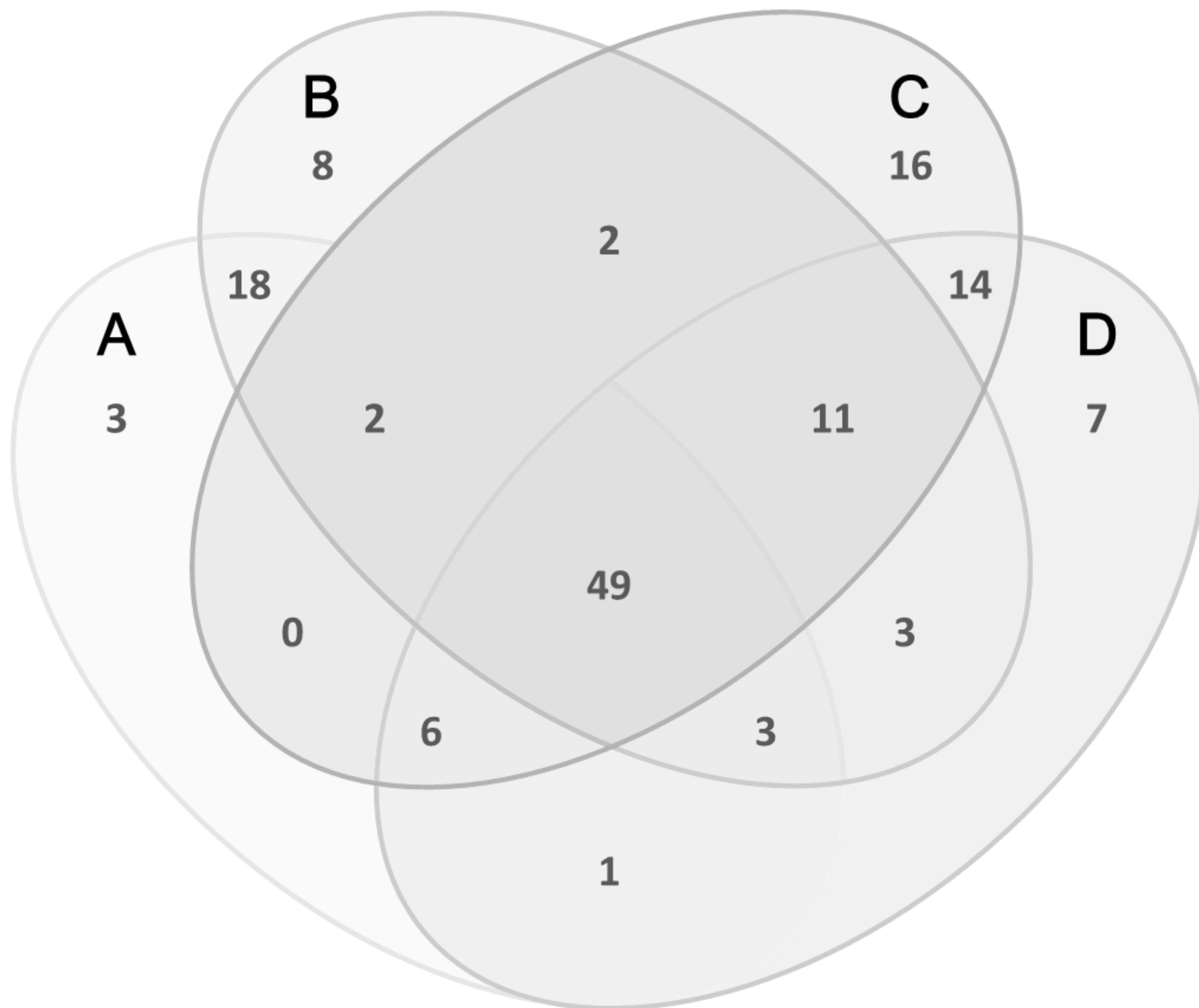

Figure 3.

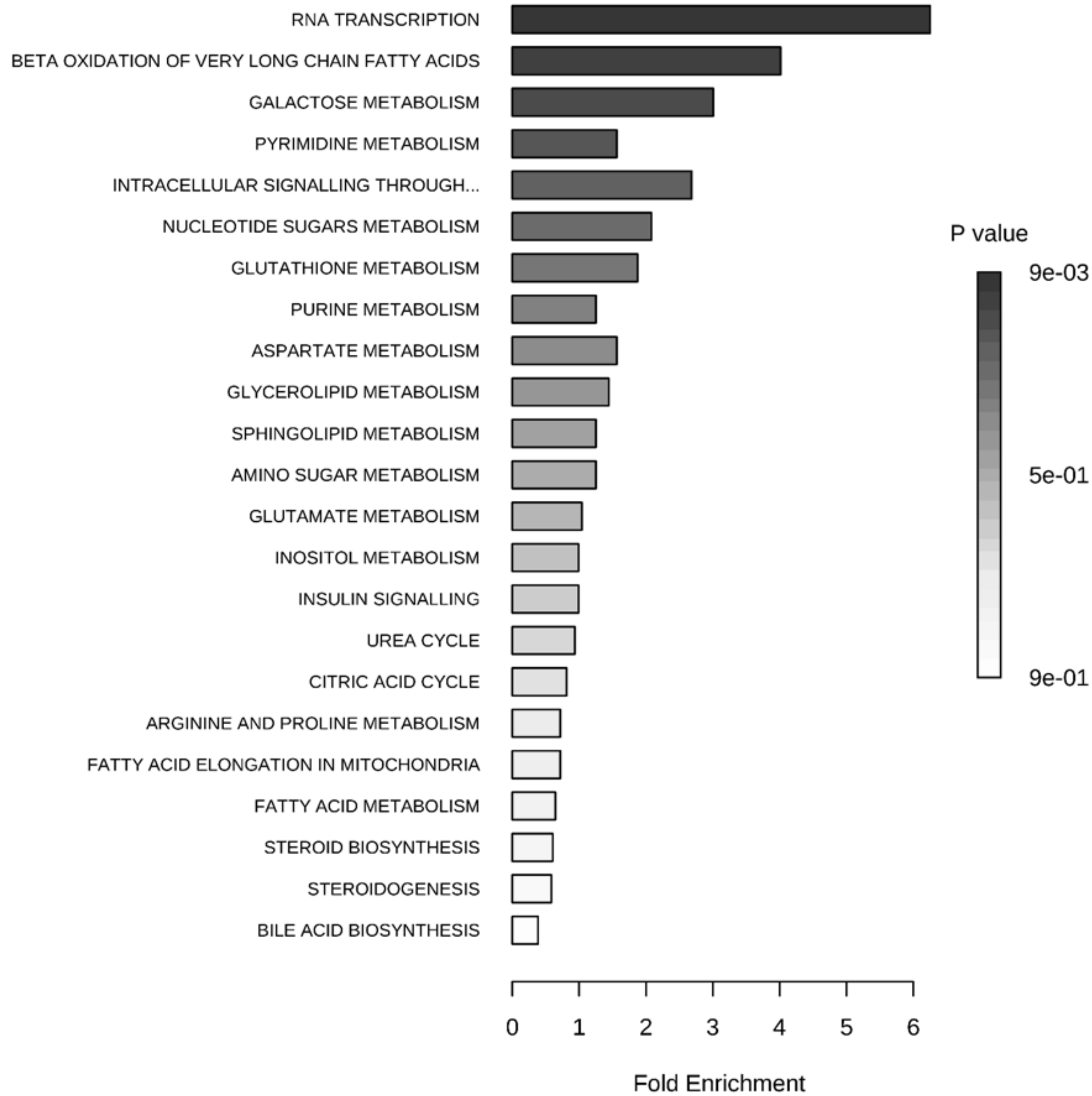

Figure 4.

