JOURNAL OF INTEGRATED OMICS
A METHODOLOGICAL JOURNAL
HTTP://WWW.JIOMICS.COM

ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v7i1.200

Comparative integrated omics approach sterically understanding hepatic metabolic dynamics in mouse model

Shoko Takahashi[1,2], Huijuan Jia[2], Kenji Saito[2] and Hisanori Kato[2]*

[1] Department of Applied Biological Chemistry, Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1, Yayoi, Bunkyo-ku, Tokyo, Japan. [2] Food for Life, Organization for Interdisciplinary Research Projects, The University of Tokyo, 1-1-1, Yayoi, Bunkyo-ku, Tokyo, Japan.

Received: 20 September 2016 **Accepted:** 24 November 2016 **Available Online**: 18 January 2017

ABSTRACT

Currently, omics-fusion, which is the combined analysis of data by employing multiple omics analyses, has been available, and it can enable a more fundamental understanding of the biological phenomena than a single omics. However, multi-layered combination of multiple omics technologies involves generation of a large amount of data, which leads to increased complexity and makes comprehension of bio-information more difficult. The objective of this study was to investigate the utility of incorporating multiple omics technologies in a multi-layered fashion. Transcriptomic, proteomic, and metabolomic analyses were carried out using a mouse model of diet-induced obesity. The present study reported the comprehensiveness of three omics analyses and the utility of using multiple omics analyses. Some uniform changes among different omis were observed, but the majority of changes were specific to each omics approach. This data supports the fact that various molecules progress through the central dogma at differing speeds. Since the time axis differs for each molecule, combining multiple omics analyses makes it possible to investigate the reactions in organisms three-dimensionally. At first glance, it simply appears that combining a number of very large data sets produces even more complexity but, if multi-layered omics data are treated with an awareness of their meaning, benefits, and limitations, then the combination of multiple omics analyses can be extremely useful for research in molecular biology.

Keywords: transcriptomics, proteomics, metabolomics, multi-omics .

Abbreviations: CE-TOF MS (Capillary electrophoresis time-of-flight mass spectrometry), FDR (False discovery rate), HF (High-fat diet), TRAQ (Isobaric tags for relative and absolute quantitation), LC-TOF MS (Liquid chromatography time-of-flight mass spectrometry), ND (Normal diet), PPARg (peroxisome proliferator-activated receptor gamma), RMA (Robust Multi-array Average).

1. Introduction

The Genome Project revealed the genomic sequences a variety of living organisms, including humans, and this information has fueled research into the comprehensive understanding of genomes. Appending the suffix 'omics' to the subject of study gives rise to the specific research area, such as genomics, transcriptomics, proteomics, and metabolomics. Such omics studies can aid the understanding of the influence of drug or food on homeostasis or the metabolic system, their role in disease prevention, and the relationship between the individual's genotype and disease [1-3]. As this type of omics approach is utilized in various research tasks, such as evaluating the functionality of drug or food, clarifying its mechanisms of action, and predicting toxicity, it can be proposed that with the spread of omics technologies, demand for such research will increase in the fields of lifescience research. Currently, omics-fusion, which is the combined analysis of data by employing multiple omics analyses, has been available, and it can enable a more fundamental understanding of the biological phenomena than a single omics. Most recently, omics analysis tool PGMiner has been published [4]. Furthermore, some integrated omics studies have been reported, for example, integration of protein, mRNA and miRNA [5], transcriptome and metagenome [6], transcriptome and proteome [7-9].

***Corresponding author:** Hisanori Kato, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan. Phone number: +81 358411607. Email Address: akatoq@mail.ecc.u-tokyo.ac.jp

However, multi-layered combination of multiple omics technologies involves generation of a large amount of data, which leads to increased complexity and makes comprehension of bio-information more difficult. The objective of this study was to investigate the utility of incorporating multiple omics technologies in a multi-layered fashion. Our group has been doing various research about the prevention of lifestyle diseases including diabetes and metabolic disorder [10]. The utility of multiple omics technologies may be a powerful tool for the research region about the prevention of lifestyle diseases like obesity and metabolic syndrome, in the field of nutriomics like this study, when the lifestyle choices are properly correlated. Therefore, transcriptomic, proteomic, and metabolomic analyses were carried out using a mouse model of diet-induced obesity, which is one of the most commonly utilized models in obesity research. Additionally, a verification experiment on multi-layered omics was performed by comparison and investigation of these data.

2. Material and Method

2.1. Animal experiments

Male C57BL/6J mice purchased from Charles River Laboratories Japan, Inc., at 7 weeks of age were divided into two groups, the normal diet group (ND group) was fed D12450B (10 kcal% fat, Research Diets) and the high-fat diet group (HF group) was fed D12492 (60 kcal% fat, Research Diets). They were housed individually at a controlled temperature of 23±1°C under a 12-h light-12-h dark cycle. After fed ad libitum for 9 weeks, on the last day of the experiment, after a 16-h food deprivation and a 1.5-h re-feeding, the mice were deeply anesthetized with an intraperitoneal injection of sodium pentobarbital. The livers were then excised and proceeded to each omics analysis. The same samples were analyzed in transcriptomics, proteomics, and metabolomics studies (n=3 for each group). All animal experiments were carried out in accordance with the guidelines of the Animal Usage Committee of the Faculty of Agriculture of the University of Tokyo.

