Application of NMR Metabolomics to Pathological and Nutritional Fields

（NMR メタボロミクスの病理学・栄養学分野への応用）

Yosuke Komatsu

A thesis submitted to the Graduation School of Life Science,

Hokkaido University for the degree of Doctor of Philosophy

March 2021

Contents

Abbreviations

ALB	Albumin
ASL	Argininosuccinate lyase
ASS1	Argininosuccinate synthase 1
CD	Crohn's disease
CDO1	Cysteine dioxygenase type 1
CE	Capillary electrophoresis
CN	Casein
CT	Control
CV-ANOVA	Cross-validated analysis of variance
ELISA	Enzyme-linked immunosorbent assays
FL	Fucosyllactose
FMO	Flavin-containing monooxygenase
GC-MS	Gas chromatography
GO	Gene ontology
HDL	High-density lipoprotein
HMO	Human milk oligosaccharide
IBD	Inflammatory bowel disease
ICR	Institute of Cancer Research
Ig	Immunoglobulin
LC	Liquid chromatography
LDL	Low-density lipoprotein
LF	Lactoferrin

LP	Low protein
LYZ	Lysozyme
MCT	Monocarboxylic acid transporter
NMR	Nuclear magnetic resonance
OPLS-DA	Orthogonal partial least squares discriminant analysis
OTUs	Operational taxonomic units
PBS	Phosphate-buffered saline
PC	Principal components
PCA	Principal component analysis
S.E.M.	Standard error of the mean
SCFA	Short-chain fatty acids
SL	Sialyllactose
TMAO	Trimethylamine N-oxide
3-(Trimethylsilyl)propionic acid-d 4 sodium salt	TSP

Chapter 1 General introduction

1.1 Introduction

Metabolomics is the technology that comprehensively analyzes low-molecular-weight metabolites in biological samples such as blood, urine, saliva, feces, and milk, and provides their profiles to help understand physiological responses [1–6]. The total number of metabolites found in a living organism is considerably smaller than the number of genes and proteins, and varies greatly depending on the species. It has been reported that the number of metabolites in each specie is about 1000 in *Escherichia coli*, about 5000 in *Arabidopsis thaliana*, and about 3000 in human [7, 8]. By definition, metabolites do not contain DNA, RNA, and protein themselves, but these degradation products and fragments are the analysis targets. Metabolites include compounds such as amino acids, amino acid derivatives, nucleic acids, sugars, lipids and organic acids that can be affected by endogenous factors (digestion, absorption, metabolism, excretion, immune response, etc.) and exogenous factors (food intake, nutrients, food additives, drugs, environmental molecules, infections, intestinal flora, etc.) [9–13], and the molecular weight of metabolites is approximately 1500 kDa or less. Conventionally, various analytical methods for individually measuring metabolites have been used, and the metabolomic analysis is to measure these compounds "as many as possible". However, metabolites are usually a mixture of various compounds with different polarities, volatility, water- and fat-solubility, so no analytical techniques which can measure all metabolites do not exist. Liquid chromatography-mass spectrometry (LC-MS), gas chromatography-mass spectrometry (GC-MS), capillary electrophoresis-mass spectrometry (CE-MS), and nuclear magnetic resonance (NMR) spectroscopy have been used as analysis methods for these metabolites. Each analysis method has its own characteristics, and there are differences in the complication of pretreatment, resolution, sensitivity, and reproducibility (**Table 1-1**). LC-MS method can detect metabolites such as neutral and fat-soluble nucleic acids, fatty acids,

phosphate compounds, and organic acids, and divide the mixture into fractions according to elution conditions. However, there are problems such as low analytical reproducibility, the complexity of solvent and column selection, and incompatibility with MS in the ionic compound analysis [14]. GC-MS method has a relatively high resolution and enables stable measurement. In particular, it is suitable for measuring volatile metabolites such as low-molecular-weight organic acids and aromatic compounds. However, when measuring non-volatile metabolites, derivatization is necessary, which may adversely affect quantification [15]. CE-MS method is capable of measuring ionic metabolites with small sample amount and is highly quantitative, with little influence of ion suppression. However, this method has the disadvantages that identified targets are limited to the ionic metabolites and the measurement of a sample with a high salt concentration is difficult [16]. Although NMR metabolomics has a lower resolution than MS-based methods, it has the advantages of simple pretreatment, higher analytical reproducibility, and non-contact and non-destructive analysis [17].

Metabolites are located in most downstream of the central dogma, and the information obtained from metabolomics is closer to the phenotype than those obtained from genomics, transcriptomics, and proteomics. Therefore, metabolomics is considered the most suitable method for obtaining the current information in the living body. The advantages of using metabolomics are that metabolite changes occur more rapidly and dramatically than changes in genomes and proteins, and that metabolites themselves have little difference among species [9]. Metabolomics has been applied to a wide range of fields such as medicine and food science by using a targeting and/or a non-targeting analysis. In particular, biomarkers that characteristically fluctuate in diseases have been searched by applying the targeting analysis of metabolomics in the medical field [18, 19]. Biomarkers are biological indicators that characterize certain diseases and biological changes, and understanding changes in the levels of metabolites in body fluids are particularly useful for diagnosis, treatment, and prevention of diseases. Most of the small molecule biomarkers discovered in past studies could be

analyzed with current metabolomics techniques. For example, urinary sarcosine, which is elevated as a result of glycine N-methyltransferase induction in prostate cancer patients, has been picked as a prostate cancer biomarker by using metabolomics technology [20]. Metabolomic analysis in a cohort study suggested that branched-chain amino acids and aromatic amino acids in the blood were significantly higher in diabetic patients than non-diabetic and could be applied as biomarkers [21]. Urine microalbumin has been used to diagnose renal disorders. However, when the standard value is exceeded, nephropathy has already progressed, so biomarkers that can detect renal damage early are needed. By metabolomic analysis, 43 types of novel metabolites were changed in plasma of patients with chronic nephritis, and 1-methyladenosine, N-acetylglucosamine, γ-butyrobetaine, sebacic acid, and cis-aconitic acid were identified as biomarker candidates [22]. Besides, metabolomics may be used to predict the onset of heart diseases and to manage their conditions. It has been reported that collagen-derived peptides are increasing in the urine of patients with coronary artery disease, and these peptides could be applied as non-invasive biomarkers. Furthermore, this suggests that collagen in the body is involved in the mechanism of arteriosclerosis progression [23]. 4-Hydroxyproline in blood was found as a biomarker for non-ST-segment acute myocardial infarction. 4-Hydroxyproline inhibits the binding of low-density lipoprotein (LDL) to lipoprotein accumulated in the blood vessel wall and has the function of removing LDL from the arteriosclerosis site, so its reduction may progress arteriosclerosis [24]. These studies show that the metabolite biomarkers of diseases may be involved in not only the prediction and diagnosis but also the onset and worsening, which are valuable. In addition to medical fields, metabolomics has also been applied to evaluate the quality and functionality of foods in the field of food science. Among them, the application of metabolomics to alcoholic beverages have been reported as an objective quality evaluation method. A study has been reported that created a quality evaluation model by combining sensory tests and metabolomic analyses of multiple types of sake [25]. This study shows that each brand contains characteristic metabolites

patterns, and the content of amino acids correlates with off-flavor. Furthermore, other study on wine have reported that each grape specie shows clear separation by non-targeting metabolomic analysis and principal component analysis (PCA). Interestingly, grapes from different harvest years showed different metabolites patterns despite the same grape specie [26]. The application of metabolomics to food science is very significant in that qualitative evaluations such as sensory tests can be replaced with quantitative ones such as changes in metabolites concentration and their pattern, this might contribute to further improvement in food quality. Furthermore, metabolomics has been utilized to elucidate the fermentation mechanism of fermented foods such as beer and cheese. A previous study on beer fermentation has reported breed improvements in beer yeast with reduced odor components while increasing the amount of antioxidant nitrite [27]. Another research group invested changes in metabolites under aging conditions of Cheddar cheese with different manufacturing processes, and constructed a model of sensory characteristics during aging [28].

As described above, studies applying metabolomics to other fields such as medicine and food science are increasing, and expectations for metabolomics research continue to rise. On the other hand, metabolomics, which captures phenomena in the living body as “snapshots,” has the limitation that the direction of the metabolic reaction is unknown. Even if the comparison targets are compared under certain specific conditions, and the difference in metabolite concentration becomes clear, the actual metabolic flow is still unknown. To solve this problem, it may be necessary to add time-varying elements by performing sample collection and metabolomic analysis over time. Alternatively, it may be necessary to add the explanation for metabolite fluctuations by using transcriptomic analysis to verify metabolic enzyme changes.

Table 1-1. Characteristics of analytical methods in metabolomics

Method	NMR	LC	GC	CE
Target	unlimited	hydrophobic compounds	volatile compounds	ionic compounds
Derivatization	unnecessary	unnecessary	necessary (non-volatile compounds)	unnecessary
Resolution	+	+++	+++	+++
Distribution	impossible	possible	impossible	impossible
Reproducibility	+++	++	++	++

1.2 Aims of thesis

The purpose of this study is to search for reliable biomarker candidates that reflect pathological and nutritional conditions, and further elucidate the mechanism to fluctuate these metabolites using NMR metabolomics. Since biological metabolites vary depending on diseases and nutritional states, they could be applied as biomarkers for diagnosing medical conditions and malnutrition. These biomarkers will enable early diagnosis and treatment of diseases as well as appropriate nutritional intervention for undernutrition. Furthermore, this thesis aims to clarify a part of the nutritional interaction through breast milk between healthy mother and child. These interactions of milk-microbiome-metabolites will help understand the effects of breastfeeding on the infant development and could also play a crucial role in the design of nutritional and functional foods for infants.

1.3 Overview of thesis

In Chapter 2, SAMP/YitFc mice were used as a spontaneous model of Crohn's disease, and ileal tissues and feces were collected over time. The ileal tissues were stained with hematoxylin–eosin for histologic characterization with CD progression. Feces samples were subjected to 16S rDNA microbial analysis and ^{1}H NMR metabolomic analysis to investigate fluctuations in gut microbiota and fecal metabolites with CD progression. In Chapter 3, urine, plasma and liver metabolomics analysis and

liver transcriptomics analysis were performed using model rats with a potential protein undernutrition. Firstly, urine metabolites that respond protein undernutrition at an early stage were identified. Subsequently, among these urinary metabolites, those supported by plasma and liver metabolites and liver gene expressions were picked up as reliable biomarkers. In Chapter 4, the effects of human nutrition on infants were observed in human clinical trial. The effects of metabolites in breast milk on gut microbiome and fecal metabolites of breast-fed infant were investigated by using ^{1}H NMR metabolomic analysis, 16S rDNA microbial analysis, and clustering analysis. Finally, the findings obtained in this study were summarized.

1.4 References

[1] Karu N, Deng L, Slae M, Chi A, Sajed T, Huynh H, et al. Analytica Chimica Acta. A review on human fecal metabolomics: Methods, applications and the human fecal metabolome database. Anal Chim Acta. 2018;1030:1–24.

[2] Duarte IF, Diaz SO, Gil AM. NMR metabolomics of human blood and urine in disease research. J Pharm Biomed Anal. 2014;93:17–26.

[3] Worley B, Powers R. Multivariate analysis in metabolomics. Curr Metabolomics. 2013;1(1):92–107.

[4] Sundekilde UK, Larsen LB, Bertram HC. NMR-based milk metabolomics. Metabolites. 2013;3(2):204–22.

[5] Sugimoto M, Wong DT, Hirayama A, Soga T, Tomita M. Capillary electrophoresis mass spectrometry-based saliva metabolomics identified oral, breast and pancreatic cancer-specific profiles. Metabolomics. 2010;6(1):78–95.

[6] Weckwerth W. Metabolomics in system biology. Annu Rev Plant Biol. 2003;54:669–89.

[7] Feist AM, Henry CS, Reed JL, Krummenacker M, Joyce AR, Karp PD, et al. A genome-scale

metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. Mol Syst Biol. 2007;3:121:1−18.

[8] Kell DB. Systems biology, metabolic modelling and metabolomics in drug discovery and development. Drug Discov Today. 2006;11(23–24):1085–92.

[9] Wishart DS. Metabolomics for investigating physiological and pathophysiological processes. Physiol Rev. 2019;99(4):1819–75.

[10] Wishart DS. Emerging applications of metabolomics in drug discovery and precision medicine. Nat Rev Drug Discov. 2016;15(7):473−84.

[11] Scalbert A, Brennan L, Manach C, Andres-lacueva C, Dragsted LO, Draper J, et al. The food metabolome: a window over dietary exposure. Am J Clin Nutr. 2014;99(6):1286–308.

[12] Antunes LCM, Arena ET, Menendez A, Han J, Ferreira RBR, Buckner MMC, et al. Impact of *Salmonella* infection on host hormone metabolism revealed by metabolomics. Infect Immun. 2011;79(4):1759–69.

[13] Kaddurah-daouk R, Kristal BS, Weinshilboum RM. Metabolomics: A global biochemical approach to drug response and disease. Annu Rev Pharmacol Toxicol. 2008;48:653–83.

[14] Dettmer K, Aronov PA, Hammock BD. Mass spectrometry-based metabolomics. Mass Spectrom Rev. 2007;26(1):51–78.

[15] Kanani H, Chrysanthopoulos PK, Klapa MI. Standardizing GC-MS metabolomics. J Chromatogr B Analyt Technol Biomed Life Sci. 2008;871(2):191–201.

[16] Ramautar R, Somsen GW, de Jong GJ. CE-MS in metabolomics. Electrophoresis. 2009;30(1):276–91.

[17] Moco S, Vervoort J, Moco S, Bino RJ, De Vos RCH, Bino R. Metabolomics technologies and metabolite identification. Trends Anal Chem. 2007;26(9):855–66.

[18] Johnson CH, Ivanisevic J, Siuzdak G. Metabolomics: beyond biomarkers and towards

mechanisms. Nat Rev Mol Cell Biol. 2016;17(7):451–9.

[19] Armitage EG, Barbas C. Metabolomics in cancer biomarker discovery: Current trends and future perspectives. J Pharm Biomed Anal. 2014;87:1–11.

[20] Sreekumar A, Poisson LM, Rajendiran TM, Khan AP, Cao Q, Yu J, et al. Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. Nature. 2009;457(7231):910–4.

[21] Wang TJ, Larson MG, Vasan RS, Cheng S, Rhee EP, Mccabe E, et al. Metabolite profiles and the risk of developing diabetes. Nat Med. 2011;17(4):448–53.

[22] Toyohara T, Akiyama Y, Suzuki T, Takeuchi Y, Mishima E, Tanemoto M, et al. Metabolomic profiling of uremic solutes in CKD patients. Hypertens Res. 2010:33(9):944–52.

[23] Muhlen C Von, Schiffer E, Zuerbig P, Kellmann M, Brasse M, Meert N, et al. Evaluation of urine proteome pattern analysis for its potential to reflect coronary artery atherosclerosis in symptomatic patients. J Proteome Res. 2009;8(1):335–45.

[24] Vallejo M, García A, Tuñón J, García-Martínez D, Angulo S, Martin-Ventura JL, et al. Plasma fingerprinting with GC-MS in acute coronary syndrome. Anal Bioanal Chem. 2009;394(6):1517–24.

[25] Sugimoto M, Koseki T, Hirayama A, Abe S, Sano T, Tomita M, et al. Correlation between sensory evaluation scores of Japanese *Sake* and metabolome profiles. J Agric Food Chem. 2010;58(1):374–83.

[26] Cuadros-inostroza A, Giavalisco P, Hummel J, Eckardt A, Willmitzer L, Pen H. Discrimination of wine attributes by metabolome analysis. Anal Chem. 2010;82(9):3573–80.

[27] Yoshida S, Imoto J, Minato T, Oouchi R, Sugihara M, Imai T, et al. Development of bottom-fermenting saccharomyces strains that produce high SO_2 levels, using integrated metabolome and transcriptome analysis. Appl Environ Microbiol. 2008;74(9):2787–96.

[28] Ochi H, Sakai Y, Koishihara H, Abe F, Bamba T, Fukusaki E. Monitoring the ripening process of Cheddar cheese based on hydrophilic component profiling using gas chromatography-mass spectrometry. J Dairy Sci. 2013;96(12):7427–41.

Chapter 2 Disease progression-associated alterations in fecal metabolites in SAMP1/YitFc mice, a Crohn's disease model

2.1 Abstract

Crohn's disease (CD) is a chronic, relapsing inflammatory bowel disease affecting the gastrointestinal tract. Although its precise etiology has not been fully elucidated, an imbalance of the intestinal microbiota has been known to play a role in CD. Fecal metabolites derived from microbiota may be related to the onset and progression of CD. This study aimed to clarify the transition of gut microbiota and fecal metabolites associated with disease progression using SAMP1/YitFc mice, a model of spontaneous CD. The ileum tissues isolated from SAMP1/YitFc mice at different ages were stained with hematoxylin-eosin for histologic characterization with CD progression. Feces from control, Institute of Cancer Research (ICR; n = 6), and SAMP1/YitFc (n = 8) mice at different ages were subjected to microbial analysis and ^{1}H nuclear magnetic resonance (NMR) analysis to investigate fluctuations in gut microbiota and fecal metabolites with CD progression. Relative abundance of the Lachnospiraceae, Ruminococcaceae, Bacteroidaceae, and Bacteroidales S24-7 at family-level gut microbiota and fecal metabolites, such as short-chain fatty acids, lactate, glucose, xylose, and choline, dramatically fluctuated with histologic progression of intestinal inflammation in SAMP1/YitFc mice. Unlike the other metabolites, fecal taurine concentration in SAMP1/YitFc mice was higher than ICR mice regardless of age. The fecal metabolites showing characteristic fluctuations may help to understand the inflammatory mechanism associated with CD, and might be utilized as potential biomarkers in predicting CD pathology.

2.2 Introduction

CD is a chronic, relapsing inflammatory bowel disease (IBD) that affects the gastrointestinal tract,

and its precise etiology has not been fully elucidated. However, CD is known to be multifactorial in origin, involving host-genetic factors [1, 2], environmental factors [3], and dysregulated immune responses [4]. Furthermore, an imbalance of gastrointestinal microbiomes in CD patients induces an adverse immune response in the host [5, 6]. Thus, gut microbiota-derived metabolites may fluctuate in conjunction with intestinal bacterial imbalance, causing dysbiosis. In fact, metabolites of intestinal bacteria, such as short-chain fatty acids (SCFAs) and amino acids, have been reported to change in the feces of CD patients and animal models [7, 8]. Although fecal metabolites could be related to the onset and progression of CD, changes in fecal metabolites accompanying disease progression have not been clarified. Metabolites showing characteristic fluctuation with CD progression may help to understand the inflammatory mechanism associated with CD.

Metabolomic approach provides a comprehensive profile of the simplest low-molecular-weight metabolites that can be considered the ultimate response of biological systems. Metabolomics has been used as a promising tool in diagnosis and onset prediction of cancer and heart diseases [9–11]. Highly informative metabolic profiles, as a potential source of biomarkers, can be used not only for prognostic objectives but also as data sets for discriminant analysis of physiological changes with diseases. NMR is an analytical technique that can be used in metabolomic approaches to identify and quantify the complex metabolites in biological samples. Mass spectrometry is also frequently used for metabolic profiling. NMR is less sensitive; however, it has many advantages such as high reproducibility, high quantitation, minimal sample requirement, and non-destructive characteristic compared with mass spectrometry methods [12]. These advantages of the NMR method have made it widely used as a metabolomics approach in various research fields. Therefore, the NMR approach was applied to identify and quantify complex metabolites in fecal samples in this study. SAMP1/YitFc mice, a CD-like ileitis model which is suitable for investigating the pathogenesis of CD intestinal inflammation [13, 14], were used to observe fluctuations in gut microbiota and fecal metabolites with CD

progression in this study. Their phenotype occurs spontaneously without chemical, genetic, or immunological manipulation, and they possess prominent similarities to the human conditions in terms of disease location, histologic characteristics, and extra-intestinal manifestations. Unlike other CD animal models, they show a time course that allows investigation of the pre-disease, acute, and chronic phases of ileitis [15]. CD is diagnosed by bloody diarrhea and endoscopic findings of chronic inflammation in the intestine [16]. Because it is difficult to study the transition of CD pathogenesis in humans from the early stage when inflammatory histological findings are not yet apparent, studies using this animal model are of great significance. This study aimed to clarify the transition of gut microbiota and fecal metabolites associated with disease progression using SAMP1/YitFc mice, a model of spontaneous CD.

2.3 Materials and methods

2.3.1 Animals and sample collection

SAMP1/YitFc mice were purchased from Charles River Laboratories Japan (Yokohama, Japan) and propagated at a laboratory in Hokkaido University. Three-week-old healthy ICR mice were purchased from CLEA Japan (Tokyo, Japan), housed in individual cages in a room maintained at 25 ± 2°C and 12-h light/dark cycle, and provided free access to a water and CE-2 diet (CLEA japan). The study design was approved by the Hokkaido University Animal Committee, and the animals were maintained in accordance with the Guide for the Care and Use of Laboratory Animals of Hokkaido University.

Fecal samples obtained from ICR (n = 6) and SAMP1/YitFc (n = 8) mice were immediately frozen at −80°C. The frozen feces samples were freeze-dried and stored at 4°C until further analysis.

2.3.2 Histological staining

The ileum (distal one-third of the whole small intestine) isolated from SAMP1/YitFc mice was washed with phosphate-buffered saline and rolled longitudinally into a Swiss-roll configuration. Tissues were fixed in 10% formalin buffer, embedded in paraffin, and sliced into 4 μm-thick sections. After deparaffinization [17], the sections were stained with hematoxylin-eosin for histologic evaluation.

2.3.3 Gut microbial analysis

2.3.3.1 DNA extraction and purification

Fecal samples obtained from SAMP1/YitFc mice were quickly snap-frozen and stored at −80°C until DNA extraction. Total DNA was extracted from 200 mg of fecal sample using QIAamp Fast DNA Stool Mini Kit (QIAGEN, Hilden, Germany). Final DNA sample concentration was determined by a NanoDrop 2000 spectrometer (Thermo Fisher Scientific, Waltham, MA).

2.3.3.2 16S rRNA sequencing

16S rRNA genes were amplified by PCR using universal primers: 341F, (5′-CCTACGGGNGGCWGCAG-3′); 805R, (5′-GACTACHVGGGTATCTAATCC-3′). PCR reaction was conducted in reaction mixtures containing 12.5 ng of template DNA, 200 nM of each universal primer, and 1× KAPA HiFi Hot Start Ready Mix (Kapa Biosystems, Wilmington, MA). The amplification program consists of one cycle at 95°C for 3 min; 25 cycles at 95°C for 30 s, 55°C for 30 s, and 72°C for 30 s; and finally one cycle at 72°C for 5 min. PCR products were purified with AMPure XP beads (Beckman Coulter, Bres, CA), and sequencing adapters containing sample specific 8 bp barcodes were added to the 3'- and 5'- ends by PCR using Nextera XT Index Kit v2 Set B (Illumina, San Diego, CA) in 50 μL reaction mixtures containing 5 μL PCR amplicon, 5 μL each indexing primer, and 1× KAPA HiFi Hot Start Ready Mix. The amplification program consists of one cycle at 95°C for

3 min; 8 cycles at 95°C for 30 s, 55°C for 30 s, and 72°C for 30 s; and finally one cycle at 72°C for 5 min. Each amplicon was quantified by a Qubit dsDNA HS Assay Kit (Invitrogen, Carlsbad, CA), and adjusted to 4 nM. Next, 4 μL pooled amplicons were subjected to quantification using KAPA Library Quantification Kit Lightcycler 480 qPCR Mix (Kapa Biosystems) and adjusted to 4 pM. The amplicon library was mixed with 5% (v/v) PhiX Control v3 (Illumina), and sequenced on a MiSeq instrument using MiSeq 600-cycle v3 kit (Illumina).

2.3.3.3 16S rDNA-based taxonomic analysis

Pair-end reads FASTQ files were obtained from the 16S rRNA sequence. Operational taxonomic units (OTUs) classification and diversity analyses were performed using QIIME version 1.9.1 2 according to method3. The adapter sequences and low-quality bases were trimmed from sequenced reads by BBDuk (http://jgi.doe.gov/data-and-tools/bbtools/), and pair-end reads were subsequently merged by BBmerge (http://jgi.doe.gov/data-and-tools/bbtools/). Furthermore, chimeric sequences were removed by UCHIME4. OTUs were classified into five categories (phylum, order, class, family, and genus) referring to the SILVA 12_8 database with a 97% similarity threshold using UCLUST5. α-diversities were calculated from the observed OTUs.

2.3.4 Sample extraction

Based on the past review [18], fecal extracts were prepared for NMR analysis. Aliquots of 50 mg powdered feces were mixed with 1.4 mL phosphate-buffered saline in highly purified light water containing 50 mM NaP, 0.5 mM 3-(Trimethylsilyl)propionic acid-d 4 sodium salt (TSP), 0.004% NaN_3, and 10% D_2O. TSP and D_2O were used as an internal standard and internal lock solvent, respectively. This mixture was shaken at 1800 rpm for 10 min at 4°C using a tube mixer (EYELA CM-1000; Wakenyaku, Kyoto, Japan). The shaken samples were centrifuged at 16,100 × g for 10 min at 4°C, and

the supernatants pH were adjusted to approximately 7.4. Five hundred and fifty microliters of the pH adjusted supernatants were transferred to a 5-mm NMR tube (Shigemi, Hachioji, Japan) for analysis.

2.3.5 NMR spectra acquisition

^{1}H NMR spectra were recorded on a Bruker 600 MHz AVANCE III spectrometer (Brucker, Rheinstetten, Germany) operating at proton frequency 600.13 MHz. Sample temperature was controlled at 298 K. Each spectrum consisted of 128 scans of 32,768 data points with spectral width of 16 ppm and 1.7 s acquisition time. noesy1d presaturation pulse sequence was used to reduce residual water signal with low-power selective pulse at the water frequency during a recycle delay [D1 (Bruker notated) = 4.0 s] and a mixing time [D8 (Bruker notated) = 0.1 s]. A 90° pulse length was automatically calculated at each sample analysis.

2.3.6 Spectral data processing

All raw spectra were manually corrected for phase and baseline distortions, and referenced to the TSP resonance at δ = 0.0 ppm using Delta 5.0.4 (JEOL, Akishima, Japan). The spectra were normalized by referring to the peak area value of TSP, an internal standard, and excluded spectral regions of residual water area (4.68–4.88 ppm) from NMR spectral data using NMR Suite 7.5 Processor (Chenomx Inc., Edmonton, Canada). The normalized spectral data were exported to NMR Suite 7.5 Profiler (Chenomx Inc.) for further processing. In the first processing, the chemical shift region of 0.0–10.0 ppm was integrated into regions with a width of 0.04 ppm. The spectral regions of residual water area (4.68–4.88 ppm) were removed from the analysis to eliminate the effects of imperfect water saturation. In the second processing, metabolite assignment and quantification were determined by referencing the 600 MHz library from Chenomx NMR Suite. Metabolites with NMR spectral peaks that were likely to be affected by the peak derived from residual water were quantified by referencing

the ones unlikely to be affected by the water area.

2.3.7 Data analysis

The spectral data matrix was exported to SIMCA-P 14.0 (Umetrics, Umeå, Sweden). Principal component analysis (PCA) and supervised classification of orthogonal partial least squares discriminant analysis (OPLS-DA) were conducted to extract significant metabolite information. Unit variance was applied on the mean centered data before PCA and OPLS-DA. The score plot was obtained from the data to visualize the clustering pattern of the feces samples along two principal components (PC1 and PC2), where each point denoted an individual spectrum of a feces sample. The loading plots indicated the metabolites responsible for the group separation. In these loading plots, increased and decreased levels of metabolites in the feces of SAMP1/YitFc mice correspond to upward- and downward-pointing peaks in the NMR spectra. Warmer colors indicate metabolites that contributed more significantly to the separation than the metabolites indicated by cooler colors. The PLS model qualities were determined by the values of R^2X, R^2Y, and Q^2, which represent the quality of the fit and predictability of the model. An additional validation tool, a permutation test, was performed for each model by randomizing the order of Y variables for a specified number of times (permutation number = 200), and the value of R^2 and Q^2 obtained from these models were compared to that of R^2 and Q^2 of the real model. If the maximum value of R^2_{max} and Q^2_{max} from the permutation test was smaller than the R^2 and Q^2 value of the real model, respectively, the model was regarded as non-overfitted and a good-predictability model [19].

JMP (version Pro 14.0.0; SAS Institute, Cary, NC) was used for statistical analysis. Values for each experimental group are presented as the mean ± standard error of the mean (S.E.M.). Student *t*-test was used to compare the means of metabolites concentration between groups at each time point. One-way analysis of variance (ANOVA) with Tukey-Kramer *post-hoc* test was used to compare the means

of metabolites concentration and α-diversity of gut microbiota between time points. *P* values below 0.05 were considered statistically significant.

2.4 Results

2.4.1 Histologic progression of intestinal inflammation in SAMP1/YitFc mice

Histologic characterization of ileitis in representative sections of SAMP1/YitFc mice in the different age groups is shown in **Figure 2-1**. No signs of tissue inflammation were observed at age 4 weeks, with preservation of the villus architecture. By comparison, the inflammatory process progressed with distortion of the villus architecture, and the muscularis propria became thick at age 10 weeks. Furthermore, infiltration of inflammatory cells was observed. At age 15 weeks, the inflammation further progressed and the villus structure collapsed, whereas the muscularis propria was even thicker, indicating CD-like pathology.

2.4.2 Microbial profiling in the feces of SAMP1/YitFc mice

Gut microbiota composition in SAMP1/YitFc mice at different age groups were shown at the family and phylum level (**Figure 2-2a, 2-2b**). At the family level, a gradual decrease in Lachnospiraceae and Ruminococcaceae relative abundance and a gradual increase in Bacteroidaceae and Bacteroidales S24-7 relative abundance were observed over time. At the phylum level, Firmicutes relative abundance decreased, whereas Bacteroidetes relative abundance increased over time. α-Diversity of gut microbiome in SAMP1/YitFc mice was significantly decreased at age 15 weeks compared with that at age 4 weeks (**Figure 2-2c**). Furthermore, we confirmed that these metabolites hardly fluctuated in a time-dependent manner after extraction (data not shown).

2.4.3 Fecal metabolomic profiling in control and SAMP1/YitFc mice

Metabolomic analysis was performed on feces from ICR and SAMP1/YitFc mice at age 4, 9, and 15 weeks. Representative ^{1}H NMR spectra of feces from ICR and SAMP1/YitFc mice at age 4 and 15 weeks are shown in **Figure 2-3**. Amino acids, SCFAs, monosaccharides, and amines were identified from comparison with the Chenomx database and from published literature [20–22]. Assigned compounds were shown in **Supplementary table 2-S1**. Changes in the NMR spectra of fecal samples from ICR and SAMP1/YitFc mice in each stage were visualized using PCA (**Figure 2-4**). For further multivariate analysis, pairwise comparisons were performed between feces samples from ICR and SAMP1/YitFc mice using OPLS-DA (**Figure 2-5a–2-5c**). R^2X, R^2Y, and Q^2 values representing the quality of the fit and predictability of the model are shown in part of **Figure 2-5a–2-5c**, respectively. OPLS-DA score plot showed clear differentiation between feces samples of ICR and SAMP1/YitFc mice. Permutation test results at each age group are shown in **Supplementary figure 2-S1**. Coefficient loading plots of ICR and SAMP1/YitFc mice are shown in **Figure 2-5d–2-5f** to clarify the metabolites responsible for the separation of ICR and SAMP1/YitFc mice. At age 4 weeks, the feces of SAMP1/YitFc mice contained less acetate, butyrate, xylose, and choline than that of ICR mice. At age 9 weeks, lower levels of acetate, butyrate, valerate, lactate, and glucose were observed in the feces of SAMP1/YitFc mice than in those of ICR mice. On the contrary, higher levels of some metabolites, including propionate, caprate, and unassigned compounds, were detected in the feces of SAMP1/YitFc mice than in those of ICR mice at age 15 weeks.

2.4.4 Age-associated fecal metabolic variations in control and SAMP1/YitFc mice

To investigate the observed changes in detail, the absolute and relative amounts of fecal metabolites responsible for the group separation in an OPLS-DA were measured; the results are shown in **Figure 2-6**. The relative amount was calculated by dividing fecal metabolite concentration of SAMP1/YitFc mice by that of ICR mice for the same metabolite. Fecal SCFAs concentrations, such

as acetate, butyrate, and valerate, and lactate, were significantly lower in SAMP1/YitFc mice than in ICR mice at age 4 weeks. The relative amount of these metabolites at age 15 weeks were significantly higher than those at other age. Fecal glucose and xylose levels of SAMP1/YitFc mice were significantly lower than those of ICR mice at age 4 weeks, and glucose level in SAMP1/YitFc mice was also significantly lower than those of ICR mice at age 9 weeks. Although the relative amount of fecal glucose at age 9 weeks was significantly lower than that at other age, that of xylose showed no significant differences. Fecal taurine concentrations of SAMP1/YitFc mice were consistently higher than those of ICR mice throughout the time course. However, the relative amount of taurine showed no significant changes between different ages. Fecal choline in SAMP1/YitFc mice at age 4 weeks was significantly lower than that in ICR mice.

