Talanta 164 (2017) 128–136

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Development of data-independent acquisition workflows for metabolomic analysis on a quadrupole-orbitrap platform

Juntuo Zhou[a], Yuhua Li[a], Xi Chen[a], Lijun Zhong[b,*], Yuxin Yin[a,c,d,**]

[a] Department of Pathology, School of Basic Medical Sciences, Peking University Health Science Center, Beijing 100191, China
[b] Medical and Health Analytical Center, Peking University Health Science Center, Beijing 100191, China
[c] Institute of Systems Biomedicine, Department of Pathology, School of Basic Medical Sciences, Peking University Health Science Center, Beijing 100191, China
[d] Beijing Key Laboratory of Tumor Systems Biology, Peking-Tsinghua Center for Life Sciences, Beijing 100191, China

ARTICLE INFO

Keywords:
Liquid chromatograph
Mass spectrometer
Data-dependent acquisition
Metabolomics
Orbitrap
Papillary thyroid carcinoma

ABSTRACT

Untargeted metabolomic profiling has been widely used in recent years. However, the low reproducibility of the data-dependent acquisition (DDA) strategy presents a major bottleneck that considerably limits the reliability of metabolomic studies in biological and clinical research. The data-independent acquisition (DIA) strategy is proposed to solve the above-mentioned problem, and it is gaining popularity. This paper presents a novel approach for performing metabolomic analysis using an untargeted, liquid chromatography–data independent-mass spectrometry (LC-DIA-MS) strategy on a quadrupole-Orbitrap platform. Using chemical standards and metabolites extracted from serum samples, we optimized the LC-DIA-MS parameters to analyze hydrophilic metabolites and lipids. The quantitative performance and analytical reliability were evaluated, and the performances of DIA, DDA, and all-ion fragmentation mode were compared. Finally, as a proof of concept, we applied the constructed DIA workflow to a comparative metabolomic study of papillary thyroid carcinoma (TC) serum samples. Several metabolites, including carnitine, trimethylamine *N*-oxide, and some amino acids, significantly changed between patients and healthy controls. This study demonstrated the feasibility and advantage of the DIA strategy on untargeted metabolomic analysis for biological study and clinical biomarker screening.

1. Introduction

Data-dependent acquisition (DDA) strategy is widely used for untargeted proteomic [1,2] and metabolomic studies [3,4]. However, reproducibly, accurately, and sensitively detecting and quantifying large fractions of the analyzed target (proteome and metabolome) across multiple samples remain challenging for the DDA strategy [5,6]. The bottleneck of the DDA strategy comes from the precursor selection strategy, in which only intense precursor is selected for mass spectrometry MS/MS analysis. To solve this problem, data-independent MS/MS acquisition methods have been proposed and developed for proteomics [5,7] and metabolomics [8,9]. In contrast to the traditional DDA strategy, data-independent acquisition (DIA) strategy can theoretically obtain all fragment ions for all precursors simultaneously, thereby increasing the coverage of observable molecules and reducing the identification of false negatives [5]. The consistent acquisition of MS/MS information can also improve analytical reproducibility and quantitative performance. However, problems exist for the DIA strategy; the main one is the complexity of MS/MS spectra because of the wide isolation window (20–50 Da or more) for precursor ion selection. Therefore, the high performance of MS platform, proper precursor isolation scheme settings, and reliable post-acquisition data-processing software are important issues for data-independent acquisition-mass spectrometry (DIA-MS) strategies.

Several isolation schemes have been developed for DIA-MS. In the field of proteomics, an isolation scheme of consecutive 20–40 Da precursor isolation windows is a common choice [5,10,11]. For metabolomics, the mostly used MS^E workflow isolates all precursors in a whole MS/MS scan instead of employing consecutive isolation windows, which can also be called all-ion fragmentation (AIF) strategy [8,12]. However, the complexity of MS/MS increases when the isolation window becomes wide. In theory, a narrow isolation window leads

* Corresponding author.
** Corresponding author at: Institute of Systems Biomedicine, Department of Pathology, School of Basic Medical Sciences, Peking University Health Science Center, No. 38 Xueyuan Road, Haidian District, Beijing 100191, China.
E-mail addresses: zhonglijun@bjmu.edu.cn (L. Zhong), yinyuxin@bjmu.edu.cn (Y. Yin).

http://dx.doi.org/10.1016/j.talanta.2016.11.048
Received 1 October 2016; Received in revised form 15 November 2016; Accepted 20 November 2016
Available online 21 November 2016
0039-9140/

to good MS/MS spectra. Nevertheless, a narrow isolation window leads to long cycle time and may result in insufficient data points for chromatography peaks to perform reliable quantification. For the DIA-MS strategy, compromise must be made between isolation window width and cycle time, and the fast scan speed of MS is crucial for retrieving good analytical results. The recently reported DIA-MS strategy [9] developed by Tsugawa et al. compared two isolation schemes using 65 and 21 Da isolation windows on a quadrupole time-of-flight (Q-TOF) MS instrument, and the 21 Da isolation scheme presented the best analytical performance for hydrophilic and lipid analysis.