2.2. Transcriptomics using DNA microarray data

DNA microarray data used in the present study were obtained in the previous experiment, which were carried out with Affymetrix GeneChip Mouse Genome 430 2.0 array (Affymetrix, Santa Clara, CA, USA), which has 45,000 probe sets and can analyze the expression level of over 39,000 transcripts and variants from over 34,000 well characterized mouse genes [10]. Briefly, total RNA was isolated from the livers of these mice using TRIzol Reagent (Invitrogen Life Technologies, Tokyo, Japan). RNA of each individual was reverse-transcribed to the first-strand complementary DNA using SuperScript II RT (Invitrogen Life Technologies, Tokyo, Japan). Second-strand complementary DNA synthesis was then carried out using a DNA polymerase. Biotinylated complementary RNA was generated from the complementary DNA using a BioArray HighYield RNA transcript labeling kit (Enzo Life Sciences, Farmingdale, NY, USA) according to the standard Affymetrix protocols. The obtained intensity files were analyzed using the statistical analysis software R. After normalization of the intensity files was performed with robust multi-array average (RMA) normalizing [11], and then a clustering analysis was done with hclust (hierarchical clustering, average linkage). Rank products was used for intergroup comparison [12]. The expression change was taken as informative when the false discovery rate (FDR) value was < 0.1. To obtain detailed molecular information and infer significant signaling pathways from global profiling results of DNA microarray data, we uploaded differentially expressed gene probes according to the above criteria to DAVID (The Database for Annotation, Visualization and Integrated Discovery, https://david.ncifcrf.gov/).

2.3. Proteomics

We performed the differential proteomic analysis of the mouse livers using Isobaric tags for relative and absolute quantitation (iTRAQ), the same methods as our previous study [13]. iTRAQ are a non-gel-based technique used to quantify proteins from different sources in a single experiment. It uses isotope-coded covalent tags. Total protein was extracted by using lysis buffer and separated by centrifugation at 12,000 × g for 30 min at 4 °C. Protein concentrations were determined using the Bradford assay, and pooled protein were proceed to iTRAQ experiment kit, performed according to the manufacture's protocol (AB SCIEX). Desalted samples were vacuum evaporated, and added 50 µl of 0.1% formic acid, and 2 µl of samples were measured with LC/MS/MS (TripleTOFTM 5600 + System with Eksigent nanoLC). ProteinPilot™ (AB SCIEX) was used to identify proteins and calculate protein expression levels by comparing *in silico* peptide data. The number of peptides used to identify proteins were shown in Table 2.

2.4. Metabolomics

Frozen mice liver samples were transferred into 500 µL of methanol containing 50 µM of external standard. After homogenization by BMS-M10N21 (bms, Tokyo) at 1,500 rpm, 120 s five times, 500 µL of chloroform and 200 µL of ultra-pure water were added to the homogenate and mixed well and centrifuged at 2,300 g for 5 min at 4°C. The resultant water phases were ultrafiltrated by the Millipore Ultrafree-MC PLHCC HMT Centrifugal Filter Device, 5 kDa (Millipore, Billerica, MA). The filtrates were dried and dissolved in 50 µL of ultra-pure water. We then subjected the samples obtained to capillary electrophoresis time-of-flight mass spectrometry (CE-TOFMS) analysis using the Agilent CE-TOFMS system (Agilent Technologies, Santa Clara, CA)

Table 1. GO enriched terms of differentially expressed genes

Increased	*p*-value	Decreased	*p*-value
Steroid biosynthesis	7.10E-21	Phosphoprotein	1.80E-09
Sterol biosynthetic process	4.20E-20	Cell fraction	3.50E-07
Steroid metabolic process	5.40E-20	Insoluble fraction	3.90E-06
Sterol biosynthesis	3.50E-19	Cytoplasm	5.20E-06
Steroid biosynthetic process	1.50E-18	Membrane fraction	2.10E-05
Cholesterol biosynthetic process	2.20E-17	Endoplasmic reticulum	3.60E-05
Sterol metabolic process	1.70E-16	Ubl conjugation	9.00E-05
Cholesterol biosynthesis	4.30E-16	Stress response	1.00E-04
Cholesterol metabolic process	8.20E-14	PPAR signaling pathway	1.90E-04
Lipid synthesis	1.50E-13	Endoplasmic reticulum	2.20E-04
Lipid biosynthetic process	1.40E-12	Microsome	2.20E-04
Oxidoreductase	1.60E-12	Membrane	2.40E-04
Endoplasmic reticulum	2.60E-12	Golgi apparatus	2.60E-04
Steroid biosynthesis	4.50E-12	Vesicular fraction	3.00E-04
Oxidation reduction	5.30E-12	Lipoprotein	4.60E-04
Endoplasmic reticulum	1.40E-10	Acetylation	5.00E-04
Terpenoid backbone biosynthesis	4.00E-08	Steroid metabolic process	5.60E-04
Transferase activity, transferring alkyl or aryl (other than methyl) groups	8.90E-08	Regulation of hydrolase activity	7.00E-04
Microsome	1.80E-07	Basolateral plasma membrane	7.10E-04
Cell fraction	2.00E-07	Cytosol	1.20E-03
Vesicular fraction	2.90E-07	Golgi apparatus	1.20E-03
Isoprenoid metabolic process	3.20E-07	Steroid dehydrogenase activity	1.50E-03
Isoprenoid biosynthetic process	7.70E-07	Nucleotide binding	2.00E-03
Insoluble fraction	9.20E-07	Small GTPase mediated signal transduction	2.10E-03
Membrane fraction	1.30E-06	Trophectodermal cell differentiation	2.50E-03
Binding site:Substrate	3.10E-06	Histone H3	2.60E-03
Peroxisome	5.90E-06	H3	2.70E-03
Microbody	5.90E-06	NADP	2.90E-03
Acetylation	1.40E-05	Leukocyte transendothelial migration	3.00E-03
Peroxisome	1.70E-05	Extrinsic to membrane	3.10E-03

*P-value is Fisher's exact test

Table 2. The list of increased or decreased proteins by high-fat diet.