2.5 Discussion

The NMR metabolomic approach in biological samples offers two crucial opportunities: first, the chance to discover metabolites that could serve as disease biomarkers; second, the obtained metabolome profile could provide an insight into disease pathogenesis. Feces can help characterize the direct association of gastrointestinal disorders with gut microbiota and microbiota-derived metabolites. We have shown that intestinal histological findings worsened and gut microbiota fluctuated with the progression of disease condition based on experiments in spontaneous CD model mice. Furthermore, we identified several metabolites that could serve as biomarkers for this disease.

Changes in intestinal histological findings of SAMP1/YitFc mice associated with disease progression were confirmed by hematoxylin-eosin staining. Although no apparent inflammation occurred in the ileum at age 4 weeks, intestinal tissue inflammation worsened with age, consistent with previous findings [14, 15]. Moreover, Firmicutes and Bacteroidetes abundance decreased and increased with age, respectively, and α-diversity of gut microbiome decreased gradually. In previous

studies [23, 24], the metagenomic approach showed that fecal Firmicutes and Bacteroidetes were less and more abundant, respectively, in CD patients than in healthy subjects, and diversity decreased. In our analysis, decreased Firmicutes, increased Bacteroidetes, and decreased microbiome diversity were consistent with increasing age, most notably at age 15 weeks. Thus, disease progression may be associated with irreversible microbiome changes.

^{1}H NMR analysis followed by OPLS-DA using the obtained NMR spectra revealed that at age 4 and 9 weeks, the R^2_{max} and Q^2_{max} value obtained from the permutation test did not exceed the R^2 and Q^2 value obtained from the real model, suggesting non-overfitted and good-predictability models. However, at 15 weeks, Q^2_{max} value from the permutation test was higher than that from the real model, implying lack of predictability and indicating the low quality of the model at age 15 weeks. Thus, the NMR spectra of the fecal metabolites of ICR and SAMP1/YitFc mice at age 15 weeks could not be distinguished.Quantitative analysis with an OPLS-DA approach of the fecal metabolites responsible for the group separation of ICR and SAMP1/YitFc mice revealed that almost no histological disorder was observed in the ileum at age 4 weeks. However, it was previously reported that production of inflammatory cytokines has already occurred in 4-week-old SAMP1/YitFc mice, and that abnormalities in the immune mechanism that do not appear histologically have occurred [14]. In our metabolomic analysis, the metabolites had already fluctuated dramatically since the mice were aged 4 weeks. These facts could account for the changes in the fecal metabolites, which had already occurred before histological damages became apparent.

The fecal concentrations of SCFAs, such as butyrate and acetate, in SAMP1/YitFc mice at age 4 and 9 weeks were significantly lower than those in ICR mice. Consistent with these results, fecal SCFA concentration in CD patients decreased compared with that in healthy subjects [25, 26]. Furthermore, in our microbial flora analysis, the Ruminococcaceae family, which includes the butyrate-producing bacteria *Faecalibacterium prausnitzii*, decreased in SAMP1/YitFc mice over time. Previous studies

reported a decrease in *Faecalibacterium prausnitzii* in the feces of CD patients [27]. Thus, it is inferred that fecal butyrate decreased owing to the reduction of the major butyrate-producing bacteria. Butyrate suppresses enteritis by inducing differentiation of regulatory T cells via epigenomic modification [28]; therefore, the butyrate deficiency that occurs early in the onset of CD might further exacerbate the disease condition. In addition, previous studies using IBD animal models have reported that acetate regulates the function of innate immune cells via GPR43 signaling [29]. It is thus possible that the reduction of acetate in the feces of SAMP1/YitFc mice at age 4 and 9 weeks caused an imbalance in the innate immune system and was associated with further deterioration of CD pathology. On the contrary, the absolute concentration and relative amount of SCFAs in SAMP1/YitFc mice at age 15 weeks tended to increase. The monocarboxylic acid transporter (MCT) is responsible for SCFA absorption [30], and its expression decreased in the digestive tract of dextran sulfate sodium-induced colitis model rats [31]. Considering these, the increase in fecal SCFAs of SAMP1/YitFc mice at age 15 weeks might be related to fecal SCFA retention due to MCT-mediated inhibition of absorption. Fecal lactate concentration of SAMP1/YitFc mice at age 4 and 9 weeks was significantly lower than that of ICR mice. Lactate protects against pathogenic microorganisms via GPR31 on macrophages in the intestinal tract [32]; therefore, decreased fecal lactate may increase the risk of secondary intestinal infections associated with CD. Fecal glucose and xylose concentrations in SAMP1/YitFc mice at age 4 weeks were significantly lower than those in ICR mice. The activities of lactase and sucrase on the brush border membrane have been reported to decrease in the intestinal tract of IBD patients [33]. These findings suggest that the carbohydrates in foods were not digested into monosaccharides and remained in the form of di-, oligo-, and polysaccharides in the intestinal tract, resulting in decreased proportion of detected monosaccharides. Unlike that of other metabolites, fecal taurine concentration of SAMP1/YitFc mice was high regardless of age. Oxidative stress has been reported to affect the onset and worsening of CD [34]. As taurine has an antioxidant effect [35], increased fecal taurine may

play an important role in CD pathogenesis; however, further studies are needed to confirm this. Fecal choline in SAMP1/YitFc mice at age 4 weeks was significantly lower than that in ICR mice. Reduced choline level in the colonocytes of patients with ulcerative colitis has been reported [36]. Choline, a cell membrane component, is involved in cell membrane maintenance [37]; therefore, decreased choline in the early stage of CD might contribute to intestinal tissue deterioration in IBD patients.

Taken together, fecal metabolites changed dramatically as mice increased in age. Despite being fed the same diet during the experimental period, the gut microbiota of SAMP1/YitFc mice changed over time compared to that of ICR mice. Fecal metabolite alteration was suggested to be partially caused by different metabolic reactions derived from different microbiota in both groups. Metabolic flux analysis may be required to further elucidate in-depth mechanisms of these metabolic changes. In addition, we must consider the possibility that the chronic inflammation negatively affected host digestibility and absorption, resulting in changes in fecal metabolite composition.

One limitation of this study was that we did not consider the metabolic changes associated with normal mouse growth. By calculating the relative amount of fecal metabolites, the fluctuations associated with growth were partially offset. As previously reported [38, 39], fecal metabolites changed over time in normal mice. Thus, SAMP1/YitFc and ICR mice of the same age may be affected by metabolic changes accompanying growth. Moreover, as CD frequently occurs in young people, metabolic changes accompanying growth might be related to its pathogenesis; thus, further investigation is necessary. Furthermore, it is still unclear whether these metabolite fluctuations are CD-specific because it has not been compared with other IBD models in this study. Comparative studies using models with other IBD pathologies or with symptoms that are more common such as diarrhea and clinical trials in patients with CD and UC are warranted to confirm that these metabolite fluctuations are specific to CD pathology.

In summary, we showed that fecal metabolites, such as SCFAs, glucose, xylose, taurine, and

choline, dramatically fluctuated with disease progression, and that the fluctuations in SAMP1/YitFc mice could be partially influenced by alterations in the gut microbiome. Several fecal metabolites of SAMP1/YitFc mice had already fluctuated at the very early stage of CD before the histological changes associated with disease progression became apparent. Therefore, those fecal metabolites showing characteristic fluctuations may be potential candidate biomarkers for predicating disease progression, and may contribute to early diagnosis of CD to improve patient's quality of life.

2.6 Figures

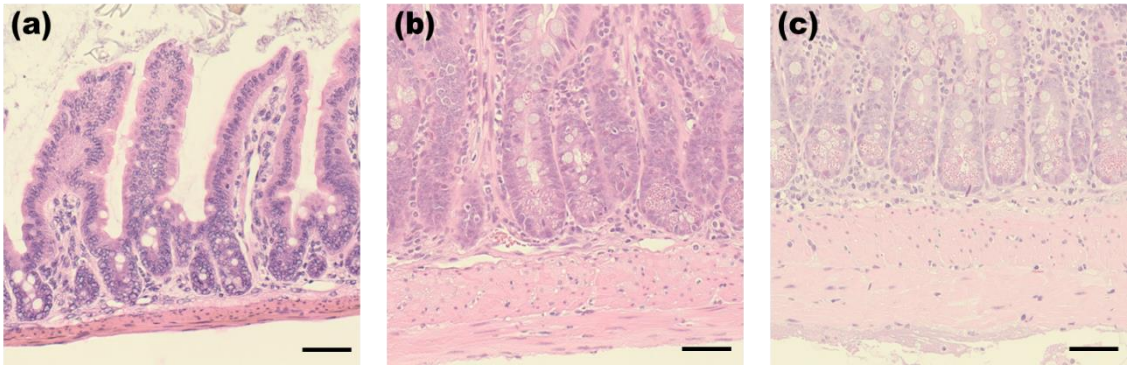


Figure 2-1. Representative histopathological images of hematoxylin and eosin-stained ileum of SAMP1/YitFc mice at age (a) 4, (b) 10, and (c) 15 weeks. Scale bars: 50 μm.

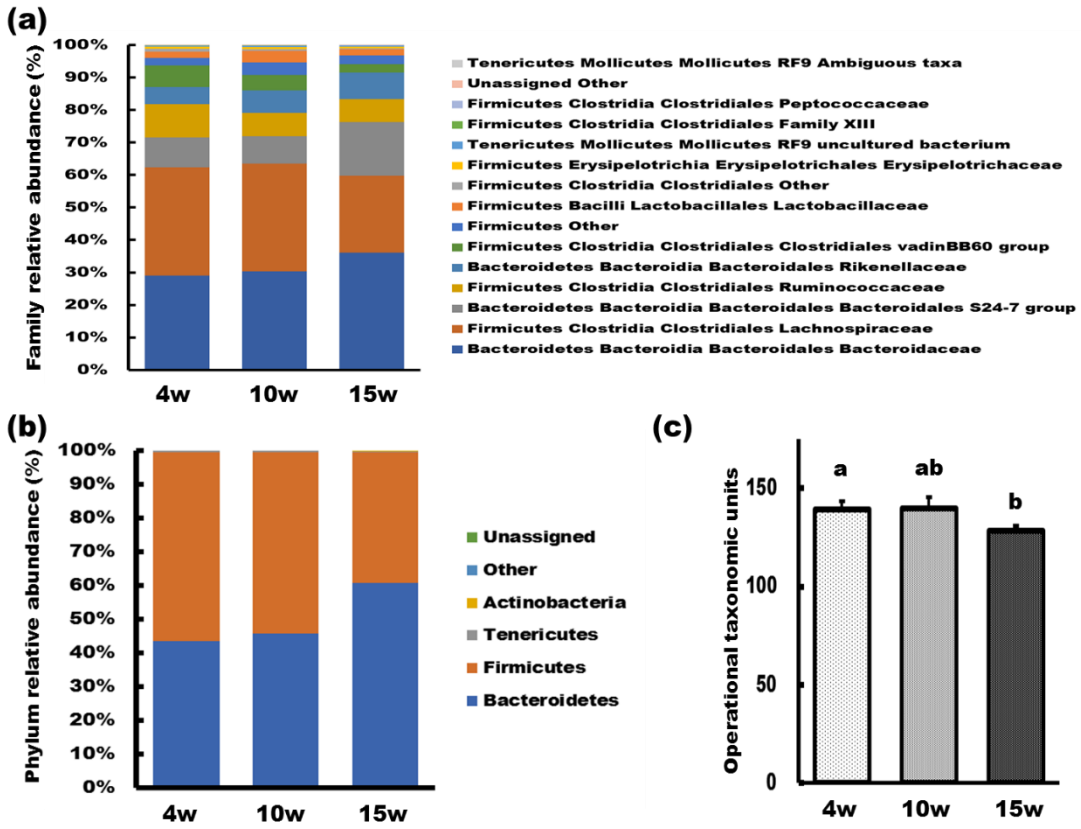


Figure 2-2. (a) Family-level gut microbiota changes in SAMP1/YitFc mice at age 4, 10, and 15 weeks (n = 8). (b) Phylum-level gut microbiota changes in SAMP1/YitFc mice at age 4, 10, and 15 weeks (n = 8). (c) Gut microbiota α-diversity in SAMP1/YitFc mice at age 4, 10, and 15 weeks. Different letters indicate $P < 0.05$, as analyzed by one-way ANOVA with Tukey-Kramer post-hoc test. Data are presented as mean ± S.E.M. (n = 8).

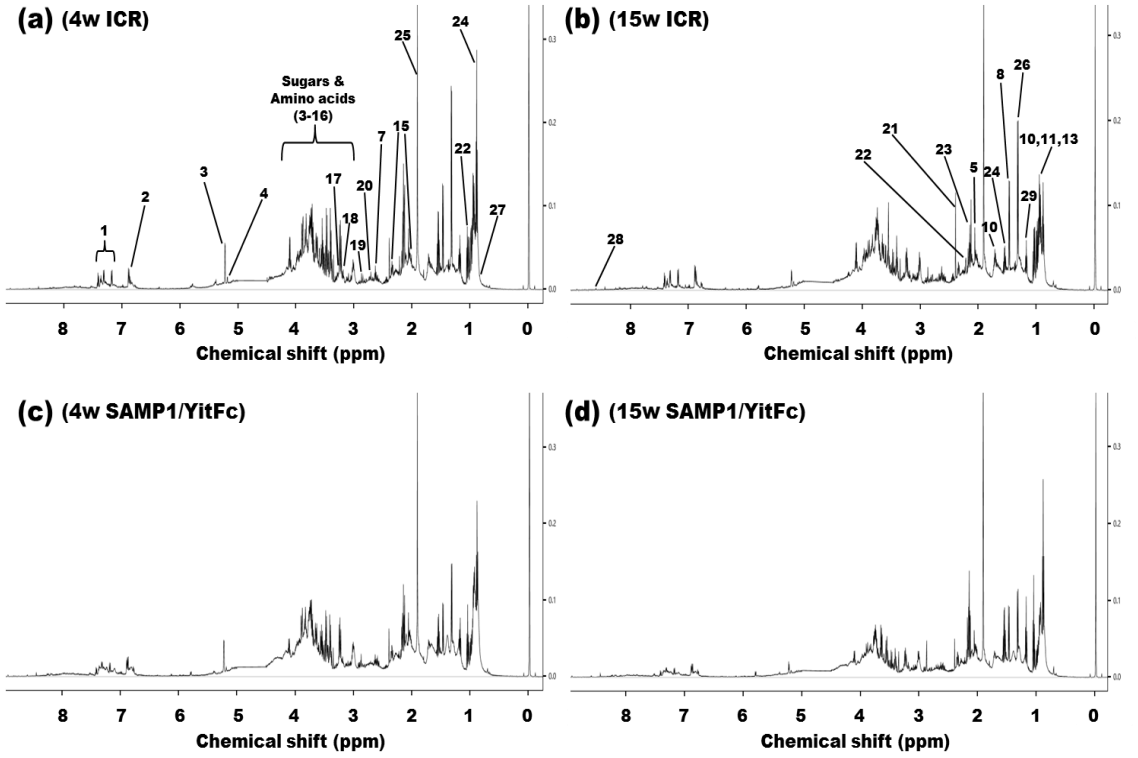


Figure 2-3. Typical 600 MHz ^{1}H NMR spectra of the feces of ICR mice at age (a) 4 and (b) 15 weeks, and of SAMP1/YitFc mice at age (c) 4 and (d) 15 weeks. Key: (1) Phenylalanine; (2) Tyrosine; (3) Glucose; (4) Xylose; (5) N-Acetylcysteine; (6) Threonine; (7) Methionine; (8) Alanine; (9) Lysine; (10) Leucine; (11) Isoleucine; (12) Glycine; (13) Valine; (14) Aspartate; (15) Glutamate; (16) Taurine; (17) Trimethylamine-N-oxide; (18) Choline; (19) Trimethylamine; (20) Dimethylamine; (21) Succinate; (22) Propionate; (23) Valerate; (24) Butyrate; (25) Acetate; (26) Lactate; (27) Caprate; (28) Formate; (29) Ethanol.

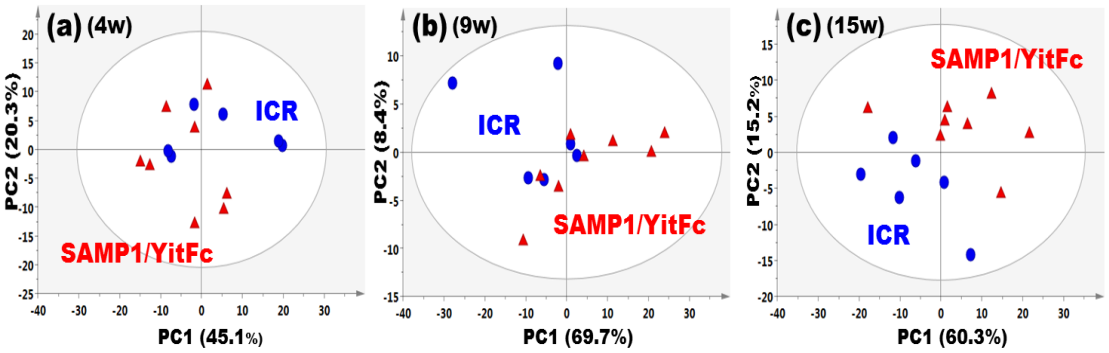


Figure 2-4. PCA score plots derived from NMR spectra of the feces of ICR (n = 6) and SAMP1/YitFc (n = 8) mice at age (a) 4, (b) 9, and (c) 15 weeks.

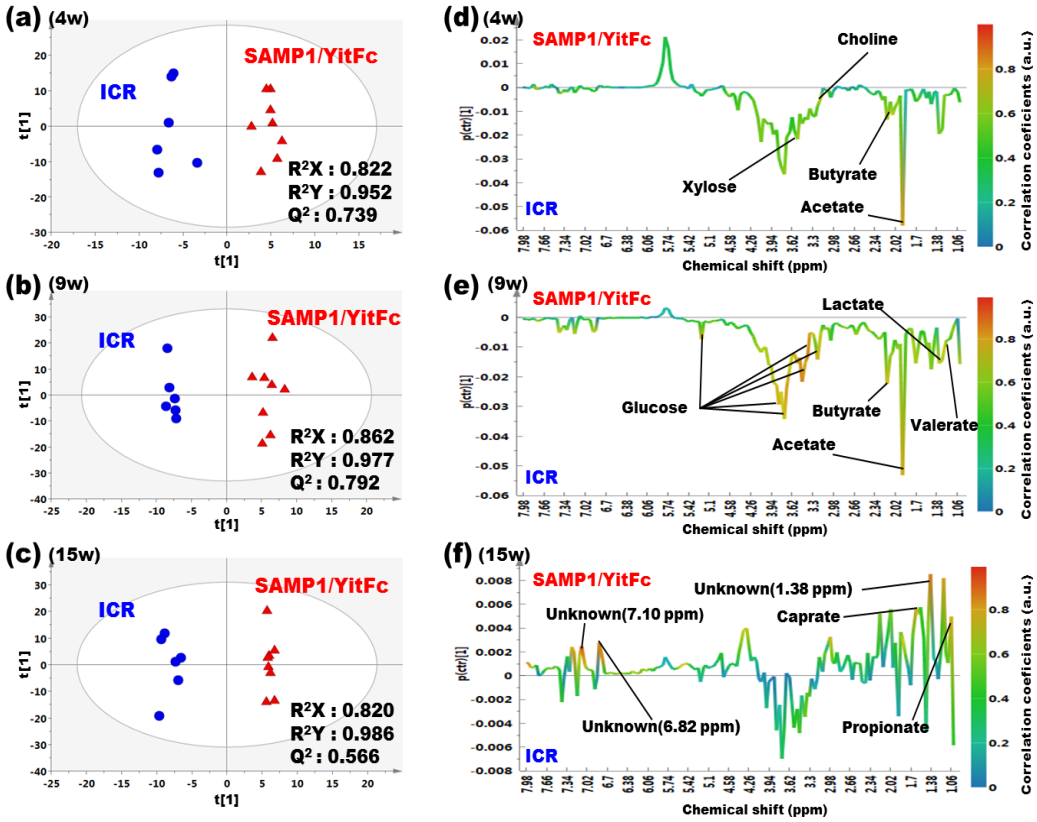


Figure 2-5. Multivariate analysis of binned NMR data. OPLS-DA score plots derived from NMR spectra of feces from ICR (n = 6) and SAMP1/YitFc (n = 8) mice at age (a) 4, (b) 9, and (c) 15 weeks. Coefficient loading plots obtained from ICR mice and SAMP1/YitFc mice at age (d) 4, (e) 9, and (f) 15 weeks. Increased and decreased fecal metabolite levels in SAMP1/YitFc mice correspond to upward- and downward-pointing peaks in the NMR spectra. Warmer colors indicate metabolites contributing more significantly to the separation than those indicated by cooler colors.

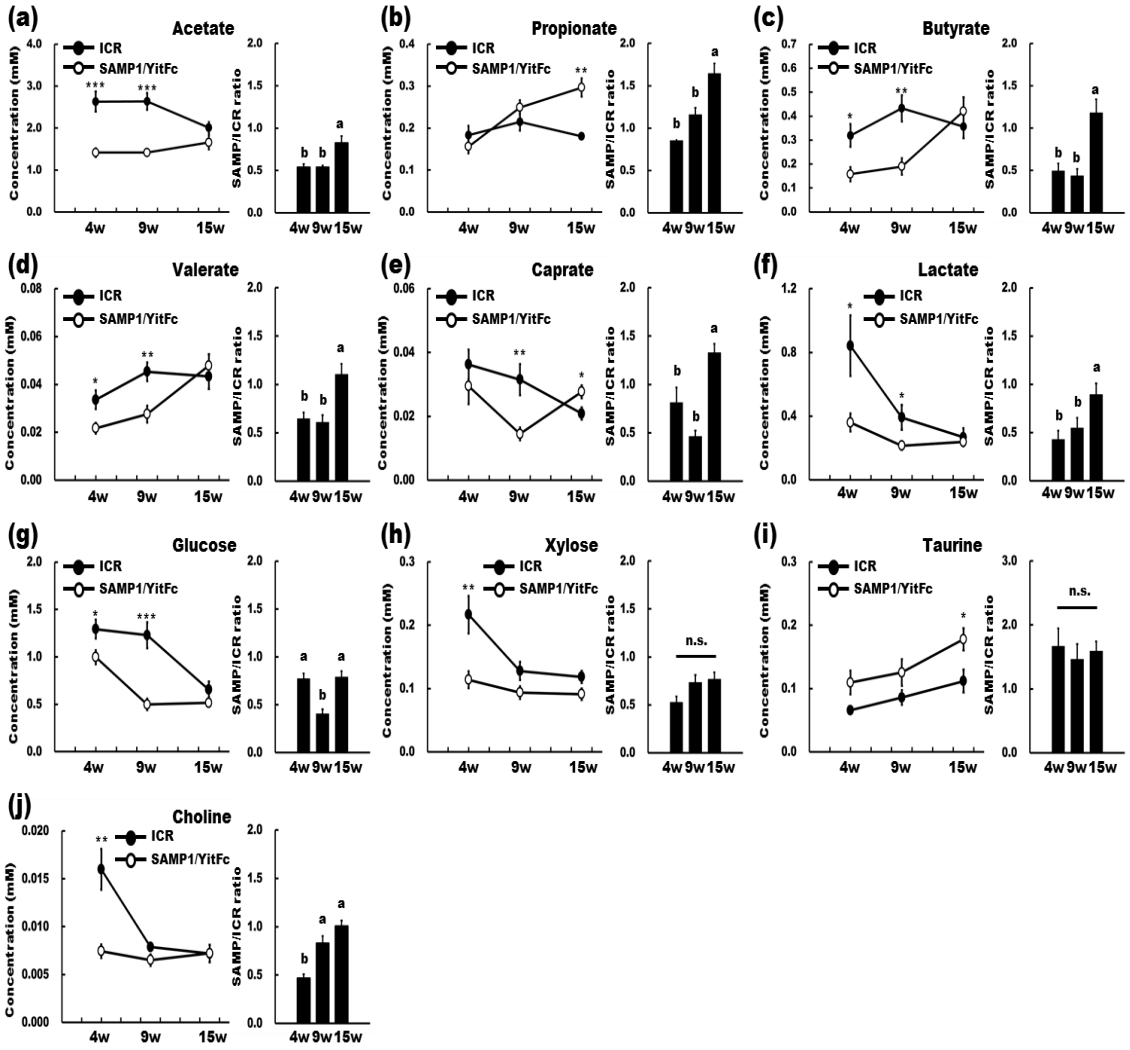


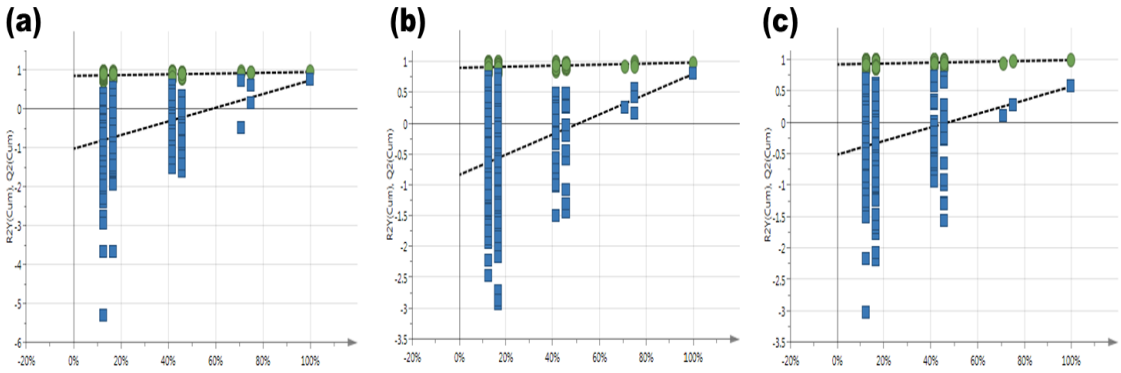
Figure 2-6. Fecal metabolite concentration changes in ICR (n = 6) and SAMP1/YitFc mice (n = 8) at age 4, 9, and 15 weeks (left; line graph). Relative fecal metabolite concentration (ratio of SAMP1/YitFc mice / ICR mice) at age 4, 9, and 15 weeks (right; bar graph). Significant differences between samples at the same time point are indicated by an asterisk symbol (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$). Different letters indicate $P < 0.05$ as analyzed by one-way ANOVA with Tukey-Kramer post-hoc test. Data are presented as mean ± S.E.M. (n = 8).

2.7 Supplementary tables and figures

Supplementary table 2-S1. ^{1}H NMR data and assignments of the metabolites in fecal extract from ICR mice and SAMP1/YitFc mice.

	Metabolite	δ^{1}H (ppm)					
1	Phenylalanine	3.11(dd)	3.27(dd)	3.99(dd)	7.32(m)	7.36(m)	7.42(m)
2	Tyrosine	3.04(dd)	3.19(dd)	3.93(dd)	6.89(m)	7.18(m)	
3	Glucose	3.23(dd)	3.40(t)	3.40(t)	3.45(m)	3.48(t)	3.52(dd)
		3.72(dd)	3.76(dd)	3.83(dd)	3.89(dd)	4.64(d)	5.22(d)
4	Xylose	3.22(dd)	3.31(t)	3.43(t)	3.52(dd)	3.60(m)	3.62(m)
		3.66(t)	3.67(t)	3.69(dd)	3.92(dd)	4.57(d)	5.19(d)
5	N-Acetylcysteine	2.06(s)	2.90(dd)	2.93(dd)	4.37(m)	8.00(s)	
6	Threonine	1.31(d)	3.58(d)	4.24(m)			
7	Methionine	2.11(m)	2.13(s)	2.18(m)	2.63(t)	3.85(dd)	
8	Alanine	1.47(d)	3.78(q)				
9	Lysine	1.43(m)	1.50(m)	1.73(m)	1.91(m)	3.02(t)	3.74(t)
10	Leucine	0.94(d)	0.95(d)	1.67(m)	1.70(m)	1.73(m)	3.74(m)
11	Isoleucine	0.93(t)	1.00(d)	1.25(m)	1.46(m)	1.97(m)	3.66(d)
12	Glycine	3.55(s)					
13	Valine	0.98(d)	1.03(d)	2.26(m)	3.60(d)		
14	Aspartate	2.66(dd)	2.80(dd)	3.89(dd)			
15	Glutamate	2.03(m)	2.12(m)	2.33(m)	2.36(m)	3.74(dd)	
16	Taurine	3.24(t)	3.41(t)				
17	Trimethylamine-N-oxide	3.25(s)					
18	Choline	3.19(s)	3.51(m)	4.06(m)			
19	Trimethylamine	2.87(s)					
20	Dimethylamine	2.71(s)					
21	Succinate	2.39(s)					
22	Propionate	1.04(t)	2.17(q)				
23	Valerate	0.88(t)	1.29(m)	1.51(m)	2.17(t)		
24	Butyrate	0.88(t)	1.55(m)	2.15(t)			
25	Acetate	1.91(s)					
26	Lactate	1.32(d)	4.11(q)				
27	Caprate	0.84(t)	1.25(m)	1.25(m)	1.26(m)	1.26(m)	1.27(m)
28	Formate	1.53(m)	2.14(t)				
		8.44(s)					
29	Ethanol	1.17(t)	3.65(q)				

Key: s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; dd, doublet of doublet.



Supplementary figure 2-S1. The permutation test plots for NMR spectra of fecal extracts from ICR mice (n = 6) and SAMP1/YitFc mice (n = 8) at (a) 4 weeks, (b) 9 weeks and (c) 15 weeks of age. The model is valid from the result of permutation test for the OPLS-DA model [number of permutations, 200; intercepts: (a) R^2 (green) = 0.855, Q^2 (blue) = -1.02, (b) R^2 (green) = 0.893, Q^2 (blue) = -0.831, (c) R^2 (green) = 0.918, Q^2 (blue) = -0.517].

2.8 References

[1] Hampe J, Franke A, Rosenstiel P, Till A, Teuber M, Huse K, et al. A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. Nat Genet. 2007;39:207–11.

[2] Hugot JP, Chamaillard M, Zouali H, Lesage S, Cézard JP, Belaiche J, et al. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. Nature. 2001;41:599–603.

[3] Ng SC, Bernstein CN, Vatn MH, Lakatos PL, Loftus EV, Tysk C, et al. Geographical variability and environmental risk factors in inflammatory bowel disease. Gut. 2013;62:630–49.

[4] Li N, Shi RH. Updated review on immune factors in pathogenesis of Crohn's disease. World J Gastroenterol. 2018;24:15–22.

[5] Brusaferro A, Cavalli E, Farinelli E, Cozzali R, Principi N, Esposito S. Gut dysbiosis and paediatric Crohn's disease. J Infect. 2019;78:1–7.

[6] Imhann F, Vich Vila A, Bonder MJ, Fu J, Gevers Di, Visschedijk MC, et al. Interplay of host genetics and gut microbiota underlying the onset and clinical presentation of inflammatory bowel disease. Gut. 2017;67:108–19.

[7] Geirnaert A, Calatayud M, Grootaert C, Laukens D, Devriese S, Smagghe G, et al. Butyrate-producing bacteria supplemented in vitro to Crohn's disease patient microbiota increased butyrate production and enhanced intestinal epithelial barrier integrity. Sci Rep. 2017;7:1–14.

[8] Sheehan D, Moran C, Shanahan F. The microbiota in inflammatory bowel disease. J Gastroenterol. 2015;50:495–507.

[9] Chen X, Yu D. Metabolomics study of oral cancers. Metabolomics. 2019;15:22.

[10] Ruiz-Canela M, Hruby A, Clish CB, Liang L, Martínez-Gonzále MA, Hu FB. Comprehensive metabolomic profiling and incident cardiovascular disease: A systematic review. J Am Heart Assoc. 2017;6:1–22.

[11] Yu L, Li K, Zhang X. Next-generation metabolomics in lung cancer diagnosis, treatment and precision medicine: mini review. Oncotarget. 2017;8:115774–86.

[12] Moco S, Bino RJ, De Vos RCH, Vervoort J. Metabolomics technologies and metabolite identification. Trends Analyt Chem. 2007;26:855–66.

[13] Kozaiwa K, Sugawara K, Smith M, Carl V, Yamschikov V, Belyea B, et al. Identification of a quantitative trait locus for ileitis in a spontaneous mouse model of Crohn's disease: SAMP1/YitFc. Gastroenterology. 2003;125:477–90.

[14] Rivera-Nieves J, Bamias G, Vidrich A, Marini M, Pizarro TT, McDuffie MJ, et al. Emergence of perianal fistulizing disease in the SAMP1/YitFc mouse, a spontaneous model of chronic ileitis. Gastroenterology. 2003;124:972–82.

[15] Pizarro TT, Pastorelli L, Bamias G, Garg RR, Reuter BK, Mercado JR, et al. SAMP1/YitFc mouse strain: A spontaneous model of Crohn's disease-like ileitis. Inflamm Bowel Dis. 2011;17:2566–84.

[16] Laass MW, Roggenbuck D, Conrad K. Diagnosis and classification of Crohn's disease. Autoimmun Rev. 2014;13:467–71.

[17] Feldman AT, Wolfe D. Tissue processing and hematoxylin and eosin staining. Histopathology: Methods and Protocols, Methods in Molecular Biology. New York: Springer 2014;1180:31–43.

[18] Karu N, Deng L, Slae M, Guo AC, Sajed T, Huynh H, et al. A review on human fecal metabolomics: Methods, applications and the human fecal metabolome database. Anal Chim Acta. 2018;1030:1–24.

[19] Mahadevan S, Shah SL, Marrie TJ, Slupsky CM. Analysis of metabolomic data using support vector machines. Anal Chem. 2008;80:7562–70.

