Vol. 5, No. 1, 2014
ISSN 2233-4203/ e-ISSN 2093-8950
LETTER www.msletters.org | Mass Spectrometry Letters

Proteome Analysis of Mouse Adipose Tissue and Colon Tissue using a Novel Integrated Data Processing Pipeline

Jong-Moon Park[1,2,*], Na-Young Han[1], Hokeun Kim[3], Injae Hwang[5], Jae Bum Kim[5], Ki-Baik Hahm[4], Sang-Won Lee[3], and Hookeun Lee[1,2]

[1]*Lee Gil Ya Cancer and Diabetes Institute, Gachon University, 7-45-ka, Songdo-dong, Yeonsu-ku, Incheon 406-840, Republic of Korea*
[2]*Gachon College of Pharmacy, Gachon University, 534-2 Yeonsu3-dong, Yeonsu-gu, Incheon 406-799, Republic of Korea*
[3]*Department of Chemistry, Korea University, 1, 5-ka, Anam-dong, Seongbuk-gu, Seoul 136-701, Republic of Korea*
[4]*CHA Cancer Prevention Research Center, Seoul and Digestive Disease Center, CHA University Bundang Medical Center, Seongnam, Korea*
[5] *Department of Biological Sciences, Institute of Molecular Biology and Genetics, Seoul National University, Seoul, Korea*

Received February 14, 2014; Revised March 7, 2014; Accepted March 11, 2014
First published on the web March 30, 2014; DOI: 10.5478/MSL.2014.5.1.16

Abstract: Liquid chromatography based mass spectrometry (LC-MS) is a key technology for analyzing highly complex and dynamic proteome samples. With highly accurate and sensitive LC-MS analysis of complex proteome samples, efficient data processing is another critical issue to obtain more information from LC-MS data. A typical proteomic data processing starts with protein database search engine which assigns peptide sequences to MS/MS spectra and finds proteins. Although several search engines, such as SEQUEST and MASCOT, have been widely used, there is no unique standard way to interpret MS/MS spectra of peptides. Each search engine has pros and cons depending on types of mass spectrometers and physicochemical properties of peptides. In this study, we describe a novel data process pipeline which identifies more peptides and proteins by correcting precursor ion mass numbers and unifying multi search engines results. The pipeline utilizes two open-source software, *i*PE-MMR for mass number correction, and iProphet to combine several search results. The integrated pipeline identified 25% more proteins in mouse epididymal adipose tissue compared with the conventional method. Also the pipeline was validated using control and colitis induced colon tissue. The results of the present study shows that the integrated pipeline can efficiently identify increased number of proteins compared to the conventional method which can be a breakthrough in identification of a potential biomarker candidate.

Key words: iPE-MMR, iProphet, Q-TOF, TPP

Introduction

Proteomics aims at comprehensive profiling of protein in tissues or cells utilizing various technical platforms such as proteome separation techniques, mass spectrometry (MS), and bioinformatics tools for data processing. In mass spectrometry based proteomics studies, bioinformatics tools are essentially required to interpret more than thousands of spectra from liquid chromatography-mass spectrometry runs. The above process faces two typical (1) error of monoisotopic mass determination and (2) loss of information due to single database search.[1] Assignments of precise precursor ion masses to MS/MS spectra is frequently debatable even when using high resolution mass spectrometers. The resultant non statistical distributions with potentially missing peaks can lead to errors in monoisotopic mass determination.[2] In order to solve the problem, *i*PEMMR has developed a method by combining reported various methods of treating MS/MS data for precursor mass refinement.[2] This method integrates steps (1) generation of refined MS/MS data by DeconMSn;[3] (2) additional refinement of the resultant MS/MS data by a modified version of PE-MMR; and (3) elimination of systematic errors of precursor masses using DtaRefinery.[4] As a result *i*PE-MMR increases sensitivity in peptide identification and provides increased accuracy when applied to complex high-throughput proteomics data. The second problem can be combated with the help of various advanced

*Reprint requests to Jong-Moon Park
E-mail: bio4647@naver.com

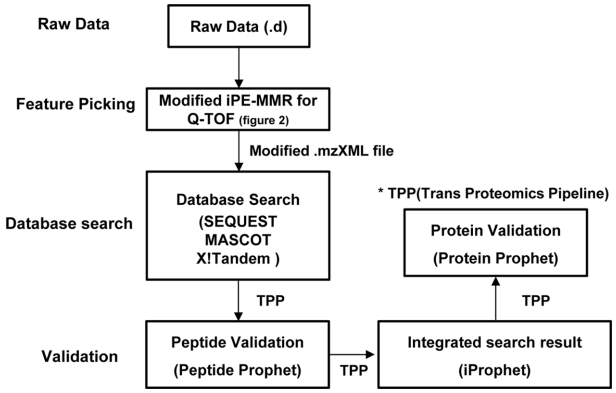

Figure 1. Overall data flowchart of modified iPE-MMR pipeline for Q-TOF data.

search engines that can identify different or overlapping sample peptides from an individual set of MS/MS spectra. Since every search engine has different MS/MS database search algorithms and statistical data methods for evaluating peptide and protein probabilities.[5] Most popular database search engines are SEQUEST,[6] Mascot,[7] and X!Tandem.[8]

