Chapter 7

LC-MS-Based Metabolomic Investigation of Chemopreventive Phytochemical-Elicited Metabolic Events

Lei Wang, Dan Yao, and Chi Chen

Abstract

Phytochemicals are under intensive investigation for their potential use as chemopreventive agents in blocking or suppressing carcinogenesis. Metabolic interactions between phytochemical and biological system play an important role in determining the efficacy and toxicity of chemopreventive phytochemicals. However, complexities of phytochemical biotransformation and intermediary metabolism pose challenges for studying phytochemical-elicited metabolic events. Metabolomics has become a highly effective technical platform to detect subtle changes in a complex metabolic system. Here, using green tea polyphenols as an example, we describe a workflow of LC-MS-based metabolomics study, covering the procedures and techniques in sample collection, preparation, LC-MS analysis, data analysis, and interpretation.

Key words Chemoprevention, Phytochemical, Metabolism, Metabolomics, LC-MS

1 Introduction

Cancer is a leading cause of disease-related mortalities over the world. In the United States, nearly one fourth of deaths are due to cancer [1]. Compared to invasive and costly surgical procedures, chemotherapy, and radiotherapy, chemoprevention is a promising approach to block and suppress carcinogenesis, especially for the people in high risk of cancer due to genetic background or environmental factors. Phytochemicals in plants (herbs and vegetables) are considered as a reliable and accessible source of chemopreventive agents since the efficacy of plant extracts against carcinogenesis has been largely attributed to specific phytochemicals, such as indole-3-carbinol in cruciferous vegetables and polyphenols in green tea [2–4]. Therefore, identifying potent chemopreventive phytochemicals and characterizing the mechanisms of their anticarcinogenic activities are the main goals of ongoing chemoprevention research.

Sabrina Strano (ed.), *Cancer Chemoprevention: Methods and Protocols*, Methods in Molecular Biology, vol. 1379,
DOI 10.1007/978-1-4939-3191-0_7,

Metabolism plays an essential role in the bioactivities of phytochemicals against carcinogenesis. On one hand, how a phytochemical is disposed in a biological system through absorption, distribution, metabolism, and excretion (ADME) determines the concentration and duration of phytochemical presence in target sites as well as whether bioactivation or detoxification biotransformation occurs to the phytochemical. On the other hand, since cancer is a metabolic disease, how a phytochemical affects the metabolism of a biological system could have major impacts on its chemopreventive activity. Studies in recent years have shown that uncontrolled proliferation of tumor cells is driven by dysregulated nutrient and energy metabolism. For example, aerobic glycolysis in cancer cells channels glucose metabolism toward lactate production in the presence of adequate oxygen [5], producing the intermediates that can be utilized for anabolic activities in growing cells, such as biosynthesis of fatty acids, nonessential amino acids, nucleic acids, and intracellular antioxidants [6, 7]. It has been shown that chemopreventive phytochemicals could affect diverse metabolic pathways, resulting in suppressing effects on tumor cells [8, 9]. Therefore, examining the metabolic interactions between phytochemicals and biological systems is essential for understanding and predicting the chemopreventive effects of phytochemicals.

Systems biology tools, such as genomics, transcriptomics, and proteomics, have been adopted to characterize the cancer prevention activities of phytochemicals due to their capacity for discovering molecular mechanisms in transcriptional and translational levels [10]. However, these platforms have clear disadvantages in elucidating the metabolic interactions between phytochemicals and biological systems since the central players of these metabolic interactions, which are phytochemical metabolites and endogenous metabolites, are not directly examined by these tools. In this regard, metabolomics, as a platform that is capable of detecting subtle metabolic changes in a complex biological system, has become a very effective tool for investigating chemical and metabolic events in chemoprevention. Untargeted metabolomics has been used to identify novel anticancer phytochemicals and examine biotransformation of phytochemicals [11, 12]. Furthermore, the values of metabolomics in characterizing chemopreventive phytochemical-induced metabolic events has also been discussed [13].