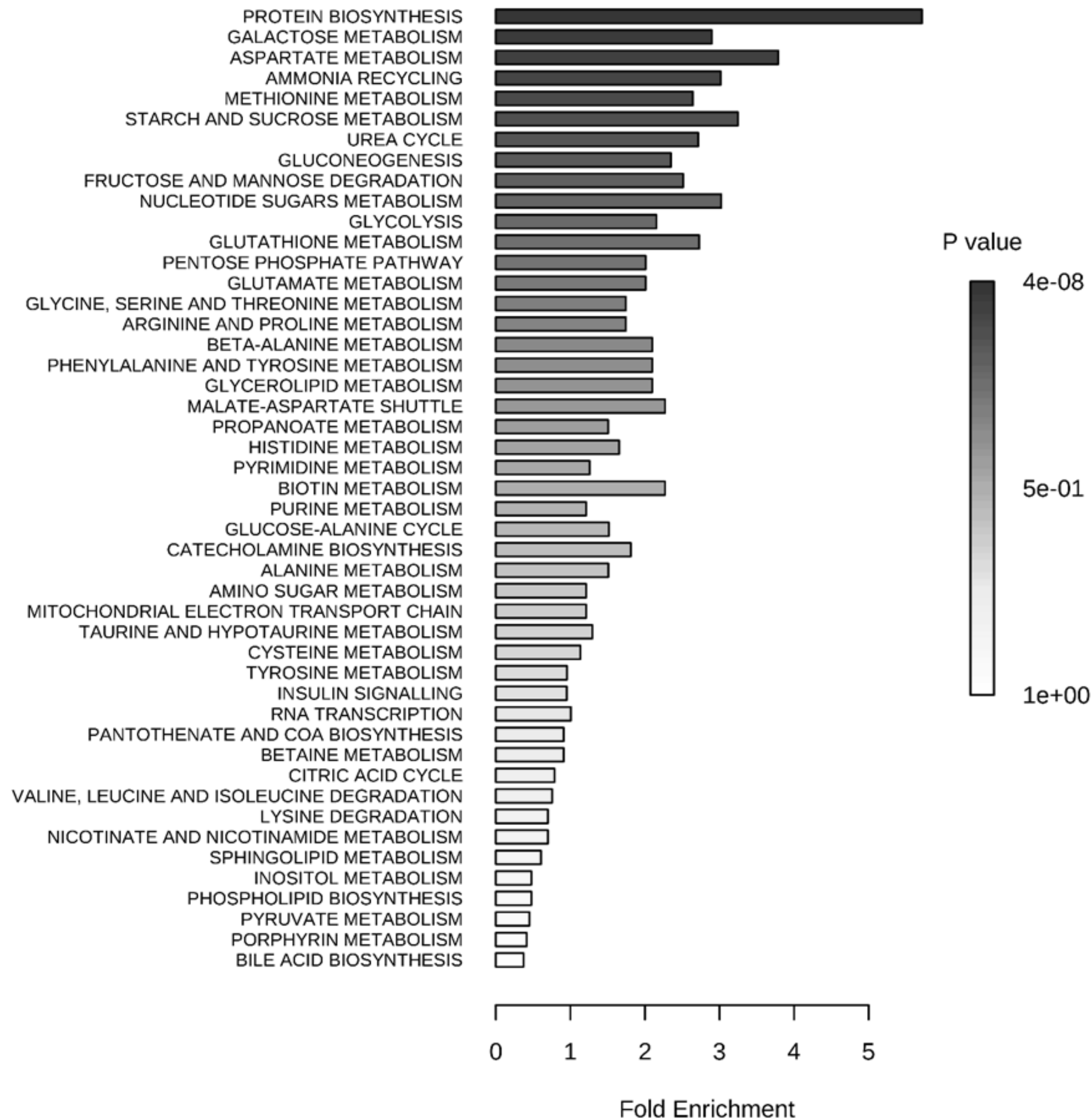

Figure 5.

Electronic Supplementary Material (online publication only)
Click here to download Electronic Supplementary Material (online publication only): Supplementary Tables and Figure legends.d

Electronic Supplementary Material (online publication only)
Click here to download Electronic Supplementary Material (online publication only): Supplementary Figures.pptx

*Revised manuscript (Clean version)
Click here to view linked References

GC-MS BASED METABOLOMICS OF COLON CANCER CELLS USING DIFFERENT EXTRACTION SOLVENTS

Clara Ibáñez[a], Carolina Simó[b], Mine Palazoglu[c], Alejandro Cifuentes[a*]

(a) Laboratory of Foodomics, CIAL (CSIC), Nicolas Cabrera 9, 28049 Madrid, Spain.

(b) Molecular Nutrition and Metabolism, CIAL (CSIC), Nicolas Cabrera 9, 28049 Madrid, Spain.

(c) University of California Davis, Genome Center, 451 E. Health Sci. Dr., Davis, CA 95616, USA.

Author for the correspondence*: Alejandro Cifuentes (a.cifuentes@csic.es). Laboratory of Foodomics, CIAL (CSIC), Nicolas Cabrera 9, 28049 Madrid, Spain. Phone: +34-91-0017955

ABSTRACT

The increasing incidence of colorectal cancer enforces the development of novel methodologies and protocols to deepen in the molecular mechanisms that govern disease pathophysiological events. The aim of this work is to deepen in the optimum metabolite extraction protocol from adherent mammalian cells of colon cancer for high throughput metabolomics using gas chromatography coupled to mass spectrometry (GC-MS). GC-MS results showed that metabolic information obtained from colon cancer cells was highly dependent on metabolite extraction selection, which at the same time is extremely influenced by the analytical platform. A further purpose of this investigation is to uncover an unexplored portion of HT-29 colon cancer cells metabolome, complementary to other already explored by CE-MS and LC-MS methods. At this respect, a total of 150 metabolites were identified in HT-29 colon cancer cells by GC-MS. The extraction protocol with acetonitrile-isopropanol-water was the most appropriate for fatty acids and related pathways analysis. Most of the metabolites involved in pathways of amino acids, glutathione, amino sugars and other polar metabolites were better extracted with acidified water, although water extraction showed the best overall reproducibility. Although pathways involving nitrogenous bases could be investigated using organic or aqueous extracts, a higher number of metabolites involved in these pathways were identified in the aqueous extracts. In addition, metabolite extraction protocol was observed to be crucial for the determination of potentially interesting clusters of metabolites.

Keywords: Colon cancer, Sample treatment, Metabolomics, gas chromatography-mass spectrometry, GC-MS.