The present study utilizes iProphet to combine the indication from multiple identifications of the same peptide sequences across different spectra, experiments, precursor ion charge states, and modified states. iProphet is a part of open-source suite of proteomic data analysis tools Trans Proteomic Pipeline (TPP).[9] Figure 1 illustrates the general MS/MS proteomics dataset workflow through the combined pipeline. The main features of the present combined pipeline include three distinct categories: (1) Feature picking, (2) Database search and (3) Validation. The raw data files from Q-TOF (quadrupole time of flight) were subjected to an integrated data analysis pipeline using modified iPE-MMR and the subsequent mzXML files were searched with three database search and results were validated using peptide prophet, iProphet and Protein prophet using TPP.[11] The present pipeline process data twice for identification of proteins. Initially the MS/MS raw data will be analyzed by database search engines. And then the un-identified result data will be adjusted to correct precursor peak mass by PE-MMR. After the PE-MMR result and identified result information will be combined and the modified data will be processed by DtaRefinery for more accurate precursor ion mass. There are several search software with different search algorithm and calculation of validation score lead to differences finding results. iProphet was found to solve this issue. Obviously the more spectra from each search tool that are to be analyzed together the better the statistics are going to be by InterProphet. Therefore, more identified proteins can be gained every though searching based on a same filter. This study shows that the novel pipeline is optimized for identification of proteome by tandem mass spectra data.

In this study, the integrated data processing pipeline was applied by mouse epididymal adipose tissue (EAT) and mouse colitis colon tissue. Obesity-induced chronic epididymal adipose tissue inflammation is considered as a crucial contributor to the above complications.[12] Ulcerative colitis (UC) was caused by chronic inflammation through diet condition.[13] Both of two metabolic diseases were developed by chronic inflammation in adipose tissue and colon tissue has not been integrated into molecular understanding. We therefore established an experimental animal model for colitic colon tissue and high fat feeding adipose tissue, and used proteomic analysis, based on LC-MS analysis and integrated data processing pipeline, to identify proteins involved in these sample.

Experimental section

Protein sample preparation and mass spectrometry

1. Sample information

All animal sample procedures were performed according to the protocols approved by the institutional ethical committee at Interdesciplinary Graduate Program in Genetic Engineering (IGPGE) of Seoul National University. Eight weeks old C57B6L/J wild type mice were used for the study. The animals were fed with high fat diet for 16 weeks. At the end of the experimental period the animals were euthanized and the whole intact epididymal adipose tissue (EAT) excluding testis were collected, and snap frozen quickly in liquid nitrogen, and then stored at -80^{o}C. Mouse tissues of repeated dextran sulfate sodium (DSS)-induced colitis-associated cancer were used for the application of the method.

2. Tryptic digestion

The snap frozen tissue was pulverized and dissolved in 6 M urea to denature the protein. The mixture was then sonicated and the protein content was determined by BCA assay. A known quantity of the denatured protein (100 μg) was reduced with TCEP (tris(2-carboxyethyl)phosphine) and incubated in 900 rpm, 37 for 30 min. Each sample was adjusted pH 8 to 9 by 1 M Tris. The samples were cysteine-blocked with 15 mM iodacetamide (IAA) at room temperature (RT) for 1 hour, 300 rpm in the dark. The urea concentration of the solution was made to below 2 M with 10 mM Tris. The proteins were then digested by Sequencing grade modified trypsin (Promega) in Trypsin resuspension buffer (Promega) for 16 hours at 37^{o}C, 300 rpm. The digested peptides were desalted using C18 spin column (Havard Apparatus) to remove interfering substances and the samples were dried using Speedvac (SCANVAC, Bio-Rad).

3. Liquid Chromatography Mass spectrometry analysis

The dried peptide pellet was re-suspended in 100 μL of 0.1% formic acid for mass spectrometric analysis using 6520 Accurate-Mass Q-TOF LC/MS coupled to Liquid

chromatogram (Agilent Technologies, DE) with a HPLC-chip cube source. The peptides separations were performed with an 1200 series High pressure liquid chromatography(HPLC) system (Agilent Technologies, DE, JP) using HPLC-chip (large capacity chip, 150 mm, 300 Å, C18 chip, w/160 nL trap column) (Agilent Technologies, DE) with a nanoflow pump. The peptides were primarily loaded and transferred to the HPLC-chip of trapping column at a flow rate 0.3 μL/min for a minute. Mobile phase A was HPLC grade water with 0.1% formic acid and B was 90% ACN, 0.1% formic acid in HPLC grade water. The sample were separated with gradient from 10% B to 45% B in 15 minutes, then to 90% B for 5 minutes, and finished in 10% B, at a flow rate of 0.3 μL/min. A blank run was carried out in between samples with similar conditions. Data were acquired in the mass range from 100 m/z to 3000 m/z with positive ion polarity, 3.7 V collision energy. Acquisition rate was set per a second of three spectra. Data acquisition of reference mass (121.050873 m/z, 922.009798 m/z) corrected to ensure high mass accuracy.