[20] Santiago GT, Contreras JIS, Camargo MEM, Vallejo LGZ. NMR-based metabonomic approach reveals changes in the urinary and fecal metabolome caused by resveratrol. J Pharm Biomed Anal.

2019;162:234–41.

[21] Le Gall G, Noor SO, Ridgway K, Scovell L, Jamieson C, Johnson IT, et al. Metabolomics of fecal extracts detects altered metabolic activity of gut microbiota in ulcerative colitis and irritable bowel syndrome. J Proteome Res. 2011;10:4208–18.

[22] Saric J, Wang Y, Li J, Coen M, Utzinger J, Marchesi JR, et al. Species variation in the fecal metabolome gives insight into differential gastrointestinal function. J Proteome Res. 2008;7:352–60.

[23] Manichanh C, Rigottier-Gois L, Bonnaud E, Gloux K, Pelletier E, Frangeul L, et al. Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. Gut. 2006;55:205–11.

[24] Andoh A, Kuzuoka H, Tsujikawa T, Nakamura S, Hirai F, Suzuki Y, et al. Multicenter analysis of fecal microbiota profiles in Japanese patients with Crohn's disease. J Gastroenterol. 2012;47:1298–307.

[25] De Preter V, Machiels K, Joossens M, Arijs I, Matthys C, Vermeire S, et al. Faecal metabolite profiling identifies medium-chain fatty acids as discriminating compounds in IBD. Gut. 2015;64:447–58.

[26] Takaishi H, Matsuki T, Nakazawa A, Takada T, Kado S, Asahara T, et al. Imbalance in intestinal microflora constitution could be involved in the pathogenesis of inflammatory bowel disease. Int J Med Microbiol. 2008;298:463–72.

[27] Sokol H, Seksik P, Furet JP, Firmesse O, Nion-Larmurier I, Beaugerie L, et al. Low counts of Faecalibacterium prausnitzii in colitis microbiota. Inflamm Bowel Dis. 2009;15(8):1183–89.

[28] Furusawa Y, Obata Y, Fukuda S, Endo TA, Nakato G, Takahashi D, et al. Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. Nature. 2013;504:446–50.

[29] Maslowski KM, Vieira AT, Ng A, Kranich J, Sierro F, Di Yu, et al. Regulation of inflammatory responses by gut microbiota and chemoattractant receptor GPR43. Nature. 2009;461:1282–86.

[30] Halestrap AP, Meredith D. The SLC16 gene family - From monocarboxylate transporters (MCTs) to aromatic amino acid transporters and beyond. Eur J Physiol. 2004;447:619–28.

[31] Thibault R, De Coppet P, Daly K, Bourreille A, Cuff M, Bonnet C, et al. Down-regulation of the monocarboxylate transporter 1 is involved in butyrate deficiency during intestinal inflammation. Gastroenterology. 2007;133:1916–27.

[32] Morita N, Umemoto E, Fujita S, Hayashi A, Kikuta J, Kimura I, et al. GPR31-dependent dendrite protrusion of intestinal CX3CR1+ cells by bacterial metabolites. Nature. 2019;566:110–14.

[33] Mourad FH, Barada KA, Saade NE. Impairment of small Intestinal function in ulcerative colitis: Role of enteric innervation. J Crohns Colitis. 2017;11:369–77.

[34] Moura FA, de Andrade KQ, dos Santos JCF, Araújo ORP, Goulart MOF. Antioxidant therapy for treatment of inflammatory bowel disease: Does it work? Redox Biol. 2015;6:617–39.

[35] Giriş M, Depboylu B, Doğru-Abbasoğlu S, Erbil Y, Olgaç V, Aliş H, et al. Effect of taurine on oxidative stress and apoptosis-related protein expression in trinitrobenzene sulphonic acid-induced colitis. Clin Exp Immunol. 2008;152:102–10.

[36] Bjerrum JT, Wang Y, Hao F, Coskun M, Ludwig C, Günther U, et al. Metabonomics of human fecal extracts characterize ulcerative colitis, Crohn's disease and healthy individuals. Metabolomics. 2015;11:122–33.

[37] Zeisel SH. Choline and human nutrition. Annu Rev Nutr. 1994;14:269–96.

[38] Li JV, Saric J, Yap IKS, Utzinger J, Holmes E. Metabonomic investigations of age- and batch-related variations in female NMRI mice using proton nuclear magnetic resonance spectroscopy. Mol BioSyst. 2013;9:3155–65.

[39] Tian Y, Zhang L, Wang Y, Tang H. Age-related topographical metabolic signatures for the rat

gastrointestinal contents. J Proteome Res. 2011;11:1397–411.

Chapter 3 ^{1}H NMR metabolomic and transcriptomic analyses reveal urinary metabolites as biomarker candidates in response to protein undernutrition in adult rats

3.1 Abstract

Protein undernutrition contributes to the development of various diseases in broad generations. Urinary metabolites may serve as non-invasive biomarkers of protein undernutrition; however, this requires further investigation. We aimed to identify novel urinary metabolites as biomarker candidates responsive to protein undernutrition. Adult rats were fed control (CT; 14% casein) or isocaloric low protein (LP; 5% casein) diets for 4 weeks. ^{1}H nuclear magnetic resonance (NMR) metabolomics was applied to urine, plasma, and liver samples to identify metabolites responsive to protein undernutrition. Liver samples were subjected to mRNA microarray and quantitative PCR analyses to elucidate the mechanisms causing fluctuations in identified metabolites. Urinary taurine levels were significantly lower in the LP group than in the CT group at week 1, and remained constant until week 4. Hepatic taurine level and gene expression level of cysteine dioxygenase type 1 were also significantly lower in the LP group than in the CT group. Urinary trimethylamine N-oxide (TMAO) levels were significantly higher in the LP group than in the CT group at week 2, and remained constant until week 4. Hepatic TMAO level and gene expression levels of flavin-containing monooxygenase 1 and 5 were also significantly higher in the LP group than in the CT group. In conclusion, urinary taurine and TMAO levels substantially responded to protein undernutrition. Furthermore, changes in hepatic levels of these metabolites and gene expressions associated with their metabolic pathways were also reflected in their fluctuating urinary levels. Thus, taurine and TMAO could act as non-invasive urinary biomarker candidates to detect protein undernutrition.

3.2 Introduction

It is estimated that an additional 50% protein will be required by 2050 as the global population increases [1], which may put individuals from all generations at an increased risk of protein undernutrition. Several studies have also shown that inadequate protein intake in broad generations, including younger individuals, might lead to reduced physical activity through the effect on muscles [2, 3]. In addition, it is speculated that the reduced physical activity in individuals from younger generation would increase the risk of diseases development later in life. Besides, several studies have reported that protein undernutrition may induce fatty liver and lead to liver disease in the long term [4–7]. Thus, because protein undernutrition has potential adverse effects on living tissues such as the muscles and liver, it would be desirable to establish minimally invasive and easy-to-collect biomarkers detecting the nutritional status as early as possible. Biomarkers that reflect protein undernutrition in bio-fluid samples such as the blood and urine could be extremely useful for the prevention of chronic protein malnutrition and the secondary diseases in a wide generation-range.

The levels of specific circulating proteins such as transthyretin, transferrin, and albumin (ALB) in blood have been reported to correlate with protein nutritional status [8, 9]. Furthermore, it has been reported that urine as well as blood may be able to detect protein nutritional status [9]. Urine could be advantageous over blood, as the sampling is non-invasive and easy to perform frequently. However, unlike proteins in blood, those in urine are difficult to apply widely as a biomarker because they are only detected in individuals with renal dysfunction [10]. Hence, low-molecular-weight metabolites, non-protein components in urine, are generally focused on. Although urinary nitrogen has been reported as the metabolite that correlates with total protein intake, nitrogen sources other than protein make a greater contribution to total nitrogen levels [11]. 3-Methylhistidine is a urinary metabolite that correlates with protein nutritional status, however, the excretion level is reported to increase with meat intake [12, 13], and the nutritional status in response to meat-derived proteins may be overestimated.

Thus, to the best of our knowledge, there are few suitable urinary metabolites that correlate with protein nutritional status.

Metabolomic approaches provide low-molecular-weight metabolite profiles to help understand biological responses. Highly informative metabolic profiles can be a potential source of biomarkers, and therefore metabolomics has been a promising tool for the diagnosis of several diseases such as cancer and heart diseases [14, 15]. Metabolomics can also be used for the discriminant analysis of physiological responses to nutritional status [16]. NMR and mass spectrometry have been used as techniques for metabolomics. The NMR method is less sensitive; however, it has many advantages such as high reproducibility, high quantitation, minimal sample requirement, and non-destructive characteristic compared with mass spectrometry methods [17]. Therefore, the NMR analysis was applied to identify and quantify complex metabolites in biological samples such as urine, plasma, and liver in this study.

We have recently investigated the effects of moderate protein undernutrition on plasma protein profiles in adult Wistar rats where rats were maintained on an AIN-93 M [14% casein (CN), CT diet] or AIN-93 M-based isocaloric diet containing 5% CN (LP diet) for 4 weeks [18]. The use of this animal model has been justified by confirming that this nutritional condition reflects the potential protein undernutrition in human [19]. Here, we expanded the investigation on the same animals, focusing on their metabolites. The aim of this study was to identify novel urinary metabolites as biomarker candidates that were responsive to protein undernutrition using NMR metabolomic and transcriptomic analyses.

3.3 Materials and methods

3.3.1 Animal experiments

The study design was approved by the Animal Research Committee of Morinaga Milk Industry

Co., Ltd. (Approval number: 18-025) and experiments were performed in accordance with their guidelines (Sep. 27th–Nov. 9th, 2018). The detail has already been described in our recent report [18]. Briefly, twelve male Wistar rats (14 weeks old; Japan SLC, Hamamatsu, Japan) were acclimated individually to polycarbonate cages with wood shavings in temperature-, humidity-, and light-controlled facility (21–25°C, 40–60%, and a 12-h light/dark cycle) by allowing them *ad libitum* access to water and the CT diet for 2 weeks. Then, animals were assigned to 2 dietary groups, being fed *ad libitum* either the CT diet or the LP diet for 4 weeks (respectively, n = 6). This sample size is the minimum necessary for reliable statistical verification. Urine samples were obtained weekly during the experimental period using metabolic cages. At the experimental period (after 4 weeks), animals were euthanized by deep anesthesia with sevoflurane and blood samples were drawn from the inferior vena cava using syringes. Blood samples were centrifuged at 1700 × *g* for 10 min at room temperature, and the upper plasma layers were collected and stored at −80°C until analysis. Liver samples were then obtained and frozen in liquid nitrogen. All the samples were stored at −80°C until analysis. Furthermore, welfare-related assessments and interventions were carried out before, during, and after the experiment by the Animal Research Committee.

3.3.2 Sample preparation for NMR analysis

Urine samples (540 μL) were mixed with 60 μL of PBS in 100% D_2O containing 500 mM NaP, 5 mM 3-(Trimethylsilyl)propionic acid-d4 sodium salt (TSP), and 0.04% (w/v) NaN_3. This mixture was centrifuged at 15,000 × *g* for 5 min at 4°C and 550 μL supernatant was transferred to a 5-mm NMR tube (Shigemi, Hachioji, Japan). Plasma samples (540 μL) were mixed with 60 μL of phosphate-buffered saline (PBS) in 100% D_2O containing 500 mM NaP, 5 mM TSP, 10 mM formate, and 0.04% (w/v) NaN_3, centrifuged at 3000 × *g* for 5 min at 4°C, and 550 μL of supernatant was transferred to a 5-mm NMR tube. Liver samples were homogenized in 1.0 mL of extraction buffer (acetonitrile:MilliQ

water = 1:1) per 100 mg of liver at 25,000 rpm for 30 s on ice. The homogenate was centrifuged at 5480 × *g* for 10 min at 4°C and 800 μL of supernatant was collected and lyophilized. The lyophilized sample was suspended in PBS in 100% D_2O containing 200 mM NaP and 2 mM TSP. The suspension was centrifuged at 15,000 × *g* for 5 min at 4°C and the 550 μL of supernatant was transferred to a 5-mm NMR tube. The pH of all supernatants was adjusted to 7.4. TSP and formate were used as internal standards. D_2O was used as an internal lock solvent.

3.3.3 NMR spectra acquisition

^1H NMR spectra were recorded using a Bruker 600 MHz AVANCE III spectrometer (Bruker, Rheinstetten, Germany) at proton frequency of 600.13 MHz with the sample temperature controlled at 298 K. The ^1H NMR spectra of urine and liver samples were recorded using a water-suppressed standard one-dimensional NOESY1D pulse sequence, while those of the plasma samples were recorded using a water-suppressed standard one-dimensional CPMG pulse sequence. Each spectrum consisted of 32,768 data points with a spectral width of 12 ppm. The acquisition time was 2.28 s and the number of scans was 128. A water-suppressed pulse sequence was used to reduce the residual water signal at the water frequency with a recycle delay [D1 (Bruker notated)] of 2.72 s and a mixing time [D8 (Bruker notated)] of 0.10 s. A 90° pulse length was automatically calculated for each sample analysis.

3.3.4 Spectral data processing and multivariate data analysis

All raw spectra were manually corrected for phase and baseline distortions against TSP resonance at δ = 0.0 ppm using Delta 5.0.4 (JEOL, Akishima, Japan). The spectra were normalized to the peak area value of the internal standards TSP or formate using the NMR Suite 7.5 Processor (Chenomx, Edomonton, Canada). Differences in urine concentration between the samples were corrected by

dividing each by their creatinine content. The normalized spectral data were exported to NMR Suite 7.5 Profiler (Chenomx). In the first round of processing, the 0.0–10.0 ppm chemical shift region was integrated into regions with a width of 0.04 ppm, while spectral regions related to residual water (4.68–5.08 ppm) and urea (5.50–6.00 ppm) were removed to eliminate their massive signals as it could affect other metabolite signals. In the second round of processing, metabolite assignment and quantification were determined by referencing the library from Chenomx NMR Suite and published literature [20–22].

3.3.5 mRNA microarray experiments

As already described in our previous report [18], total RNA was extracted from livers and purified. Complementary RNA was generated and labeled with cyanine 3 using a Low Input QuickAmp Labeling Kit (Agilent Technologies, Santa Clara, CA) and hybridized to a SurePrint G3 Rat GE microarray 8 × 60K (Agilent Technologies) according to the manufacturer's instructions. Fluorescence signals were detected using Agilent SureScan Microarray Scanner Extraction Software 12.0.3.1 (Agilent Technologies) according to the manufacturer's instructions. Raw data from the Feature Extraction Software were exported to GeneSpring GX ver. 14.9 (Agilent Technologies). To identify differentially expressed genes in the microarray hybridization and reduce noise, each data set of fluorescence signals was normalized using a median shift algorithm (shifted to 75 percentile) with background correction to the median of all samples. Only genes with normalized signals detected on all microarrays were judged to be present. Furthermore, genes detected by the multiple probes was selected one to have the largest fold change. Subsequently, in order to find potential associations between the detected hepatic genes and identified metabolites, the KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway database was used.

3.3.6 Gene Ontology (GO) data analysis

For functional analysis of differentially expressed genes in mRNA microarray, upregulated and downregulated genes were subjected to gene set enrichment analysis with GO terms using GeneSpring GX. GO term description and GO accession number were referenced to The Gene Ontology Resource, the web-available bioinformatics database (http://www. geneontology.org/). Generally, Fisher's exact test and the χ^2 test were used to classify the GO term, and the false discovery rate was calculated to correct *P* value. *P* values of < 0.05 were considered statistically significant.

3.3.7 Reverse transcription and quantitative PCR

To validate the gene fluctuations detected by mRNA microarray analysis, total RNA was reverse transcribed using a QuantiTect RT Kit (QIAGEN, Hilden, German) and DNA templates were amplified using a CFX96 Real-Time PCR (Bio-Rad Laboratories) with TaqMan Fast Advanced Master Mix and TaqMan probes (Applied Biosystems) according to the manufacturer's instructions. The primers and probes are summarized in **Supplementary table 3-S1**. Gene expression were normalized to β-actin and analyzed using CFX Manager ver. 3.1 (Bio-Rad Laboratories).

3.3.8 Statistical analysis

The NMR spectral data matrix was exported into SIMCA-P 14.0 (Umetrics, Umeå, Sweden). Principal component analysis (PCA) was performed on the area values of binned NMR spectra as a non-targeting analysis. Score plots were obtained to visualize the clustering pattern of the samples along two principal components (PC1 and PC2), with each point denoting an individual sample spectrum. Unit variance was carried out on mean centered data. Supervised classification of orthogonal partial least squares discriminant analysis (OPLS-DA) was performed following PCA, to extract significant metabolite information. Loading plots indicated the metabolites responsible for group

separation. Partial least square model quality was determined by the Q^2 value, which represents model predictability, and the model with Q^2 value of > 0.5 generally indicates a good separation [23]. Similarly, the cross-validated analysis of variance (CV-ANOVA) *P* value is also used as an indicator of model quality, and the model with the *P* value of < 0.05 indicates a statistically significant separation [24].

JMP software (version Pro 14.0.0, SAS Institute, Cary, NC,) was used for statistical analysis. Data for each dietary group are presented as the mean ± standard deviation. Two-way repeated measures ANOVA was performed for urine metabolites data, and the statistical results (effects of time, diet, and the interaction of time and diet) are presented in **Figure 3-3**. Student's *t*-tests with Welch's correction were used to compare the means of each group at each time point. *P* values of < 0.05 were considered statistically significant.

3.4 Results

3.4.1 Characteristics of the experimental animals

As already described in our previous report [18], while protein intake was significantly lower in the LP group than in the CT group (0.80 ± 0.03 vs 1.87 ± 0.08 g/d, $P < 0.0001$) and energy intake was significantly higher in the LP group than in the CT group (254 ± 9 vs 217 ± 9 kJ/d, $P < 0.0001$), no significant difference in body mass was observed between the CT and LP groups at week 4 (360 ± 7.0 vs 353 ± 7.8 g, $P = 0.12$). Despite that, the plantaris muscle mass tended to be lower in the LP group than in the CT group (1.68 ± 0.08 vs 1.75 ± 0.04 g/kg body mass, $P = 0.06$).

3.4.2 Urine metabolomic profiling

Representative ^{1}H NMR spectra of urine samples at week 4 are shown in **Supplementary figure 3-S1**. Low-molecular-weight compounds, such as amino acids (AAs), AA-derived metabolites, TCA

cycle intermediates, taurine, choline, and TMAO, were identified and their chemical shifts are shown in **Supplementary table 3-S2**. Differences in the NMR spectra between the CT or LP groups were visualized using PCA at each time point (**Figure 3-1**). The score plots derived from the urine NMR spectra of the two groups did not form separate clusters at week 0, however, the separation became apparent at weeks 1–4. Moreover, the score plots showed separation pattern by the components of PC1. In further multivariate analysis, pairwise comparisons of urine NMR spectra were performed using OPLS-DA at each time point to clarify the factors that contributed to separation by giving group information (**Figure 3-2)**. The Q^2 value, an indicator of the quality of OPLS-DA models, was maximal at week 3, and the CV-ANOVA *P* value, which assesses the validation of the OPLS-DA models, was also statistically significant at week 3. Urine metabolites with peaks containing correlation coefficients of 0.70 or more at specific time points from weeks 1–4, were illustrated in **Figure 3-2**.

Relative concentrations of 18 urine metabolites at weeks 0–4, responsible for differentiating these diet groups by OPLS-DA, were then examined (**Figure 3-3** and **Supplementary table 3-S3**). Some essential AAs (phenylalanine, threonine, tryptophan, and valine), AA-derived metabolites (isobutyrate, cadaverine, and putrescine), urea, taurine, and N-methylnicotinamide were significantly lower in the LP group than in the CT group at specific time points from weeks 1–4. In contrast, TCA cycle intermediates (citrate, cis-aconitate, 2-oxoglutarate, succinate, and fumarate), choline, tartrate, and TMAO levels were significantly higher in the LP group than in the CT group at specific time points.

3.4.3 Plasma metabolomic profiling

Representative ^{1}H NMR spectra of plasma samples at week 4 are shown in **Supplementary figure 3-S2**. Assigned compounds and their chemical shifts are shown in **Supplementary table 3-S4**. The score plots derived from the plasma NMR spectra of rats fed CT or LP diets were not clearly separated in PCA (**Supplementary figure 3-S3a**). In further multivariate analysis, pairwise comparisons of

plasma NMR spectra were performed using OPLS-DA (**Supplementary figure 3-S3b**). The Q^2 value was low (< 0.5) and the CV-ANOVA *P* value was not statistically significant. Plasma metabolites, with peaks containing correlation coefficients of 0.70 or more, were glutamine, serine, valine, and urea. The relative concentration of plasma metabolites that were detected in rats fed a CT or LP diet for 4 weeks are shown in **Supplementary table 3-S5**. Plasma betaine, fumarate, glutamine, glycine, N,N-dimethylglycine, and serine levels were significantly higher in the LP group than in the CT group at week 4, whereas isoleucine, leucine, phenylalanine, threonine, urea, and valine levels were significantly lower in the LP group than in the CT group at week 4. Among them, 5 metabolites (phenylalanine, threonine, valine, urea, and fumarate) were also identified in the urine, and these levels in plasma are shown in **Figure 3-4**.

3.4.4 Liver metabolomic profiling

Representative ^{1}H NMR spectra of liver samples at week 4 are shown in **Supplementary figure 3-S4**. Assigned compounds and their chemical shifts are shown in **Supplementary table 3-S6**. Among metabolites identified in both urine and liver, some were not identified in plasma. This may be due to these metabolites binding to proteins in the blood, which would affect the pulse sequence to eliminate signals derived from protein in NMR analysis of plasma as discussed in a previous report [25]. The score plots derived from the liver NMR spectra of the two groups were clearly separated by the components of PC2 in the PCA (**Supplementary figure 3-S5a**). In further multivariate analysis, pairwise comparisons of liver NMR spectra were performed using OPLS-DA (**Supplementary figure 3-S5b**). The Q^2 value was high (≥ 0.5) and the CV-ANOVA *P* value was statistically significant. Liver metabolites, with peaks containing correlation coefficients of 0.70 or more, were inosine, uridine, NAD^+, taurine, and hypoxanthine. The relative concentration of liver metabolites that were detected in rats fed CT or LP diets for 4 weeks are shown in **Supplementary table 3-S7**. Liver β-alanine,

choline, dimethylamine, hypoxanthine, inosine monophosphate, and TMAO levels were significantly higher in the LP group than in the CT group at week 4, whereas inosine, nicotinamide adenine, NAD^+, o-phosphocholine, taurine, and uridine levels were significantly lower in the LP group than in the CT group at week 4. Among them, 3 metabolites (taurine, TMAO, and choline) were also identified in the urine, and these levels in the liver are shown in **Figure 3-5**.

3.4.5 mRNA microarray for hepatic transcriptomic profiling

When GO analysis following hepatic mRNA microarray was performed to investigate how metabolite fluctuations occurred, the metabolic process pathway was determined to be the most significantly altered among the biological processes. The metabolic process pathway encompasses several sub-pathways, including cellular metabolic process, the organic substance metabolic process, and the biosynthetic process (**Table 3-1**).

mRNA microarray analysis revealed that a total of 19,598 gene expressions were confirmed in the livers, of which 3,056 were significantly differentially expressed between the CT and LP groups. Among them, 352 genes were significantly over-expressed in the LP group than in the CT group by more than 1.5-fold, whereas 362 genes were significantly under-expressed in the LP group than in the CT group by more than 1.5-fold (**Supplementary table 3-S8**). Gene expressions associated with 8 metabolites that were identified in both urine and plasma/liver were further analyzed by quantitative PCR to validate the microarray findings (**Figure 3-6**). The gene expression levels of argininosuccinate lyase (*Asl*), argininosuccinate synthase 1 (*Ass1*), and cysteine dioxygenase type 1 (*Cdo1*) were significantly lower in the LP group than in the CT group. In contrast, the gene expression levels of glucose-6-phosphate isomerase**,** pyruvate carboxylase, isocitrate dehydrogenase 2, flavin-containing monooxygenase 1 (*Fmo1*), and *Fmo5* were significantly higher in the LP group than in the CT group.

3.5 Discussion

We performed metabolomic and transcriptomic analyses in adult rats fed a 14% CN diet (CT diet) or 5% CN diet (LP diet) for 4 weeks. According to our previous report using the same experimental animals as this study [18], no significant differences in body mass were observed between the two groups, whereas plantaris muscle mass tended to be lower in the LP group than in the CT group (P = 0.06), implicating the manifestation of potential protein undernutrition in the LP group. Furthermore, GO analysis indicated that protein undernutrition perturbed metabolic processes such as the cellular metabolic process, the organic substance metabolic process, and the biosynthetic process, which is likely, in turn, to lead to fluctuations in metabolite levels. Thus, the novelty of this study was to identify metabolites as biomarker candidates responsive to protein undernutrition using metabolomics and transcriptomic approaches.

Metabolites identified in both urine and plasma were phenylalanine, threonine, valine, urea, and fumarate. In urine and plasma, the levels of AAs and urea were significantly lower in the LP group than in the CT group, respectively. These observations are presumed to be caused by reduced AA/protein intake. Furthermore, lowered urinary phenylalanine excretions in rats fed the LP diet were observed at weeks 1 and 2, whereas lowered threonine excretion was observed at week 3. Hence, different durations of protein undernutrition may impact the pattern of AA/protein metabolism. In line with AA excretion, hepatic expression levels of *Asl* and *Ass1*, which encode enzymes in the urea cycle, were significantly lower in the LP group than in the CT group. This is likely due to the decreased metabolism of AA-derived nitrogen, which would stagnate the urea cycle. Similarly, decreased protein intake is reported to reduce the hepatic expression of genes which encode urea cycle-related enzymes [26]. Protein insufficiency may facilitate nitrogen conservation in tissues, and decrease hepatic *Asl* and *Ass1* expressions in the LP group. In contrast, fumarate levels in urine and plasma were significantly higher in the LP group than in the CT group, respectively. The observed increase in

fumarate, a TCA intermediate, may be due to higher carbohydrate contents in the LP diet. Hepatic gene expression levels of glycolytic enzymes glucose-6-phosphate isomerase, pyruvate carboxylase, and the TCA cycle enzyme isocitrate dehydrogenase 2 were also significantly higher in the LP group than in the CT group. These suggest that glucose metabolism was activated by increased carbohydrate intake accompanied by high carbohydrate contents in the LP diet. However, it should be noted that TCA cycle acceleration following high carbohydrate intake has been linked to the progression of various diseases, such as cancer and atherosclerosis, through the overproduction of reactive oxygen species in the mitochondria [27]. Furthermore, fumarate accumulation can lead to oxidative stress by modifying glutathione metabolism [28] and induce endoplasmic reticulum stress in a diabetic model [29]. Thus, the increase in the TCA cycle activity and its intermediate fumarate associated with LP diet ingestion may increase the risk of developing specific diseases.

The OPLS-DA showed that the Q^2 value of liver samples was higher than that of plasma samples, and the CV-ANOVA *P* value of liver samples was statistically significant, suggesting that the liver reflected protein undernutrition more substantially than plasma. Thus, metabolites such as taurine, TMAO, and choline, which fluctuated not only in urine but also in the liver, may be sensitive to protein undernutrition. Notably, urinary taurine levels were significantly lower in the LP group than in the CT group at week 1 and remained constant thereafter. The liver taurine level was also significantly lower in the LP group than in the CT group at week 4. The activity of CDO1, the rate-limiting enzyme in the biosynthesis of taurine from cysteine [30], is considerably higher in the liver than other tissues [31]. Thus, decreased urinary and hepatic taurine levels could have been caused by the stagnation of cysteine metabolism associated with reduced hepatic expression of *Cdo1*. Taurine contributes to numerous biological functions through its anti-oxidative and anti-inflammatory properties [32]. It is reported that taurine deficiency induces skeletal muscle loss [33], while taurine intake improves skeletal muscle function by preventing muscle degeneration [34, 35]. Muscle degeneration is related to oxidative stress

in the muscle and promotes muscle cell apoptosis. It is reported that taurine can protect against oxidative stress as well as reducing the abundance of apoptotic signal proteins [36, 37]. In our previous study [18] examining the same experimental animals in this study, the plantaris muscle mass tended to be lower in the LP group than in the CT group ($P = 0.06$). Collectively, this might be mediated in part by decreased taurine levels in the body. Although gene expression analyses involved in muscle synthesis and degradation in muscle tissues are lacking in this study, it would be necessary to investigate the molecular mechanism to clarify the relationship between taurine concentration and muscle mass. Taurine deficiency is also reported to affect energy metabolism in adipose tissue [38]. Taurine protects against dyslipidemia by inhibiting reactive oxygen species generation [39, 40], and taurine supplementation can normalize serum high-density lipoprotein (HDL) concentrations in a high cholesterol diet-fed rats [41]. It has also been reported that protein intake and serum HDL cholesterol levels are positively correlated [42]. Furthermore, in our previous study using the same experimental animals as in the current study [18], the LP group exhibited significantly lower HDL cholesterol compared with the CT group. Considering these facts, taurine level may be involved in dyslipidemia associated with protein undernutrition. In contrast to the decreased taurine level in urine and liver of the LP group, urinary TMAO levels were significantly higher in the LP group than in the CT group at week 2 and it remained constant thereafter. The liver TMAO level was also significantly higher in the LP group than in the CT group at week 4. In our metabolomic analysis, choline in urine and liver were also higher in the LP group than in the CT group. Choline is converted to TMA by intestinal bacteria and TMAO is synthesized from TMA by the FMO family in the liver [43]. Hepatic gene expressions of *Fmo1* and *Fmo5* were also significantly higher in the LP group than in the CT group, respectively. Therefore, it is speculated that the increase in TMAO level were associated with an increase in both its substrate choline and the gene expression levels of *Fmo1* and *5*. The urinary tartrate level was also significantly higher in the LP group than in the CT group. Since choline was provided in both CT and

LP diets as choline bitartrate, the increased urinary tartrate excretion observed in the LP-fed rats may indicate the use of choline as a nitrogen source in protein undernutrition, as observed previously [44]. Besides, as protein deficiency has been suggested to alter the gut microbiota [44, 45], increased TMAO levels may also be indirectly caused by fluctuations in intestinal microbiota. Clinical trials have demonstrated that elevated TMAO level increases the risk of cardiovascular disease [43, 46]; while consumption of a high TMAO-containing diet reduces serum HDL levels in mice [47]. A recent review suggests that TMAO-induced atherosclerosis is associated with decreased HDL cholesterol, which increases the risk of cardiovascular disease [48]. Furthermore, as mentioned above, our previous study using the same animals reported that protein undernutrition reduced HDL cholesterol [18]. Thus, the increase in TMAO levels in the LP group may also be associated with decreased circulating HDL-cholesterol levels. Collectively, the increase in choline utilization as a nitrogen source induced by protein undernutrition may increase the risk of dyslipidemia and cardiovascular disorder via elevating TMAO production. It should also be noted that the plasma TMAO level is affected by the intake of fish containing TMAO and *de novo* synthesis from carnitine contained in red meat [49, 50]. Therefore, TMAO intake derived from diets should be taken into account for comprehending the TMAO level fluctuation in the body.

In conclusion, the fluctuation of several metabolites identified in urine, plasma, and liver samples were caused by the LP diet ingestion in adult rats. Among the identified metabolites, taurine and TMAO, the final metabolites that are not further metabolized, fluctuated not only in urine but also in the liver. Notably, changes in taurine and TMAO levels were detectable in the urine by week 2 after LP diet ingestion, which may reflect protein undernutrition at an early stage. Hepatic gene expressions of taurine and TMAO biosynthesis enzymes also corresponded to these fluctuations. Thus, fluctuations in urinary taurine and TMAO are considered to have occurred as a result of substantially responses to protein undernutrition in the liver. Furthermore, taurine and TMAO might also be potential mediators

of the development of skeletal muscle disorders and lipid abnormalities, which could be associated with LP nutritional status as discussed above. It follows from these results that taurine and TMAO may serve as potential biomarker candidates of protein undernutrition and targets for prevention of secondary diseases associated with protein undernutrition. To further elucidate in-depth the molecular mechanisms involved in taurine and TMAO in the context of protein nutritional status, it may be necessary to determine whether fluctuations in taurine and TMAO levels dissolve as the protein nutritional status improves. In addition, clinical trials are warranted to investigate their usefulness as protein nutrition biomarkers.

3.6 Tables and figures

Table 3-1. Gene ontology (GO) terms for functional analysis of differentially expressed genes*

GO term description		Number of genes	FDR corrected *P* value
GO:0008150 Biological process			
GO:0008152	Metabolic process	444	$P < 0.05$
GO:0044237	Cellular metabolic process	409	$P < 0.05$
GO:0071704	Organic substance metabolic process	423	$P < 0.05$
GO:0009058	Biosynthetic process	226	$P < 0.05$
GO:0044281	Small molecule metabolic process	141	$P < 0.05$
GO:0044238	Primary metabolic process	399	$P < 0.05$
GO:0071840	Cellular component organization or biogenesis	223	0.16
GO:0009987	Cellular process	562	1.00
GO:0065007	Biological regulation	393	1.00

*GO term description and GO accession number were referenced to The Gene Ontology Resource, the web-available bioinformatics database (http://www. geneontology.org/). *P* values of < 0.05 were considered statistically significant.