Protein	Name	Protein ID	# of Peptides	Fold Change	p-value
AL1A1	Retinal dehydrogenase 1	P24549	66	2.001	0.000
CAH3	Carbonic anhydrase 3	P16015	71	1.969	0.000
BHMT1	Betaine--homocysteine S-methyltransferase 1	O35490	76	1.907	0.000
HBB1	Hemoglobin subunit beta-1	P02088	79	1.879	0.018
THIL	Acetyl-CoA acetyltransferase, mitochondrial	Q8QZT1	44	1.548	0.000
AL3A2	Fatty aldehyde dehydrogenase	P47740	17	1.538	0.003
RLA2	60S acidic ribosomal protein P2	P99027	35	1.524	0.022
OAT	Ornithine aminotransferase, mitochondrial	P29758	27	1.518	0.039
RL7A	60S ribosomal protein L7a	P12970	5	1.506	0.066
PRDX6	Peroxiredoxin-6	O08709	51	1.489	0.003
NLTP	Non-specific lipid-transfer protein	P32020	70	1.475	0.001
PDIA1	Protein disulfide-isomerase	P09103	57	1.468	0.001
RL23A	60S ribosomal protein L23a	P62751	9	1.448	0.035
HCD2	3-hydroxyacyl-CoA dehydrogenase type-2	O08756	17	1.403	0.078
GSTA3	Glutathione S-transferase A3	P30115	22	1.400	0.084
APOA1	Apolipoprotein A-I	Q00623	20	1.380	0.001
GGLO	L-gulonolactone oxidase	P58710	7	1.378	0.003
PHS	Pterin-4-alpha-carbinolamine dehydra-	P61458	13	1.373	0.067
EF1A1	Elongation factor 1-alpha 1	P10126	39	1.369	0.015
DHB5	Estradiol 17 beta-dehydrogenase 5	P70694	31	1.367	0.051
DHSO	Sorbitol dehydrogenase	Q64442	45	1.359	0.003
AATM	Aspartate aminotransferase, mitochondrial	P05202	45	1.356	0.000
DECR	2,4-dienoyl-CoA reductase, mitochondrial	Q9CQ62	21	1.331	0.029
TTC36	Tetratricopeptide repeat protein 36	Q8VBW8	18	1.329	0.044
TALDO	Transaldolase	Q93092	8	1.315	0.083
GLNA	Glutamine synthetase	P15105	30	1.313	0.011
FAAA	Fumarylacetoacetase	P35505	44	1.286	0.004
CK054	Ester hydrolase C11orf54 homolog	Q91V76	10	1.275	0.054
CALR	Calreticulin	P14211	28	1.273	0.046
PDIA3	Protein disulfide-isomerase A3	P27773	50	1.263	0.042
MAOX	NADP-dependent malic enzyme	P06801	13	0.357	0.084
FAS	Fatty acid synthase	P19096	75	0.475	0.000
ACLY	ATP-citrate synthase	Q91V92	24	0.517	0.000
KPYR	Pyruvate kinase isozymes R/L	P53657	30	0.581	0.001
ACOT1	Acyl-coenzyme A thioesterase 1	O55137	4	0.589	0.014
ECHP	Peroxisomal bifunctional enzyme	Q9DBM2	49	0.610	0.000
OPLA	5-oxoprolinase	Q8K010	3	0.618	0.070
GSTP1	Glutathione S-transferase P 1	P19157	69	0.662	0.024
GPDM	Glycerol-3-phosphate dehydrogenase, mitochondrial	Q64521	8	0.686	0.014
RS13	40S ribosomal protein S13	P62301	6	0.701	0.025
APOA4	Apolipoprotein A-IV	P06728	6	0.711	0.078
ODP2	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex, mitochondrial	Q8BMF4	9	0.717	0.098
CY1	Cytochrome c1, heme protein, mitochondrial	Q9D0M3	17	0.741	0.070
PCCB	Propionyl-CoA carboxylase beta chain, mitochondrial	Q99MN9	14	0.747	0.092
DHAK	Bifunctional ATP-dependent dihydroxyacetone kinase/ FAD -AMP lyase (cyclizing)	Q8VC30	35	0.805	0.062
GLCTK	Glycerate kinase	Q8QZY2	4	0.812	0.061
PBLD2	Phenazine biosynthesis-like domain-containing protein 2	Q9CXN7	16	0.818	0.098
ALDOB	Fructose-bisphosphate aldolase B	Q91Y97	77	0.821	0.035
ANXA6	Annexin A6	P14824	14	0.839	0.086
GLGB	1,4-alpha-glucan-branching enzyme	Q9D6Y9	11	0.839	0.057
DPYD	Dihydropyrimidine dehydrogenase [NADP+]	Q8CHR6	7	0.843	0.087
PYC	Pyruvate carboxylase, mitochondrial	Q05920	73	0.870	0.041
CPSM	Carbamoyl-phosphate synthase [ammonia], mitochon-	Q8C196	360	0.872	0.004

at 4°C. The alignment of detected peaks was performed according to the m/z value and normalized migration time. The relative area value of each peak was calculated and used for the intergroup comparison. Samples that were obviously characterizing outliers were eliminated from the analysis. Metabolite extraction, MS analysis, and data analysis were performed in Human Metabolome Technologies.