Tandem Mass Spectrometry Data Analysis

Feature picking

1. Generation of lists of all monoisotopic masses observed in a whole LC/MS experiment.

The MS data set arising from the LC-MS/MS experiments were directly submitted as a batch to Decon2Ls and the isotopic distributions and charge states of the peptide ions in the mass spectra were deconvoluted by using the THRASH algorithm developed by Horn et al.[14] By using *i*PE-MMR, mass spectral peaks were grouped according to similar monoisotopic mass (within a mass tolerance of 10 ppm), different LC elution times into a unique mass class (UMC). All of the UMCs observed in an LC-MS/MS experiment were recorded in an XML-formatted file (denoted herein a "UMC list"). We used a simple approach to generate UMCs by using mass tolerance and the requirement of detection in sequential mass spectra. Figure 2A and B shows modified iPE-MMR method that modified from the conventional method to combined method of data conversion that optimized to Q-TOF data analysis.

2. Pre-Normal database search

The grouped peak list from MS/MS exported to mzXML format by Trapper [Agilent MassHunter format (.d directories) to mzXML converter]. Generated peak lists (.mzXML) was searched by SEQUEST(Sorcerer^{TM}) against International Protein Index (IPI) Mouse database(v3.73) from the European Bioinformatics Institute using the following constraints: semi tryptic peptides with up to two missed cleavage sites were allowed; 20 ppm mass tolerances for MS and 50 ppm mass tolerances for MS/MS fragment ions. Database search parameters were set for carbamidomethyl (+57.021465 Da) of cysteine residues as a fixed modification. Carbamylation (+43.005814 Da) of N-terminal and oxidation (+15.994920 Da) of methionine were specified as variable modifications. Search results were evaluated with the Trans proteomic pipeline (TPP) using Peptide Prophet (v4.4.1). And then exceptions of identified spectra were used for processing next PE-MMR step.

3. Generation of MS/MS Data

MS/MS raw data were extracted using MzXML2Search(TPP v4.4 VUVUZELA rev 1, Build 201009011732 (MinGW)) of TPP software, which determines values of the monoisotopic m/z and charge states of the precursor ions along with the values of m/z and the intensities of the fragment ions. In this step, charge states from 1 to 8 were considered, and the precursor peptide mass range was set to 400-10,000 Da.

4. Filtration, correction, and refined of tandem mass spectra

In order to obtain more missed peptide feature, among unidentified spectra, garbage MS/MS data files were

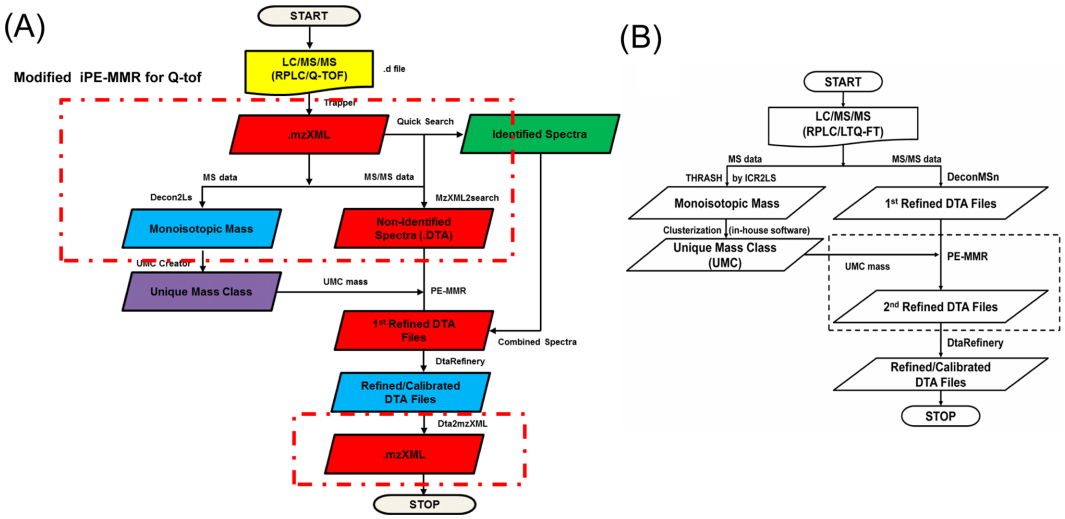

Figure 2. (A) Overall data flow chart of modified iPE-MMR for Q-TOF data (B) Conventional iPE-MMR.

filtered out by UMC mass value. The PE-MMR compares to the neutral mass (*i.e.* [M]) of the unidentified MS/MS file and the UMC mass for changing experimental [M] value to UMC mass value. If a matched precursor ion is found within a tolerance of 25 ppm, PE-MMR replaced the precursor mass of the MS/MS file with the UMC mass (mass refinement). If no match ion was found, it either uses the DTA files with no mass refinement or filters them out. We note that a single MS/MS scan can generate multiple DTA files with different precursor masses if multiple matches were found. Before processing DtaRefinery, we combined first identified MS/MS files and the finished PE-MMR unidentified MS/MS files with default settings (Figure 1A). DtaRefinery performs a preliminary database search using X!Tandem and knowledge of peptide identifications is utilized in multivariate calibration to reduce systematic instrumental mass errors.[4] And then *i*PE-MMR generates MS/MS files that are refined and calibrated, which are subsequently subjected to a several database search. In order to utilize protein database search engines for identification of peptides and proteins, whole *i*PE-MMR processed data must convert to mzXML file format.[2]