In this chapter, we describe a protocol for liquid chromatography-mass spectrometry (LC-MS)-based metabolomic investigation of chemopreventive phytochemical-induced metabolic events. Using green tea polyphenols (GTP) as an example, the procedures, techniques, and considerations in sample collection, preparation, LC-MS analysis, data analysis, and interpretation are described and discussed.

2 Materials

Following the general procedure of animal-based investigation of phytochemical-elicited metabolic events, the materials in metabolomics studies are categorized as the items for sample collection, sample preparation, LC-MS analysis, and data analysis, respectively (*see* **Note 1**). All solutions are prepared with LC-MS grade water, organic solvents, and analytical grade reagents. Unless indicated otherwise, all solutions are stored at room temperature.

2.1 Sample Collection in Animal Experiment

1. Animal: 8-week-old male C57/BL6 mice are used in this case study.
2. Chemopreventive phytochemical: The GTP extract used in this case study contains more than 50 % epigallocatechin gallate (EGCG).
3. Experimental diet: AIN93G-purified diet is used in this case study.
4. Metabolic cages (Tecniplast).
5. Lancet for submandibular bleeding.
6. Surgical apparatus for tissue collections.
7. 1.5 mL Eppendorf tubes for urine, serum, and fecal samples.
8. Cryogenic tubes for tissue collection.
9. Serum separator tubes (BD Microtainer™).
10. Liquid nitrogen.

2.2 Sample Preparation

1. Methanol for sample fractionation.
2. Chloroform for sample fractionation.
3. *n*-Butanol for dissolving lipid fraction.
4. 2 mL flat-bottom centrifuge tubes.
5. Tissue homogenizer.
6. Centrifuge.
7. Internal standards (*see* **Note 2**).
8. Reagents for derivatizing amino-containing metabolites: freshly prepared 3 mg/mL dansyl chloride (DC) in acetone, 10 mM sodium carbonate in water.
9. Reagents for derivatizing carboxylic acids, aldehydes, and ketones: freshly prepared reaction mixture containing 1 mM 2-hydrazinoquinoline (HQ), 1 mM 2,2′-dipyridyl disulfide (DPDS), and 1 mM triphenylphosphine (TPP) in acetonitrile (ACN).

2.3 LC-MS Analysis (See Note 3)

1. LC system: ACQUITY™ ultra-performance liquid chromatography (UPLC) system (Waters).
2. High-resolution MS system: SYNAPT quadrupole time-of-flight (QTOF) MS system (Waters).
3. ACQUITY UPLC BEH C18 column, 1.7 μm, 2.1 mm × 50 mm (Waters).
4. ACQUITY UPLC BEH C8 column, 1.7 μm, 2.1 mm × 50 mm (Waters).
5. ACQUITY UPLC BEH Amide column, 1.7 μm, 2.1 mm × 100 mm (Waters).
6. Mobile-phase A1 for analyzing general metabolites: H_2O containing 0.1 % formic acid (v/v).
7. Mobile-phase B1 for analyzing general metabolites: ACN containing 0.1 % formic acid (v/v).
8. Mobile-phase A2 for analyzing triglycerides and nonpolar lipids: H_2O:ACN (6:4, v:v) containing 10 mM ammonium formate and 0.1 % formic acid.
9. Mobile-phase B2 for analyzing triglycerides and nonpolar lipids: Isopropyl alcohol (IPA):ACN (9:1, v:v) containing 10 mM ammonium formate and 0.1 % formic acid.
10. Mobile-phase A3 for analyzing HQ-derivatized metabolites: H_2O containing 0.05 % acetic acid (v/v) and 2 mM ammonium acetate.
11. Mobile-phase B3 for analyzing HQ-derivatized metabolites: H_2O:ACN (5:95, v:v) containing 0.05 % acetic acid (v/v) and 2 mM ammonium acetate.
12. Lock mass: 500 pg/μL leucine encephalin in 50 % ACN (v/v) with 0.1 % formic acid (v/v).
13. LC sample vials.
14. Nitrogen gas for desolvation and ionization in MS system.
15. Argon gas for MSMS fragmentation analysis.
16. Software for operating the LC-MS system and acquiring LC-MS data: MassLynx™ software (Waters).