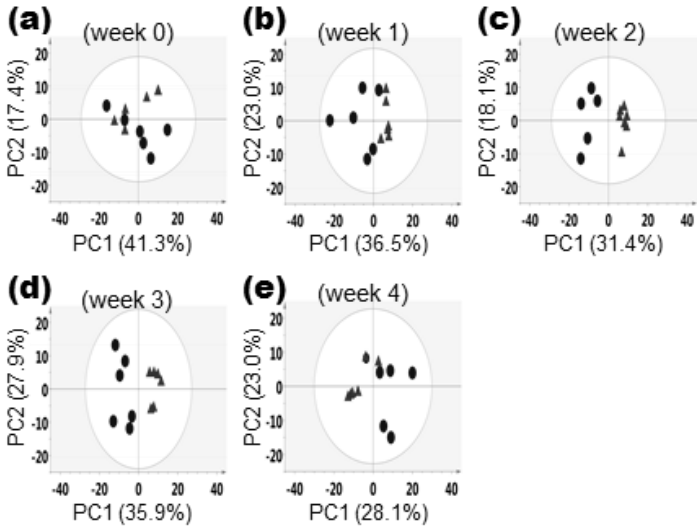


Figure 3-1. Principal component analysis (PCA) score plots derived from the NMR spectra of urine samples from rats fed control (CT) or low protein (LP) diets at (a) week 0, (b) week 1, (c) week 2, (d) week 3, and (e) week 4. Each symbol represents the urine sample from an individual animal (n = 6).

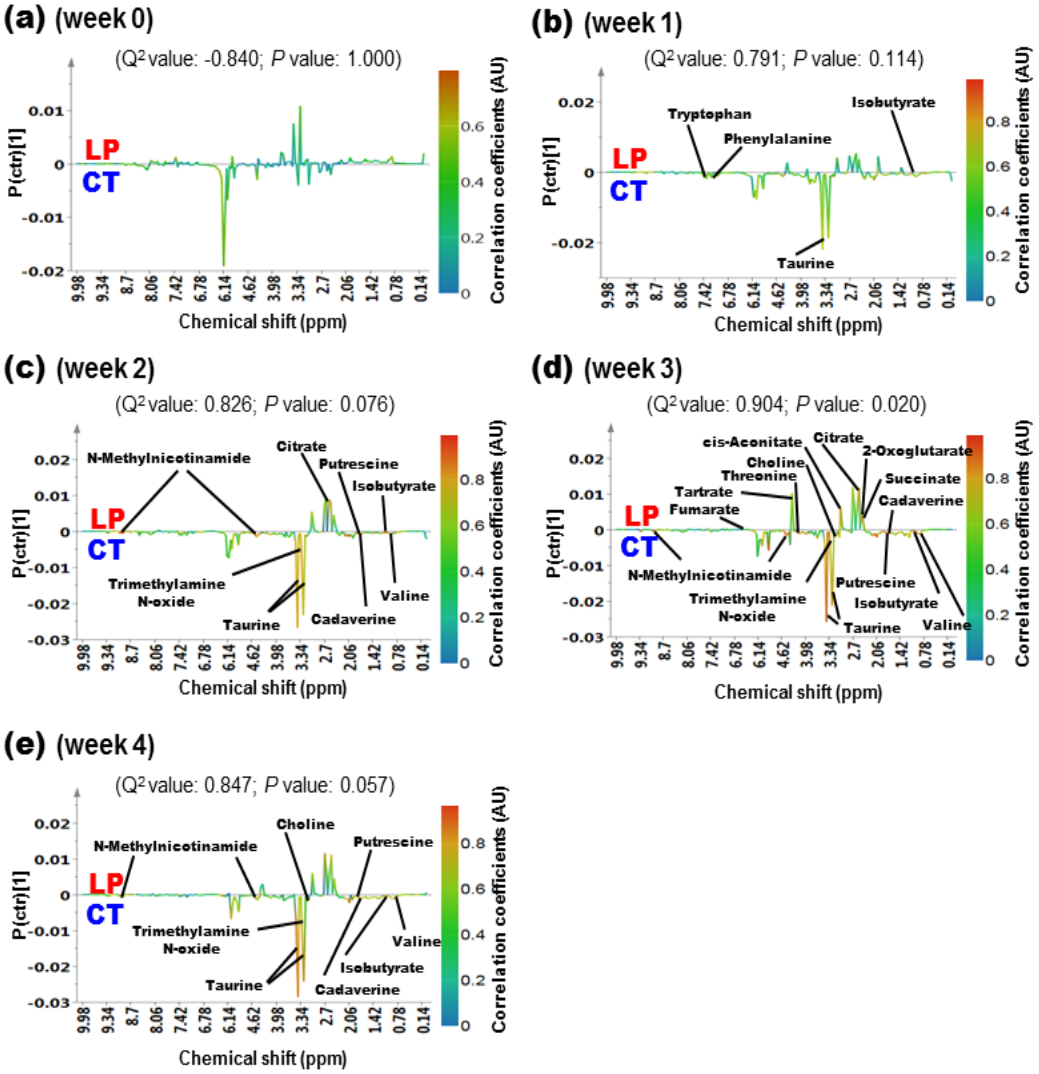


Figure 3-2. Multivariate analysis of binned NMR spectra of urine samples from rats fed control (CT) or low protein (LP) diets at weeks 0-4. Coefficient loading plots obtained from rats fed CT or LP diets at (a) week 0, (b) week 1, (c) week 2, (d) week 3, and (e) week 4. Peaks pointing upward indicate metabolites whose levels were higher in the LP group than in the CT group, whereas peaks pointing downward indicate metabolites whose levels were higher in the CT group than in the LP group. Heat colors indicate the contribution degree of metabolites that separate the CT and LP groups. Among the assigned metabolites, those with NMR spectra with correlation coefficients of 0.70 or more are annotated in the figure. Partial least square model quality was determined by Q^2 value, which represent model predictability, and the *P* value of cross-validated analysis of variance. Each value is shown in the figure.

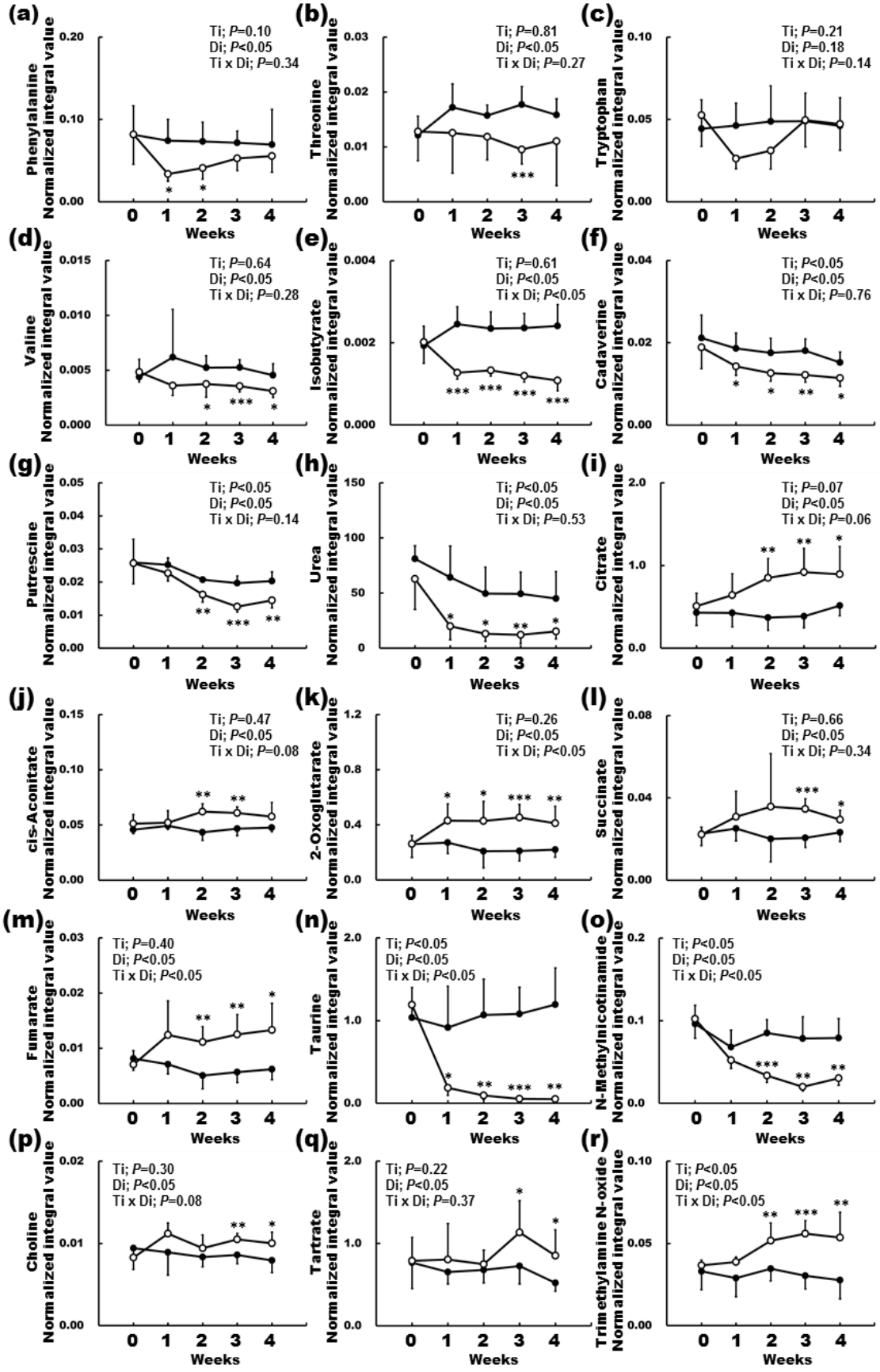

(a)
Phenylalanine
Normalized integral value
Ti; P=0.10
Di; P<0.05
Ti x Di; P=0.34
(b)
Threonine
Normalized integral value
Ti; P=0.81
Di; P<0.05
Ti x Di; P=0.27
(c)
Tryptophan
Normalized integral value
Ti; P=0.21
Di; P=0.18
Ti x Di; P=0.14
(d)
Valine
Normalized integral value
Ti; P=0.64
Di; P<0.05
Ti x Di; P=0.28
(e)
Isobutyrate
Normalized integral value
Ti; P=0.61
Di; P<0.05
Ti x Di; P<0.05
(f)
Cadaverine
Normalized integral value
Ti; P<0.05
Di; P<0.05
Ti x Di; P=0.76
(g)
Putrescine
Normalized integral value
Ti; P<0.05
Di; P<0.05
Ti x Di; P=0.14
(h)
Urea
Normalized integral value
Ti; P<0.05
Di; P<0.05
Ti x Di; P=0.53
(i)
Citrate
Normalized integral value
Ti; P=0.07
Di; P<0.05
Ti x Di; P=0.06
(j)
cis-Aconitate
Normalized integral value
Ti; P=0.47
Di; P<0.05
Ti x Di; P=0.08
(k)
2-Oxoglutarate
Normalized integral value
Ti; P=0.26
Di; P<0.05
Ti x Di; P<0.05
(l)
Succinate
Normalized integral value
Ti; P=0.66
Di; P<0.05
Ti x Di; P=0.34
(m)
Fumarate
Normalized integral value
Ti; P=0.40
Di; P<0.05
Ti x Di; P<0.05
(n)
Taurine
Normalized integral value
Ti; P<0.05
Di; P<0.05
Ti x Di; P<0.05
(o)
N-Methylnicotinamide
Normalized integral value
Ti; P<0.05
Di; P<0.05
Ti x Di; P<0.05
(p)
Choline
Normalized integral value
Ti; P=0.30
Di; P<0.05
Ti x Di; P=0.08
(q)
Tartrate
Normalized integral value
Ti; P=0.22
Di; P<0.05
Ti x Di; P=0.37
(r)
Trimethylamine N-oxide
Normalized integral value
Ti; P<0.05
Di; P<0.05
Ti x Di; P<0.05
Weeks

Figure 3-3. Relative concentrations of urinary metabolites responsible for differentiating the control (CT) and low protein (LP) diet groups by orthogonal partial least squares discriminant analysis in urine samples at weeks 0-4. The levels of (a) Phenylalanine, (b) Threonine, (c) Tryptophan, (d) Valine, (e) Isobutyrate, (f) Cadaverine, (g) Putrescine, (h) Urea, (i) Citrate, (j) cis-Aconitate, (k) 2-Oxoglutarate, (l) Succinate, (m) Fumarate, (n) Taurine, (o) N-Methylnicotinamide, (p) Choline, (q) Tartrate, and (r) Trimethylamine N-oxide are shown. The relative concentration of each metabolite is its integral value normalized to TSP, which was added to the NMR buffer as an internal standard. Data are expressed as the mean ± SD (n = 6) and were analyzed by Student's *t*-tests with Welch's correction. $*P < 0.05$, $**P < 0.01$, and $***P < 0.001$.

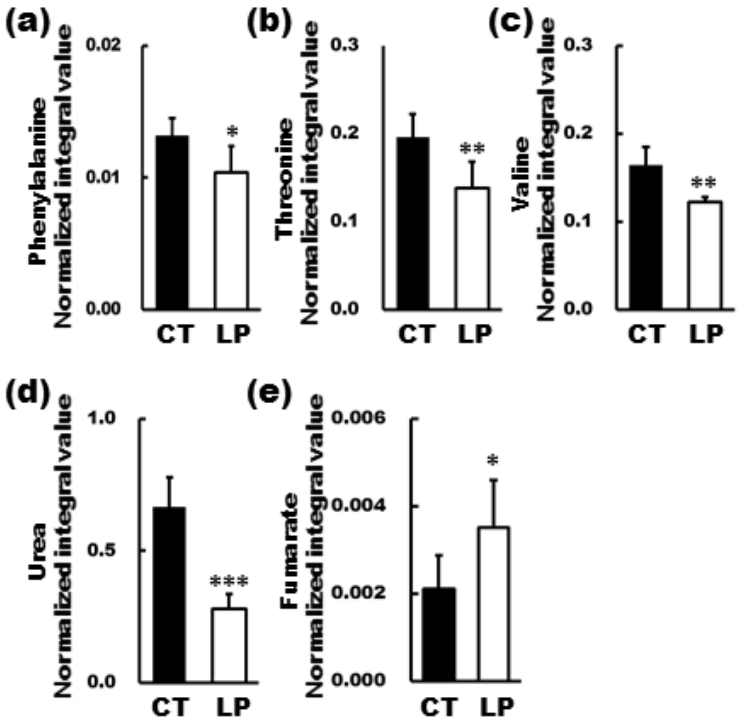


Figure 3-4. Relative concentrations of metabolites in plasma samples of rats fed control (CT) or low protein (LP) diets for 4 weeks. The levels of (a) Phenylalanine, (b) Threonine, (c) Valine, (d) Urea, and (e) Fumarate are shown. The relative concentration of each metabolite is its integral value normalized to formic acid, which was added to the NMR buffer as an internal standard. Data are expressed as the mean ± SD (n = 6) and were analyzed by Student's *t*-tests with Welch's correction. **P* < 0.05, ***P* < 0.01, and ****P* < 0.001.

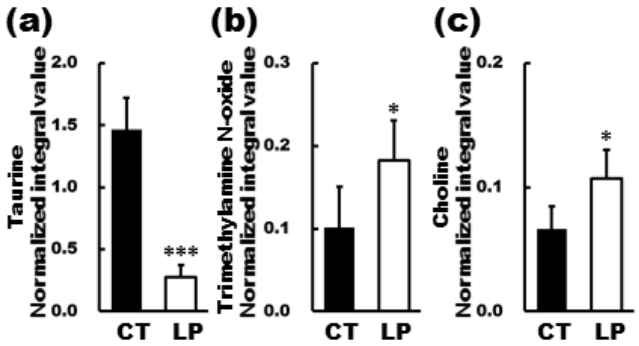


Figure 3-5. Relative concentrations of metabolites in liver samples of rats fed control (CT) or low protein (LP) diets for 4 weeks. The levels of (a) Taurine, (b) Trimethylamine N-oxide, and (c) Choline are shown. The relative concentration of each metabolite is its integral value normalized to TSP, which was added to the NMR buffer as an internal standard. Data are expressed as the mean ± SD (n = 6) and were analyzed by Student's *t*-tests with Welch's correction. **P* < 0.05 and ****P* < 0.001.

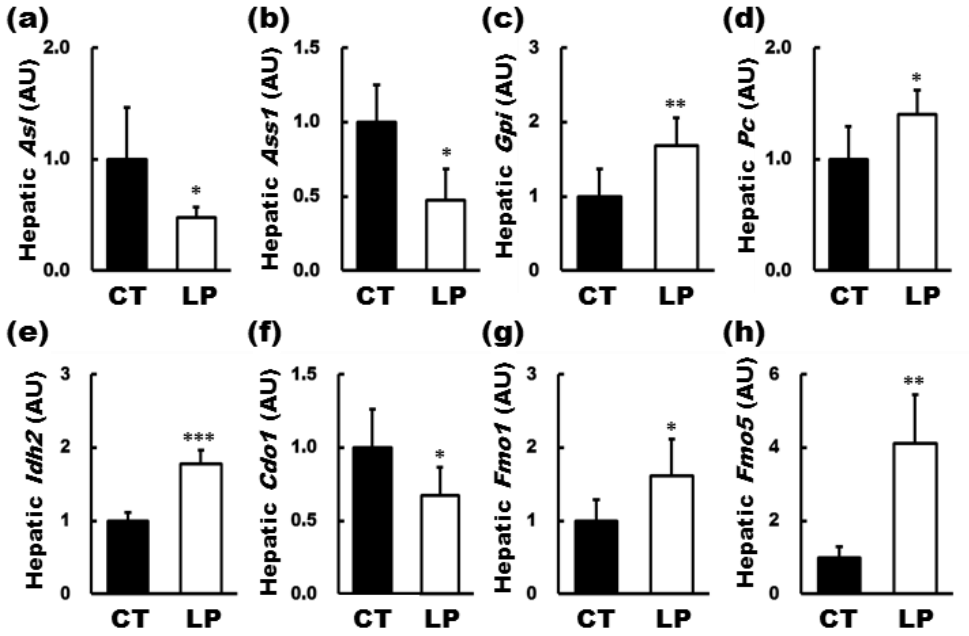


Figure 3-6. Hepatic gene expression levels in rats fed control (CT) or low protein (LP) diets for 4 weeks. Relative expression levels of (a) *Asl*, (b) *Ass1*, (c) *Gpi*, (d) *Pc*, (e) *Idh2*, (f) *Cdo1*, (g) *Fmo1*, and (h) *Fmo5* were examined at week 4. Among genes differentially expressed between the CT and LP diet groups by mRNA microarray, only those related to fluctuating metabolites were examined. Values are normalized to the CT group. Data are shown as the mean ± SD (n = 6) and were analyzed by Student's *t*-tests with Welch's correction. **P* < 0.05, ***P* < 0.01, and ****P* < 0.001. *Asl*, *Argininosuccinate lyase*; *Ass1*, *Argininosuccinate synthase 1*; *Cdo1*, *Cysteine dioxygenase type 1*; *Fmo*, *Flavin containing monooxygenase*; *Gpi*, *Glucose-6-phosphate isomerase*; *Idh2*, *Isocitrate dehydrogenase 2*; *Pc*, *Pyruvate carboxylase*.

3.7 Supplementary tables and figures

Supplementary table 3-S1. PCR primer and probe set information

NCBI Entrez Gene ID*	Gene Symbol	Gene Name	TaqMan Gene Expression Assay ID
59085	*Asl*	Argininosuccinate lyase	Rn01480437_g1
25698	*Ass1*	Argininosuccinate synthase 1	Rn00565808_g1
292804	*Gpi*	Glucose-6-phosphate isomerase	Rn01475756_m1
25104	*Pc*	Pyruvate carboxylase	Rn00562534_m1
361596	*Idh2*	Isocitrate dehydrogenase 2	Rn01478119_m1
81718	*Cdo1*	Cysteine dioxygenase type 1	Rn00583853_m1
25256	*Fmo1*	Flavin containing monooxygenase 1	Rn00562945_m1
246248	*Fmo5*	Flavin containing monooxygenase 5	Rn00595199_m1
81822	*Actb*	β-Actin	Rn00667869_m1

*Gene IDs reference The National Center for Biotechnology Information database.

Supplementary table 3-S2. Metabolite assignments and ^{1}H NMR data in urine samples of rat*

	Urine metabolite	δ^{1}H (ppm)						
1	N-Methylnicotinamide	4.48(s)	8.18(t)	8.90(d)	8.97(d)			
2	Nicotinamide N-oxide	7.73(dd)	8.12(m)	8.48(m)				
3	Formate	8.46(s)						
4	Hippurate	3.97(d)	7.55(t)	7.64(m)	7.84(dd)	8.56(s)		
5	Tryptophan	3.30(dd)	3.47(dd)	4.05(dd)	7.20(m)	7.28(m)	7.50(d)	7.71(d)
6	Phenylalanine	3.11(dd)	3.28(dd)	3.99(dd)	7.36(m)	7.42(m)		
7	Tyrosine	3.04(dd)	3.19(dd)	3.93(dd)	6.89(m)	7.17(m)		
8	Fumarate	6.53(s)						
9	Allantoin	5.39(s)	6.05(s)	7.26(s)	8.01(s)			
10	Urea	5.80(s)						
11	Glucose	3.24(dd)	3.39(t)	3.40(t)	3.45(m)	3.48(t)	3.53(dd)	3.70(t)
		3.72(dd)	3.76(dd)	3.82(m)	3.84(dd)	3.89(dd)	4.65(d)	5.24(d)
12	Tartrate	4.36(s)						
13	Threonine	1.33(d)	3.58(d)	4.24(m)				
14	Creatinine	3.05(s)	4.06(s)					
15	Creatine	3.02(s)	3.95(s)					
16	Glycine	3.57(s)						
17	Taurine	3.27(t)	3.43(t)					
18	Choline	3.20(s)	3.51(m)	4.06(m)				
19	cis-Aconitate	3.12(d)	5.69(t)					
20	Cadaverine	1.46(m)	1.72(m)	3.01(t)				
21	N,N-Dimethylglycine	2.93(s)	3.71(s)					
22	Dimethylamine	2.72(s)						
23	Citrate	2.55(d)	2.69(d)					
24	2-Oxoglutarate	2.44(t)	3.01(t)					
25	Succinate	2.41(s)						
26	Acetoacetate	2.30(s)	3.45(s)					
27	Trimethylamine N-oxide	3.27(s)						
28	Acetate	1.92(s)						
29	Putrescine	1.77(m)	3.05(m)					
30	Alanine	1.48(d)	3.78(q)					
31	Lactate	1.33(d)	4.11(q)					
32	3-Hydroxybutyrate	1.18(d)	2.30(dd)	2.39(dd)	4.14(m)			
33	Isobutyrate	1.07(d)	2.38(m)					
34	Valine	0.99(d)	1.05(d)	2.26(m)	3.60(d)			

*Key: s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; dd, doublet of doublet.

Supplementary table 3-S3. Relative concentration of the metabolites in urea samples of rats fed control (CT) or low protein (LP) diets at weeks 0-4*

Urine metabolite		CT diet group		LP diet group	*P* value	Urine metabolite		CT diet group		LP diet group	*P* value
Phenylalanine	**0w**	0.0813 ± 0.0352	**0w**	0.0819 ± 0.0366	0.98	**Cadaverine**	**0w**	0.2120 ± 0.0056	**0w**	0.0189 ± 0.0052	0.48
	1w	0.0743 ± 0.0259	**1w**	0.0336 ± 0.0089	< 0.05		**1w**	0.0186 ± 0.0037	**1w**	0.0143 ± 0.0022	< 0.05
	2w	0.0732 ± 0.0235	**2w**	0.0410 ± 0.0137	< 0.05		**2w**	0.0175 ± 0.0036	**2w**	0.0126 ± 0.0020	< 0.05
	3w	0.0715 ± 0.0144	**3w**	0.0527 ± 0.0150	0.05		**3w**	0.0180 ± 0.0029	**3w**	0.0122 ± 0.0018	< 0.01
	4w	0.0693 ± 0.0429	**4w**	0.0556 ± 0.0196	0.50		**4w**	0.0152 ± 0.0025	**4w**	0.0114 ± 0.0021	< 0.05
Threonine	**0w**	0.0121 ± 0.0034	**0w**	0.0128 ± 0.0053	0.81	**Putrescine**	**0w**	0.0259 ± 0.0070	**0w**	0.0257 ± 0.0062	0.95
	1w	0.0172 ± 0.0043	**1w**	0.0125 ± 0.0073	0.21		**1w**	0.0252 ± 0.0022	**1w**	0.0227 ± 0.0024	0.09
	2w	0.0157 ± 0.0019	**2w**	0.0118 ± 0.0042	0.07		**2w**	0.0207 ± 0.0008	**2w**	0.0162 ± 0.0024	< 0.01
	3w	0.0177 ± 0.0032	**3w**	0.0095 ± 0.0027	< 0.001		**3w**	0.0197 ± 0.0021	**3w**	0.0125 ± 0.0017	< 0.001
	4w	0.0158 ± 0.0029	**4w**	0.0110 ± 0.0081	0.22		**4w**	0.0203 ± 0.0027	**4w**	0.0144 ± 0.0023	< 0.01
Tryptophan	**0w**	0.0443 ± 0.0176	**0w**	0.0525 ± 0.0190	0.45	**Urea**	**0w**	81.164 ± 11.752	**0w**	62.853 ± 27.71	0.18
	1w	0.0463 ± 0.0135	**1w**	0.0261 ± 0.0061	< 0.01		**1w**	64.175 ± 28.652	**1w**	19.856 ± 12.237	< 0.05
	2w	0.0487 ± 0.0217	**2w**	0.0310 ± 0.0111	0.11		**2w**	49.556 ± 24.065	**2w**	13.159 ± 6.864	< 0.05
	3w	0.0490 ± 0.0171	**3w**	0.0495 ± 0.0162	0.95		**3w**	49.287 ± 19.809	**3w**	11.994 ± 7.843	< 0.01
	4w	0.0461 ± 0.0170	**4w**	0.0472 ± 0.0161	0.91		**4w**	44.974 ± 24.440	**4w**	15.073 ± 6.749	< 0.05
Valine	**0w**	0.0044 ± 0.0016	**0w**	0.0048 ± 0.0009	0.55	**Citrate**	**0w**	0.4304 ± 0.1563	**0w**	0.5085 ± 0.1560	0.41
	1w	0.0062 ± 0.0044	**1w**	0.0036 ± 0.0009	0.21		**1w**	0.4282 ± 0.1729	**1w**	0.6411 ± 0.2603	0.13
	2w	0.0052 ± 0.0011	**2w**	0.0037 ± 0.0012	< 0.05		**2w**	0.3701 ± 0.1564	**2w**	0.8490 ± 0.2354	< 0.01
	3w	0.0053 ± 0.0007	**3w**	0.0035 ± 0.0005	< 0.001		**3w**	0.3850 ± 0.1378	**3w**	0.9208 ± 0.2860	< 0.01
	4w	0.0045 ± 0.0010	**4w**	0.0031 ± 0.0006	< 0.05		**4w**	0.5150 ± 0.1242	**4w**	0.8961 ± 0.3340	< 0.05
Isobutyrate	**0w**	0.0019 ± 0.0005	**0w**	0.0020 ± 0.0005	0.77	**cis-Aconitate**	**0w**	0.0457 ± 0.0036	**0w**	0.0512 ± 0.0082	0.18
	1w	0.0025 ± 0.0004	**1w**	0.0013 ± 0.0002	< 0.001		**1w**	0.0493 ± 0.0034	**1w**	0.0521 ± 0.0108	0.56
	2w	0.0023 ± 0.0004	**2w**	0.0013 ± 0.0001	< 0.001		**2w**	0.0433 ± 0.0074	**2w**	0.0620 ± 0.0072	< 0.01
	3w	0.0024 ± 0.0004	**3w**	0.0012 ± 0.0002	< 0.001		**3w**	0.0465 ± 0.0062	**3w**	0.0609 ± 0.0054	< 0.01
	4w	0.0024 ± 0.0005	**4w**	0.0011 ± 0.0003	< 0.001		**4w**	0.0476 ± 0.0038	**4w**	0.0574 ± 0.0129	0.13

*The relative concentration of each metabolite is its integral value normalized to TSP, which was added to the NMR buffer as an internal standard. Data are expressed as the mean ± standard deviation (n = 6) and were analyzed by Student's *t*-tests with Welch's correction.

Supplementary table 3-S3, continued*

Urine metabolite		CT diet group		LP diet group	*P* value
2-Oxoglutarate	**0w**	0.2594 ± 0.0964	**0w**	0.2612 ± 0.0618	0.97
	1w	0.2720 ± 0.0801	**1w**	0.4315 ± 0.1210	< 0.05
	2w	0.2085 ± 0.1213	**2w**	0.4296 ± 0.1409	< 0.05
	3w	0.2113 ± 0.0732	**3w**	0.4534 ± 0.0953	< 0.001
	4w	0.2207 ± 0.0550	**4w**	0.4116 ± 0.1234	< 0.01
Succinate	**0w**	0.0225 ± 0.0058	**0w**	0.0222 ± 0.0035	0.90
	1w	0.0251 ± 0.0061	**1w**	0.0308 ± 0.0125	0.34
	2w	0.0201 ± 0.0112	**2w**	0.0358 ± 0.0258	0.24
	3w	0.0205 ± 0.0047	**3w**	0.0346 ± 0.0049	< 0.001
	4w	0.0232 ± 0.0044	**4w**	0.0294 ± 0.0046	< 0.05
Fumarate	**0w**	0.0081 ± 0.0021	**0w**	0.0071 ± 0.0025	0.44
	1w	0.0071 ± 0.0017	**1w**	0.0124 ± 0.0062	0.09
	2w	0.0050 ± 0.0024	**2w**	0.0111 ± 0.0028	< 0.01
	3w	0.0057 ± 0.0019	**3w**	0.0125 ± 0.0036	< 0.01
	4w	0.0062 ± 0.0019	**4w**	0.0133 ± 0.0048	< 0.05
Taurine	**0w**	1.0355 ± 0.3644	**0w**	1.1928 ± 0.1792	0.37
	1w	0.9182 ± 0.4974	**1w**	0.1861 ± 0.0909	< 0.05
	2w	1.0685 ± 0.4338	**2w**	0.0944 ± 0.0506	< 0.01
	3w	1.0803 ± 0.3228	**3w**	0.0540 ± 0.0116	< 0.001
	4w	1.1937 ± 0.4426	**4w**	0.0499 ± 0.0075	< 0.01
N-Methylnicotinamide	**0w**	0.0961 ± 0.0226	**0w**	0.1022 ± 0.0236	0.65
	1w	0.0679 ± 0.0207	**1w**	0.0523 ± 0.0103	0.13
	2w	0.0850 ± 0.0162	**2w**	0.0335 ± 0.0081	< 0.001
	3w	0.0785 ± 0.0262	**3w**	0.0200 ± 0.0049	< 0.01
	4w	0.0793 ± 0.0232	**4w**	0.0306 ± 0.0087	< 0.01

Urine metabolite		CT diet group		LP diet group	*P* value
Choline	**0w**	0.0094 ± 0.0026	**0w**	0.0083 ± 0.0010	0.35
	1w	0.0089 ± 0.0028	**1w**	0.1120 ± 0.0013	0.10
	2w	0.0084 ± 0.0012	**2w**	0.0094 ± 0.0016	0.22
	3w	0.0086 ± 0.0011	**3w**	0.0105 ± 0.0007	< 0.01
	4w	0.0079 ± 0.0015	**4w**	0.0100 ± 0.0014	< 0.05
Tartrate	**0w**	0.7685 ± 0.3176	**0w**	0.7907 ± 0.2839	0.90
	1w	0.6536 ± 0.1449	**1w**	0.8066 ± 0.4344	0.44
	2w	0.6794 ± 0.1602	**2w**	0.7488 ± 0.1689	0.48
	3w	0.7275 ± 0.2203	**3w**	1.1332 ± 0.3873	< 0.05
	4w	0.5213 ± 0.1025	**4w**	0.8518 ± 0.3132	< 0.05
Trimethylamine N-oxide	**0w**	0.0330 ± 0.0111	**0w**	0.0367 ± 0.0032	0.47
	1w	0.0289 ± 0.0114	**1w**	0.0387 ± 0.0031	0.09
	2w	0.0346 ± 0.0075	**2w**	0.0518 ± 0.0107	< 0.01
	3w	0.0303 ± 0.0081	**3w**	0.0559 ± 0.0080	< 0.001
	4w	0.0277 ± 0.0114	**4w**	0.0536 ± 0.0153	< 0.01

*The relative concentration of each metabolite is its integral value normalized to TSP, which was added to the NMR buffer as an internal standard. Data are expressed as the mean ± standard deviation (n = 6) and were analyzed by Student's *t*-tests with Welch's correction.