Database search

All raw MS data were converted to the mzXML and mgf file format,[15] and searched with X!Tandem version 8.4.4 (LabKey Sever), Sorcerer™ SEQUEST® v3.5 (Sage-N Research, Inc. Milpitas, CA), and Mascot version 2.3.0 (Matrix Science) the three most commonly used database search tools. Search parameters have set precursor ion and fragment ion tolerances were set to 20 ppm and 50 ppm each other. And fixed modification of cysteine carbamidomethylation, variable modifications of methionine oxidation and n-terminus carbamylation were set as modifications and allowing partially tryptic peptides. The tandem LC-MS data were processed and searched against mouse IPI version 3.73 protein database http://www.ebi.ac.uk/IPI/IPImouse.html and combined, validated by using Trans Proteomic Pipeline (TPP) version 4.4 VUVUSELA rev 1.

Validation

The Trans Proteomic Pipeline, which makes use of open XML file formats for storage of raw data at the peptide and protein levels. All search results were processed with PeptideProphet, iProphet and ProteinProphet, in an order. Each of database search results were processed individually by PeptideProphet. Several PeptideProphet results were combined in ProteinProphet with or without using iProphet as an intermediate step. A cutoff probability score of 0.95 was used for this study. It revealed a false positive rate of less than 1% based on a PeptideProphet probability score cutoff at 0.90. And then ProteinProphet, within the TPP, infers the simplest list of proteins consistent with the identified peptides.

Protein network analysis

Unique identified proteins were visualized and mapped into biological networks using the Ingenuity pathway analysis (IPA) tool of complex 'omics data version 14855783 (Ingenuity Systems, Inc., Redwood City, CA, USA). IPA is based on a proprietary, manually curated database of mouse protein-protein, protein-DNA, and protein-compound interactions. The differentially expressed proteins were uploaded as IPI ID Number into the IPA platform for analysis.

Results and Discussion

Increased accuracy and sensitivity of modified *i*PE-MMR method for peptide identification in Q-TOF

All spectra were searched against IPI using three different search engines - SEQUEST, X!Tandem and MASCOT. We compared efficiency of the peptide identification in three feature picking method such as conventional, *i*PE-MMR and combined method (Table 1A). All identified peptide and protein lists obtained by the conventional method, PE-MMR, *i*PE-MMR, combined *i*PE-MMR. *i*PE-MMR resulted in 10% increase in peptide identifications compared to the conventional method. Significantly, combined *i*PE-MMR method further increased the number of peptide identifications from 1050 to 1203 (~15% increase), within the same cut-off value of PeptideProphet error rate of 1%. Other search engines result is likewise increased number of identified peptides and proteins. In SEQUEST result, combined *i*PE-MMR method resulted in 25% increase in protein identifications compared to the conventional method (Table 1B). Distribution of ppm in Q-TOF data, improved the mass measurement accuracy (MMA) of peptide identification, with conventional extracted spectrum method, and with *i*PE-MMR method analysis (Figure 3). We compared same scan number's Xcorr values between the database search results from identified peptide of *i*PE-MMR-filtered data Xcorr values (Figure 4B) and those from the non-identified peptide of conventional method processed data (Figure 4A). After *i*PE-MMR, the distribution of unique identified peptide's Xcorr has

Table 1. Number of identified peptide (A) and protein (B).

(A)			
EAT_FAT	Conventional	iPE-MMR	Combined
SEQUEST	1050	1154	1203
X!Tandem	898	847	953
Mascot	291	315	364
iProphet	1235	1295	1402
(B)			
SEQUEST	238	308	296
X!Tandem	206	197	213
Mascot	111	112	122
iProphet	264	328	322

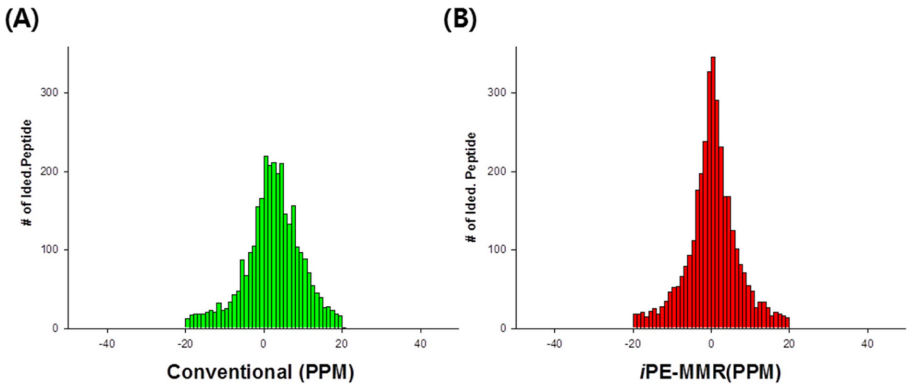

Figure 3. MMA distribution of Conventional (A), iPE-MMR (B) in Q-TOF data.