2.4 Data Analysis (See Note 4)

1. Software for processing LC-MS data:MassLynx™ (Waters).
2. Software for deconvoluting LC-MS data: Markerlynx™ (Waters).
3. Software for multivariate data analysis (MDA): SIMCA-P+™ (Umetrics).
4. Chemical standards for confirming the structures of interested metabolites.

3 Methods

The methods described here aim to detect the metabolic differences between control and phytochemical-treated animals. Following the general procedure of animal-based investigation of phytochemical-elicited metabolic events, the methods in metabolomics studies are categorized as the methods in animal treatment and sample collection, sample preparation, LC-MS analysis, and data analysis, respectively.

3.1 Animal Treatment and Sample Collection (See Notes 5 and 6)

In this case study, mice are housed under controlled temperature and lighting conditions (20-22°C and a 14-h/10-h light/dark cycle). Two groups of mice are acclimated for 3 days on control diet (AIN93G diet) before the treatment. Then one group of mice is switched to GTP diet (AIN93G + 0.12 % GTP) for 2 weeks.

1. Urine and fecal samples: Urine and fecal samples are collected by housing the mice in metabolic cages for 24 h, and then transferred to 1.5 mL Eppendorf tubes. All urine and fecal samples are stored at -80°C.
2. Serum samples: Blood is collected in serum separator tubes by submandibular bleeding. After clotting at room temperature, blood samples are centrifuged at 3000 × *g* to separate serum and blood cells. All serum samples are stored at -80°C.
3. Tissue collection: After the mice are euthanized by carbon dioxide, the liver and other tissue samples are harvested into cryogenic tubes and then snap-frozen in liquid nitrogen. All tissue samples are stored at -80°C.

3.2 Sample Preparation (See Note 7)

In this case study, the metabolism of GTP is determined by metabolomic analysis of urine and fecal samples while the influences of GTP on endogenous metabolism are examined by metabolomic analysis of serum and tissue samples.

1. Precipitation: To remove proteins and particles in biofluid samples through solvent denaturation and centrifugation.
   (a) Urine: Mix 40 μL of urine sample with 160 μL of 50 % aqueous ACN (v/v) in a 1.5 mL Eppendorf tube, and then centrifuge at 18,000 × *g* for 10 min. Transfer supernatant to a LC vial for LC-MS analysis.
   (b) Serum: Mix 5 μL of serum sample with 195 μL of 66 % aqueous ACN (v/v) in a 1.5 mL Eppendorf tube, and then centrifuge at 18,000 × *g* for 10 min. Transfer supernatant to a LC vial for LC-MS analysis.
2. Fractionation: To prepare aqueous and lipid fractions of serum and tissue samples (*see* **Note 8**).

(a) Tissue: Homogenize 100 mg of tissue sample with 0.5 mL of methanol in a 2 mL flat-bottom centrifuge tube using a mechanical homogenizer. The homogenate is then mixed with 0.5 mL of chloroform and 0.4 mL of water. After vortex and 10-min centrifugation at 18,000 × *g*, aqueous and lipid phases are separated by tissue debris. The upper aqueous phase is ready for direct LC-MS analysis or further chemical derivatization. The lower lipid phase is dried under nitrogen and then reconstituted in 0.5 mL of *n*-butanol for LC-MS analysis.

(b) Serum: Mix 20 μL of serum sample with 100 μL of methanol in a 1.5 mL Eppendorf tube, and then add 100 μL of chloroform and 80 μL of water. After vortex and 10-min centrifugation at 18,000 × *g*, aqueous and lipid phases are separated. The upper aqueous phase is ready for direct LC-MS analysis or further chemical derivatization. The lower lipid phase is dried under nitrogen and then reconstituted in 100 μL of *n*-butanol for LC-MS analysis.

3. Derivatization: To facilitate the detection of metabolites that have poor retention in LC system or poor ionization in MS system under general analytical conditions, samples are derivatized prior to LC-MS analysis.