2.5. Integrated analysis of transcriptomics and metabolomics

We used the web-based tool Keggle (http://keggle.jp), a novel tool for the visualization of omics-data created by the author's group [14]. Transcriptome and metabolome data were mapped onto KEGG (Kyoto Encyclopedia of Genes and Genomes, http://www.genome.jp/kegg/) pathways on the web.

3. Results

3.1. Animal experiments

The final body weights were higher in the HF group compared to the ND group. The C57BL/6J mouse used in this study is regarded as relatively sensitive to high-fat diets and suitable as a non-genetic obese animal model when fed a high-fat diet. The body weight gain of the mice fed a high-fat diet (60% energy) for 9 weeks was 17.2 g, whereas that of the mice fed a normal diet was 4.59 g, suggesting that the mice fed a high-fat diet in this study can be regarded as a valid obesity model for research (Supplemental Figure S1).

3.2. Results of DNA microarray data

Among the total 817 differentially expressed gene probes, 480 gene probes showed increases, and 337 gene probes showed decreases in HF group. These differentially expressed genes were uploaded to DAVID functional annotation tools and the enrichment analysis was performed (Table 1). The top 10 of enriched terms of increased genes were Steroid biosynthesis, Sterol biosynthetic process, Steroid metabolic process, Sterol biosynthesis, Steroid biosynthetic process, Cholesterol biosynthetic process, Sterol metabolic process, Cholesterol biosynthesis, Cholesterol metabolic process, Lipid synthesis, lipid biosynthetic process.

3.3. Results of proteome analysis using iTRAQ method

The proteome of mouse liver were analyzed using iTRAQ labeling and LC/MS/MS. Among identified 1043 proteins, 50, 23 proteins were significantly increased and decreased by HF, respectively (p-value < 0.1). In these proteins which showed changes, the most differentially expressed proteins were shown in Table 2. Among the increased proteins, ALDH1A1, HSD17B10, APOA1, AKR1C4, GOT2 and ACSL1 are related to the oxidation of fat. Furthermore, lipid transport proteins, including SCP2, APOA1, GOT2 and ACSL1, and oxidative stress proteins including PRDX2 and PRDX6 were increased. In the decreased proteins, there were APOA4, ACLY, FASN and DECR1, which related to fatty acid synthesis, PKLR and PCX, which related to glucose metabolism, ME1 and ACLY, which related to TCA cycle. These alterations mean that lipid oxidation, transport and oxidative stress were increased and TCA cycle, fatty acid synthesis and glucose metabolism were decreased in the liver tissue of obese mouse.

3.4. Results of metabolome analysis

We performed a metabolome analysis to explore the hepatic metabolic alterations underlying the effects of HF diet. Among the peaks obtained from the CE- and LC-TOF MS analysis, 385 peaks were identified according to the value of m/z and MT from metabolite database. Of these metabolites, 26 showed changes, 10 were increased and 16 were decreased (Table 3). N-acetylglutamate (N-AcGlu), which is a positive regulator of the urea cycle, was decreased by the high-fat diet. Similar changes were found in the abundance of ornithine, citruline, argininosuccinate (ArgSuccinate), and arginine, which are the intermediates of the urea cycle, and the final product, urea. In diet-indused obese animal, it is well known that the urea production is decreased by the dysfunction of urea cycle related enzymes [15].

Furthermore, the abundance of choline was decreased by HF diet. Choline is known to be a nutrient, which prevent the accumulation of hepatic triglyceride, and reported to be decreased in hepatosteatosis mouse liver [16].

4. Discussion

It is widely accepted amongst molecular biologists that analysis based on these omics analyses is extremely effective, and rapid technological innovation such as high-throughput DNA sequencing and high-precision electron impact mass spectrometry as well as improvements in DNA microarray chips have led to its frequent utilization in various fields. Prior to the popularization of omics technology, research was carried out by focusing on a target molecule from a wider range of biomolecules, but the utilization of omics has led to a significant increase in the number of molecules that can be captured. Omics analysis is not a tool for arriving at conclusions, but is rather an approach for unbiased screening. Even though currently DNA microarray probably would've been easier to analyze data through many analysis tools, using RNA-seq technology would become less biased. Furthermore, taking into account the complexity of the interactions between food, pharmaceuticals, and the organism, methods that provide a complete overview of the influences on the biomolecules should lead to rapid research progress. However, while the use of omics technology has