Supplementary table 3-S4. Metabolite assignments and ^{1}H NMR data in plasma samples of rat*

	Plasma metabolite	δ^{1}H (ppm)						
1	3-Hydroxybutyrate	1.22(d)	2.32(dd)	2.37(dd)	4.14(m)			
2	Acetate	1.91(s)						
3	Acetoacetate	2.22(s)	3.43(s)					
4	Alanine	1.47(d)	3.78(q)					
5	Betaine	3.26(s)	3.92(s)					
6	Choline	3.21(s)	3.47(m)	4.06(m)				
7	Creatine	3.02(s)	3.93(s)					
8	Creatine phosphate	3.03(s)	3.94(s)					
9	Formate	8.44(s)						
10	Fumarate	6.52(s)						
11	Glucose	3.24(dd)	3.40(t)	3.41(t)	3.46(m)	3.48(t)	3.53(dd)	3.71(t)
		3.72(dd)	3.77(dd)	3.82(m)	3.84(dd)	3.90(dd)	4.64(d)	5.23(d)
12	Glutamine	2.12(m)	2.14(m)	2.42(m)	2.46(m)	3.77(t)	6.87(s)	
13	Glycine	3.57(s)						
14	Histidine	3.12(dd)	3.22(dd)	3.99(dd)	7.06(s)	7.82(s)		
15	Isoleucine	0.93(t)	1.00(d)	1.25(m)	1.46(m)	1.97(m)	3.66(d)	
16	Lactate	1.32(d)	4.11(q)					
17	Leucine	0.95(d)	0.96(d)	1.68(m)	1.70(m)	1.73(m)	3.73(dd)	
18	Methionine	2.11(m)	2.12(s)	2.19(m)	2.63(t)	3.84(dd)		
19	N,N-Dimethylglycine	2.91(s)	3.71(s)					
20	Phenylalanine	3.12(dd)	3.28(dd)	3.99(dd)	7.31(d)	7.35(m)	7.41(m)	
21	Pyruvate	2.36(s)						
22	Serine	3.83(dd)	3.94(dd)	3.98(dd)				
23	Succinate	2.40(s)						
24	Threonine	1.32(d)	3.58(d)	4.25(m)				
25	Tyrosine	3.04(dd)	3.19(dd)	3.93(dd)	6.88(m)	7.17(m)		
26	Urea	5.84(s)						
27	Valine	0.98(d)	1.03(d)	2.26(m)	3.61(d)			

*Key: s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; dd, doublet of doublet.

Supplementary table 3-S5. Relative concentration of the metabolites in plasma samples of rats fed control (CT) or low protein (LP) diets for 4 weeks*

Plasma metabolite	CT diet group	LP diet group	*P* value
Significantly higher in the LP diet group			
Betaine	0.050 ± 0.016	0.078 ± 0.014	< 0.05
Fumarate	0.002 ± 0.001	0.004 ± 0.001	< 0.05
Glutamine	0.439 ± 0.046	0.514 ± 0.062	< 0.05
Glycine	0.036 ± 0.006	0.045 ± 0.004	< 0.05
N,N-Dimethylglycine	0.003 ± 0.001	0.007 ± 0.002	< 0.001
Serine	0.141 ± 0.023	0.327 ± 0.072	< 0.001
Significantly lower in the LP diet group			
Isoleucine	0.062 ± 0.008	0.047 ± 0.005	< 0.01
Leucine	0.086 ± 0.013	0.062 ± 0.014	< 0.05
Phenylalanine	0.013 ± 0.001	0.010 ± 0.002	< 0.05
Threonine	0.196 ± 0.028	0.138 ± 0.032	< 0.01
Urea	0.667 ± 0.122	0.281 ± 0.063	< 0.001
Valine	0.164 ± 0.023	0.122 ± 0.007	< 0.01
Not significantly			
3-Hydroxybutyrate	0.013 ± 0.005	0.014 ± 0.003	0.80
Acetate	0.092 ± 0.085	0.095 ± 0.063	0.94
Acetoacetate	0.022 ± 0.009	0.022 ± 0.017	0.98
Alanine	0.404 ± 0.099	0.514 ± 0.092	0.07
Choline	0.051 ± 0.024	0.058 ± 0.018	0.59
Creatine	0.024 ± 0.008	0.032 ± 0.012	0.17
Creatine phosphate	0.005 ± 0.002	0.007 ± 0.002	0.15
Formate	1.069 ± 0.026	1.073 ± 0.026	0.80
Glucose	3.880 ± 0.500	4.731 ± 1.378	0.20
Histidine	0.053 ± 0.006	0.059 ± 0.007	0.15
Lactate	2.398 ± 0.815	2.716 ± 0.200	0.39
Methionine	0.014 ± 0.006	0.013 ± 0.007	0.77
Pyruvate	0.042 ± 0.012	0.047 ± 0.009	0.38
Succinate	0.007 ± 0.006	0.023 ± 0.025	0.18
Tyrosine	0.019 ± 0.004	0.023 ± 0.004	0.10

*The relative concentration of each metabolite is its integral value normalized to formic acid, which was added to the NMR buffer as an internal standard. Data are expressed as the mean ± standard deviation (n = 6) and were analyzed by Student's t-tests with Welch's correction.

Supplementary table 3-S6. Metabolite assignments and ^{1}H NMR data in liver sample of rat *

	Liver metabolite	δ^{1}H (ppm)					
1	3-Hydroxybutyrate	1.20(d)	2.31(dd)	2.39(dd)	4.14(m)		
2	Acetate	1.92(s)					
3	Alanine	1.49(d)	3.77(q)				
4	β-alanine	2.56(t)	3.18(t)				
5	Betaine	3.24(s)	3.88(s)				
6	Choline	3.21(s)	3.51(m)	4.07(m)			
7	Creatine	3.03(s)	3.93(s)				
8	Creatine phosphate	3.04(s)	3.94(s)				
9	Cytidine	3.80(dd)	3.92(dd)	4.12(m)	4.20(t)	4.30(dd)	5.92(d)
		6.07(d)	7.85(d)				
10	Dimethylamine	2.72(s)					
11	Formate	8.46(s)					
12	Fumarate	6.52(s)					
13	Glucose	3.26(dd)	3.41(t)	3.42(t)	3.47(m)	3.50(t)	3.54(dd)
		3.72(t)	3.73(dd)	3.78(dd)	3.85(m)	3.86(dd)	3.90(dd)
		4.66(d)	5.24(d)				
14	Glutamate	2.05(m)	2.16(m)	2.33(m)	2.38(m)	3.75(dd)	
15	Glutamine	2.05(m)	2.13(m)	2.43(m)	2.48(m)	3.77(t)	6.92(s)
16	Glycerol	3.56(dd)	3.65(dd)	3.78(m)			
17	Histidine	3.13(dd)	3.22(dd)	3.98(dd)	7.11(s)	7.92(s)	
18	Hypoxanthine	8.20(s)	8.22(s)				
19	Inosine	3.83(dd)	3.90(dd)	4.29(q)	4.45(dd)	4.77(t)	6.11(s)
		8.25(s)	8.36(s)				
20	Inosine monophosphate (IMP)	4.00(m)	4.36(m)	4.50(m)	6.13(d)	8.22(s)	8.59(s)
21	Isoleucine	0.94(t)	1.02(d)	1.25(m)	1.46(m)	1.97(m)	3.66(d)
22	Lactate	1.33(d)	4.12(q)				
23	Leucine	0.96(d)	0.97(d)	1.67(m)	1.70(m)	1.73(m)	3.73(dd)
24	Lysine	1.43(m)	1.50(m)	1.71(m)	1.87(m)	1.91(m)	3.06(t)
25	Methionine	2.11(m)	2.14(s)	2.19(m)	2.65(t)	3.85(dd)	
26	Niacinamide	7.60(m)	8.26(tt)	8.72(dd)	8.95(dd)		
27	Nicotinamide adenine dinucleotide (NAD^+)	4.20(m)	4.23(m)	4.25(m)	4.36(m)	4.37(m)	4.42(m)
		4.48(t)	4.50(m)	4.54(m)	4.76(t)	6.03(d)	6.09(d)
		8.16(s)	8.19(q)	8.42(s)	8.81(m)	9.10(m)	9.35(m)
28	O-Phosphocholine	3.23(s)	3.59(m)	4.17(m)			
29	Phenylalanine	3.14(dd)	3.28(dd)	3.99(dd)	7.33(m)	7.38(m)	7.43(m)
30	Succinate	2.41(s)					
31	Taurine	3.28(t)	3.42(t)				
32	Threonine	1.34(d)	3.57(d)	4.25(m)			
33	Trimethylamine N-oxide	3.26(s)					
34	Tryptophan	3.30(dd)	3.47(dd)	4.05(dd)	7.19(m)	7.28(m)	7.31(s)
		7.54(d)	7.74(d)				
35	Tyrosine	3.04(dd)	3.19(dd)	3.93(dd)	6.91(m)	7.20(m)	
36	Uracil	5.81(d)	7.55(d)				
37	Uridine	3.80(dd)	3.90(dd)	4.12(m)	4.24(t)	4.36(t)	5.91(d)
		5.93(d)	7.89(d)				
38	Valine	1.00(d)	1.05(d)	2.26(m)	3.60(d)		

*Key: s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; dd, doublet of doublet.

Supplementary table 3-S7. Relative concentration of the metabolites in liver samples of rats fed control (CT) or low protein (LP) diets for 4 weeks*

Liver metabolite	CT diet group	LP diet group	*P* value
Significantly higher in the LP diet group			
β-alanine	0.087 ± 0.009	0.110 ± 0.016	< 0.05
Choline	0.066 ± 0.020	0.107 ± 0.025	< 0.05
Dimethylamine	0.002 ± 0.0005	0.005 ± 0.002	< 0.05
Hypoxanthine	0.101 ± 0.015	0.250 ± 0.105	< 0.05
Inosine monophosphate (IMP)	0.017 ± 0.004	0.027 ± 0.006	< 0.05
Trimethylamine N-oxide	0.101 ± 0.054	0.183 ± 0.052	< 0.05
Significantly lower in the LP diet group			
Inosine	0.176 ± 0.037	0.092 ± 0.025	< 0.001
Nicotinamide adenine dinucleotide (NAD^+)	0.012 ± 0.002	0.008 ± 0.002	< 0.01
o-Phosphocholine	0.311 ± 0.062	0.192 ± 0.055	< 0.01
Taurine	1.468 ± 0.278	0.278 ± 0.106	< 0.001
Uridine	0.142 ± 0.031	0.102 ± 0.022	< 0.05
Not significantly			
3-Hydroxybutyrate	0.059 ± 0.018	0.054 ± 0.020	0.67
Acetate	0.229 ± 0.089	0.176 ± 0.078	0.30
Alanine	0.983 ± 0.209	1.180 ± 0.197	0.12
Betaine	0.084 ± 0.065	0.104 ± 0.027	0.50
Creatine	0.030 ± 0.005	0.032 ± 0.010	0.72
Creatine phosphate	0.026 ± 0.005	0.028 ± 0.007	0.56
Cytidine	0.023 ± 0.003	0.029 ± 0.009	0.13
Formate	0.020 ± 0.019	0.012 ± 0.009	0.35
Fumarate	0.025 ± 0.006	0.019 ± 0.007	0.13
Glucose	9.115 ± 3.622	12.564 ± 3.498	0.12
Glutamate	0.417 ± 0.097	0.542 ± 0.209	0.21
Glutamine	0.823 ± 0.182	0.807 ± 0.093	0.85
Glycerol	0.289 ± 0.037	0.317 ± 0.070	0.40
Histidine	0.104 ± 0.015	0.089 ± 0.016	0.14
Isoleucine	0.062 ± 0.007	0.060 ± 0.022	0.82
Lactate	1.798 ± 0.310	2.316 ± 0.539	0.07
Leucine	0.147 ± 0.024	0.130 ± 0.035	0.37
Lysine	0.073 ± 0.006	0.076 ± 0.027	0.80
Methionine	0.049 ± 0.009	0.043 ± 0.012	0.38
Niacinamide	0.125 ± 0.032	0.145 ± 0.012	0.21
Phenylalanine	0.071 ± 0.009	0.071 ± 0.020	0.99
Succinate	0.107 ± 0.034	0.120 ± 0.063	0.67
Threonine	0.176 ± 0.040	0.153 ± 0.033	0.30
Tryptophan	0.014 ± 0.002	0.014 ± 0.004	0.71
Tyrosine	0.078 ± 0.015	0.078 ± 0.020	1.00
Uracil	0.010 ± 0.004	0.032 ± 0.022	0.06
Valine	0.126 ± 0.017	0.114 ± 0.035	0.45

*The relative concentration of each metabolite is its integral value normalized to TSP, which was added to the NMR buffer as an internal standard. Data are expressed as the mean ± standard deviation (n = 6) and were analyzed by Student's *t*-tests with Welch's correction.

Supplementary table 3-S8. Hepatic genes showing statistically significant expression changes and up- or down-regulation by more than 1.5-fold by mRNA microarray in rats fed a low-protein (LP) diet compared to a control (CT) diet for 4 weeks

NCBI Entrez Gene ID [a]	Gene Symbol	Gene Name	Fold Change	CT diet group (raw intensity) [b]	LP diet group (raw intensity)	*P* value
58835	Phgdh	phosphoglycerate dehydrogenase	123.5	166.2	21998.5	0.0000
399475	Adm2	adrenomedullin 2	42.9	17.9	824.4	0.0000
25612	Asns	asparagine synthetase (glutamine-hydrolyzing)	27.4	193.0	5676.8	0.0000
293820	Psat1	phosphoserine aminotransferase 1	20.1	283.2	6088.2	0.0000
170580	Fgf21	fibroblast growth factor 21	17.5	277.5	5196.9	0.0000
680308	Mthfd2	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2	14.3	115.6	1773.9	0.0000
25334	Bhlha15	basic helix-loop-helix family, member a15	11.5	105.6	1297.2	0.0000
361637	Acsm5	acyl-CoA synthetase medium-chain family member 5	11.3	194.3	2343.8	0.0020
81919	Fut1	fucosyltransferase 1	9.1	43.8	427.5	0.0000
361755	Aldh18a1	aldehyde dehydrogenase 18 family, member A1	8.1	15.2	131.9	0.0000
287877	Pycr1	pyrroline-5-carboxylate reductase 1	7.3	105.3	820.5	0.0000
494500	Gsta3	glutathione S-transferase alpha 3	7.1	23.6	180.0	0.0012
300719	Cib2	calcium and integrin binding family member 2	6.0	735.3	4728.2	0.0004
140942	Ddit4	DNA-damage-inducible transcript 4	5.9	255.9	1604.9	0.0000
361424	RGD1309651	similar to 1190005I06Rik protein	5.4	35.1	205.0	0.0016
685612	Col26a1	collagen type XXVI alpha 1 chain	5.0	12.9	69.4	0.0000
29142	Vnn1	vanin 1	5.0	308.2	1654.5	0.0000
100910829	LOC100910829	probable N-acetyltransferase CML2-like	4.8	458.3	2334.6	0.0000
362835	Reep6	receptor accessory protein 6	4.3	151.8	703.6	0.0001
26760	Akr7a3	aldo-keto reductase family 7 member A3	4.3	2360.6	10788.7	0.0001
302507	Efhc2	EF-hand domain containing 2	4.1	15.6	68.6	0.0000
100363433	Srrm5	serine/arginine repetitive matrix 5	3.9	11.7	49.1	0.0020
113992	Ugt1a6	UDP glucuronosyltransferase family 1 member A6	3.8	63.6	261.4	0.0000
24188	Aldh1a1	aldehyde dehydrogenase 1 family, member A1	3.8	4897.0	19893.7	0.0001
50567	Slc3a2	solute carrier family 3 member 2	3.8	1037.0	4192.6	0.0000
498416	Grb10	growth factor receptor bound protein 10	3.7	13.2	53.0	0.0000
305166	Gpat3	glycerol-3-phosphate acyltransferase 3	3.5	53.9	202.0	0.0003
290277	Cryl1	crystallin, lambda 1	3.3	943.9	3385.8	0.0002
246248	Fmo5	flavin containing monooxygenase 5	3.3	182.1	645.6	0.0001
500237	Cml2	camello-like 2	3.2	466.7	1620.6	0.0011
54193	Pbsn	probasin	3.2	33.0	113.6	0.0059
361659	Plekha1	pleckstrin homology domain containing A1	3.1	1667.9	5620.4	0.0000
246273	Trib3	tribbles pseudokinase 3	3.1	212.8	707.1	0.0000
361042	Pck2	phosphoenolpyruvate carboxykinase 2 (mitochondrial)	3.0	622.0	1986.2	0.0000
81806	Serpina7	serpin family A member 7	3.0	215.6	682.6	0.0042
499010	Syne1	spectrin repeat containing nuclear envelope protein 1	2.9	10.3	32.3	0.0002
64679	Tgm4	transglutaminase 4	2.8	27.4	83.3	0.0000
171026	Akap5	A-kinase anchoring protein 5	2.8	34.9	105.8	0.0018
100360982	Relb	RELB proto-oncogene, NF-kB subunit	2.8	20.1	60.2	0.0001
360857	Rgs16	regulator of G-protein signaling 16	2.8	4531.3	13540.0	0.0193
64570	Nat8	N-acetyltransferase 8	2.8	912.7	2700.9	0.0060
102556148	LOC102556148	probable N-acetyltransferase CML2-like	2.7	437.1	1265.2	0.0014
24443	Hdc	histidine decarboxylase	2.7	320.2	926.0	0.0046
297096	Snx10	sorting nexin 10	2.7	192.8	557.4	0.0002
117271	Hrk	harakiri, BCL2 interacting protein	2.6	27.7	77.2	0.0199
690130	Rhof	ras homolog family member F, filopodia associated	2.6	62.7	174.2	0.0006
294708	Sgtb	small glutamine rich tetratricopeptide repeat containing beta	2.5	31.9	86.4	0.0000

[a] Gene IDs reference The National Center for Biotechnology Information database.

[b] Data are expressed as the mean (n = 6), analyzed by Moderated t-tests using GeneSpring GX ver. 14.9.

Supplementary table 3-S8, continued

NCBI Entrez Gene ID [a]	Gene Symbol	Gene Name	Fold Change	CT diet group (raw intensity) [b]	LP diet group (raw intensity)	*P* value
246263	Acsm2	acyl-CoA synthetase medium-chain family member 2	2.5	859.4	2314.3	0.0006
297113	Gars	glycyl-tRNA synthetase	2.5	7947.2	21151.8	0.0000
360642	Kansl1	KAT8 regulatory NSL complex subunit 1	2.5	22.5	59.4	0.0123
690315	Derl3	derlin 3	2.5	193.4	508.1	0.0102
85332	Cavin3	caveolae associated protein 3	2.4	6734.9	17630.1	0.0000
116636	Eif4ebp1	eukaryotic translation initiation factor 4E binding protein 1	2.4	6325.6	16485.5	0.0000
362899	Nipal2	NIPA-like domain containing 2	2.4	18.9	48.6	0.0001
619561	Acsf2	acyl-CoA synthetase family member 2	2.4	16.3	42.0	0.0006
679091	Mettl1	methyltransferase like 1	2.4	49.4	124.6	0.0001
313560	Ctps1	CTP synthase 1	2.3	623.4	1557.8	0.0000
307395	Ablim3	actin binding LIM protein family, member 3	2.3	238.3	591.9	0.0359
685448	Pcp4l1	Purkinje cell protein 4-like 1	2.3	181.7	450.1	0.0004
287450	Slc16a11	solute carrier family 16, member 11	2.3	949.3	2348.0	0.0004
84422	Gfra3	GDNF family receptor alpha 3	2.3	33.2	81.5	0.0004
83509	Slc7a7	solute carrier family 7 member 7	2.3	10.9	26.6	0.0000
314462	Trmt61a	tRNA methyltransferase 61A	2.3	533.8	1297.9	0.0003
59320	Adcy10	adenylate cyclase 10 (soluble)	2.3	40.3	98.0	0.0086
57033	Adam22	ADAM metallopeptidase domain 22	2.3	24.7	59.6	0.0026
311428	RGD1311739	similar to RIKEN cDNA 1700037H04	2.3	688.9	1661.7	0.0000
291541	Cidea	cell death-inducing DFFA-like effector a	2.2	136.9	329.3	0.0409
293638	Cars	cysteinyl-tRNA synthetase	2.2	463.5	1109.8	0.0000
685679	Nme4	NME/NM23 nucleoside diphosphate kinase 4	2.2	25.6	61.2	0.0001
288499	Slc29a4	solute carrier family 29 member 4	2.2	171.7	411.0	0.0004
313047	Yars	tyrosyl-tRNA synthetase	2.2	311.4	740.1	0.0000
287818	Cd300lf	Cd300 molecule-like family member F	2.2	72.0	170.0	0.0000
306804	Iars	isoleucyl-tRNA synthetase	2.2	1620.4	3814.8	0.0000
100912108	Nupr1	nuclear protein 1, transcriptional regulator	2.2	4555.4	10689.6	0.0000
25586	Alpl	alkaline phosphatase, liver/bone/kidney	2.2	207.0	485.8	0.0000
363465	Pir	pirin	2.2	480.0	1117.1	0.0001
294072	Slc35g1	solute carrier family 35, member G1	2.2	56.3	130.4	0.0001
681337	Acot4	acyl-CoA thioesterase 4	2.2	2724.6	6311.8	0.0000
652928	Rbm12	RNA binding motif protein 12	2.2	26.6	61.5	0.0125
365901	Gnat2	G protein subunit alpha transducin 2	2.2	41.6	96.2	0.0055
304429	Psph	phosphoserine phosphatase	2.1	554.0	1258.7	0.0000
29455	Gdf15	growth differentiation factor 15	2.1	44.8	101.5	0.0228
310376	Nim1k	NIM1 serine/threonine protein kinase	2.1	587.3	1317.1	0.0095
291317	Rpp38	ribonuclease P/MRP subunit p38	2.1	103.2	229.6	0.0000
116509	Slc6a9	solute carrier family 6 member 9	2.1	489.8	1086.4	0.0000
65035	Nr1i3	nuclear receptor subfamily 1, group I, member 3	2.1	994.3	2198.9	0.0002
308584	Plekha4	pleckstrin homology domain containing A4	2.1	167.8	369.4	0.0001
366593	Dus4l	dihydrouridine synthase 4-like	2.0	30.0	65.8	0.0000
29739	Gclm	glutamate cysteine ligase, modifier subunit	2.0	92.9	203.6	0.0000
305826	Samd4a	sterile alpha motif domain containing 4A	2.0	119.4	261.3	0.0011
25146	Cyp17a1	cytochrome P450, family 17, subfamily a, polypeptide 1	2.0	566.3	1236.7	0.0072
364838	Reep5	receptor accessory protein 5	2.0	21.1	46.1	0.0000
499129	Kctd15	potassium channel tetramerization domain containing 15	2.0	17.2	37.6	0.0000
64526	Ech1	enoyl-CoA hydratase 1	2.0	3614.4	7880.1	0.0001
309122	H19	H19, imprinted maternally expressed transcript (non-protein coding)	2.0	26.8	58.4	0.0014
60630	Unc5b	unc-5 netrin receptor B	2.0	276.9	599.6	0.0080
308602	Tmem86a	transmembrane protein 86A	2.0	1643.0	3552.8	0.0145

[a] Gene IDs reference The National Center for Biotechnology Information database.

[b] Data are expressed as the mean (n = 6), analyzed by Moderated t-tests using GeneSpring GX ver. 14.9.

Supplementary table 3-S8, continued

NCBI Entrez Gene ID [a]	Gene Symbol	Gene Name	Fold Change	CT diet group (raw intensity) [b]	LP diet group (raw intensity)	*P* value
314304	Acot3	acyl-CoA thioesterase 3	2.0	95.6	206.6	0.0001
299851	Mars	methionyl-tRNA synthetase	2.0	2599.5	5617.0	0.0000
24536	Lepr	leptin receptor	2.0	202.2	432.8	0.0001
26989	Cadps	calcium dependent secretion activator	2.0	40.1	85.9	0.0000
64316	Srpx	sushi-repeat-containing protein, X-linked	2.0	161.2	343.5	0.0049
64157	Ddah1	dimethylarginine dimethylaminohydrolase 1	2.0	5189.5	10943.4	0.0037
361441	Fam89a	family with sequence similarity 89, member A	2.0	2886.9	6063.2	0.0006
315759	Igdcc3	immunoglobulin superfamily, DCC subclass, member 3	2.0	102.0	213.7	0.0362
192267	Obp1f	odorant binding protein I f	2.0	84.5	177.0	0.0041
291624	Lars	leucyl-tRNA synthetase	1.9	3421.6	7112.7	0.0000
291773	Taf4b	TATA-box binding protein associated factor 4b	1.9	39.8	82.5	0.0000
680493	Cd3eap	CD3e molecule associated protein	1.9	268.7	557.4	0.0000
24552	Me1	malic enzyme 1	1.9	7733.7	16005.7	0.0031
100911757	Tshz2	teashirt zinc finger homeobox 2	1.9	117.7	242.3	0.0006
298699	Ppan	peter pan homolog (Drosophila)	1.9	1636.8	3358.7	0.0001
192272	Acot2	acyl-CoA thioesterase 2	1.9	35.6	72.4	0.0007
60356	Csad	cysteine sulfinic acid decarboxylase	1.9	1396.0	2833.6	0.0001
24627	Pde4d	phosphodiesterase 4D	1.9	9.6	19.5	0.0036
309145	RGD1311946	similar to RIKEN cDNA 1810055G02	1.9	87.9	177.8	0.0007
102553031	LOC102553031	zinc finger and SCAN domain-containing protein 20-like	1.9	102.8	207.3	0.0055
499709	Gar1	GAR1 ribonucleoprotein	1.9	116.2	233.9	0.0001
89820	Prmt3	protein arginine methyltransferase 3	1.9	24.0	48.2	0.0002
294007	Pprc1	peroxisome proliferator-activated receptor gamma, coactivator-related 1	1.9	209.8	419.9	0.0000
289533	Ugt2a3	UDP glucuronosyltransferase family 2 member A3	1.9	185.0	369.8	0.0005
29740	Eci1	enoyl-CoA delta isomerase 1	1.8	5682.5	11214.2	0.0011
297436	Chchd6	coiled-coil-helix-coiled-coil-helix domain containing 6	1.8	60.0	118.3	0.0001
298031	Susd1	sushi domain containing 1	1.8	34.0	66.5	0.0000
171093	Shank2	SH3 and multiple ankyrin repeat domains 2	1.8	20.8	40.6	0.0014
365192	Zfp324	zinc finger protein 324	1.8	20.2	39.3	0.0111
298541	Dhdds	dehydrodolichyl diphosphate synthase subunit	1.8	594.1	1154.2	0.0043
102552326	LOC102552326	late cornified envelope protein 5A-like	1.8	24.0	46.6	0.0001
286989	Ugt2b7	UDP glucuronosyltransferase family 2 member B7	1.8	574.1	1113.9	0.0013
81008	Itga7	integrin subunit alpha 7	1.8	35.7	69.2	0.0004
295228	Gpatch4	G patch domain containing 4	1.8	160.8	311.7	0.0081
298594	Rcc2	regulator of chromosome condensation 2	1.8	152.2	294.9	0.0000
79255	Atf4	activating transcription factor 4	1.8	9581.1	18524.8	0.0000
300862	Irak1bp1	interleukin-1 receptor-associated kinase 1 binding protein 1	1.8	31.4	60.6	0.0392
500707	Tc2n	tandem C2 domains, nuclear	1.8	142.4	274.3	0.0000
117024	Crym	crystallin, mu	1.8	1624.3	3114.5	0.0010
364827	Zscan30	zinc finger and SCAN domain containing 30	1.8	11.8	22.7	0.0105
54261	Kcnn1	potassium calcium-activated channel subfamily N member 1	1.8	14.3	27.3	0.0172
85255	Hacl1	2-hydroxyacyl-CoA lyase 1	1.8	8519.2	16277.1	0.0072
288053	Ropn1	rhophilin associated tail protein 1	1.8	47.9	91.1	0.0287
314405	Otub2	OTU deubiquitinase, ubiquitin aldehyde binding 2	1.8	2675.5	5065.4	0.0079
362287	Slc17a9	solute carrier family 17 member 9	1.8	216.4	409.5	0.0031
171347	Mat2a	methionine adenosyltransferase 2A	1.8	2123.6	4005.8	0.0000
266975	Sars	seryl-tRNA synthetase	1.8	665.4	1255.0	0.0000
85262	Slc25a27	solute carrier family 25, member 27	1.8	15.1	28.3	0.0049
308911	Rrp8	ribosomal RNA processing 8	1.7	73.6	136.6	0.0000
288622	Zbed5	zinc finger, BED-type containing 5	1.7	30.1	55.7	0.0079

[a] Gene IDs reference The National Center for Biotechnology Information database.

[b] Data are expressed as the mean (n = 6), analyzed by Moderated t-tests using GeneSpring GX ver. 14.9.

Supplementary table 3-S8, continued

NCBI Entrez Gene ID [a]	Gene Symbol	Gene Name	Fold Change	CT diet group (raw intensity) [b]	LP diet group (raw intensity)	*P* value
114483	Cdk6	cyclin-dependent kinase 6	1.7	16.7	30.8	0.0051
84596	Srm	spermidine synthase	1.7	1421.4	2626.0	0.0000
292757	Fbxo17	F-box protein 17	1.7	49.7	91.8	0.0349
288741	Ficd	FIC domain containing	1.7	322.6	594.3	0.0063
313610	Extl1	exostosin-like glycosyltransferase 1	1.7	2199.9	4052.3	0.0010
59115	Ninj2	ninjurin 2	1.7	118.0	217.2	0.0008
679127	Rrp12	ribosomal RNA processing 12 homolog	1.7	117.0	215.4	0.0001
362703	Wdr43	WD repeat domain 43	1.7	730.1	1342.8	0.0001
60335	Tgm1	transglutaminase 1	1.7	11.8	21.7	0.0166
299316	Slc25a47	solute carrier family 25, member 47	1.7	13494.0	24704.3	0.0076
363328	Ticam1	toll-like receptor adaptor molecule 1	1.7	25.1	45.9	0.0093
498196	Cyp3a85-ps	cytochrome P450, family 3, subfamily a, polypeptide 85, pseudogene	1.7	2821.9	5158.2	0.0005
25283	Gclc	glutamate-cysteine ligase, catalytic subunit	1.7	5831.2	10647.5	0.0004
500292	Cidec	cell death-inducing DFFA-like effector c	1.7	1103.4	2014.7	0.0001
300211	Fkbp11	FK506 binding protein 11	1.7	978.7	1786.5	0.0001
498206	Tnfrsf11a	TNF receptor superfamily member 11A	1.7	15.1	27.6	0.0165
300888	Tmed3	transmembrane p24 trafficking protein 3	1.7	458.9	835.0	0.0035
362976	Trmu	tRNA 5-methylaminomethyl-2-thiouridylate methyltransferase	1.7	91.9	166.9	0.0029
690297	Pwp2	PWP2, small subunit processome component	1.7	432.2	783.3	0.0000
171145	Eif2b3	eukaryotic translation initiation factor 2B subunit gamma	1.7	229.4	415.5	0.0001
316421	Aox2	aldehyde oxidase 2	1.7	11.3	20.4	0.0021
362979	Panx2	pannexin 2	1.7	25.1	45.4	0.0011
295384	Prmt6	protein arginine methyltransferase 6	1.7	16.1	28.9	0.0001
311849	Crat	carnitine O-acetyltransferase	1.7	3571.8	6408.8	0.0008
25458	Gss	glutathione synthetase	1.7	620.1	1112.7	0.0020
289338	Disp1	dispatched RND transporter family member 1	1.7	146.0	261.9	0.0038
498678	Ccdc71	coiled-coil domain containing 71	1.7	17.6	31.5	0.0057
314128	RGD1304624	similar to RIKEN cDNA 2700097O09	1.7	38.5	69.0	0.0036
291553	Nedd4l	neural precursor cell expressed, developmentally down-regulated 4-like	1.7	38.3	68.4	0.0383
29508	Nfya	nuclear transcription factor Y subunit alpha	1.7	14.6	26.1	0.0036
290686	Tma16	translation machinery associated 16 homolog	1.7	159.9	284.8	0.0000
313245	Polr1e	RNA polymerase I subunit E	1.7	97.3	173.3	0.0001
689232	Nxnl2	nucleoredoxin-like 2	1.7	23.3	41.5	0.0159
360482	Tfap4	transcription factor AP-4	1.7	93.0	165.2	0.0000
60373	Nop58	NOP58 ribonucleoprotein	1.7	923.9	1637.4	0.0010
685474	Zfp385a	zinc finger protein 385A	1.7	1169.8	2069.9	0.0054
25062	Gpd2	glycerol-3-phosphate dehydrogenase 2	1.6	393.9	696.6	0.0072
309673	Rrp1b	ribosomal RNA processing 1B	1.6	13.5	23.9	0.0105
171179	Glyatl2	glycine-N-acyltransferase-like 2	1.6	5039.5	8904.1	0.0006
25413	Cpt2	carnitine palmitoyltransferase 2	1.6	3975.0	7008.0	0.0002
303569	Rundc3a	RUN domain containing 3A	1.6	19.4	34.3	0.0192
300994	Ifrd2	interferon-related developmental regulator 2	1.6	276.2	486.5	0.0036
300120	Fam118a	family with sequence similarity 118, member A	1.6	38.8	68.4	0.0009
140665	Rab3d	RAB3D, member RAS oncogene family	1.6	40.2	70.6	0.0044
361221	Nol8	nucleolar protein 8	1.6	249.2	437.5	0.0001
362495	Ndufaf4	NADH:ubiquinone oxidoreductase complex assembly factor 4	1.6	41.0	71.8	0.0046
100360406	Tsr1	TSR1, ribosome maturation factor	1.6	111.1	194.8	0.0044
291057	Eef1e1	eukaryotic translation elongation factor 1 epsilon 1	1.6	15.8	27.6	0.0075
294091	Ifit2	interferon-induced protein with tetratricopeptide repeats 2	1.6	335.5	587.2	0.0155
114555	Casp4	caspase 4	1.6	777.0	1357.8	0.0000

[a] Gene IDs reference The National Center for Biotechnology Information database.

[b] Data are expressed as the mean (n = 6), analyzed by Moderated t-tests using GeneSpring GX ver. 14.9.