Abbreviations: CRC, colorectal cancer; MSEA, metabolite set enrichment analysis; MPA, metabolic pathway analysis; MSTFA, N-methyl-N-(trimethylsilyl)-trifluoroacetamide; ORA, overrepresentation analysis; PBS, phosphate buffered saline; PCA, principal component analysis; RIM, internal retention index markers; TIC, total ion chromatogram

1. INTRODUCTION

In the past few years, cancer biology research has increasingly been focused on metabolism in cancer cells [1]. The complexity and diversity of biological alterations inherently linked to cancer metabolism is in accordance with the increasing unresolved issues to determine specific causes in cancer development, to assess the progression and to unravel molecules or pathways to target in cancer therapy and/or prevention. It is known that cancer cells show different metabolism from healthy cells. One of the well-known metabolic alterations in cancer cells is the glycolysis utilization over the oxidative phosphorylation [2, 3]. In the past years, other perturbations in specific metabolic pathways have been addressed linked to energy metabolism [4] such as pyruvate production [5] and tricarboxylic acid (TCA) cycle [6]. The progress of Metabolomics offers valuable opportunity to better understand biochemical changes produced in cancer metabolism for improving early detection, progression and therapy monitoring of cancer disease [7]. Metabolomics is able to meet the challenge of detecting hundreds of metabolites in short time analysis thanks to the evolution of analytical technologies and software tools for data processing and statistics [8]. With the evolution of metabolomics platforms, rapid and comprehensive analyses of the metabolome complexity can be performed in a high-throughput manner with improved sensitivity, accuracy and resolution [1].Typical metabolomic studies are based on mass spectrometry (MS) or nuclear magnetic resonance (NMR) whose advantages and limitations have been discussed elsewhere [9, 10]. It is now assumed that the coverage of the human metabolome is impossible to achieve with a single analytical methodology and an increasingly common practice is the combination of analytical techniques to achieve complementary information [11]. Metabolomics potential has already been shown with the

detection of novel biomarkers involved in different metabolic pathways related to breast [12], liver [13], prostate [14], colon [15] or lung cancer [16] among the vast amount of investigations. From those publications it can be assumed that different cancer subtypes exhibit different phenotypes and therefore show different metabolic alterations and biomarkers. This diversity enforces the investigation of cancer with special attention to most worldwide harmful cancer types.

Colorectal cancer is the second leading cause of cancer mortality in the USA [17] and Europe [18] with a continuous increasing incidence. Identifying soon the colon cancer apparition and/or progression will increase our knowledge on efficient drug discovery research and prevention. With this aim, in the last decade, Metabolomics has emerged to understand pathophysiological processes related to colon cancer. From the about 30 works related to colon cancer Metabolomics published so far, a minor part have been accomplished using NMR [19-23] with a clear trend toward the use of MS-based analytical platforms either alone [24-26] or in combination with separation techniques such as capillary electrophoresis (CE) [27-30], liquid chromatography (LC) [24, 27-29, 31-34] and gas chromatography (GC) [15, 31-33, 35-44].

Metabolomics investigations of colorectal cancer by GC-MS are summarized in Table S1. As can be seen in Table S1 a variety of samples including biofluids, tissues and cell culture models have been under scrutiny. The use of cell cultures is included in common reference model systems and is considered an invaluable biomedical research tool. However, the application of cell culture in metabolomics requires further development and standardization of study design steps, metabolism quenching method selection and optimization of metabolite extraction protocols, among others [45, 46]. Namely, quenching

1
2
3
4 89 step at the time of harvesting aims to prevent metabolic content to be altered before the
5
6
7 90 analysis so that the metabolic state of the cell is preserved. There is a great controversy and
8
9 91 continuous research in the determination of the optimum quenching method. Until date the
10
11
12 92 most widely spread techniques to quench the metabolism in cells is a shock maintaining the
13
14 93 cells below -20 ºC [47, 48], the addition of cold solvents such as cold methanol [49], use of
15
16
17 94 cold isotonic PBS [49-51] or the combination in a single step of quenching and extraction
18
19 95 procedures in mammalian adherent cells [45, 52].
20
21
22
23 96 As can be observed in Table S1, three cell lines (namely SW, HT and Caco-2) have been
24
25 97 used to investigate colon cancer following metabolomics approaches using GC-MS. All
26
27
28 98 these cell lines are very well established for in-vitro studies of colorectal cancer and vary in
29
30 99 phenotype, growth rate, differentiation and morphology among other characteristics [53,
31
32
33 100 54]. Metabolic fingerprinting by GC- MS from two SW cell lines (SW-1116 and SW-480)
34
35 101 was first reported by Zimmermann et al. [35]. More recently LC-MS and CE-MS based
36
37
38 102 metabolomics have been used to investigate HT-29 cell line revealing significant metabolic
39
40 103 information in colon cancer metabolism after treatment with different polyphenol-rich
41
42 104 extracts [28, 29, 55].
43
44
45
46 105 Considering that no single method is appropriate for the determination of all intracellular
47
48
49 106 metabolites, metabolic information obtained from cell cultures will depend on multiple
50
51 107 factors and metabolite extraction selection is highly influenced by the analytical platform
52
53 108 [56]. Optimization and selection of the protocol for metabolite extraction from HT-29 cells
54
55
56 109 [57] and evaluation of the subsequent cytosol sample treatment [30] for CE-MS
57
58 110 metabolomics have been previously carried out. CE-MS is particularly suited for the rapid
59
60
61 111 separation of ionic, weakly ionic, and/or highly polar metabolites and aqueous solvent was
62
63
64
65

the preferred for metabolite extraction of HT-29 cells [57]. On the contrary, GC-MS is preferred for the analysis of less polar, volatile (or amenable to chemical derivatization) compounds and requires thermal stability of the analytes. Considering that no single analytical platform is able to determine all intracellular metabolites, in this work, a GC-MS method has been developed due to the high complementarity nature of GC-MS used in this work and the CE-MS [57] and LC-MS [28, 29] methods already developed by our group. Further, taking into account all the above considerations, in this work GC-MS has been used to evaluate the coverage and reproducibility of four metabolic extraction solvents selected according to the physicochemical diversity of metabolites including the large variation on solubility (in aqueous or organic solvents) and based on previous published works [30, 57-63]. To our knowledge this is the first time that HT-29 colon cancer cells are examined using GC-MS metabolomics which could be of special importance in future studies to detect biochemical alterations due to colon cancer or to explore new preventive interventions.

2. MATERIALS AND METHODS

2.1. Chemicals

All reagents were of analytical grade. Phosphate buffered saline (PBS) was purchased from Lonza (Barcelona, Spain). For cell counting, trypan blue was purchased from Sigma Aldrich (St. Louis, MO, USA). Metabolite extraction solvents were of MS grade: formic acid and 2-propanol were from Riedel-de Haën (Seelze, Germany) while water and acetonitrile were from Labscan (Gliwice, Poland). Reagents for derivatization included methoxyamine hydrochloride from Sigma Aldrich, pyridine (silylation grade) and N-

methyl-N-(trimethylsilyl)-trifluoroacetamide (MSTFA) purchased in 1 mL bottles from Pierce (Rockford IL USA).

For GC-MS analysis, a mixture of internal retention index markers (RIM) was prepared using fatty acid methyl esters of C8, C9, C10, C12, C14, C16, C18, C20, C22, C24, C26, C28 and C30 linear chain length, dissolved in chloroform (from Mallinckrodt Baker Inc., Phillipsburg. NJ, USA) at a concentration of 0.8 mg/mL (C8-C16) and 0.4 mg/mL (C18-C30). Fatty acid methyl esters were from Sigma Aldrich.FC43 (Perfluorotributylamine) was acquired from Leco (Stockport, Cheshire, UK) for MS mass calibration.