Table 3. The list of increased or decreased metabolites by high-fat diet

KEGG ID	Compound name	Fold change	p-value
No ID	1H-Imidazole-4-propionic acid	3.693	0.006
C00167	UDP-glucuronic acid	2.565	0.018
C00149,C00497, C00711	Malic acid	2.048	0.029
C00122	Fumaric acid	2.014	0.037
C01879	5-Oxoproline	1.863	0.010
No ID	Heptadecanoic acid	1.649	0.037
C00307	CDP-choline	1.503	0.032
C00262	Hypoxanthine	1.503	0.034
No ID	Stearoyl ethanolamide	1.435	0.016
C01081	Thiamine phosphate	1.376	0.017
No ID	2-Hydroxyisobutyric acid	0.852	0.014
C00780	Serotonin	0.831	0.010
C00086	Urea	0.786	0.022
C00114	Choline	0.753	0.032
No ID	Homoserinelactone	0.669	0.014
C03425	Methyl oleate	0.643	0.029
C00042	Succinic acid	0.641	0.049
C00601	Phenylacetaldehyde	0.632	0.036
C01026	N,N-Dimethylglycine	0.569	0.029
C05568	Imidazolelactic acid	0.465	0.007
C03406	Argininosuccinic acid	0.440	0.007
C02835	Imidazole-4-acetic acid	0.424	0.042
No ID	AC(22:0)	0.412	0.008
C02592	Taurolithocholic acid	0.321	0.030
C00489	Glutaric acid	0.273	0.015
C01921	Glycocholic acid	0.241	0.035

spread to molecular biology and numerous studies utilizing omics technology have been conducted, this technology was not employed effectively in a large number of studies. Therefore, it is important to have a clear understanding of the problems that omics can solve.

Fundamentally, when the research objective or task (the 'why') is clear, the four aspects of 'where', 'when', 'what', and 'how' on which the analysis will be carried out are of importance. For example, in an experiment conducted in this study, these aspects are—what will be fed and on which organ or cell the changes will be observed (where), when will the analysis take place in the experiment schedule (when), which molecule is observed in the selected cell or time (what), and how the data will be analyzed (how). If, the research design is broken down into these four aspects, the factor that omics can solve is the 'what' factor, i.e., a large number of biomolecules can be captured at once. In molecular biology research, omics technology does not solve the 'where' or 'when', overwhelmingly solves the 'what', and makes solving the 'how' more difficult. Employing multiple omics in a multi-layered fashion will broaden the field of observation for the 'what' and make the 'how' more difficult due to increased complexity. Because a large amount of data on biomolecules can be gathered, knowledge about bioinformatics becomes necessary and a situation arises where the acquired data is so vast that it becomes difficult to analyze. The discernment of false-negatives and false-positives also becomes a problem. Even though a variety of tools for performing omics analysis have become prevalent, and tools such as the Ingenuity pathway analysis (IPA, http://www.ingenuity.com), KeyPathwayMiner (http://tomcat.compbio.sdu.dk/keypathwayminer/) [17, 18] and DAVID are used in most research labs, the problem of 'how' still remains a challenge for the future. Despite the presence of these kinds of benefits and challenges, when the complexity of the central dogma is taken into consideration, it can be proposed that there would be large benefits from employing omics in a multi-layered fashion at multiple levels such as mRNA, protein, and metabolite.

In this experiment, omics analysis was performed at three levels—gene expression, protein expression, and metabolite formation. When considering the comprehensiveness of these omics approaches, it is first necessary to consider that the number and characteristics of the molecules, which are the targets of these omics, are all different. Figure 1 represents the molecule count for the subject of analysis and the number of molecules that were detectable in this experiment for the three stages (molecule count is presented as logarithm). The number of genes in the mouse used in this animal experiment is said to be approximately 30,000 [19], with over 20,000 being detectable using DNA microarray. DNA micro array is said to have a higher comprehensiveness than other omics, but the reason for this is that the molecules of interest all have the same properties. Additionally, it is said that there are between 500,000 to 1 million types of proteins, with approximately 1,000 being able to be detected by proteomics in this experiment (Figure 1)—the lowest comprehensiveness as compared to

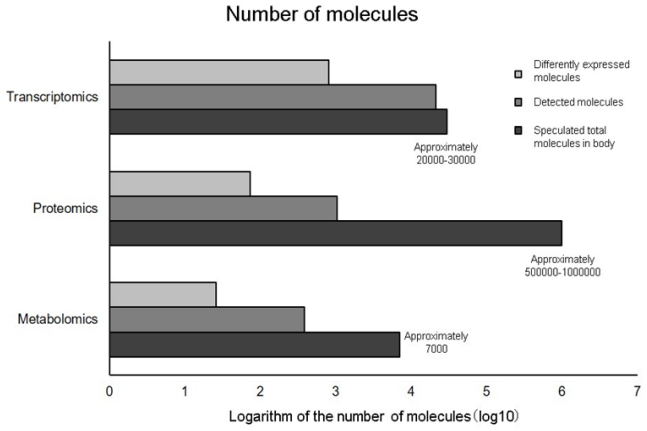

Figure 1. The comprehensiveness of each omics technique.
Molecule count for the subject of analysis and the number of molecules that were detectable in this experiment for the three stages are represented (molecule count is presented as logarithm).