Supplementary table 3-S8, continued

NCBI Entrez Gene ID [a)]	Gene Symbol	Gene Name	Fold Change	CT diet group (raw intensity) [b)]	LP diet group (raw intensity)	*P* value
317395	Spin2a	spindlin family, member 2A	1.6	55.7	97.2	0.0001
296302	Eif2s2	eukaryotic translation initiation factor 2 subunit beta	1.6	818.6	1426.6	0.0001
314879	Xpot	exportin for tRNA	1.6	1156.4	2010.5	0.0000
117020	Mcf2l	MCF.2 cell line derived transforming sequence-like	1.6	38.4	66.7	0.0002
291556	Nars	asparaginyl-tRNA synthetase	1.6	372.2	644.6	0.0000
297508	Nampt	nicotinamide phosphoribosyltransferase (NAMPT)	1.6	21.8	37.7	0.0414
501582	Gspt2	G1 to S phase transition 2	1.6	26.8	46.3	0.0018
292804	Gpi	glucose-6-phosphate isomerase	1.6	3634.9	6271.4	0.0011
362214	Nop56	NOP56 ribonucleoprotein	1.6	872.0	1504.1	0.0001
500595	Smim1	small integral membrane protein 1	1.6	1564.2	2694.2	0.0126
290558	Nt5dc2	5'-nucleotidase domain containing 2	1.6	169.8	292.3	0.0001
501283	Plin5	perilipin 5	1.6	16.0	27.6	0.0053
311441	Trmt6	tRNA methyltransferase 6	1.6	157.5	270.9	0.0000
58822	Csnk1e	casein kinase 1, epsilon	1.6	52.7	90.6	0.0010
26203	ND6	NADH dehydrogenase subunit 6	1.6	807.5	1385.7	0.0211
303372	Nle1	notchless homolog 1	1.6	181.4	311.0	0.0001
287451	Slc16a13	solute carrier family 16, member 13	1.6	371.6	636.9	0.0001
307253	Katnal2	katanin catalytic subunit A1 like 2	1.6	40.6	69.5	0.0004
316533	Chpf	chondroitin polymerizing factor	1.6	371.3	635.7	0.0048
83582	Polr1b	RNA polymerase I subunit B	1.6	98.6	168.6	0.0003
289827	Ugp2	UDP-glucose pyrophosphorylase 2	1.6	14721.6	25168.6	0.0024
689933	Syce3	synaptonemal complex central element protein 3	1.6	62.4	106.6	0.0067
81522	Mtr	5-methyltetrahydrofolate-homocysteine methyltransferase	1.6	44.7	76.3	0.0006
266764	Tbkbp1	TBK1 binding protein 1	1.6	620.6	1057.8	0.0008
266770	Soat2	sterol O-acyltransferase 2	1.6	289.0	491.7	0.0001
117242	Tenm2	teneurin transmembrane protein 2	1.6	57.6	97.9	0.0009
682457	Txlna	taxilin alpha	1.6	22.8	38.7	0.0025
499155	Fancf	Fanconi anemia, complementation group F	1.6	17.9	30.4	0.0001
170509	Cyp3a62	cytochrome P450, family 3, subfamily a, polypeptide 62	1.6	434.7	736.3	0.0206
298096	Wdr31	WD repeat domain 31	1.6	69.3	117.2	0.0003
290228	Ipo4	importin 4	1.6	1972.8	3334.9	0.0000
25009	Vars	valyl-tRNA synthetase	1.6	10.3	17.4	0.0010
114021	Ebna1bp2	EBNA1 binding protein 2	1.6	2795.8	4723.8	0.0000
361205	Lect2	leukocyte cell-derived chemotaxin 2	1.6	4932.7	8330.2	0.0001
292792	Tmem147	transmembrane protein 147	1.6	1600.1	2701.6	0.0000
313200	Hsdl2	hydroxysteroid dehydrogenase like 2	1.6	1456.2	2456.6	0.0002
317382	Foxp3	forkhead box P3	1.6	13.0	21.9	0.0046
295394	Dph5	diphthamide biosynthesis 5	1.6	141.4	238.0	0.0002
305540	Slc1a4	solute carrier family 1 member 4	1.6	14.7	24.7	0.0067
362850	Angptl4	angiopoietin-like 4	1.6	18435.2	31008.6	0.0070
361935	Spata5	spermatogenesis associated 5	1.6	33.3	56.0	0.0187
303211	Mmgt2	membrane magnesium transporter 2	1.6	66.1	111.1	0.0050
685560	Mogat3	monoacylglycerol O-acyltransferase 3	1.6	30.2	50.7	0.0153
289352	Eprs	glutamyl-prolyl-tRNA synthetase	1.6	5622.7	9450.4	0.0000
501396	LOC501396	hypothetical protein LOC501396	1.6	59.8	100.4	0.0005
314969	Nop2	NOP2 nucleolar protein	1.6	1177.2	1975.4	0.0001
315212	Alg12	ALG12, alpha-1,6-mannosyltransferase	1.6	64.4	108.0	0.0013
79116	Apex1	apurinic/apyrimidinic endodeoxyribonuclease 1	1.6	2077.3	3483.9	0.0002
171134	Lgals2	galectin 2	1.6	89.3	149.7	0.0148
60571	Mybbp1a	MYB binding protein 1a	1.6	253.9	424.7	0.0000

[a)] Gene IDs reference The National Center for Biotechnology Information database.

[b)] Data are expressed as the mean (n = 6), analyzed by Moderated t-tests using GeneSpring GX ver. 14.9.

Supplementary table 3-S8, continued

NCBI Entrez Gene ID [a]	Gene Symbol	Gene Name	Fold Change	CT diet group (raw intensity) [b]	LP diet group (raw intensity)	*P* value
309646	Slc26a8	solute carrier family 26 member 8	1.6	32.8	54.9	0.0217
171139	Timm9	translocase of inner mitochondrial membrane 9	1.6	1091.6	1824.3	0.0004
117544	Ppat	phosphoribosyl pyrophosphate amidotransferase	1.6	276.0	461.2	0.0003
363017	Taf1d	TATA-box binding protein associated factor, RNA polymerase I subunit D	1.6	96.5	161.0	0.0187
314243	Wdr89	WD repeat domain 89	1.6	103.0	171.6	0.0000
685697	Supt3h	SPT3 homolog, SAGA and STAGA complex component	1.6	71.9	119.5	0.0351
298384	RGD1559786	similar to RIKEN cDNA 0610037L13	1.6	24.7	41.2	0.0011
300652	Sorl1	sortilin related receptor 1	1.6	16.2	27.0	0.0447
102553962	LOC102553962	KRAB domain-containing protein ZNF747-like	1.5	507.7	842.9	0.0016
361783	Zfp57	zinc finger protein 57	1.5	39.5	65.5	0.0004
117041	Nln	neurolysin	1.5	807.0	1336.9	0.0002
84484	Slc4a4	solute carrier family 4 member 4	1.5	804.4	1331.2	0.0076
688966	Borcs8	BLOC-1 related complex subunit 8	1.5	141.6	234.3	0.0013
362224	Kat14	lysine acetyltransferase 14	1.5	27.1	44.8	0.0435
25623	Gys2	glycogen synthase 2	1.5	13653.3	22551.8	0.0040
315994	Mapkapk3	mitogen-activated protein kinase-activated protein kinase 3	1.5	474.9	783.6	0.0204
25256	Fmo1	flavin containing monooxygenase 1	1.5	3997.0	6590.7	0.0033
171047	Dnph1	2'-deoxynucleoside 5'-phosphate N-hydrolase 1	1.5	1947.5	3210.5	0.0000
300114	Rtl6	retrotransposon Gag like 6	1.5	281.6	464.2	0.0004
301544	Farsb	phenylalanyl-tRNA synthetase, beta subunit	1.5	173.7	285.1	0.0011
308652	Smox	spermine oxidase	1.5	24.0	39.4	0.0222
291469	Srfbp1	serum response factor binding protein 1	1.5	135.0	221.6	0.0001
367384	B3gntl1	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase-like 1	1.5	174.7	285.7	0.0004
100911224	LOC100911224	zinc finger protein interacting with ribonucleoprotein K-like	1.5	46.8	76.5	0.0005
81683	Mif	macrophage migration inhibitory factor (glycosylation-inhibiting factor)	1.5	10071.7	16460.6	0.0078
297784	Rrs1	ribosome biogenesis regulator homolog	1.5	157.3	256.7	0.0253
293738	Ccdc86	coiled-coil domain containing 86	1.5	311.0	507.1	0.0021
245981	Rpl15	ribosomal protein L15	1.5	165.6	270.0	0.0001
308568	Fv1	Friend virus susceptibility 1	1.5	13.6	22.1	0.0152
291540	Cep76	centrosomal protein 76	1.5	224.3	365.1	0.0000
29681	C1qbp	complement C1q binding protein	1.5	13227.7	21531.0	0.0000
24214	Atp1b2	ATPase Na+/K+ transporting subunit beta 2	1.5	51.0	83.1	0.0040
361596	Idh2	isocitrate dehydrogenase (NADP(+)) 2, mitochondrial	1.5	5218.3	8478.2	0.0013
503049	Acot5	acyl-CoA thioesterase 5	1.5	2522.3	4094.8	0.0107
499756	Lcn10	lipocalin 10	1.5	10.8	17.6	0.0240
300839	Lysmd2	LysM domain containing 2	1.5	498.8	808.4	0.0040
360569	Mrm3	mitochondrial rRNA methyltransferase 3	1.5	104.1	168.6	0.0013
363134	Rrp9	ribosomal RNA processing 9, U3 small nucleolar RNA binding protein	1.5	169.3	273.7	0.0006
362058	Miga1	mitoguardin 1	1.5	55.9	90.3	0.0012
310625	Rhbg	Rh family B glycoprotein	1.5	31.8	51.4	0.0100
500981	Zfp202	zinc finger protein 202	1.5	85.3	137.8	0.0001
303343	Abhd15	abhydrolase domain containing 15	1.5	15.4	24.8	0.0177
29477	Mapt	microtubule-associated protein tau	1.5	85.4	137.6	0.0030
288176	Cmss1	cms1 ribosomal small subunit homolog (yeast)	1.5	73.2	117.9	0.0001
363237	Wdr12	WD repeat domain 12	1.5	1478.2	2380.9	0.0014
287273	Nhp2	NHP2 ribonucleoprotein	1.5	1103.1	1776.5	0.0000
362983	Dennd6b	DENN domain containing 6B	1.5	18.6	29.9	0.0031
363029	Swsap1	SWIM-type zinc finger 7 associated protein 1	1.5	90.3	145.3	0.0003
681037	Prelid2	PRELI domain containing 2	1.5	1133.5	1822.8	0.0003
681023	Qsox2	quiescin sulfhydryl oxidase 2	1.5	90.1	144.9	0.0000

[a] Gene IDs reference The National Center for Biotechnology Information database.

[b] Data are expressed as the mean (n = 6), analyzed by Moderated t-tests using GeneSpring GX ver. 14.9.

Supplementary table 3-S8, continued

NCBI Entrez Gene ID [a]	Gene Symbol	Gene Name	Fold Change	CT diet group (raw intensity) [b]	LP diet group (raw intensity)	*P* value
102551435	LOC102551435	endothelin-converting enzyme 2-like	1.5	215.3	345.6	0.0001
64846	Slc13a3	solute carrier family 13 member 3	1.5	2678.5	4297.0	0.0024
307989	Ablim1	actin-binding LIM protein 1	1.5	778.0	1247.7	0.0002
282840	Atf5	activating transcription factor 5	1.5	9153.1	14678.5	0.0004
29224	Cbr1	carbonyl reductase 1	1.5	573.5	919.5	0.0000
25315	Ephx1	epoxide hydrolase 1	1.5	1705.2	2732.0	0.0100
360463	Mpv17	MpV17 mitochondrial inner membrane protein	1.5	3929.7	6295.5	0.0000
311391	Cep152	centrosomal protein 152	1.5	88.9	142.5	0.0331
497978	Dgke	diacylglycerol kinase epsilon	1.5	30.1	48.2	0.0076
266682	Cyp3a2	cytochrome P450, family 3, subfamily a, polypeptide 2	1.5	50933.2	81387.4	0.0007
294230	Mrps18b	mitochondrial ribosomal protein S18B	1.5	371.0	592.5	0.0001
685545	Mcrip2	MAPK regulated co-repressor interacting protein 2	1.5	2788.3	4451.6	0.0072
364755	Ero1b	endoplasmic reticulum oxidoreductase 1 beta	1.5	68.4	109.1	0.0004
303333	Slc46a1	solute carrier family 46 member 1	1.5	426.0	680.0	0.0215
306768	Nop16	NOP16 nucleolar protein	1.5	2990.0	4767.5	0.0000
24908	Dnajb9	DnaJ heat shock protein family (Hsp40) member B9	1.5	991.0	1575.9	0.0239
114300	Gtpbp4	GTP binding protein 4	1.5	358.0	569.2	0.0045
84491	Qsox1	quiescin sulfhydryl oxidase 1	1.5	1269.8	2018.2	0.0000
140946	Paics	phosphoribosylaminoimidazole carboxylase	1.5	4619.0	7331.1	0.0000
362883	Slc35e3	solute carrier family 35, member E3	1.5	23.2	36.7	0.0029
300554	Olr1229	olfactory receptor 1229	1.5	25.4	40.2	0.0338
360811	Naa25	N(alpha)-acetyltransferase 25, NatB auxiliary subunit	1.5	30.4	48.1	0.0016
103693432	Ptges3l	prostaglandin E synthase 3 like	1.5	20.4	32.2	0.0012
29660	Pnck	pregnancy up-regulated nonubiquitous CaM kinase	1.5	35.1	55.5	0.0040
311279	Arl14ep	ADP-ribosylation factor like GTPase 14 effector protein	1.5	704.1	1112.8	0.0013
315554	Pus3	pseudouridylate synthase 3	1.5	243.1	384.2	0.0002
290783	Tusc3	tumor suppressor candidate 3	1.5	169.4	267.6	0.0039
171439	Bzw2	basic leucine zipper and W2 domains 2	1.5	766.5	1209.4	0.0002
24763	Acsm3	acyl-CoA synthetase medium-chain family member 3	1.5	199.4	314.4	0.0434
304054	Prdm15	PR/SET domain 15	1.5	13.5	21.2	0.0038
311257	Nat10	N-acetyltransferase 10	1.5	55.7	87.7	0.0013
500909	Ccdc134	coiled-coil domain containing 134	1.5	484.8	763.2	0.0020
311835	Ddx31	DEAD-box helicase 31	1.5	22.3	35.1	0.0084
363924	Rad9b	RAD9 checkpoint clamp component B	1.5	128.9	202.7	0.0009
116684	Oplah	5-oxoprolinase (ATP-hydrolysing)	1.5	727.8	1143.3	0.0064
499106	RGD1560854	similar to FLJ41131 protein	1.5	113.3	177.9	0.0003
289758	Pold2	DNA polymerase delta 2, accessory subunit	1.5	275.2	431.8	0.0021
308850	Klhl35	kelch-like family member 35	1.5	34.1	53.5	0.0054
24468	Hspa8	heat shock protein family A (Hsp70) member 8	1.5	32107.1	50365.3	0.0031
310244	Zc2hc1a	zinc finger, C2HC-type containing 1A	1.5	19.0	29.7	0.0020
303772	Slc16a6	solute carrier family 16, member 6	1.5	36.4	57.0	0.0211
290596	Ghitm	growth hormone inducible transmembrane protein	1.5	3324.9	5210.9	0.0011
302680	Txlng	taxilin gamma	1.5	920.7	1442.3	0.0007
360970	Cabp7	calcium binding protein 7	1.5	12.3	19.3	0.0002
688637	LOC688637	similar to WD repeat domain 36	1.5	39.3	61.6	0.0012
58958	Nup210	nucleoporin 210	1.5	650.0	1016.8	0.0001
29407	Tas1r1	taste 1 receptor member 1	1.5	96.8	151.3	0.0041
498385	Lcorl	ligand dependent nuclear receptor corepressor-like	1.5	14.4	22.5	0.0311
116641	Lgals8	galectin 8	1.5	926.4	1444.8	0.0011
498297	Cnst	consortin, connexin sorting protein	1.5	90.2	140.5	0.0013

[a] Gene IDs reference The National Center for Biotechnology Information database.

[b] Data are expressed as the mean (n = 6), analyzed by Moderated t-tests using GeneSpring GX ver. 14.9.

Supplementary table 3-S8, continued

NCBI Entrez Gene ID [a]	Gene Symbol	Gene Name	Fold Change	CT diet group (raw intensity) [b]	LP diet group (raw intensity)	*P* value
294436	Rpf2	ribosome production factor 2 homolog	1.5	198.4	309.0	0.0016
691414	LOC691414	hypothetical protein LOC691414	1.5	250.0	389.2	0.0008
300050	Bop1	block of proliferation 1	1.5	233.1	362.7	0.0003
363472	Prrg1	proline rich and Gla domain 1	1.5	31.8	49.4	0.0295
687147	Ric3	RIC3 acetylcholine receptor chaperone	1.5	17.1	26.5	0.0082
362374	Vopp1	vesicular, overexpressed in cancer, prosurvival protein 1	-1.5	48.8	36.1	0.0169
56765	Plvap	plasmalemma vesicle associated protein	-1.5	1309.6	967.4	0.0020
59085	Asl	argininosuccinate lyase	-1.5	17819.0	13150.2	0.0217
60586	Clcn4	chloride voltage-gated channel 4	-1.5	219.0	161.6	0.0047
25642	Cyp3a23/3a1	cytochrome P450, family 3, subfamily a, polypeptide 23/polypeptide 1	-1.5	511.0	377.0	0.0240
288264	Ifnar1	interferon alpha and beta receptor subunit 1	-1.5	41.3	30.5	0.0005
307726	Chd9	chromodomain helicase DNA binding protein 9	-1.5	321.8	237.3	0.0327
290529	Fam213a	family with sequence similarity 213, member A	-1.5	1028.8	758.1	0.0006
291359	Ly86	lymphocyte antigen 86	-1.5	411.5	303.2	0.0121
25164	Cdkn2b	cyclin-dependent kinase inhibitor 2B	-1.5	75.6	55.7	0.0120
301252	Hsp90ab1	heat shock protein 90 alpha family class B member 1	-1.5	8485.9	6249.7	0.0011
493909	Aox3	aldehyde oxidase 3	-1.5	1182.2	870.5	0.0100
25668	Cd38	CD38 molecule	-1.5	54.2	39.9	0.0287
313166	Nfx1	nuclear transcription factor, X-box binding 1	-1.5	1673.2	1231.3	0.0003
289175	Slc19a2	solute carrier family 19 member 2	-1.5	1188.5	874.5	0.0025
362479	Tcea1	transcription elongation factor A (SII) 1	-1.5	1921.2	1413.3	0.0000
311642	Sulf2	sulfatase 2	-1.5	1261.4	927.5	0.0065
303029	Ranbp17	RAN binding protein 17	-1.5	161.6	118.8	0.0066
367181	Gadl1	glutamate decarboxylase-like 1	-1.5	218.2	160.4	0.0304
304971	Tomm40l	translocase of outer mitochondrial membrane 40 like	-1.5	77.1	56.6	0.0018
102549509	LOC102549509	MLV-related proviral Env polyprotein-like	-1.5	27.1	19.9	0.0001
362455	Fam234b	family with sequence similarity 234, member B	-1.5	39.0	28.6	0.0100
78969	Trib1	tribbles pseudokinase 1	-1.5	120.1	88.0	0.0351
89811	Vegfb	vascular endothelial growth factor B	-1.5	165.5	121.0	0.0003
315658	Zc3h12c	zinc finger CCCH type containing 12C	-1.5	237.8	173.8	0.0035
361313	Rell2	RELT-like 2	-1.5	64.0	46.7	0.0003
309891	Septin10	septin 10	-1.5	82.3	60.1	0.0001
304988	Mnda	myeloid cell nuclear differentiation antigen	-1.5	61.1	44.6	0.0023
679154	Garem1	GRB2 associated regulator of MAPK1 subtype 1	-1.5	59.1	43.1	0.0267
100362283	Tmem52	transmembrane protein 52	-1.5	1060.6	773.2	0.0115
29500	Slc10a2	solute carrier family 10 member 2	-1.5	32.9	23.9	0.0376
25211	Lyz2	lysozyme 2	-1.5	4557.8	3320.6	0.0034
361135	Marchf1	membrane associated ring-CH-type finger 1	-1.5	135.0	98.4	0.0042
305452	Fam193a	family with sequence similarity 193, member A	-1.5	61.8	45.0	0.0011
24811	Tap1	transporter 1, ATP binding cassette subfamily B member	-1.5	168.4	122.6	0.0038
362833	Mob3a	MOB kinase activator 3A	-1.5	27.9	20.3	0.0076
295647	Gca	grancalcin	-1.5	46.2	33.6	0.0051
499358	Scd4	stearoyl-coenzyme A desaturase 4	-1.5	12664.0	9199.0	0.0145
313838	Cdc42ep3	CDC42 effector protein 3	-1.5	105.6	76.7	0.0015
29197	Il18	interleukin 18	-1.5	729.5	529.1	0.0001
291705	Iws1	IWS1, SUPT6H interacting protein	-1.5	69.1	50.1	0.0033
24162	Acp2	acid phosphatase 2, lysosomal	-1.5	5406.1	3916.5	0.0001
367289	Kansl1l	KAT8 regulatory NSL complex subunit 1-like	-1.5	22.0	15.9	0.0058
299566	Cyp4f39	cytochrome P450, family 4, subfamily f, polypeptide 39	-1.5	240.9	174.4	0.0035
301549	Wdfy1	WD repeat and FYVE domain containing 1	-1.5	699.5	506.2	0.0001

[a] Gene IDs reference The National Center for Biotechnology Information database.

[b] Data are expressed as the mean (n = 6), analyzed by Moderated t-tests using GeneSpring GX ver. 14.9.

Supplementary table 3-S8, continued

NCBI Entrez Gene ID [a]	Gene Symbol	Gene Name	Fold Change	CT diet group (raw intensity) [b]	LP diet group (raw intensity)	*P* value
678796	LOC678796	similar to acetyl-Coenzyme A acetyltransferase 2	-1.5	1628.4	1178.5	0.0031
681458	LOC681458	similar to stearoyl-coenzyme A desaturase 3	-1.5	4026.8	2913.7	0.0186
116637	Ccl4	C-C motif chemokine ligand 4	-1.5	86.4	62.4	0.0278
286918	Mx2	MX dynamin like GTPase 2	-1.5	77.4	55.9	0.0071
58935	Gas6	growth arrest specific 6	-1.5	1661.1	1198.9	0.0116
25023	Prkcb	protein kinase C, beta	-1.5	62.9	45.4	0.0195
100910616	Hmgn5	high mobility group nucleosome binding domain 5	-1.5	413.1	297.7	0.0001
114247	Slfn4	schlafen 4	-1.5	357.2	257.1	0.0053
300807	Rora	RAR-related orphan receptor A	-1.5	106.3	76.5	0.0067
292710	Ethe1	ETHE1, persulfide dioxygenase	-1.5	3727.3	2679.5	0.0040
294273	RT1-DMb	RT1 class II, locus DMb	-1.5	602.7	433.2	0.0020
24230	Tspo	translocator protein	-1.5	908.3	652.8	0.0001
308350	Lilrb3	leukocyte immunoglobulin like receptor B3	-1.5	1749.4	1257.3	0.0017
54272	Mcpt9	mast cell protease 9	-1.5	369.1	265.3	0.0000
298296	Pcsk9	proprotein convertase subtilisin/kexin type 9	-1.5	10062.8	7222.6	0.0047
83502	Cdh1	cadherin 1	-1.5	2989.5	2144.6	0.0281
291770	Chst9	carbohydrate sulfotransferase 9	-1.5	35.1	25.2	0.0011
81670	Gpt	glutamic--pyruvic transaminase	-1.5	6475.7	4641.5	0.0021
25045	Dbi	diazepam binding inhibitor, acyl-CoA binding protein	-1.5	82323.4	58982.2	0.0000
246266	Aspg	asparaginase	-1.5	19013.7	13613.9	0.0002
500561	Themis2	thymocyte selection associated family member 2	-1.5	61.6	44.1	0.0042
192256	Pja2	praja ring finger ubiquitin ligase 2	-1.5	335.7	240.3	0.0001
24298	Cyp21a1	cytochrome P450, family 21, subfamily a, polypeptide 1	-1.5	159.0	113.7	0.0008
286888	Wfdc2	WAP four-disulfide core domain 2	-1.5	2905.9	2075.2	0.0000
288289	Chodl	chondrolectin	-1.5	218.5	155.9	0.0312
291840	Slc38a7	solute carrier family 38, member 7	-1.5	54.0	38.5	0.0018
501779	Cd200r1l	CD200 receptor 1-like	-1.5	224.2	159.8	0.0175
681820	Erich5	glutamate-rich 5	-1.5	91.9	65.5	0.0259
29618	Btg1	BTG anti-proliferation factor 1	-1.5	30110.7	21437.2	0.0016
689069	Parvg	parvin, gamma	-1.5	70.9	50.4	0.0140
302492	Mbnl3	muscleblind-like splicing regulator 3	-1.5	690.9	491.6	0.0028
171164	Gbp2	guanylate binding protein 2	-1.5	1201.0	854.4	0.0253
361663	Lhpp	phospholysine phosphohistidine inorganic pyrophosphate phosphatase	-1.5	34.9	24.8	0.0019
24629	Pdgfrb	platelet derived growth factor receptor beta	-1.5	290.2	206.4	0.0460
498407	Purb	purine rich element binding protein B	-1.5	267.2	190.0	0.0000
293645	Ppfia1	PTPRF interacting protein alpha 1	-1.5	197.9	140.5	0.0004
287532	Slc43a2	solute carrier family 43 member 2	-1.5	333.1	236.4	0.0028
314751	Tmcc3	transmembrane and coiled-coil domain family 3	-1.5	324.0	229.9	0.0107
25618	Acadsb	acyl-CoA dehydrogenase, short/branched chain	-1.5	107.7	76.4	0.0025
108348057	NEWGENE1591402	pleckstrin homology like domain family B member 2	-1.5	86.9	61.6	0.0317
287362	Nlrp3	NLR family, pyrin domain containing 3	-1.5	23.3	16.5	0.0160
294993	Hspa4l	heat shock protein family A (Hsp70) member 4 like	-1.5	175.9	124.4	0.0136
24575	Mx1	myxovirus (influenza virus) resistance 1	-1.5	1085.5	767.0	0.0009
500709	Itpk1	inositol-tetrakisphosphate 1-kinase	-1.5	29.8	21.0	0.0284
54308	Slk	STE20-like kinase	-1.5	243.1	171.5	0.0001
450235	Adgre4	adhesion G protein-coupled receptor E4	-1.5	300.7	212.2	0.0233
25107	Avpr1a	arginine vasopressin receptor 1A	-1.5	5005.1	3530.4	0.0467
58962	Hpgds	hematopoietic prostaglandin D synthase	-1.5	91.8	64.8	0.0022
688749	Tulp3	tubby-like protein 3	-1.5	680.1	479.4	0.0006
81805	Suox	sulfite oxidase	-1.5	4175.4	2942.4	0.0000

[a] Gene IDs reference The National Center for Biotechnology Information database.

[b] Data are expressed as the mean (n = 6), analyzed by Moderated t-tests using GeneSpring GX ver. 14.9.

Supplementary table 3-S8, continued

NCBI Entrez Gene ID [a]	Gene Symbol	Gene Name	Fold Change	CT diet group (raw intensity) [b]	LP diet group (raw intensity)	*P* value
316775	Ankrd12	ankyrin repeat domain 12	-1.5	77.3	54.4	0.0015
316009	Plxnb1	plexin B1	-1.5	149.8	105.4	0.0060
406008	Olr1694	olfactory receptor 1694	-1.5	26.9	18.9	0.0313
309375	Marveld1	MARVEL domain containing 1	-1.5	794.8	558.8	0.0099
246208	Ifi47	interferon gamma inducible protein 47	-1.5	142.6	100.3	0.0097
29269	Mcpt8	mast cell protease 8	-1.5	777.4	545.7	0.0000
296925	Aass	aminoadipate-semialdehyde synthase	-1.5	527.0	369.8	0.0081
691416	Aknad1	AKNA domain containing 1	-1.5	666.2	467.4	0.0001
294691	Naip5	NLR family, apoptosis inhibitory protein 5	-1.5	184.8	129.6	0.0018
288866	Olr1057	olfactory receptor 1057	-1.5	22.7	15.9	0.0230
307414	MGC108823	similar to interferon-inducible GTPase	-1.5	140.0	98.0	0.0016
24962	Cth	cystathionine gamma-lyase	-1.5	6930.5	4844.9	0.0060
361069	Dmtn	dematin actin binding protein	-1.5	113.3	79.1	0.0004
83801	Ptms	parathymosin	-1.5	4174.4	2915.4	0.0001
304549	Oasl2	2'-5' oligoadenylate synthetase-like 2	-1.5	197.4	137.8	0.0025
100361457	LOC100361457	actin, gamma 1 propeptide-like	-1.5	622.8	433.9	0.0026
246142	Bmf	Bcl2 modifying factor	-1.5	5309.2	3698.1	0.0005
361625	Nrip3	nuclear receptor interacting protein 3	-1.5	28.9	20.1	0.0034
641316	Aldh4a1	aldehyde dehydrogenase 4 family, member A1	-1.5	325.1	226.1	0.0003
100911537	LOC100911537	5'-3' exoribonuclease 1-like	-1.5	38.8	27.0	0.0078
293749	Ms4a6bl	membrane-spanning 4-domains, subfamily A, member 6B-like	-1.5	1548.4	1075.1	0.0005
161476	Hspb2	heat shock protein family B (small) member 2	-1.5	9970.5	6917.9	0.0009
362832	Izumo4	IZUMO family member 4	-1.5	180.9	125.4	0.0003
294274	RT1-DMa	RT1 class II, locus DMa	-1.5	245.6	170.1	0.0086
310917	Gbp4	guanylate binding protein 4	-1.6	17.5	12.1	0.0249
170841	Mutyh	mutY DNA glycosylase	-1.6	405.8	280.3	0.0003
25017	Myo5a	myosin VA	-1.6	23.6	16.3	0.0148
114490	Cited2	Cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy-terminal domain, 2	-1.6	332.5	229.3	0.0003
100910577	LOC100910577	zinc finger protein 709-like	-1.6	88.5	61.0	0.0020
361422	Cotl1	coactosin-like F-actin binding protein 1	-1.6	18.9	13.0	0.0301
304061	Brwd1	bromodomain and WD repeat domain containing 1	-1.6	92.0	63.4	0.0037
171497	Sgk2	SGK2, serine/threonine kinase 2	-1.6	1060.6	730.6	0.0286
296653	Phf19	PHD finger protein 19	-1.6	836.3	573.1	0.0030
60371	Birc2	baculoviral IAP repeat-containing 2	-1.6	52.1	35.6	0.0082
497840	Cps1	carbamoyl-phosphate synthase 1	-1.6	4255.3	2907.6	0.0006
313033	Xkr8	XK related 8	-1.6	2050.3	1400.0	0.0000
83791	Fdps	farnesyl diphosphate synthase	-1.6	13595.1	9277.0	0.0022
306115	RGD1560797	similar to glyceraldehyde-3-phosphate dehydrogenase	-1.6	47.3	32.2	0.0010
316638	Sned1	sushi, nidogen and EGF-like domains 1	-1.6	99.2	67.5	0.0402
81718	Cdo1	cysteine dioxygenase type 1	-1.6	9016.4	6127.9	0.0015
102549391	LOC102549391	zinc finger protein 879-like	-1.6	37.2	25.2	0.0137
290032	Eddm3b	epididymal protein 3B	-1.6	74.7	50.7	0.0002
24950	Srd5a1	steroid 5 alpha-reductase 1	-1.6	6613.1	4488.2	0.0237
311569	Acss2	acyl-CoA synthetase short-chain family member 2	-1.6	1048.9	711.4	0.0047
362300	Agap3	ArfGAP with GTPase domain, ankyrin repeat and PH domain 3	-1.6	191.7	130.0	0.0001
59296	Dnase2b	deoxyribonuclease 2 beta	-1.6	1653.8	1120.4	0.0382
298033	Snx30	sorting nexin family member 30	-1.6	22.2	15.0	0.0046
309312	Gldc	glycine decarboxylase	-1.6	6090.9	4114.3	0.0002
116699	Inpp4b	inositol polyphosphate-4-phosphatase type II B	-1.6	100.2	67.6	0.0030
293624	Irf7	interferon regulatory factor 7	-1.6	1532.0	1033.9	0.0020

[a] Gene IDs reference The National Center for Biotechnology Information database.

[b] Data are expressed as the mean (n = 6), analyzed by Moderated t-tests using GeneSpring GX ver. 14.9.