(a) DC derivatization for analyzing the metabolites with amino group. Mix 5 μL of sample (serum, urine, or tissue extract) with 5 μL of 100 μM *p*-chlorol-L-phenylalanine (internal standard), 50 μL of 10 mM sodium carbonate, and 100 μL of DC acetone solution (3 mg/mL) in sequence. After 10-min incubation at 60°C, the reaction mixture is centrifuged at 18,000 × *g* for 10 min, and the supernatant is transferred to an LC vial.

(b) HQ derivatization for analyzing carboxylic acids, aldehydes, and ketones. Mix 5 μL of test sample with 100 μL of freshly prepared reaction mixture containing 1 mM DPDS, 1 mM TPP, and 1 mM HQ in ACN. Incubate at 60°C for 30 min, quickly chill on ice and then mix with 100 μL of H_2O. After centrifugation at 18,000 × *g* for 10 min, transfer the supernatant into an LC vial for LC-MS analysis.

3.3 LC-MS Analysis (See Note 9)

In this case study, LC-MS analysis of control and GTP treatment samples is conducted using a UPLC-QTOFMS system (*see* **Note 10**).

1. LC system: In general, 5 μL of processed sample is injected into a UPLC system and separated by a gradient of mobile phase over a 10-min run at flow rate 0.5 mL/min (*see* **Note 11**).

(a) Using mobile-phase A1 and B1 at 40°C, nonpolar metabolites in urine, fecal extract, and DC-derivatized samples

could be separated in C18 column while polar metabolites could be retained in amide column (*see* **Note 12**).

(b) Using mobile-phase A2 and B2 at 60°C, lipids, including phospholipids and triglycerides, in serum and tissue extracts, could be separated in C18 or C8 columns.

(c) Using mobile-phase A3 and B3 at 40°C, HQ-derivatized samples could be separated in C18 column.

2. MS system:

(a) General parameters of MS analysis: Capillary voltage and cone voltage for electrospray ionization (ESI) are maintained at 3 kV and 30 V for positive-mode detection, and at -3 kV and -35 V for negative-mode detection, respectively. Source temperature and desolvation temperature are set at 120°C and 350°C, respectively. Nitrogen is used as both cone gas (50 L/h) and desolvation gas (600 L/h) and argon as collision gas. Tandem MS (MS/MS) fragmentation is performed with collision energy ranging from 15 to 40 eV.

(b) Calibration for accurate mass measurement: The QTOF mass spectrometer is calibrated with sodium formate solution (range m/z 50–1000) and monitored by the intermittent injection of the lock mass leucine enkephalin ($[M+H]^+ = 556.2771$ m/z and $[M-H]^- = 554.2615$ m/z) in real time.

3.4 Data Analysis (See Note 13)

In this case study, untargeted metabolomics approach is adopted to identify GTP metabolites in urine.

1. Data deconvolution: Chromatographic and spectral data of LC-MS analysis are deconvoluted by MarkerLynx™ software (*see* **Note 14**) to construct a data matrix that comprises samples, metabolites (represented by retention time and m/z ratio), and signal intensity.
2. MDA: The data matrix is further exported into SIMCA-P+™ software, and transformed by mean-centering and *Pareto* scaling. Based on the complexity and quality of the data, either unsupervised such as principal components analysis (PCA), or supervised MDA, such as partial least squares-discriminant analysis (PLS-DA) and orthogonal partial least squares (OPLS), are adopted to analyze the data matrix. Major latent variables in the data matrix are defined in a scores scatter plot of defined multivariate model.
3. Marker identification and structural analysis: Potential biomarkers are identified by analyzing ions contributing to the principal components in the loadings plot. The chemical identities of biomarkers are determined by accurate mass measurement, elemental composition analysis, MS/MS fragmentation, database search (*see* **Note 15**), and comparisons with authentic standards.

Using the materials and methods described in this protocol, urine samples from control and GTP-treated mice are harvested, prepared, and analyzed by a UPLC-QTOFMS system in negative mode. The LC-MS data are deconvoluted to a data matrix (Fig. 1a), which contains the information on samples, metabolites, and signals. The data matrix is further processed by MDA to generate a PLS-DA model, in which the urine samples from control and GTP-treated mice are clearly separated (Fig. 1b). The urinary metabolites affected by GTP treatment are identified in a loadings S-plot, which reveals the metabolites contributing to the separation of two sample groups in a multivariate model (Fig. 1c). Two urinary metabolites increased by GTP treatment are further characterized as 3-hydroxyphenylvaleric acid sulfate (I), a general bacterial metabolite of catechins, and epicatechin sulfate (II), a metabolite of EGCG and other polyphenols in GTP (Fig. 1d, e).