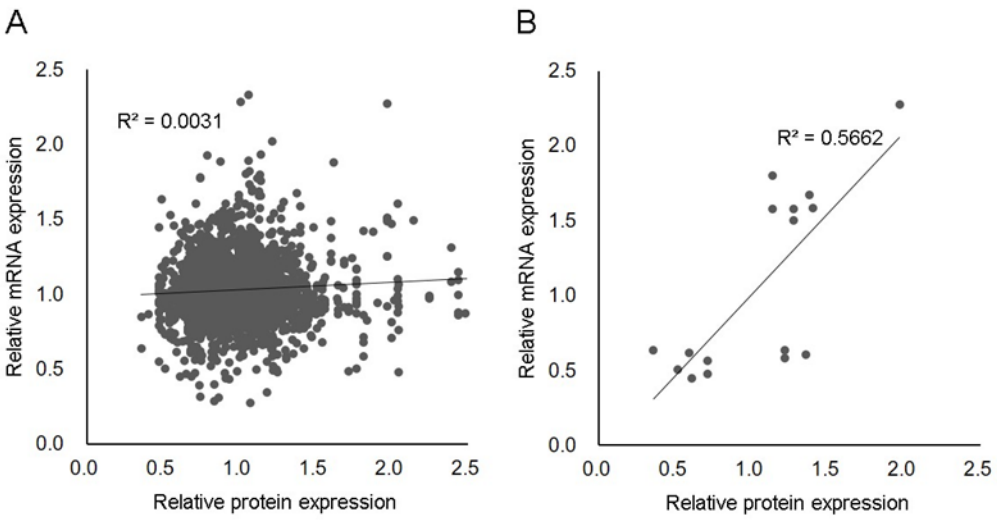

Figure 2. A correlation plot of the change in scaling factors of the transcriptomics and proteomics data.
(A) The plot of the whole data, **(B)** Extracting and plotting the significant changes in the transcriptomics and proteomics data.

transcriptomics. The reason for this is that physical properties of each protein differ, and from a technical point of view, a method with high comprehensiveness has not yet been established. Furthermore, according to a study by Wishart et al. (2007) [20], approximately 7,000 molecules (approximately 2500 metabolites, 1200 pharmaceuticals, and 3500 food based elements) existed as target metabolite molecules in metabolomics. As the physical properties of metabolites differ by each molecule, there is no one technique, which can comprehend everything at the same time. In this experiment, metabolomics was carried out through the combination of CE-TOF/MS and LC-TOF/MS, resulting in the successful detection of 385 metabolites (Figure 1). Comparing the comprehensiveness of the three omics analyses, we noted that proteomics has the lowest comprehensiveness. However, there are striking innovations in current proteome analysis technology including the iTRAQ method employed in this study, which is a shotgun method in proteomics that uses stable radioactive isotopes. Currently, differences in comprehensiveness do exist but they will be resolved over time as technology improves.

When conducting research involving multiple omics, it is necessary to be aware that the molecule counts will differ and for technological reason, large differences in comprehensiveness will exist between various omics analyses. Specifically, attention should be given to the method of applying multiple omics data. It is essential to understand that there is a reason for the common change that exists in the multiple omics data and to approach the issue from this point. As noted above, the reason for this is that there are large differences in the comprehensiveness between omics; to have a common difference means to be tied to the weakness of the statistical power of the omics approach with the lowest comprehensiveness. The next section discusses the presence of uniform changes for different omics approaches at different stages when dealing with multiple omics data.

This study compared the transcriptomics and proteomics data in order to consider the uniformity between results of different omics. First, a correlation plot of the change in scaling factors of the transcriptomics and proteomics data was created (Figure 2). The plot of the whole data does not show any particular correlation (Figure 2-A). Extracting and plotting the significant changes in the transcriptomics and proteomics data reveals a high level of correlation (Figure 2-B). Molecules exhibiting changes in the same direction (increase/decrease in expression) were extracted from the data and those exhibiting common changes for both transcriptomics and proteomics, were listed in Supplemental Table S1. There were eleven molecules where a common change was noted in the mRNA and protein levels, summing up to 15.1% of the number of changes in protein count observed through proteomics.

As is the case for the molecules in Supplemental Table S1, a number of things can be said about the fact that a common change can be observed among multiple stages. Because they possess a quality, which makes it possible to observe the same change at multiple stages at the exact moment the analysis occurs, it can be said that there is no need to conduct analyses at multiple stages for molecules with the same observed change. However, in fact, the same change can be observed which will increase the reliability of the data. The fact that the changes in the mRNA level can also be observed in the protein level (which catalyzes the actual response in the organism) makes it easier to understand the meaning of the mRNA change. Further, the change in mRNA also provides support as to why the change in protein level occurs. Obtaining a single change from