2.2. Cell culture

The human HT-29 colorectal carcinoma cell line was purchased from the American Type Culture Collection (ATCC- LGC, Promochem, UK). HT-29 cells were cultured in McCoy's 5A medium supplemented with 10% heat-inactivated fetal calf serum, 50 U/mL penicillin G, and 50 U/mL streptomycin, at 37 ºC in humidified atmosphere with 5% CO_2. Once cells reached about 50% confluency, they were trypsinized, neutralized with culture medium and plated at 10000 cells/cm^2 density in different P150 culture dishes. Cell cultures were allowed to adhere overnight at 37 ºC and then plates were pooled. Next, cell count was performed using the trypan blue exclusion test to examine the viability and to estimate the number of cells. For that purpose, 5 µL cell suspension pool were further diluted 1:1 in trypan blue stock solution (0.4%, w/v in sterile PBS) and counted in a Neubauer counting chamber using a light microscope (ID3, Carl Zeiss, Jena, Germany). The estimated volume to have 10 million cells from the culture suspension was added to 20 aliquots. After aliquots were centrifuged (300 xg, 10 min) and supernatant was discarded, pellets containing 10 x 10^6 cells were subjected to metabolite extraction.

2.3. Metabolite extraction and derivatization

The 20 aliquots were divided in four solvent extraction groups: A, ACN; B, ACN-IsopOH-Water (3:3:1, v/v/v); C, Water; and D, 5% formic acid in water. Five independent metabolite extractions (with10 million cells each) were performed for each extraction solvent group. Metabolic content extraction from HT-29 cells have been previously optimized in our laboratory [57] and applied in this work with slight modifications. Briefly, 300 µL of extraction solvent and 0.3 g glass beads (212-300 µm) from Sigma-Aldrich were added to 10 x 10^6 cell pellets. Then metabolism quenching and cell disruption was accomplished by using three cycles of snap-freeze (liquid N2, 1 min), thawing in ultrasound bath (50 Hz, 3 min) (ultrasonic cleaning bath Ultrasons from JP Selecta, Barcelona, Spain) and grinding using a mixer mill for 3 min at 30/s (MM 400, Retsch GmbH, Germany). After this, cell insoluble material and glass beads were removed by centrifugation (24000 xg for 10 min at 4ºC) and the supernatant was collected, and 50 µL metabolic extracts were vacuum-dried and kept at -80ºC until derivatization. The supernatants obtained in A and B extraction groups are protein-free considering organic extraction solvents lead to protein precipitation. On the contrary, supernatants obtained from aqueous-based extractions (C and D extraction groups) were submitted to ultrafiltration with 3 kDa centrifugal filters (Amicon Ultra 0.5 mL, Millipore,Billerica, MA, USA) to remove proteins according to the manufacturer's protocol. 50 µL of both metabolic extracts were vacuum-dried and kept at -80ºC until derivatization.

The best sensitivity was achieved with the following optimized derivatization protocol. First, the dried extracts were dissolved in 10 µL of a solution consisting of 40 mg/mL methoxyamine in pyridine and shaken for 15 min at 60 ºC to protect aldehyde and ketone

groups. Then 50 µL of MSTFA with 1% RIM was added for trimethylsilylation of acidic protons and shaken at 37 °C for 30 min. Derivatized samples were immediately transferred to 2 mL clear glass autosampler vials with microinserts (Agilent, Santa Clara CA), closed by 11 mm T/S/T crimp caps (MicroLiter, Suwanee GA) and analyzed by GC-MS.

2.4. GC-TOF MS analysis

An Agilent 6890 gas chromatograph (Santa Clara CA) and a Leco Pegasus IV time of flight mass (TOF) spectrometer were controlled by the Leco ChromaTOF software vs. 2.32 (St. Joseph, MI). A 30 m long, 0.25 mm ID Rtx-5Sil MS column with 0.25 µm 95% dimethyl 5% diphenyl polysiloxane film and additional 10 m integrated guard column was used (Restek, Bellefonte PA). High purity helium with built-in purifier (Airgas, Radnor PA) was set at constant flow of 1 mL/min. The oven temperature was held constant at 50°C for 1 min and then ramped at 20°C/min to 330°C at which it was held constant for 5 min. The transfer line temperature between gas chromatograph and mass spectrometer was set to 280°C. A Gerstel automatic liner exchange system with multipurpose sample MPS2 dual rail and two derivatization stations was used in conjunction with a Gerstel CIS cold injection system (Gerstel, Muehlheim, Germany). The injector operated in splitless mode, opening the split vent after 25 s. Electron impact ionization at 70V was applied maintaining ion source temperature at 250°C. For every 10 samples, a fresh multibaffled liner was inserted (Gerstel) using the Maestro1 Gerstel software vs. 1.1.4.18 (Gerstel). Before and after each injection, the 10 µL injection syringe was washed three times with 10 µL ethyl acetate. 1 µL sample was filled using 39 mm vial penetration at 1 µL/s fill speed, injecting 0.5 µl at 10 µL/s injection speed. GC oven temperature was maintained at 50°C during one

minute followed by a temperature ramp at 20°C/min until 330°C at which it was held constant for 5 min.

TOF-MS analysis in the positive ion mode was accomplished applying 1775 V detector voltage. Data was recorded in the 85-500 m/z range for 1200 s after 290 s solvent delay. Mass resolving power was 600 and acquisition rate was set at 17 spectra/s. Daily quality controls were used to assure quality of data. The instrument performed autotuning for mass calibration using FC43 before each sequence analysis using the following m/z values: 93, 95,100, 101, 112, 113, 114, 119, 131, 132, 145, 150, 164, 169, 176, 181, 214, 219, 220, 226, 264, 265, 314, 326, 352, 376, 414, 415, 426, 448, 464, 502 and 503.

2.5. Data mining

GC-TOF MS data were acquired as ChromaTOF specific .peg files, and stored also in .txt and .cdf generic formats. Data mining was performed as already described [60]. Briefly, absolute spectra intensities from all sample files were automatically processed and filtered by BinBase algorithm using previously optimized parameters [64]. Then post-processing module to replace missing values was carried out. The resulting report contained row identifier, mass spectra ions, quantification by means of unique ion height, retention index and identification using defined databases [65]. Additional mass signal filtering was manually performed to remove not identified and low confident ions from the table. Then MS signals belonging to the same ion were grouped and data was normalized based on total ion height for each sample. The table with high-confident metabolite information from all extraction replicates obtained with the four solvents (five replicates per extraction solvent) was then examined in detail. Principal component analysis (PCA) and hierarchical

clustering was performed to explore possible clustering, similarities and differences in metabolic coverage obtained among the four extraction protocols. Then comparison of metabolome coverage among the four extracts was performed by metabolite set enrichment analysis (MSEA) and metabolic pathway analysis (MPA). MSEA was performed by means of overrepresentation analysis (ORA). ORA assess whether a set of metabolites is enriched in a particular metabolic pathway compared to the total of annotated metabolites for that pathway (i.e. if the metabolites identified in the study represent a high percentage of the metabolites theoretically annotated in that pathway) [66]. Statistical analysis, MSEA and MPA were performed using metaboanalyst [67]. MBRole [68] was used to assist in the biological interpretation.