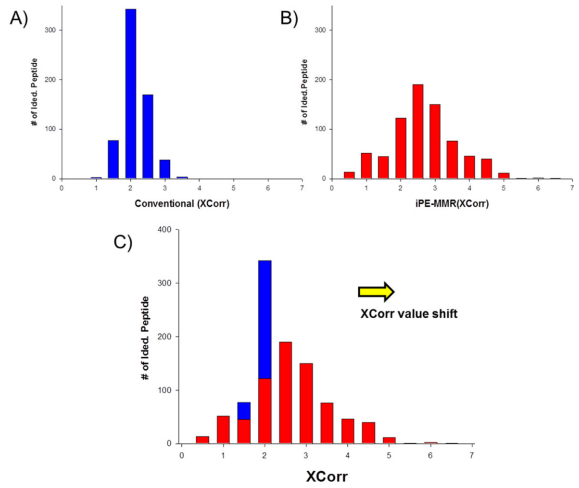

Figure 4. Xcorr score distribution of Conventional (A), iPE-MMR (B) and Comparison of Conventional and iPE-MMR (C) in Q-TOF data.

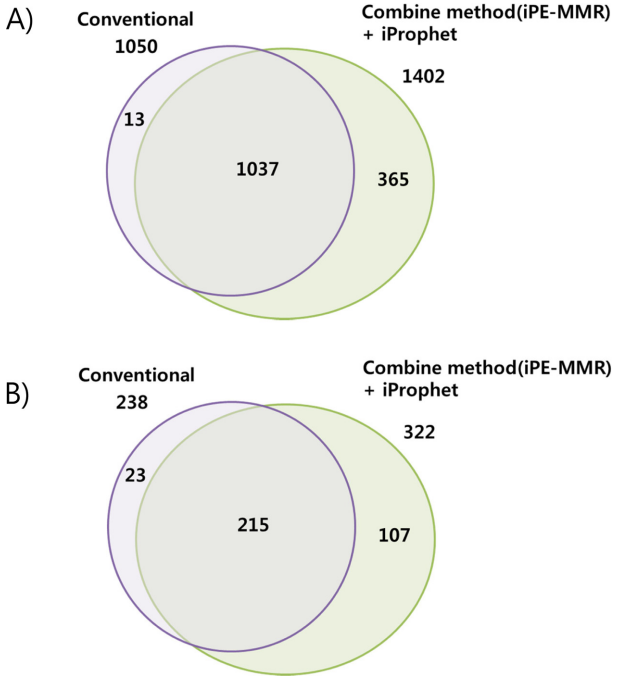

Figure 5. Venn diagram of identified peptide (A) and protein (B) of conventional and combined method.

moved to high Xcorr score, as a result, from 2.0 xcorr to 2.5 xcorr. The most crucial point of this result, we modified logic of *i*PE-MMR pipeline and applied Q-TOF MS data was processed by already developed *i*PE-MMR with suitable Fourier Transform Mass Spectrometry.

Increased sensitivity and Identification number of Peptides and Proteins by iProphet.

In order to combine identified peptides, the result was analyzed by three kinds of search engines (SEQUEST, X!Tandem and Mascot). Then PeptideProphet validates individual MS/MS spectra matches to peptide, and assign the probability of correctness for the MS/MS spectra. The functionality of 'iProphet' is combining of multiple PeptideProphet results. Number of identified proteins were compared from searched results by iProphet and conventional method for the sample of EAT tissue sample. All identified peptide and protein lists obtained by the SEQUEST, X!Tandem, MASCOT and iProphet. iProphet resulted in 18% increase in peptide identifications compared to the only SEQUEST. And the combined *i*PE-MMR method further increased the number of peptide identifications from 1050 to 1402 (~35% increase) (Figure 1A). In combined *i*PE-MMR result, iProphet method resulted in 35% increase in protein identifications compared to the only SEQUEST search engine (Table 1B). As Figure 5, mostly conventional search result was overlapped with result of iProphet. The combined method identified 365 unique peptides (Figure 5A) and 107 unique proteins (Figure 5B) whereas the conventional method showed only 13 and 23 unique peptides and proteins respectively. Which can be considered as an important factor for identification of biomarker candidate by the combined method. The combined method search results shows higher cutoff, whereas the conventional method showed lower conventional method cutoff value for the identified unique proteins.

Increase identified efficiency of low abundance protein in cell through advanced pipeline

Overall number of identified result of combined method was increased more than conventional method result in gene level. Figure 6 shows distribution of identified proteins of cellular localization. The majority of proteins were placed in cytoplasm. A few proteins were spread throughout the other parts (*i.e.* membrane, extracellular and