4 Notes

1. Enlisted are the items used in the GTP case study. The selection of animal, phytochemical, reagent, instrument, and software should be based on the experiment design of each metabolomics study.
2. Spiking internal standards to the samples could facilitate the efforts to monitor the efficacy of sample preparation and the performance of LC-MS system. The choices of internal standards includes table isotope-labeled metabolites, halogenated metabolites, or other unnatural analogs of phytochemicals or endogenous metabolites.
3. A UPLC-QTOF system is used in this case study, which could be changed to other types of high-resolution LC-MS systems. Column and mobile phases could also be changed according to the experiment design and instrument availability.
4. Other commercial software and free public platforms are also available for processing LC-MS data [14, 15].
5. Considerations on animal experiment: (a). Avoid or minimize the influences of confounding factors, such as gender, age, strain, and environment, when selecting animals for control and phytochemical treatment. (b). Basal components of animal diets used in control and phytochemical treatments should be the same.
6. Considerations on sample collection: (a). Different types of samples contain different types of metabolites, and can reflect different aspects of phytochemical-elicited changes in the metabolome. In general, urine and fecal samples could be used for identifying or profiling the metabolites of chemopreventive

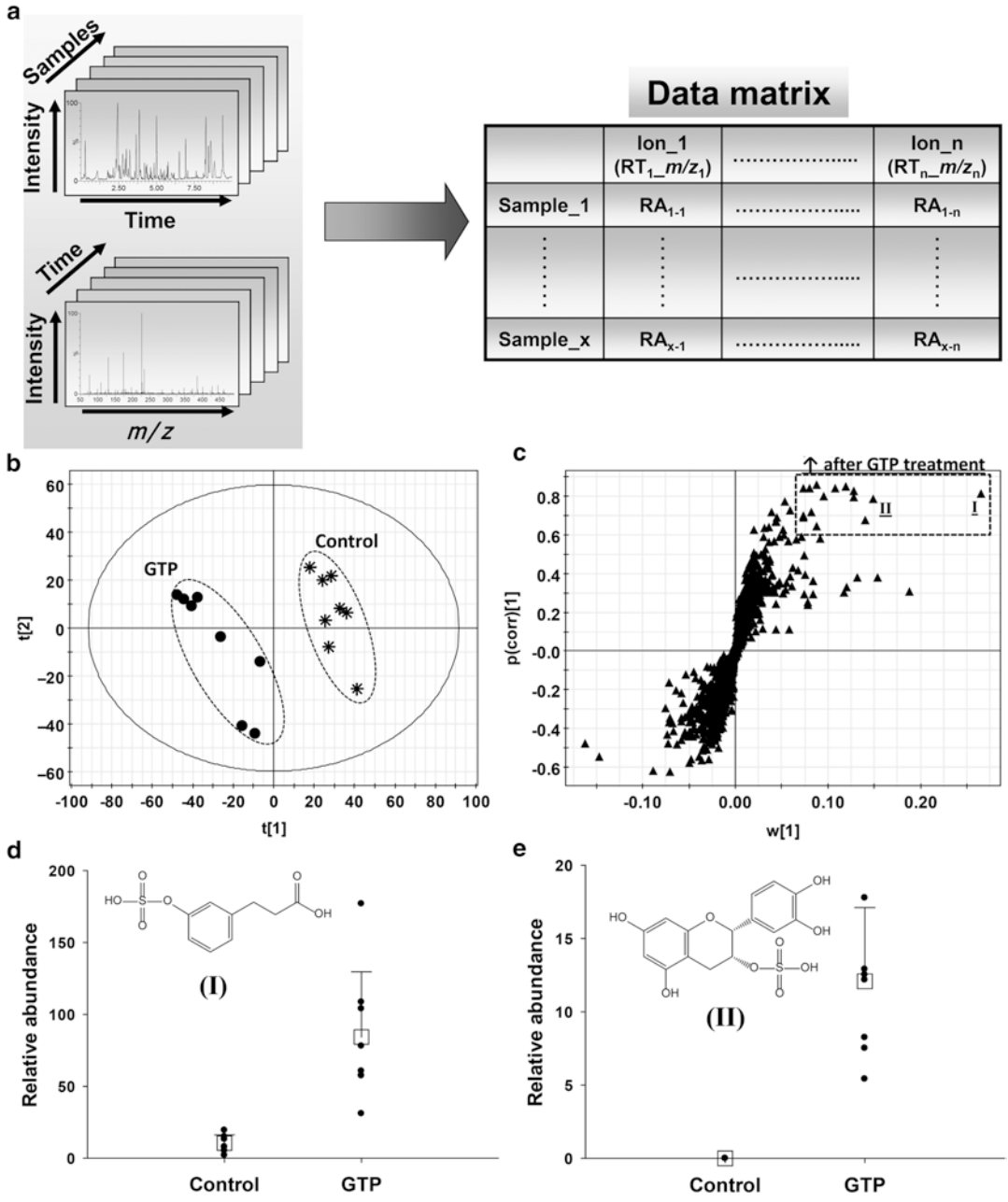

Fig. 1 LC-MS-based metabolomic investigation of GTP-induced changes in urine metabolome. (**a**) Deconvolution of LC-MS data to a multivariate data matrix. RA stands for the relative abundance of single ion count in total ion count of a chromatogram; RT stands for the retention time in LC column. (**b**) Scores plot of a PLS-DA model on the urine samples from control and GTP-treated mice ($n = 8$). The t[1] and t[2] values represent the respective scores of each sample in the principal component 1 and 2 of the model. (**c**) S loadings plot for identifying the metabolites contributing to the separation of control and GTP samples in the model. Two metabolites increased by GTP treatment are labeled (I: 245.106 *m*/*z* and II: 369.0279 *m*/*z*). (**d**, **e**) Structures of metabolites I and II and their relative abundances (individual values and mean ± S.D.) in control and GTP samples