3. RESULTS AND DISCUSSION

3.1. Metabolic extraction assessment

Typical total ion chromatograms (TIC) for the cell extracts obtained using the four different extraction solvents tested in this work are represented in Fig. 1. As can be observed, the four metabolite profiles were very different in terms of intensity and distribution of the peaks along the chromatogram. Although aqueous extracts (Fig. 1C and 1D) presented greater overall intensity, peaks were mainly placed at the beginning of the chromatogram.

After data processing, 364 metabolic signals were detected. Signal grouping and filtering processes (i.e. unidentified metabolites and metabolites not found in at least 3 out of 5 replicates of each extraction solvent) rendered a total of 290 metabolites from which a group of 150 metabolites could be identified (see Table 1 and Table S2). Then, each of the metabolic extracts were investigated in detail in terms of detected metabolites and

identified metabolites found in each extract as well as other parameters to evaluate metabolite extraction performance and reproducibility. An average variation of the internal standards of 15% RSD was observed. Given that value the number of identified metabolites with RSD < 15% was also determined for the four extraction solvents.

Among the two aqueous based extractions C (water) and D (acidified water), C showed higher overall metabolome coverage in terms of number of detected (172 vs. 168) and identified metabolites (100 vs. 94). However, deepening in the identified metabolites in both aqueous extracts, D showed a vast gain over C with respect to metabolite recovery performance (76 vs. 25 metabolites found at maximum intensity in D and C respectively) but the inverse was observed in terms of reproducibility (RSD <15% in 62 vs. 20 metabolites in C and D, respectively) (see Table 1). As can be observed in Table 1, a more evident enhancement was observed in B metabolic coverage when compared with A extract among the two organic extracts. Namely, 96 metabolites were identified out of the 182 detected in B, while for A only 82 metabolites could be identified out of the 141 detected. Moreover, triple metabolite recovery performance (i.e. metabolites at maximum intensity) from cell culture was observed for B with respect to A. In addition, metabolic content in A was extracted with less reproducibility among the 5 extraction replicates analysed, including only 17 identified metabolites with RSD values lower than 15%. Thus, among the organic solvents, B extract clearly offered the best overall performance when compared to A.

3.2. Colon cancer cell metabolome identified by GC-MS

Metabolic information obtained from the four extraction protocols was then statistically investigated in order to find patterns and possible clustering in metabolic coverage obtained applying the four extraction solvents. Fig. 2 represents the unsupervised hierarchical clustering of the samples (i.e. metabolic extracts) according to the relative content of each metabolite in each sample (represented by the colour of the cell). As can be observed from the clustering information of Fig. 2, the four extracts were mainly differentiated according to the polarity of the solvent used in the extraction. Thus, as expected two clear clusters were observed corresponding to A with B (i.e. organic based solvents), and C with D (i.e. aqueous based solvents). This information was also observed and corroborated by means of PCA (see Fig. S1) and Venn diagram (see Fig. 3). In addition, PCA also showed that the metabolic information obtained from both organic solvents was more similar (presented less variation in their distribution in the plane defined by the main principal components from PC1 to PC5) than the one more heterogeneous obtained from both aqueous based extracts (more dispersion of aqueous samples in the plane defined by the PCs) (see Fig. S1).

Venn representation (Fig. 3) corroborated that the highest similarity of metabolic coverage was according to the extraction solvent polarity. Namely, organic based extracts (A and B) shared 72 metabolites, among which, 18 metabolites were exclusively detected in A and B. Similarly 80 metabolic species were common in both aqueous based extracts (C and D) from which 14 were only found in C and D. In addition, it could be observed that ACN and water extraction solvents (A and C, respectively) lead to the most different metabolic

extracts with only 57 metabolites in common (from which 49 were found in all the four extracts) showing high complementarity.

To deepen in the coverage of the different extracts according to the nature of the extracted metabolites and pathways involved, two web servers were used. Information about the biological role and nature of the identified metabolites was derived from MBRole web server analysis. Thus, the total set of metabolites with available KEGG identifiers in the four extracts (107 out of 150 identified metabolites) were submitted to analysis. Most important groups of metabolites identified were amino acids and peptides followed by fatty acids and lipids, nucleic acids and carbohydrates. Other groups of metabolites were represented to a lesser extent such as vitamins and cofactors (data not shown). Main metabolic pathways involving metabolites preferentially extracted (i.e. showing maximum levels) using the organic-based (A and B) and aqueous-based (C and D) protocols were analyzed using Metaboanalyst (Table S3 and S4 respectively). Fatty acid and nitrogenous bases (i.e. pyrimidine and purine) related pathways are well represented in organic based extracts. Among the two organic extracts, contribution of B extract is clearly superior to A whose metabolites are almost exclusively involved in nucleotide sugars metabolism. The main reason for this inequality could be explained for the lowest number of metabolites at highest intensity in A (see Table 1). On the other hand, pathways involving biosynthesis and metabolism of amino acids, glutathione, amino sugars and nitrogenous bases among others, are highly represented in aqueous extracts (C and D). Most of the metabolites involved in those pathways present higher performance of extraction using acidified water extraction (D) when compared with C (water extraction). Although in lower relative concentration, metabolic species in C showed the best reproducibility.

Finally, identified metabolites were submitted to enrichment analysis (MSEA) by means of overrepresentation analysis (ORA) to determine the more represented metabolic pathways in the defined set of metabolites. As can be observed in Fig. S2, "protein biosynthesis" is the metabolic pathway in which a higher number of the metabolites are involved. However, more useful information can be obtained if solvents are subdivided regarding polarity. When each protocol for metabolic extraction is investigated in detail, different trends are observed. MSEA results of metabolites preferentially extracted (i.e. showing maximum levels) in organic and aqueous based extracts are shown in Fig. 4 and Fig. 5 respectively. As can be observed depending on the nature and polarity of the solvents used in the extraction process diverse metabolic pathways are more suitable to study due to the very different metabolic information. Some of the most relevant metabolic pathways that can be studied with more efficiency in organic based solvents are "RNA transcription" and "beta oxidation of very long fatty acids" (Fig. 4), while pathways most suitable to be evaluated in aqueous based extracts include protein, sugars and amino acids biosynthesis and metabolism (Fig. 5). Interestingly, pathways involving metabolism of glutathione and nucleotides can be approached using both organic and/or aqueous extracts.

CONCLUDING REMARKS

To date it is not possible to study the whole metabolome of a biological system and this is partially due to the selection of a certain sample treatment and analytical platform. In this work, four reliable sample preparation protocols have been evaluated to extract metabolites from HT-29 colon cancer cells and to uncover a portion of this cell line metabolome by GC-MS.

The unsupervised statistical analysis showed that the four extracts evaluated were mainly differentiated according to the polarity of the solvent used in the extraction process and that the overall GC-MS variability of aqueous extracts was higher compared to the organic ones.