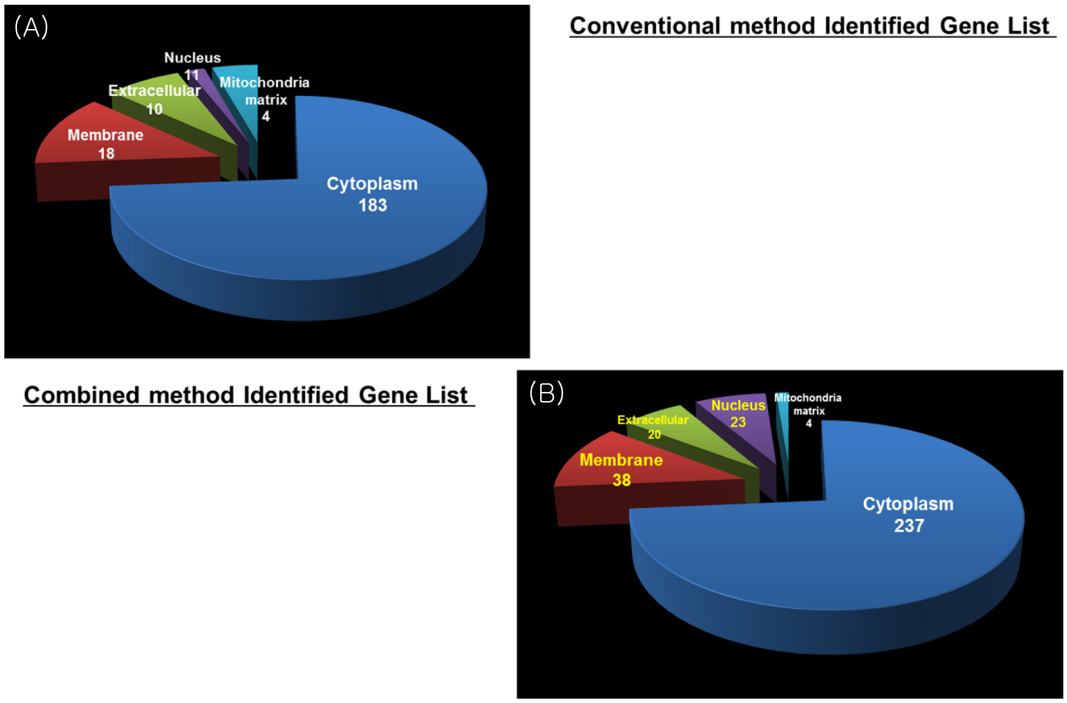

Figure 6. Distribution of identified protein cell localization in conventional (A) and combined method (B).

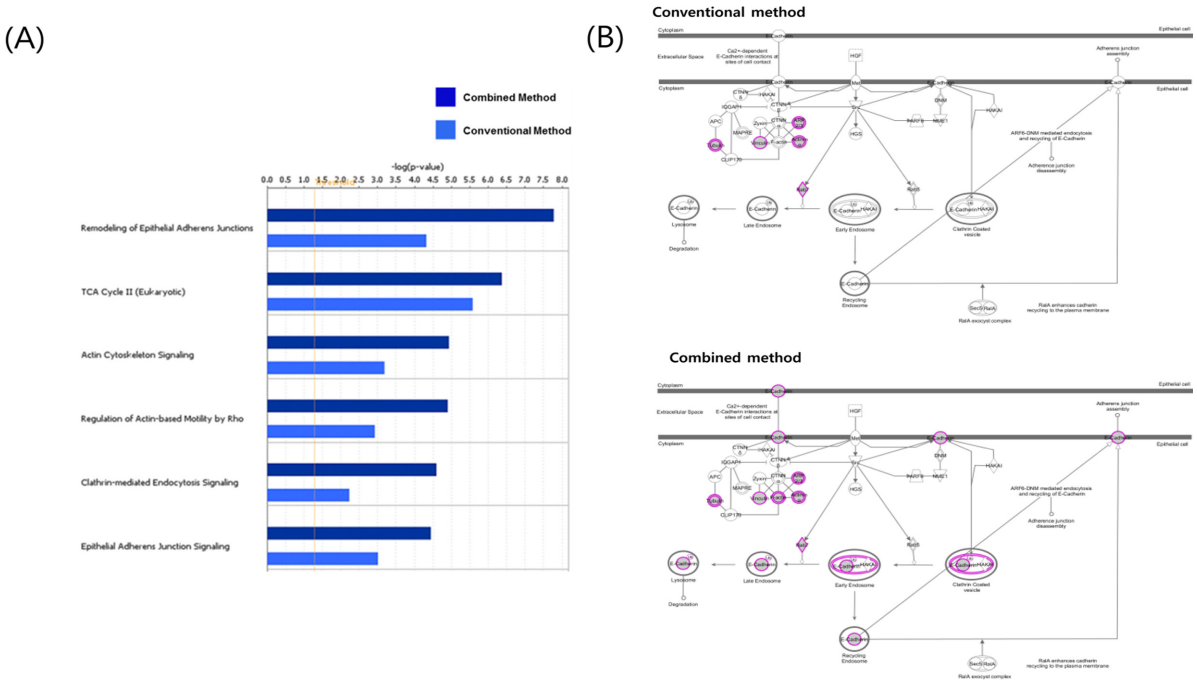

Figure 7. Comparison of conventional method and combined method of identified proteins in IPA. Related with functional ontology canonical pathway list from experiment data (A), top-ranking canonical pathway (Remodeling of Epithelial Adherens Junctions) in ranked pathway list and identified protein were marked by shaded symbol and red edge (B).

nucleus). Through data analysis of combined method, low abundance proteins were identified almost twice as many as number of result of conventional method.