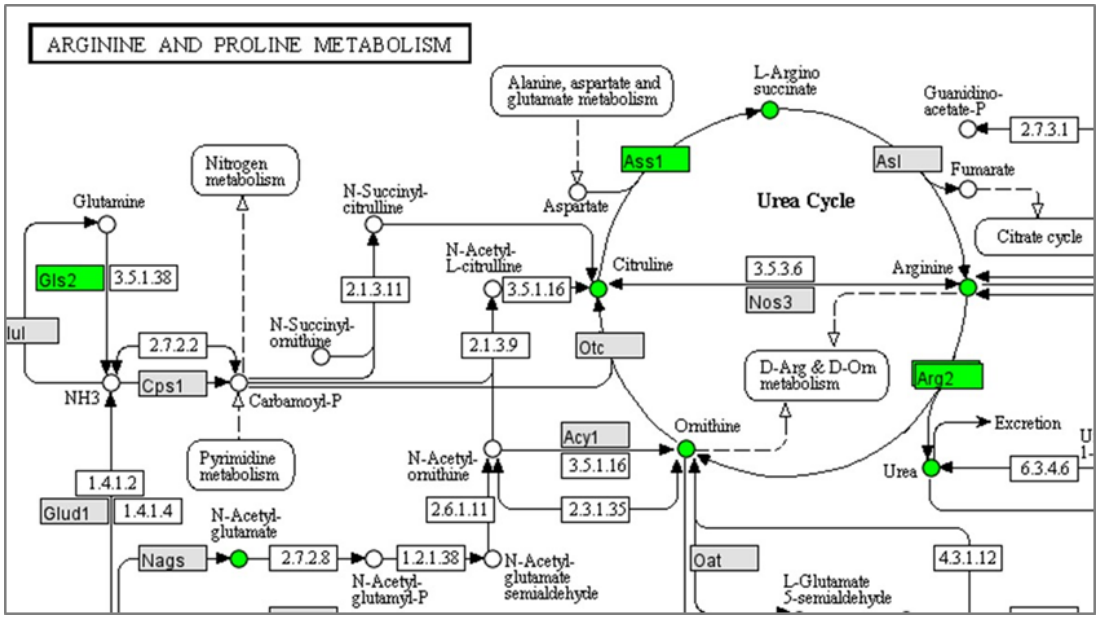

Figure 3. The metabolic and transcriptomic change in the urea cycle.
The metabolic and transcriptomic change in the urea cycle were represented by mapping simultaneously on a KEGG metabolism map using Keggle analysis tool. Decreased expression of transcripts or metabolites showed green color.

multiple stages enables observation from a three-dimensional perspective. One needs to be aware of the above point regarding the comprehensiveness when considering uniformity. While it is true that uniformity in multiple omics data leads to higher reliability, this will not mean that high reliability of data will lead to uniform results. The molecules given in Supplemental Table S1 were not extracted because of the high reliability of their changes; they also were not extracted because they were the most meaningful out of all those present in the reactions in the organism. In fact, the molecules were extracted because they constituted an overlapping portion of the low comprehensiveness of the proteomics with the comprehensiveness of the transcriptomics. Certainly, it is a type of important information, but it is not useful in terms of sensitivity of screening. Evidence exists that carrying out multiple omics analyses enables observation. For example, in this study, performing metabolomics made it possible to capture the metabolite change in the urea cycle. Figure 3 represents the metabolic and transcriptomic change in the urea cycle, mapped simultaneously on a KEGG metabolism map. There was no up-regulation in the urea cycle pathway. It was not possible to focus only on the urea cycle throughout the results of transcriptomics and proteomics, but being able to observe the metabolite change enabled the detection of a new change. The fact that there are aspects that can only be observed when multiple omics approaches are applied can be viewed as a disadvantage of using a single omics technique. However, in this context, multi-layered omics data should be interpreted with an understanding of its meaning, benefits, and limitations.

In contrast to the above results, the results for multi-layered omics approaches are not uniform. This section discusses the meaning of a non-uniform change. The transcriptomics and proteomics data given above reveal that they are almost non-uniform. While high correlation was observed between changes that were statistically significant, in almost all of the cases, one was not significantly different. The protein levels observed to be changed in the proteomics data were 15.1%, which were also observed to be changed in the transcriptomics data. From a reverse perspective, approximately 85% changes are specific to proteomics and are non-uniform between the two sets of omics data. In this case, when different results are derived from multiple omics, the differences in detection sensitivity—post-translation modification, oxidization of the protein, the stability of the protein, or metabolite—are the focus of debate and it is definitely true that such causes also play a large part. However, the essential meaning can be understood as the deviation in the time axis of the central dogma for each of the molecules. For example, in the case of molecule A and molecule B reacting to some kind of stimulus, the speed of the response to the stimulus and the speed of the transcription will differ. In the case of an immediate early gene, which rapidly responds to stimulus, the amount of mRNA will peak at an early stage after receiving the stimulus. However, a two-dimensional gene, which is transcriptionally controlled by some kind of transcriptional product, will slow down. In a study published in Nature, analysis of the speed of translation to protein was carried out on individual genes; the results showed that the speed of translation differed for each molecule [21]. Accordingly, for multiple omics data analyzed at the same time point it is impossible for all molecules to exhibit the same change. This is the paradox of multi-layered omics. Even though we tried to capture multiple stages of the biomolecule comprehensively, the fact that each molecule has a different speed means that it is impossible to comprehensively capture

the central dogma of the biomolecule.

However, this paradoxical aspect of multi-layered omics may actually be its largest benefit. If all molecules exhibited the same change, then there would be no need to use multiple omics. Precisely, because various molecules progress through the central dogma at differing speeds, it is extremely unlikely that the changes in all molecules would be uniform under multiple omics approaches and that the changes should not be uniform. This is why there is a necessity to use multiple omics analyses. The fundamental importance of using multiple omics approaches is not because it is possible to observe differing stages, it is because the time axis differs for each molecule that it is possible to capture the meaning of an organism which cannot be grasped using a single omics analysis.

5. Concluding Remarks

The present study reported the comprehensiveness of three omics analyses (transcriptomics, proteomics, and metabolomics) and the utility of using multiple omics analyses. Because comprehensiveness differs widely across these omics approaches, it is necessary to be aware of this when using multiple omics approaches. Furthermore, uniform changes were observed among changes at all stages but the majority of these specific to the omics approach. This data supports the fact that various molecules progress through the central dogma at differing speeds. Because the time axis differs for each molecule, combining multiple omics analyses makes it possible, for the first time, to investigate the reactions in organisms three-dimensionally. At first glance, it simply appears that combining a number of very large data sets produces even more complexity but, as discussed in this paper, if multi-layered omics data are treated with an awareness of their meaning, benefits, and limitations, then the combination of multiple omics analyses can be extremely useful for research in molecular biology. We hope that the knowledge shared in this paper can be of value in future research utilizing multiple omics analyses.