In summary, the protocol involving ACN-Isopropanol-water (3:3:1, v/v/v) was shown to be the most appropriate extraction procedure for fatty acid and/or related pathways analysis among the four metabolic extraction procedures. Metabolic extraction with ACN yielded the lowest reproducibility of extraction together with the smallest metabolite coverage. On the other hand, metabolic coverage was similar in both aqueous extracts. However, overall performance of extraction of individual metabolites seemed to be higher using acidified water (5% formic acid) compared to just water, which on the contrary showed better GC-MS reproducibility. Most of the metabolites involved in pathways of amino acids, glutathione, amino sugars and other polar metabolites present higher performance of extraction in the acidified water extract, although metabolic species in water extract show the best overall reproducibility. Although pathways involving nitrogenous bases could be investigated using organic or aqueous extracts, more metabolites involved in those pathways can be identified in the aqueous extracts.

In light of the results obtained, qualitative information about the metabolic coverage and related pathways exhibited high complementarity among the extracts. This information is useful for future hypothesis-driven studies for the evaluation of a certain metabolic pathway or a group of metabolites (i.e. if they are thought to be altered in HT29 colorectal cells by the action of a drug or a natural bioactive compound).

Conflict of interest statement

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

This work was supported the projects AGL2011-29857-C03-01 (Ministerio de Educación y Ciencia, Spain) and AGL2014-53609- P (Ministerio de Economía y Competitividad, Spain). CI thanks to Ministerio de Economía y Competitividad for her stay abroad grant (EEBB-I-12-03995) and for her Juan de la Cierva postdoctoral research contract (FJCI-2014-19601). Authors thank Prof. Oliver Fiehn (UC Davis, CA, USA) for his help with this work. Current address of CI: IMDEA Food Institute, Ctra. Cantoblanco 4, 28049 Madrid, Spain.

References

[1] R.A. Cairns, I.S. Harris, T.W. Mak, Regulation of cancer cell metabolism, Nat. Rev. Cancer 11 (2011) 85-95.

[2] O. Warburg, On the origin of cancer cells, Science 123 (1956) 309-314.

[3] M.D. Williams, R. Reeves, L.S. Resar, H.H. Hill Jr, Metabolomics of colorectal cancer: past and current analytical platforms, Anal. Bioanal. Chem. 405 (2013) 5013-5030.

[4] D. Hanahan, R.A. Weinberg, Hallmarks of cancer: the next generation, Cell 144 (2011) 646-674.

[5] M.G. Vander Heiden, J.W. Locasale, K.D. Swanson, H. Sharfi, G.J. Heffron, D. Amador-Noguez, H.R. Christofk, G. Wagner, J.D. Rabinowitz, J.M. Asara, L.C. Cantley, Evidence for an alternative glycolytic pathway in rapidly proliferating cells, Science 329 (2010) 1492-1499.

[6] H. Kondoh, M.E. Lleonart, J. Gil, J. Wang, P. Degan, G. Peters, D. Martinez, A. Carnero, D. Beach, Glycolytic enzymes can modulate cellular life span, Cancer Res. 65 (2005) 177-185.

[7] N.J. Serkova, K. Glunde, Metabolomics of cancer, Methods Mol. Biol. 520 (2009) 273-295.

[8] R.D. Beger, A review of applications of metabolomics in cancer, Metabolites 3 (2013) 552-574.

[9] K. Dettmer, P.A. Aronov, B.D. Hammock, Mass spectrometry-based metabolomics, Mass Spectrom. Rev. 26 (2007) 51-78.

[10] A. Smolinska, L. Blanchet, L.M. Buydens, S.S. Wijmenga, NMR and pattern recognition methods in metabolomics: from data acquisition to biomarker discovery: a review, Anal. Chim. Acta 750 (2012) 82-97.

[11] W.B. Dunn, D.I. Ellis, Metabolomics: current analytical platforms and methodologies, TrAC-Trend. Anal. Chem. 24 (2005) 285-294.

[12] C. Denkert, E. Bucher, M. Hilvo, R. Salek, M. Orešič, J. Griffin, S. Brockmöller, F. Klauschen, S. Loibl, D.K. Barupal, Metabolomics of human breast cancer: new approaches for tumor typing and biomarker discovery, Genome Med. 4 (2012) 37.

[13] Q. Huang, Y. Tan, P. Yin, G. Ye, P. Gao, X. Lu, H. Wang, G. Xu, Metabolic characterization of hepatocellular carcinoma using nontargeted tissue metabolomics, Cancer Res. 73 (2013) 4992-5002.

[14] T. Zhang, D.G. Watson, L. Wang, M. Abbas, L. Murdoch, L. Bashford, I. Ahmad, N. Lam, A.C. Ng, H.Y. Leung, Application of holistic liquid chromatography-high resolution mass spectrometry based urinary metabolomics for prostate cancer detection and biomarker discovery, PloS one 8 (2013) e65880.

[15] S. Nishiumi, T. Kobayashi, A. Ikeda, T. Yoshie, M. Kibi, Y. Izumi, T. Okuno, N. Hayashi, S. Kawano, T. Takenawa, A novel serum metabolomics-based diagnostic approach for colorectal cancer, PloS one 7 (2012) e40459.

[16] S. Hori, S. Nishiumi, K. Kobayashi, M. Shinohara, Y. Hatakeyama, Y. Kotani, N. Hatano, Y. Maniwa, W. Nishio, T. Bamba, A metabolomic approach to lung cancer, Lung Cancer 74 (2011) 284-292.

[17] American Cancer Society, Colorectal Cancer Facts & Figures 2011-2013, Atlanta: American Cancer Society (2011).

[18] F. Bray, The burden of cancer in Europe, Responding to the challenge of cancer in Europe.. in: M.P. Coleman, D.M. Alexe, T. Albreht, M. McKee (Eds.) Responding to the challenge of cancer in Europe. Institute of Public Health of the Republic of Slovenia, Ljubljana, 2008, pp. 7-40.

[19] D. Monleon, J.M. Morales, A. Barrasa, J.A. Lopez, C. Vazquez, B. Celda, Metabolite profiling of fecal water extracts from human colorectal cancer, NMR Biomed. 22 (2009) 342-348.

[20] M. Tessem, K.M. Selnæs, W. Sjursen, G. Tranø, G.F. Giskeødegård, T.F. Bathen, I.S. Gribbestad, E. Hofsli, Discrimination of patients with microsatellite instability colon cancer using 1H HR MAS MR spectroscopy and chemometric analysis, J. Proteome Res. 9 (2010) 3664-3670.

[21] H. Wang, D.E. Schiller, V. Tso, C. Slupsky, C.K. Wong, R.N. Fedorak, A novel highly sensitive test for detecting colon cancer using spot urine metabolomics, Gastroenterology 140 (2011) S-40.

[22] I. Bertini, S. Cacciatore, B.V. Jensen, J.V. Schou, J.S. Johansen, M. Kruhoffer, C. Luchinat, D.L. Nielsen, P. Turano, Metabolomic NMR fingerprinting to identify and predict survival of patients with metastatic colorectal cancer, Cancer Res. 72 (2012) 356-364.

[23] B. Jiménez, R. Mirnezami, J. Kinross, O. Cloarec, H.C. Keun, E. Holmes, R.D. Goldin, P. Ziprin, A. Darzi, J.K. Nicholson, 1H HR-MAS NMR spectroscopy of tumor-induced local metabolic “field-effects” enables colorectal cancer staging and prognostication, J. Proteome Res. 12 (2013) 959-968.