More effectual data mining through protein network analysis

In the study, protein profiling was processed by two kinds of data analysis method from mouse adipose tissue. Conventional method and combined method results compared gene information of identified protein list and defines common and unique subsets followed by functional ontology enrichment, calculation of p-value and network analysis (Figure 7). As shown Figure 7A, it arranges canonical pathway list in order of high ranking of p-value. Top-ranking pathway is remodeling of epithelial adherens junctions in the list. E-cadherin is a calcium-dependent transmembrane protein involved in homotypic cell-cell interactions.[15] E-cadherin at the surface of islet beta-cells is controlled by secretagogues including glucose, correlates with insulin secretion, and can serve as a surface marker of beta-cell function.[16] As shown in Figure 7B, marked shaded proteins were identified in data from mouse adipose tissue. This result demonstrates convincingly that combined method is more efficient for finding biomarker candidates than result of single search engine. Only combined method result data appeared known biomarker candidates such as E-cadherin. Two kinds of analysis results were shown similar pattern in signaling pathway analysis. However identified proteins of combined method were demonstrated higher p-value than conventional analysis method. The main reason is that analyzed proteins were more accurate to interesting biomarker candidates in combined method. As efficiency of combined analysis method evidence, E-cadherin is potential clinical diagnostic biomarker of diabetic nephropathy and may participate in the pathogenesis of diabetes.[17]

(A)

Sample	#	Conventional	iPE-MMR	Combined
Normal	Peptide	1364	1409	1829
	Protein	404	440	557
RDSS	Peptide	**1252**	1270	**1796**
	Protein	**453**	469	**598**

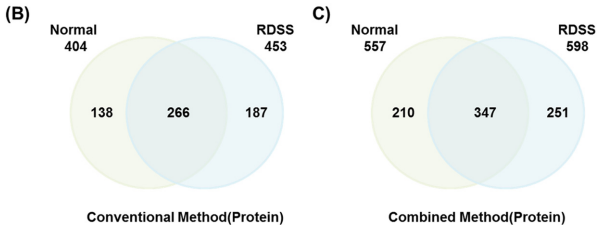

Figure 8. Number of identified peptides and proteins in mouse colon tissue (A), Venn diagram of identified peptide and protein of conventional (B) and combined method (C).

Application of colon tissue of RDSS induced ulcerative colitis sample

Mouse colon tissue was analyzed by conventional method and combined method. Two groups of sample are control and treated RDSS for induced ulcerative colitis. Increased number of peptides and proteins were identified in control and treated samples in combined method compared to the conventional method (Figure 8). Figure 8A shows that identified proteins were distributed between conventional method and combined method. Figure 8B and Figure 8C show that combined method was able to obtain missed proteins in conventional method. These identified proteins built related with

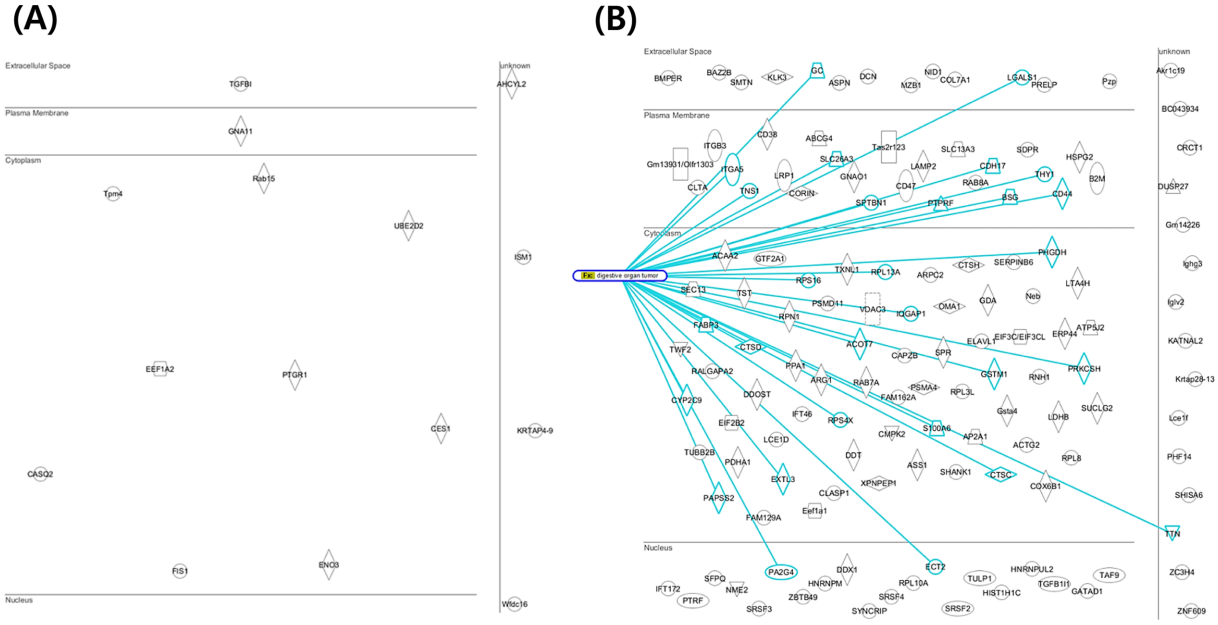

Figure 9. Comparison of network of unique identified proteins between conventional method and combined method related with disease of digestive organ tumor, (A) conventional method and (B) combined method.

inflammatory bowel disease of protein network for mining biological meaning. Figure 9 shows unique identified protein network of related with digestive organ tumor were compared between conventional method and combined method. Figure 9A and Figure 9B shows that the results of combined method showed significant involvement digestive disease. The result shows that the combined method is more efficient for discovering biomarker.