phytochemicals, while blood and tissue samples could reveal the phytochemical-induced changes in endogenous metabolism. (b). The procedures in sample collection should maintain the chemical integrity of biological samples and avoid significant degradation or changes.

7. Considerations on sample preparation: (a). For untargeted metabolomics, the procedures of sample preparation should aim to maintain the integrity of metabolome in acquired samples through avoiding or minimizing the formation of new chemical species or the degradation of existing metabolites [16]. (b). For targeted metabolomics, appropriate procedures, such as solid-phase extraction, could be adopted to enrich the interested metabolites.
8. A modified Folch method is used to separate aqueous and lipid fractions [17].
9. Considerations on LC-MS analysis: (a). For untargeted metabolomics, a MS system with high-resolution capacity to determine accurate mass and broad dynamic range to measure signal intensity is preferred for structural elucidation of interested metabolites and multivariate data analysis. (b). Selections of LC column, mobile phase, column temperature, and ionization condition are based on the chemical properties of samples and metabolites, such as polarity, reactivity, and ionization efficiency.
10. Enlisted conditions and parameters are specifically for the UPLC-QTOF system used in this case study. Different conditions and parameters are expected for other LC-MS systems.
11. The gradient usually starts with low percent of organic phase (mobile phase B), and then gradually increases to high percent of B. At the end of 10-min run, the gradient returns to the starting gradient for the next sample. For example, A1-B1 mobile phase is used to separate urine samples in the GTP case study. The gradient profile starts at 0.5 % B1 with a flow rate of 0.5 mL/min for 0.5 min and then rises to 20 % B1 at 4 min and 95 % B1 at 8 min. At 8.1 min, the gradient is increased to 100 % B1. At 9.1 min, the gradient returns to 0.5 % B1.
12. Different to C8 and C18 column, Amide column uses hydrophilic interaction chromatography (HILIC) to retain metabolites that are too polar to be retained by reversed-phase chromatography.
13. Considerations on data analysis: (a). General procedure of data analysis in untargeted metabolomics include data deconvolution, multivariate data analysis, marker identification, bioinformatics, structural confirmation, and potential mechanistic investigation. (b). Targeted metabolomics focuses on identification and quantification of targeted metabolites.