[24] S.A. Ritchie, P.W. Ahiahonu, D. Jayasinghe, D. Heath, J. Liu, Y. Lu, W. Jin, A. Kavianpour, Y. Yamazaki, A.M. Khan, Reduced levels of hydroxylated, polyunsaturated ultra long-chain fatty acids in the serum of colorectal cancer patients: implications for early screening and detection, BMC Med. 8 (2010) 1.

[25] F. Li, X. Qin, H. Chen, L. Qiu, Y. Guo, H. Liu, G. Chen, G. Song, X. Wang, F. Li, Lipid profiling for early diagnosis and progression of colorectal cancer using direct-infusion electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry, Rapid Commun. Mass Sp. 27 (2013) 24-34.

[26] E.G. Armitage, H.L. Kotze, J.S. Fletcher, A. Henderson, K.J. Williams, N.P. Lockyer, J.C. Vickerman, Time-of-flight SIMS as a novel approach to unlocking the hypoxic properties of cancer, Surf. Interface Anal. 45 (2013) 282-285.

[27] A. Hirayama, K. Kami, M. Sugimoto, M. Sugawara, N. Toki, H. Onozuka, T. Kinoshita, N. Saito, A. Ochiai, M. Tomita, H. Esumi, T. Soga, Quantitative metabolome profiling of colon and stomach cancer microenvironment by capillary electrophoresis time-of-flight mass spectrometry, Cancer Res. 69 (2009) 4918-4925.

[28] C. Ibáñez, C. Simó, V. García-Cañas, Á Gómez-Martínez, J.A. Ferragut, A. Cifuentes, CE/LC-MS multiplatform for broad metabolomic analysis of dietary polyphenols effect on colon cancer cells proliferation, Electrophoresis 33 (2012) 2328-2336.

[29] C. Ibáñez, A. Valdés, V. García-Cañas, C. Simó, M. Celebier, L. Rocamora-Reverte, Á Gómez-Martínez, M. Herrero, M. Castro-Puyana, A. Segura-Carretero, Global Foodomics strategy to investigate the health benefits of dietary constituents, J. Chromatogr A 1248 (2012) 139-153.

[30] C. Simó, C. Ibáñez, Á Gómez-Martínez, J.A. Ferragut, A. Cifuentes, Is metabolomics reachable? Different purification strategies of human colon cancer cells provide different CE-MS metabolite profiles, Electrophoresis 32 (2011) 1765-1777.

[31] B. Tan, Y. Qiu, X. Zou, T. Chen, G. Xie, Y. Cheng, T. Dong, L. Zhao, B. Feng, X. Hu, Metabonomics identifies serum metabolite markers of colorectal cancer, J. Proteome Res. 12 (2013) 3000-3009.

[32] D.C. Montrose, X.K. Zhou, L. Kopelovich, R.K. Yantiss, E.D. Karoly, K. Subbaramaiah, A.J. Dannenberg, Metabolic profiling, a noninvasive approach for the detection of experimental colorectal neoplasia, Cancer. Prev. Res. (Phila) 5 (2012) 1358-1367.

[33] Y. Cheng, G. Xie, T. Chen, Y. Qiu, X. Zou, M. Zheng, B. Tan, B. Feng, T. Dong, P. He, Distinct urinary metabolic profile of human colorectal cancer, J. Proteome Res. 11 (2011) 1354-1363.

[34] Y. Ma, H. Qin, W. Liu, J. Peng, L. Huang, X. Zhao, Y. Cheng, Ultra-high performance liquid chromatography–mass spectrometry for the metabolomic analysis of urine in colorectal cancer, Dig. Dis. Sci. 54 (2009) 2655-2662.

[35] D. Zimmermann, M. Hartmann, M.P. Moyer, J. Nolte, J.I. Baumbach, Determination of volatile products of human colon cell line metabolism by GC/MS analysis, Metabolomics 3 (2007) 13-17.

[36] E.C.Y. Chan, P.K. Koh, M. Mal, P.Y. Cheah, K.W. Eu, A. Backshall, R. Cavill, J.K. Nicholson, H.C. Keun, Metabolic profiling of human colorectal cancer using high-resolution magic angle spinning nuclear magnetic resonance (HR-MAS NMR) spectroscopy and gas chromatography mass spectrometry (GC/MS), J. Proteome Res. 8 (2008) 352-361.

[37] C. Denkert, J. Budczies, W. Weichert, G. Wohlgemuth, M. Scholz, T. Kind, S. Niesporek, A. Noske, A. Buckendahl, M. Dietel, Metabolite profiling of human colon carcinoma–deregulation of TCA cycle and amino acid turnover, Mol. Cancer 7 (2008) 72.

[38] Y.L. Ma, W.J. Liu, J.Y. Peng, P. Zhang, H.Q. Chen, H.L. Qin, Study on specific metabonomic profiling of serum from colorectal cancer patients by gas chromatography-mass spectrometry, Zhonghua Wei Chang Wai Ke Za Zhi 12 (2009) 386-390.

[39] Y. Ma, W. Liu, J. Peng, L. Huang, P. Zhang, X. Zhao, Y. Cheng, H. Qin, A pilot study of gas chromatograph/mass spectrometry-based serum metabolic profiling of colorectal cancer after operation, Mol. Biol. Rep. 37 (2010) 1403-1411.

[40] Y. Qiu, G. Cai, M. Su, T. Chen, Y. Liu, Y. Xu, Y. Ni, A. Zhao, S. Cai, L.X. Xu, Urinary metabonomic study on colorectal cancer, J. Proteome Res. 9 (2010) 1627-1634.

[41] Y. Kondo, S. Nishiumi, M. Shinohara, N. Hatano, A. Ikeda, T. Yoshie, T. Kobayashi, Y. Shiomi, Y. Irino, T. Takenawa, Serum fatty acid profiling of colorectal cancer by gas chromatography/mass spectrometry, Biomark. Med. 5 (2011) 451-460.

[42] M. Mal, P.K. Koh, P.Y. Cheah, E.C.Y. Chan, Metabotyping of human colorectal cancer using two-dimensional gas chromatography mass spectrometry, Anal. Bioanal. Chem. 403 (2012) 483-493.

[43] T. Yoshie, S. Nishiumi, Y. Izumi, A. Sakai, J. Inoue, T. Azuma, M. Yoshida, Regulation of the metabolite profile by an APC gene mutation in colorectal cancer, Cancer Sci. 103 (2012) 1010-1021.

[44] L.C. Phua, M. Mal, P.K. Koh, P.Y. Cheah, E.C.Y. Chan, H.K. Ho, Investigating the role of nucleoside transporters in the resistance of colorectal cancer to 5-fluorouracil therapy, Cancer Chemother. Pharmacol. 71 (2013) 817-823.

[45] M.A. Lorenz, C.F. Burant, R.T. Kennedy, Reducing time and increasing sensitivity in sample preparation for adherent mammalian cell metabolomics, Anal. Chem. 83 (2011) 3406-3414.

[46] C.A. Sellick, D. Knight, A.S. Croxford, A.R. Maqsood, G.M. Stephens, R. Goodacre, A.J. Dickson, Evaluation of extraction processes for intracellular metabolite profiling of mammalian cells: matching extraction approaches to cell type and metabolite targets, Metabolomics 6 (2010) 427-438.