Supplementary table 3-S8, continued

NCBI Entrez Gene ID [a]	Gene Symbol	Gene Name	Fold Change	CT diet group (raw intensity) [b]	LP diet group (raw intensity)	*P* value
100302372	LOC100302372	hypothetical protein LOC100302372	-1.6	123.0	83.0	0.0333
100912573	Ms4a6e	membrane spanning 4-domains A6E	-1.6	24.1	16.3	0.0051
301363	Map4k4	mitogen-activated protein kinase kinase kinase kinase 4	-1.6	303.6	204.4	0.0002
307526	Ammecr1l	AMMECR1 like	-1.6	997.0	670.3	0.0132
360733	Rtp4	receptor (chemosensory) transporter protein 4	-1.6	2268.9	1523.8	0.0242
679600	Xaf1	XIAP associated factor 1	-1.6	57.0	38.3	0.0002
688047	Lyc2	lysozyme C type 2	-1.6	949.6	636.9	0.0013
29412	Adra1a	adrenoceptor alpha 1A	-1.6	35.7	24.0	0.0264
29191	Tac3	tachykinin 3	-1.6	62.6	42.0	0.0309
361735	Ms4a6a	membrane spanning 4-domains A6A	-1.6	215.1	144.0	0.0005
366669	Syne2	spectrin repeat containing nuclear envelope protein 2	-1.6	681.9	456.3	0.0002
64322	Dap	death-associated protein	-1.6	4095.0	2740.0	0.0001
679692	Lpgat1	lysophosphatidylglycerol acyltransferase 1	-1.6	323.1	216.0	0.0001
680409	Prodh1	proline dehydrogenase 1	-1.6	1085.2	724.5	0.0300
140868	Fabp5	fatty acid binding protein 5, epidermal	-1.6	12215.0	8146.1	0.0160
89787	Lrp3	LDL receptor related protein 3	-1.6	1751.4	1167.2	0.0128
680656	Fndc10	fibronectin type III domain containing 10	-1.6	56.4	37.5	0.0010
29210	Epha3	Eph receptor A3	-1.6	118.1	78.7	0.0469
362454	Hebp1	heme binding protein 1	-1.6	10468.6	6956.0	0.0013
367391	Marco	macrophage receptor with collagenous structure	-1.6	361.2	240.0	0.0380
309962	Pde8b	phosphodiesterase 8B	-1.6	241.8	160.5	0.0015
64191	Dhcr7	7-dehydrocholesterol reductase	-1.6	3334.6	2206.1	0.0133
361604	Sytl2	synaptotagmin-like 2	-1.6	182.8	120.8	0.0468
305235	Art3	ADP-ribosyltransferase 3	-1.6	149.4	98.8	0.0339
24934	Klrk1	killer cell lectin like receptor K1	-1.6	19.8	13.0	0.0010
683415	Prss55	protease, serine, 55	-1.6	40.9	27.0	0.0030
54269	Mcpt10	mast cell protease 10	-1.6	284.5	187.3	0.0000
64552	Mpeg1	macrophage expressed 1	-1.6	83.6	55.1	0.0018
304645	Tbc1d9	TBC1 domain family member 9	-1.6	239.8	157.8	0.0002
312560	Xpc	XPC complex subunit, DNA damage recognition and repair factor	-1.6	526.6	346.6	0.0052
366352	Mob3b	MOB kinase activator 3B	-1.6	58.6	38.5	0.0016
367975	LOC367975	similar to B-cell translocation gene 1	-1.6	618.0	405.8	0.0019
291699	Stard4	StAR-related lipid transfer domain containing 4	-1.6	24.5	16.1	0.0022
338475	Nrep	neuronal regeneration related protein	-1.6	82.6	54.2	0.0134
295625	Pkp4	plakophilin 4	-1.6	1267.2	830.2	0.0000
24895	Cyp2a2	cytochrome P450, family 2, subfamily a, polypeptide 2	-1.6	11440.7	7488.7	0.0000
170641	Adgrl3	adhesion G protein-coupled receptor L3	-1.6	34.2	22.4	0.0066
289594	Nipal1	NIPA-like domain containing 1	-1.6	585.7	383.3	0.0042
303141	Rapgef6	Rap guanine nucleotide exchange factor 6	-1.6	58.7	38.3	0.0004
287472	Tlcd1	TLC domain containing 1	-1.6	337.4	220.3	0.0009
360942	Pcdh7	protocadherin 7	-1.6	36.4	23.7	0.0206
306600	Atp11a	ATPase phospholipid transporting 11A	-1.6	66.3	43.1	0.0014
84424	Tle3	transducin-like enhancer of split 3	-1.7	71.9	46.7	0.0037
81869	Gstm7	glutathione S-transferase, mu 7	-1.7	2244.1	1451.8	0.0025
688429	Arhgap10	Rho GTPase activating protein 10	-1.7	28.9	18.7	0.0175
500200	Atoh8	atonal bHLH transcription factor 8	-1.7	209.7	135.2	0.0018
317213	Brwd3	bromodomain and WD repeat domain containing 3	-1.7	89.3	57.6	0.0000
316626	Hes6	hes family bHLH transcription factor 6	-1.7	7287.5	4688.1	0.0012
368088	Dgkd	diacylglycerol kinase, delta	-1.7	363.6	233.8	0.0000
366518	Tnfrsf14	TNF receptor superfamily member 14	-1.7	74.1	47.6	0.0052

[a] Gene IDs reference The National Center for Biotechnology Information database.

[b] Data are expressed as the mean (n = 6), analyzed by Moderated t-tests using GeneSpring GX ver. 14.9.

Supplementary table 3-S8, continued

NCBI Entrez Gene ID [a]	Gene Symbol	Gene Name	Fold Change	CT diet group (raw intensity) [b]	LP diet group (raw intensity)	*P* value
287734	Atxn7l3	ataxin 7-like 3	-1.7	55.4	35.6	0.0003
171105	Lnpep	leucyl and cystinyl aminopeptidase	-1.7	31.7	20.3	0.0000
312275	Ephb6	Eph receptor B6	-1.7	1256.5	801.3	0.0009
306038	Enox1	ecto-NOX disulfide-thiol exchanger 1	-1.7	32.6	20.8	0.0322
24778	Slc2a1	solute carrier family 2 member 1	-1.7	420.4	267.8	0.0250
288665	Cux2	cut-like homeobox 2	-1.7	25.5	16.2	0.0113
24912	Sult2a1	sulfotransferase family 2A member 1	-1.7	864.1	549.2	0.0357
308797	Cemip	cell migration-inducing hyaluronan binding protein	-1.7	88.9	56.4	0.0081
360509	Dock2	dedicator of cytokinesis 2	-1.7	525.6	333.1	0.0002
687064	Col25a1	collagen type XXV alpha 1 chain	-1.7	27.8	17.6	0.0150
192268	Gls2	glutaminase 2	-1.7	28655.7	18142.3	0.0001
102556096	LOC102556096	pyrin domain-containing protein 3-like	-1.7	44.6	28.2	0.0001
304507	Oas1i	2 ' -5 ' oligoadenylate synthetase 1I	-1.7	471.5	296.4	0.0006
293587	Fuom	fucose mutarotase	-1.7	279.7	175.4	0.0003
29536	Rmrp	RNA component of mitochondrial RNA processing endoribonuclease	-1.7	918.6	575.9	0.0036
308100	Acat2	acetyl-CoA acetyltransferase 2	-1.7	2234.8	1400.4	0.0023
362282	Pck1	phosphoenolpyruvate carboxykinase 1	-1.7	13784.0	8630.6	0.0260
309622	RT1-Bb	RT1 class II, locus Bb	-1.7	505.0	315.3	0.0136
494198	Oas1k	2 ' -5 ' oligoadenylate synthetase 1K	-1.7	516.7	321.1	0.0004
312607	Pdzrn3	PDZ domain containing RING finger 3	-1.7	188.5	117.1	0.0316
25427	Cyp51	cytochrome P450, family 51	-1.7	10134.1	6271.0	0.0004
294945	RGD1562890	RGD1562890	-1.7	92.4	57.1	0.0000
246771	Slc25a25	solute carrier family 25 member 25	-1.7	23072.5	14258.9	0.0157
114554	Pi4k2a	phosphatidylinositol 4-kinase type 2 alpha	-1.7	163.0	100.5	0.0007
294251	Ncr3	natural cytotoxicity triggering receptor 3	-1.7	195.5	120.3	0.0000
192281	Oas1a	2'-5' oligoadenylate synthetase 1A	-1.7	565.7	347.2	0.0001
108352751	LOC108352751	S-antigen protein-like	-1.7	2403.6	1474.3	0.0143
315644	Bco2	beta-carotene oxygenase 2	-1.7	421.9	258.8	0.0009
170573	Slc38a4	solute carrier family 38, member 4	-1.7	8866.3	5432.8	0.0004
25081	Apoa1	apolipoprotein A1	-1.8	58117.5	35596.9	0.0002
684984	Cd300lg	Cd300 molecule-like family member G	-1.8	81.0	49.5	0.0152
312907	Slco5a1	solute carrier organic anion transporter family, member 5A1	-1.8	38.4	23.5	0.0001
81681	Lss	lanosterol synthase	-1.8	1292.9	788.2	0.0001
100360180	Pgd	phosphogluconate dehydrogenase	-1.8	813.7	494.8	0.0044
171400	Elovl5	ELOVL fatty acid elongase 5	-1.8	203.9	123.5	0.0000
501145	Satb2	SATB homeobox 2	-1.8	57.4	34.8	0.0378
280670	Lzts3	leucine zipper tumor suppressor family member 3	-1.8	471.6	285.1	0.0000
500011	RGD1563091	similar to OEF2	-1.8	44.0	26.5	0.0006
24932	Cd4	Cd4 molecule	-1.8	1819.7	1097.4	0.0000
304077	Dopey2	dopey family member 2	-1.8	169.5	102.1	0.0000
365352	Mical2	microtubule associated monooxygenase, calponin and LIM domain containing 2	-1.8	42.6	25.6	0.0151
316729	Fam114a1l1	family with sequence similarity 114, member A1-like 1	-1.8	86.7	52.1	0.0003
29157	Ccng2	cyclin G2	-1.8	373.7	224.2	0.0007
301444	LOC301444	pseudogene for diazepam binding inhibitor 1	-1.8	1493.3	893.9	0.0000
288280	Ncam2	neural cell adhesion molecule 2	-1.8	68.3	40.8	0.0025
84575	Fads1	fatty acid desaturase 1	-1.8	4985.4	2979.8	0.0000
25599	Cd74	CD74 molecule	-1.8	2962.9	1770.4	0.0115
498728	Elovl2	ELOVL fatty acid elongase 2	-1.8	2321.8	1381.0	0.0039
298579	Pla2g2d	phospholipase A2, group IID	-1.8	208.7	124.0	0.0020
84012	Slc1a6	solute carrier family 1 member 6	-1.8	85.1	50.5	0.0041

[a] Gene IDs reference The National Center for Biotechnology Information database.

[b] Data are expressed as the mean (n = 6), analyzed by Moderated t-tests using GeneSpring GX ver. 14.9.

Supplementary table 3-S8, continued

NCBI Entrez Gene ID [a]	Gene Symbol	Gene Name	Fold Change	CT diet group (raw intensity) [b]	LP diet group (raw intensity)	*P* value
84426	Wnt4	Wnt family member 4	-1.8	38.3	22.7	0.0380
266729	Dab1	DAB1, reelin adaptor protein	-1.8	163.0	96.5	0.0455
298077	RGD1305807	hypothetical LOC298077	-1.8	1579.4	934.7	0.0000
361049	RGD1563302	RGD1563302	-1.8	23.3	13.8	0.0337
294269	RT1-Da	RT1 class II, locus Da	-1.8	1508.2	891.0	0.0088
445271	Jph4	junctophilin 4	-1.8	56.8	33.6	0.0257
115768	Zfp37	zinc finger protein 37	-1.8	73.4	43.4	0.0429
363143	Gnat1	G protein subunit alpha transducin 1	-1.8	180.1	106.4	0.0011
24482	Igf1	insulin-like growth factor 1	-1.8	2330.6	1375.7	0.0000
293166	Insc	INSC, spindle orientation adaptor protein	-1.8	438.2	258.0	0.0040
362319	Tmbim7	transmembrane BAX inhibitor motif containing 7	-1.8	38.3	22.6	0.0016
360406	Ptprf	protein tyrosine phosphatase, receptor type, F	-1.8	20956.3	12290.0	0.0037
25060	Hk3	hexokinase 3	-1.8	105.0	61.5	0.0032
304445	Glt1d1	glycosyltransferase 1 domain containing 1	-1.8	136.6	79.9	0.0015
361834	RGD1306739	similar to RIKEN cDNA 1700040L02	-1.8	37.7	22.0	0.0035
83512	Fads2	fatty acid desaturase 2	-1.8	2488.3	1444.6	0.0013
286954	Ugt2b1	UDP glucuronosyltransferase 2 family, polypeptide B1	-1.8	3105.0	1800.5	0.0304
294270	RT1-Db1	RT1 class II, locus Db1	-1.9	1283.4	741.0	0.0061
58975	Klrg1	killer cell lectin like receptor G1	-1.9	30.4	17.5	0.0055
114024	Acsl3	acyl-CoA synthetase long-chain family member 3	-1.9	372.2	214.4	0.0001
64554	Slc7a2	solute carrier family 7 member 2	-1.9	607.6	349.0	0.0143
24981	RT1-Db2	RT1 class II, locus Db2	-1.9	464.8	266.3	0.0076
78968	Srebf1	sterol regulatory element binding transcription factor 1	-1.9	1128.5	646.0	0.0159
25317	Fgf1	fibroblast growth factor 1	-1.9	112.6	64.4	0.0038
81677	Itpka	inositol-trisphosphate 3-kinase A	-1.9	397.8	227.2	0.0002
314349	Ston2	stonin 2	-1.9	44.6	25.4	0.0021
25698	Ass1	argininosuccinate synthase 1	-1.9	40916.8	23228.6	0.0005
24237	C6	complement C6	-1.9	3798.0	2148.4	0.0000
502776	Scrn1	secernin 1	-1.9	85.0	47.8	0.0372
60391	Nrxn1	neurexin 1	-1.9	51.0	28.7	0.0055
299131	Dhrs7l1	dehydrogenase/reductase (SDR family) member 7-like 1	-1.9	1955.2	1092.0	0.0265
309866	Crybg1	crystallin beta-gamma domain containing 1	-1.9	1748.8	970.7	0.0281
300158	Kif21a	kinesin family member 21A	-1.9	29.6	16.4	0.0000
114856	Dusp1	dual specificity phosphatase 1	-1.9	302.5	167.7	0.0247
289392	Plxna2	plexin A2	-1.9	2412.0	1335.2	0.0004
295649	Fign	fidgetin, microtubule severing factor	-1.9	371.0	205.3	0.0002
60582	Il1rn	interleukin 1 receptor antagonist	-1.9	1823.2	1008.1	0.0020
103694826	LOC103694826	P2Y purinoceptor 2-like	-1.9	172.9	95.5	0.0000
685940	Samd3	sterile alpha motif domain containing 3	-1.9	65.6	36.1	0.0003
29619	Btg2	BTG anti-proliferation factor 2	-1.9	475.1	261.5	0.0013
29597	P2ry2	purinergic receptor P2Y2	-2.0	1180.4	644.6	0.0000
306534	Unc5d	unc-5 netrin receptor D	-2.0	29.6	16.1	0.0009
25117	Hsd11b2	hydroxysteroid 11-beta dehydrogenase 2	-2.0	298.8	162.1	0.0006
25453	Gdnf	glial cell derived neurotrophic factor	-2.0	96.0	51.6	0.0000
316940	Spetex-2F	Spetex-2F protein	-2.0	47.1	25.3	0.0000
102556734	LOC102556734	igE-binding protein-like	-2.0	25.2	13.5	0.0068
312381	Ppm1k	protein phosphatase, Mg2+/Mn2+ dependent, 1K	-2.0	138.4	74.0	0.0054
498834	Myo7b	myosin VIIb	-2.0	131.6	70.0	0.0009
300691	Nnmt	nicotinamide N-methyltransferase (NNMT)	-2.0	530.6	281.9	0.0135
498545	Tsc22d1	TSC22 domain family, member 1	-2.0	4514.2	2390.4	0.0015

[a] Gene IDs reference The National Center for Biotechnology Information database.

[b] Data are expressed as the mean (n = 6), analyzed by Moderated t-tests using GeneSpring GX ver. 14.9.

Supplementary table 3-S8, continued

NCBI Entrez Gene ID [a]	Gene Symbol	Gene Name	Fold Change	CT diet group (raw intensity) [b]	LP diet group (raw intensity)	*P* value
362790	Inf2	inverted formin, FH2 and WH2 domain containing	-2.0	694.8	367.6	0.0001
311676	Pmepa1	prostate transmembrane protein, androgen induced 1	-2.0	195.7	103.3	0.0002
24267	Comt	catechol-O-methyltransferase	-2.0	5385.9	2838.0	0.0002
300679	Mpzl2	myelin protein zero-like 2	-2.1	130.8	67.8	0.0457
498950	Marveld3	MARVEL domain containing 3	-2.1	158.2	81.9	0.0001
296787	Sema3c	semaphorin 3C	-2.1	74.8	38.6	0.0054
24401	Got1	glutamic-oxaloacetic transaminase 1	-2.1	5452.7	2803.3	0.0085
29298	Cyp2c7	cytochrome P450, family 2, subfamily c, polypeptide 7	-2.1	7306.0	3721.2	0.0001
362720	Rrm2	ribonucleotide reductase regulatory subunit M2	-2.1	191.4	97.4	0.0037
691517	Megf11	multiple EGF-like-domains 11	-2.1	92.1	46.8	0.0197
25598	Fabp2	fatty acid binding protein 2	-2.1	659.9	334.2	0.0064
309810	RGD1304770	similar to Na+ dependent glucose transporter 1	-2.1	39.7	20.1	0.0000
24706	Rarb	retinoic acid receptor, beta	-2.1	38.1	19.2	0.0005
314640	Tjp3	tight junction protein 3	-2.1	4881.6	2456.0	0.0000
366275	Ptk6	protein tyrosine kinase 6	-2.2	43.9	21.5	0.0139
310645	Pmvk	phosphomevalonate kinase	-2.2	2329.1	1137.4	0.0001
363974	Nckap5	NCK-associated protein 5	-2.2	97.8	47.3	0.0009
171352	Cyp3a9	cytochrome P450, family 3, subfamily a, polypeptide 9	-2.2	1907.6	912.3	0.0016
24249	S100g	S100 calcium binding protein G	-2.3	112.9	52.9	0.0009
192213	Rasgrf1	RAS protein-specific guanine nucleotide-releasing factor 1	-2.3	2464.0	1151.5	0.0002
360228	Wfdc21	WAP four-disulfide core domain 21	-2.3	26589.1	12407.0	0.0013
308739	Rgma	repulsive guidance molecule family member A	-2.3	214.6	99.9	0.0177
304537	Bicdl1	BICD family like cargo adaptor 1	-2.3	75.4	34.8	0.0001
171385	Acmsd	aminocarboxymuconate semialdehyde decarboxylase	-2.3	337.4	155.6	0.0003
65984	Aacs	acetoacetyl-CoA synthetase	-2.3	92.4	42.6	0.0261
298079	Aldh1b1	aldehyde dehydrogenase 1 family, member B1	-2.3	4537.7	2079.2	0.0011
303336	Wsb1	WD repeat and SOCS box-containing 1	-2.4	2466.6	1118.6	0.0007
316052	Eomes	eomesodermin	-2.4	33.6	15.2	0.0002
362378	Fam13a	family with sequence similarity 13, member A	-2.4	653.0	295.7	0.0032
361749	Il33	interleukin 33	-2.4	1863.8	836.6	0.0000
311396	Atp8b4	ATPase phospholipid transporting 8B4 (putative)	-2.4	96.0	42.5	0.0001
29517	Sgk1	serum/glucocorticoid regulated kinase 1	-2.4	983.3	435.1	0.0051
293688	Tm7sf2	transmembrane 7 superfamily member 2	-2.4	2661.4	1165.4	0.0001
116719	Acacb	acetyl-CoA carboxylase beta	-2.5	72.8	31.7	0.0000
361206	Slc25a48	solute carrier family 25, member 48	-2.5	35.2	15.3	0.0002
315179	Efcab6	EF-hand calcium binding domain 6	-2.5	114.0	48.7	0.0000
681319	Ido2	indoleamine 2,3-dioxygenase 2	-2.5	995.9	420.9	0.0000
114097	Ltc4s	leukotriene C4 synthase	-2.6	217.4	90.2	0.0022
24913	Pla2g16	phospholipase A2, group XVI	-2.7	7969.8	3173.9	0.0000
171516	Akr1c3	aldo-keto reductase family 1, member C3	-2.8	1399.4	545.4	0.0000
100359918	Lnp1	leukemia NUP98 fusion partner 1	-2.8	40.8	15.5	0.0000
361008	Spetex-2H	Spetex-2H protein	-2.8	140.9	53.5	0.0000
306490	Mtmr7	myotubularin related protein 7	-2.8	366.0	137.8	0.0008
266998	Slc13a5	solute carrier family 13 member 5	-3.0	875.1	315.9	0.0001
29168	Ubd	ubiquitin D	-3.0	10397.5	3745.7	0.0050
362662	Rbp7	retinol binding protein 7	-3.0	243.9	87.8	0.0000
25650	Atp1b1	ATPase Na+/K+ transporting subunit beta 1	-3.1	4422.6	1546.0	0.0003
305816	Ddhd1	DDHD domain containing 1	-3.2	78.1	26.1	0.0001
25304	Comp	cartilage oligomeric matrix protein	-3.4	94.9	30.3	0.0034
25011	Cyp2c12	cytochrome P450, family 2, subfamily c, polypeptide 12	-3.4	2837.2	892.0	0.0022

[a] Gene IDs reference The National Center for Biotechnology Information database.

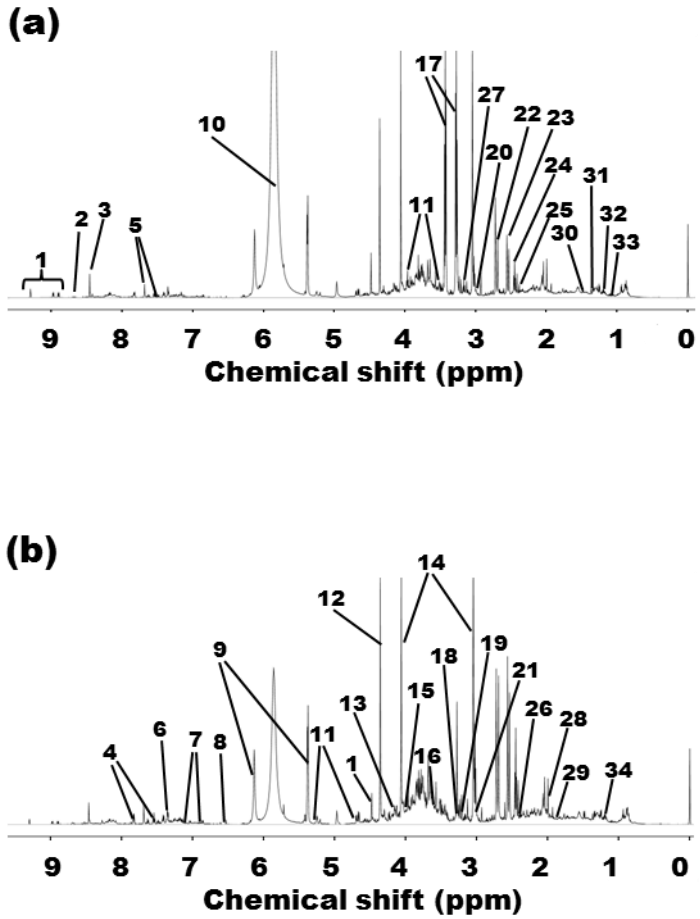
[b] Data are expressed as the mean (n = 6), analyzed by Moderated t-tests using GeneSpring GX ver. 14.9.

Supplementary table 3-S8, continued

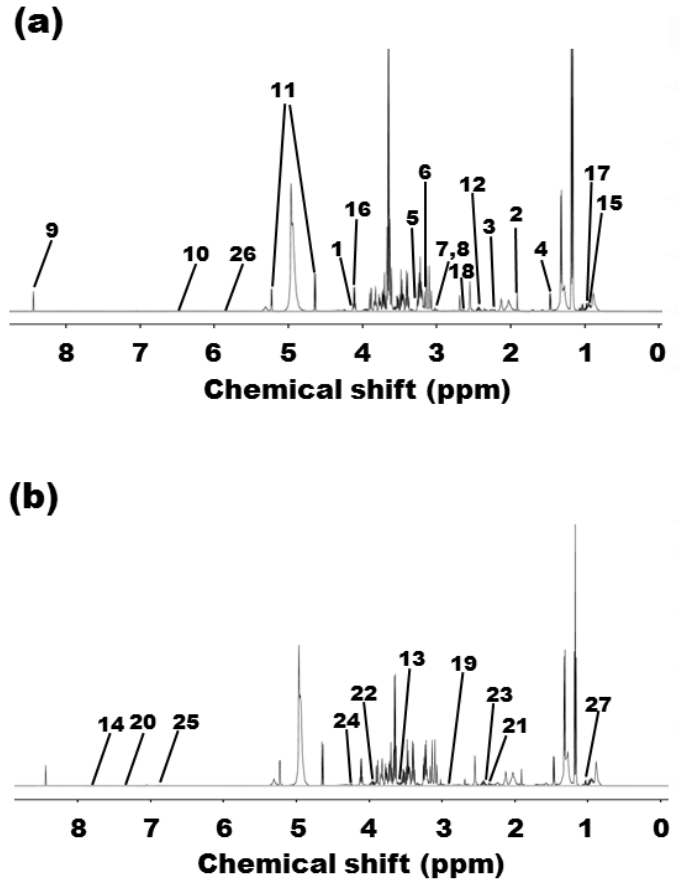
NCBI Entrez Gene ID [a]	Gene Symbol	Gene Name	Fold Change	CT diet group (raw intensity) [b]	LP diet group (raw intensity)	*P* value
24377	G6pd	glucose-6-phosphate dehydrogenase	-3.5	3157.9	963.0	0.0000
306748	Cxcl14	C-X-C motif chemokine ligand 14	-3.6	283.2	85.1	0.0002
108350694	LOC108350694	insulin-induced gene 1 protein pseudogene	-3.6	1333.2	392.1	0.0005
170566	Slc16a10	solute carrier family 16 member 10	-3.9	1443.7	401.0	0.0000
24813	Tat	tyrosine aminotransferase	-4.0	1745.4	473.3	0.0001
64194	Insig1	insulin induced gene 1	-4.1	13598.1	3518.4	0.0027
64313	Oat	ornithine aminotransferase	-4.4	21073.4	5079.9	0.0000
315962	Ky	kyphoscoliosis peptidase	-4.5	97.4	23.4	0.0014
84021	Irs3	insulin receptor substrate 3	-4.9	99.4	21.8	0.0000
680110	Rprm	reprimo, TP53 dependent G2 arrest mediator homolog	-4.9	430.5	94.1	0.0007
298062	Plppr1	phospholipid phosphatase related 1	-5.2	124.2	25.4	0.0079
300108	Pnpla5	patatin-like phospholipase domain containing 5	-5.7	6676.0	1262.6	0.0000
102557436	LOC102557436	cell surface A33 antigen-like	-7.3	442.8	64.7	0.0000
29464	Slc6a6	solute carrier family 6 member 6	-10.9	361.8	35.5	0.0000
362972	Pnpla3	patatin-like phospholipase domain containing 3	-15.8	929.4	62.9	0.0002
368066	Inmt	indolethylamine N-methyltransferase	-16.0	998.7	66.8	0.0000
25044	Sds	serine dehydratase	-35.7	2284.8	68.6	0.0000

[a] Gene IDs reference The National Center for Biotechnology Information database.

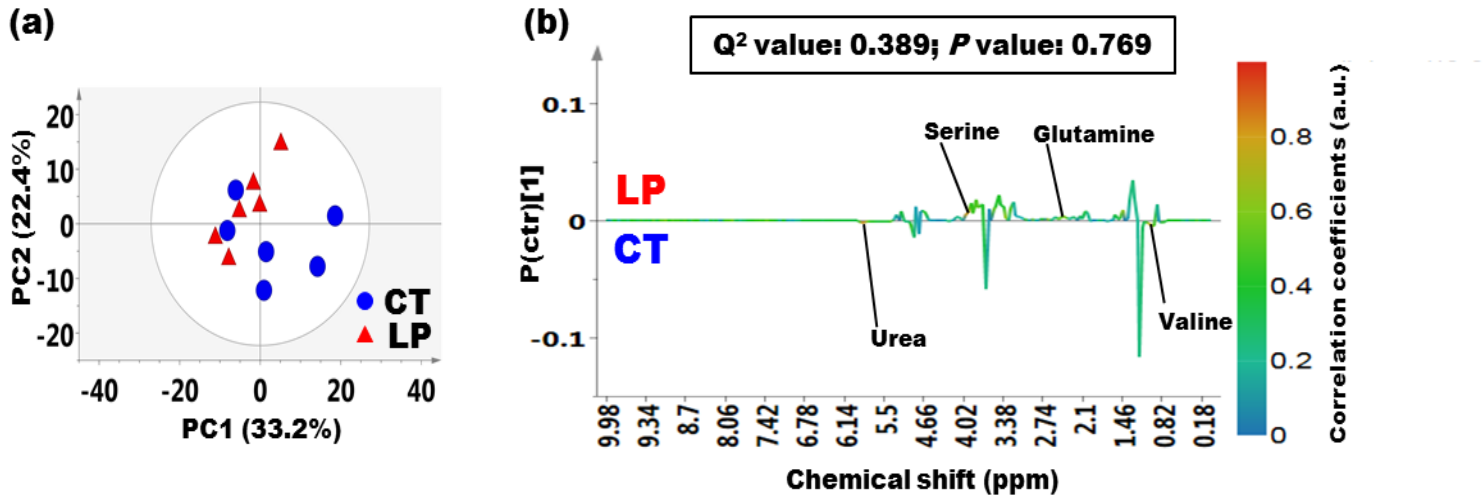
[b] Data are expressed as the mean (n = 6), analyzed by Moderated t-tests using GeneSpring GX ver. 14.9.



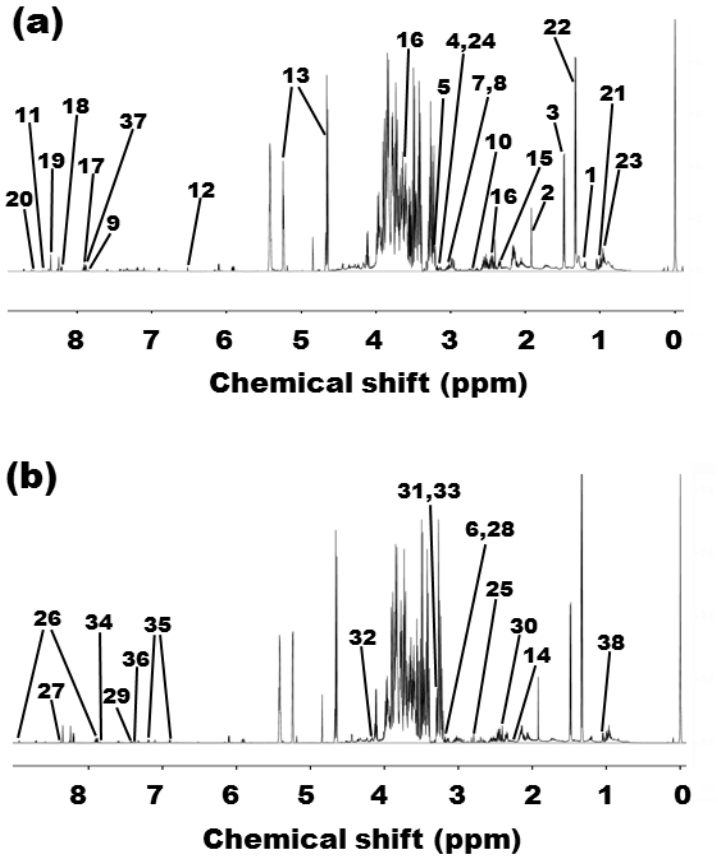
Supplementary figure 3-S1. ^{1}H NMR spectra of urine samples from rats fed control (CT) or low-protein (LP) diets at week 4. Typical 600 MHz ^{1}H NMR spectra of urine samples from rats fed (a) CT or (b) LP diets at week 4. Identified metabolites: 1: N-methylnicotinamide; 2: nicotinamide N-oxide; 3: formate; 4: hippurate; 5: tryptophan; 6: phenylalanine; 7: tyrosine; 8: fumarate; 9: allantoin; 10: urea; 11: glucose; 12: tartrate; 13: threonine; 14: creatinine; 15: creatine; 16: glycine; 17: taurine; 18: choline; 19: cis-aconitate; 20: cadaverine; 21: N,N-dimethylglycine; 22: dimethylamine; 23: citrate; 24: 2-oxoglutarate; 25: succinate; 26: acetoacetate; 27: trimethylamine N-oxide; 28: acetate; 29: putrescine; 30: alanine; 31: lactate; 32: 3-hydroxybutyrate; 33: isobutyrate; 34: valine.