Conclusion

This novel integrated data processing pipeline can assist in improved discovery of potential biomarker candidates. In proteomic analysis based on mass spectrometry, we still missed many low abundance proteins. Most of biomarker candidates are low abundance and hence present low intensity peaks in MS analysis. Hence, the pipeline was demonstrated in disease pathway analysis by enriching the pathways more specifically. Because this pipeline could be identified not only more protein numbers but also more accurate result.

Acknowledgment

Funds provided by the Ministry of Health and Welfare (HI11V-0005-010013), Fountain Research Section of National Research Foundation (No. 2013044311) through Green Technology of the Ministry of Education, Science and Technology, Korea and the National Research Foundation (No. 2013035851) through Ministry of Education, Science and Technology, Korea.

References

1. Shin, B.; Jung, H. -J.; Hyung, S. -W.; Kim, H.; Lee, D.; Lee, C.; Yu, M. -H.; Lee, S. -W. *Mol. Cell. Proteomics* **2008**, 7, 1124.
2. Jung, H. -J.; Purvine, S. O.; Kim, H.; Petyuk, V. A.; Hyung, S. -W.; Monroe, M. E.; Mun, D. -G.; Kim, K. -C.; Park, J. -M.; Kim, S. -J.; Tolic, N.; Slysz, G. W.; Moore, R. J.; Zhao, R.; Adkins, J. N.; Anderson, G. A.; Lee, H.; Camp, D. G.; Yu, M. -H.; Smith, R. D.; Lee, S. -W. *Anal. Chem.* **2010**. 82, 8510.
3. Mayampurath, A. M.; Jaitly, N.; Purvine, S. O.; Monroe, M. E.; Auberry, K. J.; Adkins, J. N.; Smith, R. D. *Bioinforma. Oxf. Engl.* **2008**, 24, 1021.
4. Petyuk, V. A.; Mayampurath, A. M.; Monroe, M. E.; Polpitiya, A. D.; Purvine, S. O.; Anderson, G. A.; Camp, D. G.; Smith, R. D. *Mol. Cell. Proteomics* **2010**, 9, 486.
5. Shteynberg, D.; Deutsch, E. W.; Lam, H.; Eng, J. K.; Sun, Z.; Tasman, N.; Mendoza, L.; Moritz, R. L.; Aebersold, R.; Nesvizhskii, A. I. *Mol. Cell. Proteomics* **2011**, 10, M111.007690.
6. Eng, J. K.; McCormack, A. L.; Yates, J. R. *J. Am. Soc. Mass Spectrom.* **1994**, 5, 976.
7. Pappin, D. J.; Hojrup, P.; Bleasby, A. J. *Curr. Biol.* **1993**, 3, 327.
8. Bjornson, R. D.; Carriero, N. J.; Colangelo, C.; Shifman, M.; Cheung, K. -H.; Miller, P. L.; Williams, K. *J. Proteome Res.* **2008**, 7, 293.
9. Keller, A.; Shteynberg, D. *Methods Mol. Biol.* **2011**, 694, 169.
10. Pedrioli, P. G. A.; Eng, J. K.; Hubley, R.; Vogelzang, M.; Deutsch, E. W.; Raught, B.; Pratt, B.; Nilsson, E.; Angeletti, R. H.; Apweiler, R.; Cheung, K.; Costello, C. E.; Hermjakob, H.; Huang, S.; Julian, R. K.; Kapp, E.; McComb, M. E.; Oliver, S. G.; Omenn, G.; Paton, N. W.; Simpson, R.; Smith, R.; Taylor, C. F.; Zhu, W.; Aebersold, R. *Nat. Biotechnol.* **2004**, 22, 1459.
11. Keller, A.; Nesvizhskii, A. I.; Kolker, E.; Aebersold, R. *Anal. Chem.* **2002**, 74, 5383.
12. Bao, B.; Chen, Y.G.; Zhang, L.; Na, Xu Y.L.; Wang, X.; Liu, J.; Qu, W. *PLoS One* **2013**, 8, e84075.
13. Yeo, M.; Kim, D. K.; Park, H. J.; Oh, T. Y.; Kim, J. H.; Cho, S. W.; Paik, Y. K.; Hahm, K. B. *Proteomics* **2006**, 6, 1158.
14. Jaitly, N.; Mayampurath, A.; Littlefield, K.; Adkins, J. N.; Anderson, G. A.; Smith, R. D. *BMC Bioinformatics* **2009**, 10, 87.
15. Angst, B. D.; Marcozzi, C.; Magee, A. I. *J. Cell Sci.* **2001**, 114, 629.
16. Bosco, D.; Rouiller, D. G.; Halban, P. A. *J. Endocrinol.* **2007**, 194, 21.
17. Jiang, H.; Guan, G.; Zhang, R.; Liu, G.; Cheng, J.; Hou, X.; Cui, Y. *Diabetes Metab. Res. Rev.* **2009**, 25, 232.