[47] B. Álvarez-Sánchez, F. Priego-Capote, M.L. de Castro, Metabolomics analysis II. Preparation of biological samples prior to detection, TrAC-Trend. Anal. Chem. 29 (2010) 120-127.

[48] B. Gonzalez, J. François, M. Renaud, A rapid and reliable method for metabolite extraction in yeast using boiling buffered ethanol, Yeast 13 (1997) 1347-1355.

[49] S. Dietmair, N.E. Timmins, P.P. Gray, L.K. Nielsen, J.O. Krömer, Towards quantitative metabolomics of mammalian cells: Development of a metabolite extraction protocol, Anal. Biochem. 404 (2010) 155-164.

[50] S. Dietmair, M.P. Hodson, L. Quek, N.E. Timmins, P. Chrysanthopoulos, S.S. Jacob, P. Gray, L.K. Nielsen, Metabolite profiling of CHO cells with different growth characteristics, Biotechnol. Bioeng. 109 (2012) 1404-1414.

[51] J. Kronthaler, G. Gstraunthaler, C. Heel, Optimizing high-throughput metabolomic biomarker screening: a study of quenching solutions to freeze intracellular metabolism in CHO cells, Omics 16 (2012) 90-97.

[52] Q. Teng, W. Huang, T.W. Collette, D.R. Ekman, C. Tan, A direct cell quenching method for cell-culture based metabolomics, Metabolomics 5 (2009) 199-208.

[53] I. Chantret, A. Barbat, E. Dussaulx, M.G. Brattain, A. Zweibaum, Epithelial polarity, villin expression, and enterocytic differentiation of cultured human colon carcinoma cells: a survey of twenty cell lines, Cancer Res. 48 (1988) 1936-1942.

[54] D. Ahmed, P. Eide, I. Eilertsen, S. Danielsen, M. Eknæs, M. Hektoen, G. Lind, R. Lothe, Epigenetic and genetic features of 24 colon cancer cell lines, Oncogenesis 2 (2013) e71.

[55] S. Fernández-Arroyo, A. Gómez-Martínez, L. Rocamora-Reverte, R. Quirantes-Piné, A. Segura-Carretero, A. Fernández-Gutiérrez, J. Ferragut, Application of nanoLC-ESI-TOF-MS for the metabolomic analysis of phenolic compounds from extra-virgin olive oil in treated colon-cancer cells, J. Pharm. Biomed. Anal. 63 (2012) 128-134.

[56] M. Čuperlović-Culf, D.A. Barnett, A.S. Culf, I. Chute, Cell culture metabolomics: applications and future directions, Drug Discov. Today 15 (2010) 610-621.

[57] C. Ibáñez, C. Simó, A. Valdés, L. Campone, A.L. Piccinelli, V. García-Cañas, A. Cifuentes, Metabolomics of adherent mammalian cells by capillary electrophoresis-mass spectrometry: HT-29 cells as case study, J. Pharm. Biomed. Anal. 110 (2015) 83-92.

[58] H. Meyer, H. Weidmann, M. Lalk, Methodological approaches to help unravel the intracellular metabolome of Bacillus subtilis, Microb. Cell. Fact. 12 (2013) 69.

[59] A.D. Patterson, H. Li, G.S. Eichler, K.W. Krausz, J.N. Weinstein, A.J. Fornace, F.J. Gonzalez, J.R. Idle, UPLC-ESI-TOFMS-based metabolomics and gene expression dynamics inspector self-organizing metabolomic maps as tools for understanding the cellular response to ionizing radiation, Anal. Chem. 80 (2008) 665-674.

[60] D.Y. Lee, O. Fiehn, High quality metabolomic data for Chlamydomonas reinhardtii, Plant Methods 4 (2008) 7.

[61] J.L. Au, M.H. Su, M.G. Wientjes, Extraction of intracellular nucleosides and nucleotides with acetonitrile, Clin. Chem. 35 (1989) 48-51.

[62] L. von Stechow, A. Ruiz-Aracama, B. van de Water, A. Peijnenburg, E. Danen, A. Lommen, Identification of cisplatin-regulated metabolic pathways in pluripotent stem cells, PloS one 8 (2013) e76476.

[63] O. Fiehn, T. Kind, Metabolite profiling in blood plasma, Method. Mol. Biol. 358 (2007) 3-17.

[64] O. Fiehn, G. Wohlgemuth, M. Scholz, Setup and Annotation of Metabolomic Experiments by Integrating Biological and Mass Spectrometric Metadata. In: Ludäscher B., Raschid L. (eds) Data Integration in the Life Sciences. DILS 2005. Lecture Notes in Computer Science, vol 3615. Springer, Berlin, Heidelberg.

541 [65] M. Scholz, O. Fiehn, SetupX--a public study design database for metabolomic projects, Pac.
542 Symp. Biocomput. (2007) 169-180.

543 [66] M. Chagoyen, F. Pazos, Tools for the functional interpretation of metabolomic experiments,
544 Brief Bioinform. 14 (2013) 737-744.

545 [67] J. Xia, N. Psychogios, N. Young, D.S. Wishart, MetaboAnalyst: a web server for metabolomic
546 data analysis and interpretation, Nucleic Acids Res. 37 (2009) W652-60.

547 [68] M. Chagoyen, F. Pazos, MBRole: enrichment analysis of metabolomic data, Bioinformatics 27
548 (2011) 730-731.

1
2
3
4 549 **FIGURE CAPTIONS**
5
6 550
7
8
9
10 551 **Figure 1**. GC-MS total ion chromatograms (TIC) of metabolite extracts obtained using
11
12 552 ACN (A), ACN-IsopOH-water (3:3:1, v/v/v) (B), water (C), and 5% formic acid in water
13
14 553 (D). GC-MS conditions are described in the text.
15
16
17
18 554 **Figure 2**. Hierarchical clustering and heat map of total identified metabolites (rows) in all
19
20 555 samples (columns). A (ACN), B (ACN-IsopOH-water (3:3:1, v/v/v)), C (water) and D (5%
21
22 556 formic acid) extracts are represented in red, green, dark blue and light blue respectively.
23
24 557 Cell color represents metabolite relative content.
25
26
27
28 558 **Figure 3**. Venn diagram representation of number of HT-29 metabolites identified in A, B,
29
30 559 C and D extracts.
31
32
33
34 560 **Figure 4.** Summary plot for Metabolite Set Enrichment Analysis (MSEA) of metabolites
35
36 561 identified in the organic extracts (A and B) from HT-29 cells.
37
38
39
40 562 **Figure 5.** Summary plot for Metabolite Set Enrichment Analysis (MSEA) of metabolites
41
42 563 identified in the aqueous extracts (C and D) from HT-29 cells.
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Table 1. Summary of metabolite extraction performance of the four extraction solvents used in this work.

Metabolic extract	Extraction solvent	Detected metabolites	Identified metabolites	Metabolites at maximum intensity [a]	Metabolites RSD<15% [a]
A	ACN	141	82	13	17
B	ACN-IsopOH-H_2O (3:3:1, v/v/v)	182	96	36	51
C	H_2O	172	100	25	62
D	5% HCOOH	168	94	76	20

[a]Metabolites at maximum intensity (5th column) and with RSD <15% (6th column) out of the total identified metabolites.