6. Supplementary material

Supplemental Figure S1. The food intake and the body weight of each group. (A) total food intake, (B) body weight.

Supplemental Table S1. The list of molecules that showed common changes between transcriptomics and proteomics.

References

1] G.A. Evans, Nat Biotechnol, 18 (2000) 127. doi: 10.1038/72480

2] G. N., Curr Opin Biotechnol., 13 (2002) 4

3] H. Kato, Asia Pac J Clin Nutr, 17 Suppl 1 (2008) 12-15

4] C. Has, J. Allmer, Information Sciences, (2016). doi: 10.1016/j.ins.2016.08.005

5] M. Schouten, P. Bielefeld, S.A. Fratantoni, C.J. Hubens, S.R. Piersma, T.V. Pham, R.A. Voskuyl, P.J. Lucassen, C.R. Jimenez, C.P. Fitzsimons, Sci Data, 3 (2016) 160068. doi: 10.1038/sdata.2016.68

6] F. Turroni, C. Milani, S. Duranti, L. Mancabelli, M. Mangifesta, A. Viappiani, G.A. Lugli, C. Ferrario, L. Gioiosa, A. Ferrarini, J. Li, P. Palanza, M. Delledonne, D. van Sinderen, M. Ventura, ISME J, 10 (2016) 1656-1668. doi: 10.1038/ismej.2015.236

7] D. Kumar, A.K. Yadav, X. Jia, J. Mulvenna, D. Dash, Mol Cell Proteomics, 15 (2016) 329-339. doi: 10.1074/mcp.M114.047126

8] E. Montague, I. Janko, L. Stanberry, E. Lee, J. Choiniere, N. Anderson, E. Stewart, W. Broomall, R. Higdon, N. Kolker, E. Kolker, Nucleic Acids Res, 43 (2015) D1145-1151. doi: 10.1093/nar/gku1175

9] M.P. van Iersel, M. Sokolovic, K. Lenaerts, M. Kutmon, F.G. Bouwman, W.H. Lamers, E.C. Mariman, C.T. Evelo, J Integr Bioinform, 11 (2014) 235. doi: 10.2390/biecoll-jib-2014-235

10] S. Takahashi, K. Egashira, K. Saito, H. Jia, K. Abe, H. Kato, Journal of Functional Foods, 6 (2014) 157-167. doi: 10.1016/j.jff.2013.10.002

11] R.A. Irizarry, B. Hobbs, F. Collin, Y.D. Beazer-Barclay, K.J. Antonellis, U. Scherf, T.P. Speed, Biostatistics, 4 (2003) 249-264. doi: 10.1093/biostatistics/4.2.249

12] K. Kadota, Y. Nakai, K. Shimizu, Algorithms Mol Biol, 3 (2008) 8. doi: 10.1186/1748-7188-3-8

13] H. Jia, W. Aw, M. Hanate, S. Takahashi, K. Saito, H. Tanaka, M. Tomita, H. Kato, Journal of Functional Foods, 11 (2014) 438-448. doi: 10.1016/j.jff.2014.09.018

14] H. Kato, Takahashi, S. and Saito K. J., Trad. Comp. Med., (2011) 6

15] T. Barber, J.R. Vina, J. Vina, J. Cabo, Biochem J, 230 (1985) 675-681

16] J.F. Cobbold, Q.M. Anstee, R.D. Goldin, H.R. Williams, H.C. Matthews, B.V. North, N. Absalom, H.C. Thomas, M.R. Thursz, R.D. Cox, S.D. Taylor-Robinson, I.J. Cox, Clin Sci (Lond), 116 (2009) 403-413. doi: 10.1042/CS20080159

17] N. Alcaraz, J. Pauling, R. Batra, E. Barbosa, A. Junge, A.G. Christensen, V. Azevedo, H.J. Ditzel, J. Baumbach, BMC Syst Biol, 8 (2014) 99. doi: 10.1186/s12918-014-0099-x

18] M. List, N. Alcaraz, M. Dissing-Hansen, H.J. Ditzel, J. Mollenhauer, J. Baumbach, Nucleic Acids Res, 44 (2016) W98-W104. doi: 10.1093/nar/gkw373

19] C. Mouse Genome Sequencing, R.H. Waterston, et al, Nature, 420 (2002) 520-562. doi: 10.1038/nature01262

20] D.S. Wishart, D. Tzur, C. Knox, R. Eisner, A.C. Guo, N. Young, D. Cheng, K. Jewell, D. Arndt, S. Sawhney, C. Fung, L. Nikolai, M. Lewis, M.A. Coutouly, I. Forsythe, P. Tang, S. Shrivastava, K. Jeroncic, P. Stothard, G. Amegbey, D. Block, D.D. Hau, J. Wagner, J. Miniaci, M. Clements, M. Gebremedhin, N. Guo, Y. Zhang, G.E. Duggan, G.D. Macinnis, A.M. Weljie, R. Dowlatabadi, F. Bamforth, D. Clive, R. Greiner, L. Li, T. Marrie, B.D. Sykes, H.J. Vogel, L. Querengesser, Nucleic Acids Res, 35 (2007) D521-526. doi: 10.1093/nar/gkl923

21] B. Schwanhausser, D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen, M. Selbach, Nature, 495 (2013) 126-127. doi: 10.1038/nature11848