For DIA data processing, the acquired original MS/MS spectra consist of fragments of several precursors, leading to great complexity. Data need to be processed by the deconvolution algorithm to reveal the MS/MS spectra for each precursor. Therefore, the performance of software used for DIA data processing is important. Considerable software, such as Open-SWATH [13], Skyline [14], DIA-Umpire [15] and Group-DIA [16] for proteomics, has been developed for DIA data processing. However, no freely available third-party software for DIA data processing of metabolomics was available until MS-DIAL software was developed by Tsugawa et al. [9]. This newly developed software can process data acquired by DIA-MS using different isolation schemes (consecutive isolation windows or AIF) and support raw MS data files from several common instrument suppliers. With the help of MS-DIAL software, Tsugawa et al. developed a DIA-MS workflow on a Q-TOF MS platform to enable reliable analysis for untargeted metabolomics. Nonetheless, no DIA-MS workflow has been developed and optimized using the quadrupole-Orbitrap (Q-Orbitrap) platform, which is also a popularly used high-resolution liquid chromatography-mass spectrometry (LC-MS) platform. Constructing a reliable workflow to optimize the DIA strategy is urgent, providing various choices for untargeted metabolomic study.

In this paper, on a Q-Orbitrap LC-MS platform, we describe two workflows of hydrophilic interaction liquid chromatography (HILIC)-DIA-MS and reversed phase liquid chromatography (RPLC)-DIA-MS methodologies for untargeted metabolomic analysis to expand metabolome coverage and quantification performance. We optimized the parameters for applying the DIA-MS strategy to serum samples, which could provide unique information on physiological or pathological mechanisms and biomarker discovery. We also evaluated the feasibility, analytical performance, stability, and dynamic range of DIA strategies. Finally, we applied the proposed HILIC-DIA-MS workflow to profile serum from patients with papillary thyroid carcinoma (TC), which is a common endocrine malignancy worldwide and accounts for 1% of all diagnosed cancers and approximately 91.5% of the malignancies of head and neck [17], to further demonstrate the performance of DIA-MS workflows.

2. Experimental section

2.1. Sample preparations

Primary serum specimens from 30 papillary thyroid carcinoma (TC) patients or healthy people were obtained from the 222th hospital of PLA. This study was approved by the Institute Research Ethics Committee for clinical studies. Only serum from individuals who agreed to give samples for the purpose of laboratory research was used. For metabolites extraction, 100 μl serum was extracted by 4-fold cold chloroform: methanol (2:1). The mixture was centrifuged at 13,000 rpm for 15 min and then the upper aqueous phase (hydrophilic metabolites) and lower organic phase (hydrophobic metabolites) were separately collected and evaporated at room temperature under vacuum. Detailed information may be found in the Supplementary materials.

2.2. High-performance liquid chromatography

For LC-MS analysis, the Ultimate 3000 UHPLC system was coupled to Q-Exactive MS (Thermo Scientific) for metabolite separation and detection. For HILIC-DIA-MS workflow, an Xbridge amide column (100×4.6 mm, 3.5 μm; Waters) was employed for compound separation at 30 °C. For RPLC-DIA-MS workflow, an XSELECT CSH C18 column (100×4.6 mm, 2.5μm, Waters) was employed for compound separation at 50 °C.

2.3. Mass spectrometry

Data-dependent acquisition (DDA) and data-independent acquisition (DIA) assays were performed using the Q-Exactive MS (Thermo Scientific).

2.3.1. HILIC-DIA-MS

For DDA mode, the assay consisted of 1 survey scan (MS1 scan) at 35,000 resolution from 50 to 750 *m/z* and followed by 10 MS/MS scans in HCD mode at 17,500 resolution. The DIA method consisted of a survey scan at 35,000 resolution from 50 to 750 *m/z*. Then, 14 DIA windows were acquired at 17,500 resolution using an isolation window of 50 *m/z*.

2.3.2. HILIC-DIA-MS

For DDA mode, the assay consisted of 1 survey scan (MS1 scan) at 35,000 resolution from 200 to 1200 *m/z* and followed by 10 MS/MS scans in HCD mode at 17,500 resolution. The DIA method consisted of a survey scan at 35,000 resolution from 200 to 1200 *m/z*. Then, 20 DIA windows were acquired at 17,500 resolution using an isolation window of 50 *m/z*. Detailed information may be found in the Supplementary materials.

2.4. Raw data processing

After data acquisition, raw data of DDA and DIA were converted from the vendor file format (.raw) into the common file format of Reifycs Inc. (Analysis Base File format *.abf) using the freely available Reifycs ABF converter (http://www.reifycs.com/AbfConverter/index.html). After conversion, the MS-DIAL software was used for feature detection, spectra deconvolution, metabolite identification and peak alignment between samples. The parameters used in MS-DIAL is included in Supplementary materials.

2.5. Statistical analysis

Response curves were calculated with linear regression. The r^2 distribution was calculated in Microsoft Office Excel 2013 and represented as histogram using GraphPad 6.0. Principal components analysis (PCA) and the Student's *t*-test were carried out with the MetaboAnalyst 3.0 web service (http://www.metaboanalyst.ca/). Significance was determined by the Student's *t*-test. *$p < 0.05$; **$p < 0.01$; NS, not significant.

3. Results

3.1. Overview of the DIA-MS workflow

We developed two DIA strategies for measuring hydrophilic and hydrophobic metabolites. The main difference of these two strategies is the different scan mass range and different isolation schemes used for DIA. The entire workflow is shown in Fig. 1. Hydrophilic metabolites and lipids were extracted separately from serum samples using liquid–liquid extraction. The metabolites were then submitted to RPLC-DIA-MS (for lipids) or HILIC-DIA-MS (for hydrophilic metabolites). An isolation scheme of 50 Da consecutive isolation window was used, and