14. The multivariate data matrix is generated through centroiding, deisotoping, filtering, peak recognition and integration. The signal intensity of each ion is calculated by normalizing the single ion counts (SIC) versus the total ion counts (TIC) in the whole chromatogram.
15. Databases for metabolite identification and structural analysis: Human Metabolome Database (http://www.hmdb.ca/), Kyoto Encyclopedia of Genes and Genomes (KEGG, http://www.genome.jp/kegg/), METLIN database (http://metlin.scripps.edu/), Lipid Maps (http://www.lipidmaps.org/), BioCyc (http://biocyc.org/), Spectral Database for organic compounds (http://sdbs.riodb.aist.go.jp).

Acknowledgements

Research projects in Dr. Chi Chen's lab are supported in part by an Agricultural Experiment Station project MIN-18-082 from the United States Department of Agriculture (USDA). We thank all the members in Dr. Chi Chen's lab for their help in preparing this protocol.

References

1. Jemal A, Siegel R, Xu J et al (2010) Cancer statistics, 2010. CA Cancer J Clin 60(5):277–300
2. Chen C, Kong ANT (2005) Dietary cancer-chemopreventive compounds: from signaling and gene expression to pharmacological effects. Trends Pharmacol Sci 26(6):318–326
3. Yang CS, Wang X, Lu G et al (2009) Cancer prevention by tea: animal studies, molecular mechanisms and human relevance. Nat Rev Cancer 9(6):429–439
4. Ahmad A, Sakr WA, Rahman KM (2010) Anticancer properties of indole compounds: mechanism of apoptosis induction and role in chemotherapy. Curr Drug Targets 11(6):652–666
5. Vander Heiden MG, Cantley LC, Thompson CB (2009) Understanding the Warburg effect: the metabolic requirements of cell proliferation. Science 324(5930):1029–1033
6. Gatenby RA, Gillies RJ (2004) Why do cancers have high aerobic glycolysis? Nat Rev Cancer 4(11):891–899
7. Lunt SY, Vander Heiden MG (2011) Aerobic glycolysis: meeting the metabolic requirements of cell proliferation. Annu Rev Cell Dev Biol 27:441–464
8. Tan AC, Konczak I, Sze DM et al (2011) Molecular pathways for cancer chemoprevention by dietary phytochemicals. Nutr Cancer 63(4):495–505
9. Lee KW, Bode AM, Dong Z (2011) Molecular targets of phytochemicals for cancer prevention. Nat Rev Cancer 11(3):211–218
10. Urich-Merzenich G, Zeitler H, Jobst D et al (2007) Application of the "-omic-" technologies in phytomedicine. Phytomedicine 14(1):70–82
11. Kersten RD, Dorrestein PC (2009) Secondary metabolomics: natural products mass spectrometry goes global. ACS Chem Biol 4(8):599–601
12. Chen V, Staub RE, Baggett S et al (2012) Identification and analysis of the active phytochemicals from the anti-cancer botanical extract Bezielle. PLoS One 7(1):e30107
13. Wang L, Chen C (2013) Emerging applications of metabolomics in studying chemopreventive phytochemicals. AAPS J 15(4):941–950

14. Sugimoto M, Kawakami M, Robert M et al (2012) Bioinformatics tools for mass spectroscopy-based metabolomic data processing and analysis. Curr Bioinform 7(1):96–108
15. Blekherman G, Laubenbacher R, Cortes DF et al (2011) Bioinformatics tools for cancer metabolomics. Metabolomics 7(3):329–343
16. Villas-Boas SG, Mas S, Akesson M et al (2005) Mass spectrometry in metabolome analysis. Mass Spectrom Rev 24(5):613–646
17. Folch J, Lees M, Sloane Stanley GH (1957) A simple method for the isolation and purification of total lipides from animal tissues. J Biol Chem 226(1):497–509