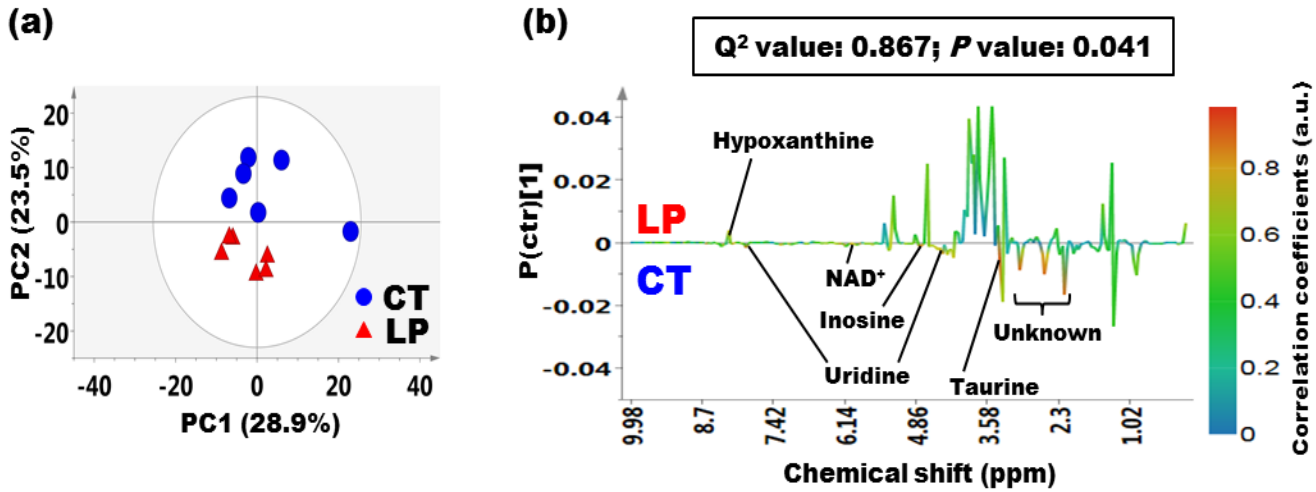
Supplementary figure 3-S2. ^{1}H NMR spectra of plasma samples from rats fed control (CT) or low-protein (LP) diets at week 4. Typical 600 MHz ^{1}H NMR spectra of plasma samples from rats fed (a) CT or (b) LP diets at week 4. Identified metabolites: 1: 3-hydroxybutyrate; 2: acetate; 3: acetoacetate; 4: alanine; 5: betaine; 6: choline; 7: creatine; 8: creatine phosphate; 9: formate; 10: fumarate; 11: glucose; 12: glutamine; 13: glycine; 14: histidine; 15: isoleucine; 16: lactate; 17: leucine; 18: methionine; 19: N,N-dimethylglycine; 20: phenylalanine; 21: pyruvate; 22: serine; 23: succinate; 24: threonine; 25: tyrosine; 26: urea; 27: valine.



Supplementary figure 3-S3. Principal component analysis (PCA) and Orthogonal partial least squares discriminant analysis of binned NMR data of plasma samples from rats fed control (CT) or low protein (LP) diets in at week 4. (a) PCA score plots derived from the NMR spectra of plasma samples from rats fed CT or LP diets at week 4. Each symbol represents the plasma sample from an individual animal (n = 6). (b) Coefficient loading plots obtained of plasma samples from rats fed CT or LP diets at week 4. Peaks pointing upward indicate metabolites whose levels were higher in the LP group than in the CT group, whereas peaks pointing downward indicate metabolites whose levels were higher in the CT group than in the LP group. Heat colors indicate the contribution degree of metabolites that separate the CT and LP diet groups. Among the assigned metabolites, those whose NMR spectra with a correlation coefficient of 0.70 or more are annotated in the figure. Partial least square model quality was determined by Q^2 value, which represent model predictability, and the *P* value of cross-validated analysis of variance. Each value is shown in the figure.



Supplementary figure 3-S4. ^{1}H NMR spectra of liver samples from rats fed control (CT) or low-protein (LP) diets at week 4. Typical 600 MHz ^{1}H NMR spectra of liver samples from rats fed (a) CT or (b) LP diets at week 4. Identified metabolites: 1: 3-hydroxybutyrate; 2: acetate; 3: alanine; 4: β-alanine; 5: betaine; 6: choline; 7: creatine; 8: creatine phosphate; 9: cytidine; 10: dimethylamine; 11: formate; 12: fumarate; 13: glucose; 14: glutamate; 15: glutamine; 16: glycerol; 17: histidine; 18: hypoxanthine; 19: inosine; 20: inosine monophosphate (IMP); 21: isoleucine; 22: lactate; 23: leucine; 24: lysine; 25: methionine; 26: niacinamide; 27: nicotinamide adenine dinucleotide (NAD^{+}); 28: o-phosphocholine; 29: phenylalanine; 30: succinate; 31: taurine; 32: threonine; 33: trimethylamine N-oxide; 34: tryptophan; 35: tyrosine; 36: uracil; 37: uridine; 38: valine.



Supplementary figure 3-S5. Principal component analysis (PCA) and Orthogonal partial least squares discriminant analysis of binned NMR data of liver samples from rats fed control (CT) or low protein (LP) diets in at week 4. (a) PCA score plots derived from the NMR spectra of liver samples from rats fed CT or LP diets at week 4. Each symbol represents the liver sample from an individual animal (n = 6). (b) Coefficient loading plots obtained of liver samples from rats fed CT or LP diets at week 4. Peaks pointing upward indicate metabolites whose levels were higher in the LP group than in the CT group, whereas peaks pointing downward indicate metabolites whose levels were higher in the CT group than in the LP group. Heat colors indicate the contribution degree of metabolites that separate the CT and LP diet groups. Among the assigned metabolites, those whose NMR spectra with a correlation coefficient of 0.70 or more are annotated in the figure. Partial least square model quality was determined by Q^2 value, which represent model predictability, and the *P* value of cross-validated analysis of variance. Each value is shown in the figure.

3.8 References

[1] Hernández MJJ, Gómez MC, Morillas-Ruiz JM. Dietary factors associated with frailty in old adults: A review of nutritional interventions to prevent frailty development. Nutrients. 2019;11:102.

[2] Deutz NE, Bauer JM, Barazzoni R, Biolo G, Boirie Y, Bosy-Westphal A, et al. Protein intake and exercise for optimal muscle function with aging: recommendations from the ESPEN Expert Group. Clin Nutr. 2014;33:929–36.

[3] Landi F, Calvani R, Tosato M, Martone AM, Ortolani E, Savera G, et al. Protein intake and muscle health in old age: From biological plausibility to clinical evidence. Nutrients. 2016;8:295.

[4] Yamada R, Odamaki S, Araki M, Watanabe T, Matsuo K, Uchida K, et al. Dietary protein restriction increases hepatic leptin receptor mRNA and plasma soluble leptin receptor in male rodents. PLoS One. 2019;14:e0219603.

[5] Nishi H, Yamanaka D, Kamei H, Goda Y, Kumano M, Toyoshima Y, et al. Importance of serum amino acid profile for induction of hepatic steatosis under protein malnutrition. Sci Rep. 2018;8:5461.

[6] Taguchi Y, Toyoshima Y, Tokita R, Kato H, Takahashi S-I, Minami S. Triglyceride synthesis in hepatocytes isolated from rats fed a low-protein diet is enhanced independently of upregulation of insulin signaling. Biochem Biophys Res Commun. 2017;490,800–805.

[7] Toyoshima Y, Tokita R, Taguchi Y, Akiyama-Akanishi N, Takenaka A, Kato H, et al. Tissue-specific effects of protein malnutrition on insulin signaling pathway and lipid accumulation in growing rats. Endocr J. 2014;61:499–512.

[8] Zhang Z, Pereira SL, Luo M, Matheson EM. Evaluation of blood biomarkers associated with risk of malnutrition in older adults: A systematic review and meta-analysis. Nutrients.

2017;9:E829.

[9] Omran ML, Morley JE. Assessment of protein energy malnutrition in older persons, Part II: Laboratory evaluation. Nutrition. 2000;16:131–40.

[10] Gowda S, Desai PB, Kulkarni SS, Hull VV, Math AA, Vernekar SN. Markers of renal function tests. N Am J Med Sci. 2010;2:170–3.

[11] Bingham SA. Urine nitrogen as a biomarker for the validation of dietary protein intake. J Nutr. 2003;3:921S–4S.

[12] Marliss EB, Wei CN, Dietrich LL. The short-term effects of protein intake on 3-methylhistidine excretion. Am J Clin Nutr. 1979;32:1617–21.

[13] Stella C, Beckwith-Hall B, Cloarec O, Holmes E, Lindon JC, Powell J, et al. Susceptibility of human metabolic phenotypes to dietary modulation. J Proteome Res. 2006;5:2780–8.

[14] Abate-Shen C, Shen MM. Diagnostics: The prostate-cancer metabolome. Nature. 2009;457:799–800.

[15] Brindle JT, Antti H, Holmes E, Tranter G, Nicholson JK, Bethell HWL, et al. Rapid and noninvasive diagnosis of the presence and severity of coronary heart disease using ^{1}H-NMR-based metabonomics. Nat Med. 2002;8:1439–44.

[16] Jones DP, Park Y, Ziegler TR. Nutritional metabolomics: progress in addressing complexity in diet and health. Annu Rev Nutr. 2012;32:183–202.

[17] Moco S, Bino RJ, De Vos RCH, Vervoort J. Metabolomics technologies and metabolite identification. Trends Anal Chem. 2007;26:855–66.

[18] Wada Y, Izumi H, Shimizu T, Takeda Y. A more oxidized plasma albumin redox state and lower plasma HDL particle number reflect low-protein diet ingestion in adult rats. J Nutr. 2020;150:256–66.

[19] Wada Y, Komatsu Y, Izumi H, Shimizu T, Takeda Y, Kuwahata M. Increased ratio of non-

mercaptalbumin-1 among total plasma albumin demonstrates potential protein undernutrition in adult rats. Front Nutr. 2018;5:64.

[20] Xiong Z, Lang L, Gao X, Xiao W, Wang Z, Zhao L. An integrative urinary metabolomic study of the therapeutic effect of Guizhi Fuling capsule on primary dysmenorrheal rats based ^{1}H NMR and UPLC-MS. J Pharm Biomed Anal. 2019;164:750–8.

[21] Wang LF, Hu XJ, Peng RY, Wang SM, Gao YB, Dong J, et al. Application of ^{1}H-NMR-based metabolomics for detecting injury induced by long-term microwave exposure in Wistar rats' urine. Anal Bioanal Chem. 2012;404:69–78.

[22] Ebbels TM, Holmes E, Lindon JC, Nicholson JK. Evaluation of metabolic variation in normal rat strains from a statistical analysis of ^{1}H NMR spectra of urine. J Pharm Biomed Anal. 2004;36:823–33.

[23] MKS Umetrics. User guide to SIMCA Version 13. 2012.

[24] Triba MN, Le Moyec L, Amathieu R, Goossens C, Bouchemal N, Nahon P, et al. PLS/OPLS models in metabolomics: the impact of permutation of dataset rows on the K-fold cross-validation quality parameters. Mol BioSyst. 2015;11:13–9.

[25] de Graaf RA, Beharn KL. Quantitative ^{1}H NMR spectroscopy of blood plasma metabolites. Anal Chem. 2003;75:2100–4.

[26] Young VR, Marchini JS. Mechanisms and nutritional significance of metabolic responses to altered intakes of protein and amino acids, with reference to nutritional adaptation in humans. Am J Clin Nutr. 1990;51:270–89.

[27] Matsuda M, Shimomura I. Increased oxidative stress in obesity: implications for metabolic syndrome, diabetes, hypertension, dyslipidemia, atherosclerosis, and cancer. Obes Res Clin Pract. 2013;7:e330–41.

[28] Zheng L, Cardaci S, Jerby L, MacKenzie ED, Sciacovelli M, Johnson TI, et al. Fumarate induces

redox-dependent senescence by modifying glutathione metabolism. Nat Commun. 2015;6:6001.

[29] You YH, Quach T, Saito R, Pham J, Sharma K. Metabolomics reveals a key role for fumarate in mediating the effects of NADPH oxidase 4 in diabetic kidney disease. J Am Soc Nephrol. 2016;27:466–81.

[30] Hayes KC, Sturman JA. Taurine in metabolism. Annu Rev Nutr. 1981;1:401–25.

[31] Stipanuk MH. Role of the liver in regulation of body cysteine and taurine levels: a brief review. Neurochem Res. 2004;29:105–10.

[32] Marcinkiewicz J, Kontny E. Taurine and inflammatory diseases. Amino Acids. 2014;46:7–20.

[33] De Luca A, Pierno S, Camerino DC. Taurine: the appeal of a safe amino acid for skeletal muscle disorders. J Transl Med. 2015;13:243.

[34] Seidel U, Huebbe P, Rimbach G. Taurine: A regulator of cellular redox homeostasis and skeletal muscle function. Mol Nutr Food Res. 2019;63:e1800569.

[35] Terrill JR, Pinniger GJ, Graves JA, Grounds MD, Arthur PG. Increasing taurine intake and taurine synthesis improves skeletal muscle function in the mdx mouse model for Duchenne muscular dystrophy. J Physiol. 2016;594:3095–110.

[36] Khalil RM, Abdo WS, Saad A, Khedr EG. Muscle proteolytic system modulation through the effect of taurine on mice bearing muscular atrophy. Mol Cell Biochem. 2018;444:161–168.

[37] Takatani T, Takahashi K, Uozumi Y, Shikata E, Yamamoto Y, Ito T, et al. Taurine inhibits apoptosis by preventing formation of the Apaf-1/caspase-9 apoptosome. Am J Physiol Cell Physiol. 2004;287:C949–53.

[38] Murakami S. Role of taurine in the pathogenesis of obesity. Mol Nutr Food Res. 2015;59:1353–63.

[39] Wen C, Li F, Zhang L, Duan Y, Guo Q, Wang W, et al. Taurine is involved in energy metabolism in muscles, adipose tissue, and the liver. Mol Nutr Food Res. 2019;63:e1800536.

[40] Scicchitano BM, Sica G. The beneficial effects of taurine to counteract sarcopenia. Curr Protein Pept Sci. 2018;19:673–80.

[41] Yokogoshi H, Mochizuki H, Nanami K, Hida Y, Miyachi F, Oda H. Dietary taurine enhances cholesterol degradation and reduces serum and liver cholesterol concentrations in rats fed a high-cholesterol diet. J Nutr. 1999;129:1705–12.

[42] Pasiakos SM, Lieberman HR, Fulgoni VL. Higher-protein diets are associated with higher HDL cholesterol and lower BMI and waist circumference in US adults. J Nutr. 2015;145:605–14.

[43] Tang WH, Wang Z, Levison BS, Koeth RA, Britt EB, Fu X, et al. Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk. N Engl J Med. 2013;368:1575–84.

[44] Mayneris-Perxachs J, Bolick DT, Leng J, Medlock GL, Kolling GL, Papin JA, et al. Protein- and zinc-deficient diets modulate the murine microbiome and metabolic phenotype. Am J Clin Nutr. 2016;104:1253–62.

[45] Ma N, Tian Y, Wu Y, Ma X. Contributions of the interaction between dietary protein and gut microbiota to intestinal health. Curr Protein Pept Sci. 2017;18:795–808.

[46] Randrianarisoa E, Lehn-Stefan A, Wang X, Hoene M, Peter A, Heinzmann SS, et al. Relationship of serum trimethylamine N-oxide (TMAO) levels with early atherosclerosis in humans. Sci Rep. 2016;6:26745.

[47] Hu Y, Zhao Y, Yuan L, Yang X. Protective effects of tartary buckwheat flavonoids on high TMAO diet-induced vascular dysfunction and liver injury in mice. Food Funct. 2015;6:3359–72.

[48] Canyelles M, Tondo M, Cedó L, Farràs M, Escolà-Gil JC, Blanco-Vaca F. Trimethylamine N-oxide: A link among diet, gut microbiota, gene regulation of liver and intestine cholesterol homeostasis and HDL function. Int J Mol Sci. 2018;19:E3228.

[49] Cho CE, Taesuwan S, Malysheva OV, Bender E, Tulchinsky NF, Yan J, et al. Trimethylamine-

N-oxide (TMAO) response to animal source foods varies among healthy young men and is influenced by their gut microbiota composition: A randomized controlled trial. Mol Nutr Food Res. 2017;61:1600324.

[50] Koeth RA, Wang Z, Levison BS, Buffa JA, Org E, Sheehy BT, et al. Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis. Nat Med. 2013;19:576–85.

Chapter 4 Association of milk components with fecal microbiome and metabolites in a mother-infant dyad

4.1 Abstract

The gut microbiome of breastfed infants changes during lactation, and breast milk contains bioactive components that affect it. However, the relationships between these bioactive components and the infant gut microbiome and fecal metabolites throughout lactation remains obscure. Breast milk and infant feces were collected for 5 months after delivery, from a mother and an exclusively breastfed infant, respectively. Composition of the fecal microbiome was determined with 16S rRNA sequencing. Low-molecular-weight metabolites, including human milk oligosaccharides (HMOs), and antibacterial proteins were measured in feces and milk using ^{1}H NMR metabolomics and enzyme-linked immunosorbent assays. The association of milk bioactive components with the infant gut microbiome and fecal metabolites was determined with Python clustering and correlation analyses. The HMOs in milk did not fluctuate throughout the lactation period. However, they began to disappear in infant feces at the beginning of month 4. Notably, at this time-point, a bifidobacterium species switching (from *breve* to *longum* subsp. *infantis*) occurred, accompanied by fluctuations in several metabolites including acetate and butyrate in infant feces. Milk bioactive components, such as HMOs, might play different roles in the exclusively breastfed infants depending on the lactation period.

4.2 Introduction

The gut microbiome plays an important role in the normal growth and development of infants, and it is less diversely and biased toward pathogens in premature infants with necrotizing enterocolitis [1]. Perturbation of the gut microbiome could enhance the inflammatory response and

lead to the development of necrotizing enterocolitis in infants [2]. In contrast, the gut microbiome of healthy infants delivered at full term is dominated by the genus *Bifidobacterium*, which confers benefits such as the prevention of bacterial infections and the maturation of immune functions, upon infants [3]. The infant gut microbiome changes during lactation, and the ingestion of breast milk influences the composition of the infant microbiome [4]. However, it is not fully understood which breast milk components are associated with changes in the gut microbiome of exclusively breastfed infants throughout lactation.

Breastfeeding is recommended by the World Health Organization and numerous pediatric associations, as the best way of feeding infants during at least six months of life; breastfeeding is extremely beneficial for the health and development of infants [5]. In fact, breast milk, a dynamic biofluid, is recognized as the only food able to meet all of the newborn's nutritional needs. Moreover, breast milk includes various bioactive compounds, such as immunoglobulins (Igs), cytokines, chemokines, hormones, and growth factors, important to prevent infection [6–8], and to promote the development of organs and systems (e.g., the immune system) [9, 10]. Additionally, low-molecular-weight metabolites up to a molecular size of approximately 1500 Da are also contained in breast milk. Recent studies have reported that breast milk components are affected by several factors, including diet, genetic background, lifestyle, and body mass index of the mother, among which the lactation stage has been reported to be the most influential [11–13].

Some of these fluctuating breast milk components are bioactive and affect the gut microbiome of infants. Human milk oligosaccharides are unconjugated glycans with a lactose core varying in chain length from 3 to 15 carbohydrates, and are assimilated by specific bacteria [14]. The most abundant HMOs in breast milk are 2'-fucosyllactose (2'-FL), 3-FL, 3'-sialyllactose (3'-SL), and 6'-SL. The HMOs might play essential roles in supporting the bifidobacteria-predominant gut microbiome in breastfed infants [15]. Furthermore, breast milk contains antibacterial proteins such

as IgA, lactoferrin (LF), and lysozyme (LYZ), which can contribute to establishing the gut microbiome of infants fed with breast milk [16–18]. The gut microbiome modified by milk bioactive components confers various benefits on infants via its metabolites. In fact, a study using an animal model has reported that bifidobacteria could inhibit *Escherichia coli* infection via the production of acetate and the improvement of the intestinal barrier, due to the use of indigestible oligosaccharides [19]. In addition, short-chain fatty acids (SCFAs) produced by the gut microbiome have been reported to increase resistance to inflammation and promote sympathetic nervous system maturation [20, 21]. However, the relationship between milk bioactive components, such as HMOs and antibacterial proteins, and the gut microbiome and fecal metabolites of breastfed infants remains unclear.

This study aimed to clarify whether the gut microbiome and fecal metabolites are associated with milk bioactive components during lactation in an exclusively breastfed infant. We collected breast milk and the resultant infant feces during the lactation period of an exclusively breastfed infant until the 5th month after delivery once every few days. We sequenced 16S rRNA from infant fecal samples to determine the microbiome composition at the species level, and measured low-molecular-weight metabolites, including HMOs, and antibacterial proteins, such as IgA, LF, and LYZ, in infant feces and breast milk, using ^{1}H NMR metabolomics and enzyme-linked immunosorbent assays (ELISA). The data were standardized by Z-scoring and those with similar properties were categorized into the same group via clustering analysis using the Python software. Factors that positively or negatively correlated with the gut microbiome were extracted from categorized components using Spearman rank correlation coefficients. We believe that establishing interactions between milk bioactive components and fecal metabolites in the gut microbiome will facilitate better understanding of how breastfeeding affects infant development, and the design of nutritional and functional foods for infants at various stages of lactation.

4.3 Materials and methods

4.3.1 Subject and sample collection

The institutional review board at the Japan Conference of Clinical Research approved this study (approval number: BONYU-01), which proceeded in accordance with the latest Declaration of Helsinki in 2013. The healthy mother of a newborn infant provided written, informed consent to the participation of herself and her infant in all procedures associated with the study. Breast milk samples were collected every few days until the 5th month after delivery. Then, whenever the infant breastfed, the first excreted feces was collected. Of note, the infant was exclusively breastfed during the experimental period. The collected samples were temporarily stored at −20°C until a month and were subsequently transferred to −80°C storage conditions until analysis.

4.3.2 DNA extraction

Total DNA was extracted from the feces samples, as described previously [22]. Briefly, 200 µl of fecal sample in GuSCN solution was lysed with glass beads (300 mg, 0.1 mm diameter) and 300 µl of lysis buffer (No. 10 buffer, Kurabo Industries Ltd., Osaka, Japan) by FastPrep-245G (MP Biomedicals LLC, Santa Ana, CA) at a 5 power level for 45 sec with 5-min cooling intervals on ice. After centrifugation at 12,000 rpm for 5 min, DNA was extracted from 200 µl of the supernatant using GENE PREP STAR PI-480 (Kurabo Industries Ltd.) following the manufacturer's protocol.

4.3.3 Microbiome analysis

After DNA extraction, PCR amplification and DNA sequencing of the V3-V4 region of the bacterial 16S rRNA gene was performed in an Illumina MiSeq instrument (Illumina, San Diego, CA) as described previously [23].

4.3.4 Measurement of macronutrients in breast milk

Fat, protein, carbohydrate, and total solid contents in homogenized human milk samples (3 mL) were measured using the Human Milk Analyzer (MIRIS AB, Uppsala, Sweden) according to the manufacturer's instructions. All samples were homogenized using the MIRIS sonicator (MIRIS AB) and were maintained at 40°C prior to measurement.

4.3.5 Measurement of antibacterial proteins in breast milk and infant feces

The total IgA, LF, and LYZ in breast milk and infant feces were measured using a human ELISA kit (Abcam, Cambridge, UK) according to manufacturer's protocols. Extracts of infant feces were prepared for ELISA using the same method as the preparation for NMR. All samples were diluted within the measurement range with sample buffer included in the ELISA kit and were assayed in duplicate.

4.3.6 Sample preparation for NMR analysis

Breast milk samples (1 mL) were centrifuged at 3000 × g for 30 min at 4°C, and the aqueous layer was carefully removed and filtered through a 5-kDa cutoff filter (UltrafreeMC-PLHCC filter; HMT, Tsuruoka, Japan) at 9000 × g at 4°C to remove lipids and proteins. The filtrate (540 μL) was mixed with 60 μL 100% D_2O (Cambridge Isotope Laboratories, Inc., Andover, MA) containing 500 mM NaP, 5 mM 3-(trimethylsilyl)propionic acid-d4 sodium salt (TSP), and 0.04% (w/v) NaN_3, and 550 μL of the mixture was transferred to a 5-mm NMR tube (Shigemi, Hachioji, Japan).

Fecal extracts were prepared for NMR analysis as follows: fecal samples were mixed with MilliQ water (Millipore, Billerica, MA) containing 50 mM NaP, 0.5 mM TSP, 0.004% (w/v) NaN_3, and 10% D_2O (per 100 mg of feces). This mixture was homogenized at 1800 rpm for 15 min at 4°C using a

tube mixer (EYELA CM-1000; Wakenyaku, Kyoto, Japan). The shaken samples were then centrifuged at 15,000 × g for 10 min at 4°C. The supernatant pH was adjusted to 7.4 ± 0.1 by adding small amounts NaOH or HCl to minimize pH-based peak movement. The 550 µL aliquot was transferred to a 5-mm NMR tube.

4.3.7 NMR spectra acquisition

^{1}H NMR spectra were recorded using a Bruker 600 MHz AVANCE spectrometer (Bruker, Billerica, MA) with equipped with cryoprobe at a proton frequency of 600.13 MHz. The sample temperature controlled at 298 K. As per the analysis conditions reported in our previous report [24], ^{1}H NMR spectra were recorded using a water-suppressed standard one-dimensional NOESY1D pulse sequence. Each spectrum consisted of 32,768 data points with a spectral width of 12 ppm. The acquisition time was 2.28 s, and the number of scans was 128. A water-suppressed pulse sequence was used to reduce the residual water signal at the water frequency with a recycle delay [D1 (Bruker notated)] of 2.72 s and a mixing time [D8 (Bruker notated)] of 0.10 s. A 90° pulse length was automatically calculated in the analysis of each sample.

4.3.8 NMR spectral data processing

All raw spectra were manually corrected for phase and baseline distortions against TSP resonance at δ = 0.0 ppm using the Delta 5.0.4 software (JEOL, Tokyo, Japan) and then analyzed. Spectral binning, multivariate analysis, or direct metabolite relative quantification was performed. The spectra were normalized to the peak area value of the internal standard TSP using the NMR Suite 7.5 Processor (Chenomx, Edmonton, Canada), and the normalized spectral data were further processed. Briefly, in the first round of processing, the 0.0–10.0 ppm chemical shift region was integrated into regions with a width of 0.04 ppm, while the spectral regions related to residual water area (4.68–5.08

ppm) were removed from the multivariate analysis to eliminate the effects of imperfect water suppression. In the second round of processing, metabolite assignment and quantification were determined using the 600 MHz library from Chenomx NMR Suite as reference.

4.3.9 Statistical analyses

The spectral data matrix was obtained using the SIMCA-P 14.0 software (Umetrics, Umeå, Sweden). Concentrations of breast milk components and binned NMR spectra in infant feces during each lactation period were assessed using principal component analysis (PCA). Score plots were obtained to visualize the clustering pattern of the samples in the context of two principal components (PC1 and PC2), with each point denoting an individual sample. For microbiome data, weighted UniFrac distance metrics analysis was performed using operational taxonomic units (OTUs) for each sample, and PCA was conducted according to the matrix of distance using the SIMCA-P 14.0 software. The α-diversity of the microbiome in feces at various lactation points is represented by the Shannon index.

A clustering analysis using Python version 3.7.3 was performed based on a previous study, to verify the similarity between milk bioactive components, the infant gut microbiome, and fecal metabolites [25]. After logarithmizing the numerical data, standardization was performed via Z-scoring with the mean and standard deviation set to 0 and 1, respectively. Furthermore, milk bioactive components, infant gut microbiome, and fecal metabolites showing similar changes over the lactation period were categorized into the same group using the scientific computing package SciPy and the visualization package Seaborne. The association of milk bioactive components with the infant gut microbiome and fecal metabolites was explored using Spearman rank correlation coefficient analyses and heatmaps generated using SciPy and Seaborne.

4.4 Results

4.4.1 Profiling of microbiome and metabolites in infant feces

The weighted UniFrac distances based on the OTUs for the microbiome in feces of breastfed infant during each lactation period were visualized using PCA (**Figure 4-1a**). The score plots in months 1–2 and 3–5 after delivery formed separate clusters, respectively. Changes in microbiome α-diversity in infant feces at each lactation point are represented by the Shannon index (**Figure 4-1b**). As the lactation period progressed, the diversity increased, especially after month 3. The transitions of 15 bacterial species with an average abundance of 0.5% or more during the lactation period was shown in **Figure 4-1c**. Furthermore, microbiome composition of breastfed infant feces at the all 40 species-level at each lactation point is shown in **Supplementary Table 4-S1**.

Representative ^{1}H NMR spectra of infant feces samples at months 1 and 5 after delivery are shown in **Supplementary Figure 4-S1**. Low-molecular-weight compounds, such as monosaccharides, disaccharides, oligosaccharides, SCFAs, amino acids, and their metabolites were identified, and their chemical shifts are shown in **Supplementary Table 4-S2**. Changes in the feces NMR spectra during each lactation period were visualized using PCA (**Figure 4-1d**). The score plots derived from the feces NMR spectra in months 1–2 and 3–5 after delivery formed separate clusters. The relative concentrations of infant feces metabolites in months 1–5 after delivery are shown in **Supplementary Table 4-S3**. In consideration of the direct supply of antibacterial proteins from breast milk, the concentrations of IgA, LF, and LYZ in infant feces were measured in months 1–5 after delivery (**Supplementary Table 4-S3**).

4.4.2 Profiling of components in breast milk

The representative ^{1}H NMR spectra of breast milk samples at months 1 and 5 after delivery are shown in **Supplementary Figure 4-S2**. Low-molecular-weight compounds such as monosaccharides,

disaccharides, oligosaccharides, SCFAs, amino acids, and their metabolites were identified, and their chemical shifts are shown in **Supplementary Table 4-S4**. **Supplementary Table 4-S5** shows the concentrations of metabolites, macronutrients (fats, proteins, carbohydrates, and total solids), and antibacterial proteins (IgA, LF, and LYZ) in breast milk at 1–5 months postpartum. Changes in breast milk components during each lactation period were visualized using PCA (**Figure 4-2a**). Furthermore, **Figure 4-2b** shows fluctuations in the levels of bioactive HMOs and IgA, LF, and LYZ in breast milk. Although the score plots derived from the milk components in months 1–2 and 3–5 after delivery formed moderately separate clusters, these bioactive components did not characteristically vary throughout lactation as shown **Figure 4-2b**.

4.4.3 Clustering analysis of milk bioactive components, microbiome, and metabolites in infant feces

When milk bioactive compounds act on the infant gut microbiome via breast milk, bacterial-derived metabolites in infant feces should also fluctuate in association with bacterial changes. Therefore, association of milk bioactive components with the microbiome (at the species level; average abundance >0.5%), and fecal metabolites of the infant was assessed with clustering analysis using the Python software. Three major clusters were defined, showing similar fluctuations over time (**Figure 4-3**): Values for cluster A were high and low, whereas those for cluster B were low and high during early and late-lactation, respectively, and a trend for cluster C was not evident throughout lactation. Among compounds and bacterial species classified into the same cluster, those with a particularly high similarity were visually selected. **Figure 4-4a** shows fluctuations in the levels of four HMOs, antibacterial proteins, urea, and hypoxanthine, and **Figure 4-4b** shows transitions in the abundance of *B. breve* and *Veillonella ratti* in infant feces categorized into cluster A. **Figure 4-5a** shows fluctuations in the levels of acetate, butyrate, 5-aminopentanoate, and fucose in infant feces categorized into cluster B. **Figure 4-5b** shows transitions in the abundance of *B. longum* subsp.

infantis, *E. coli*, *Clostridium neonatale*, and *Clostridioides mangenotii* in infant feces categorized into cluster B.

4.4.4 Spearman correlation analysis of milk bioactive components, infant gut microbiome, and fecal metabolites

We used Spearman rank correlation coefficient analysis to explore associations between milk bioactive HMOs and antibacterial proteins in breast milk and those in infant feces that were clearly distinguished by PCA, at 1–2 and 3–5 months postpartum, and visualized them using heatmaps (**Figure 4-6a**). The results showed that 3-FL, 3′-SL, and 6′-SL did not correlate at 1–2 months (ρ = +0.00 to -0.25), but positively correlated at 3–5 months (ρ = +0.40 to +0.74). A weak positive correlation for LF was observed at 1–2 months (ρ = +0.49), which subsequently disappeared at 3–5 months postpartum (ρ = -0.12). We also used Spearman rank correlation coefficients to extract factors with significant positive or negative correlations from milk bioactive components, bacterial species, and metabolites with high similarity in clusters A and B from infant feces. The results are shown, in heatmaps, at 1–2 and 3–5 months postpartum, respectively (**Figure 4-6b and 4-6c**). Few correlations were identified between milk bioactive components and the microbiome and fecal metabolites at 1–2 months postpartum. *Bifidobacterium breve* correlated positively with 3-FL and 3′-SL (ρ = +0.79 and +0.76), and negatively with IgA and LF (ρ = -0.72 and -0.92). Correlations were much closer at 3–5 than at 1–2 months. *Bifidobacterium breve* correlated significantly and negatively with *B. longum* subsp. *infantis* and *E. coli*, respectively, in infant feces in months 3–5 (ρ = -0.97 and -0.88). *Bifidobacterium longum* subsp. *infantis* also significantly correlated with *E. coli* (ρ = +0.85). Correlations (ρ = +0.82 to +0.98) were significantly positive among four HMOs in infant feces. These HMOs were significantly and positively correlated (ρ = +0.69 to +0.94) with *B. breve* in cluster A in infant feces. *Bifidobacterium longum* subsp. *infantis* and *E. coli* in cluster B in infant feces

significantly and negatively correlated with HMOs (ρ = -0.65 to -0.96). No bacterial species significantly correlated with the antibacterial proteins IgA, LF, and LYZ, hypoxanthine, or fucose. Urea levels correlated positively with *B. breve* (ρ = 0.78) and negatively with *B. longum* subsp. *infantis* and *E. coli* (ρ = -0.81 and -0.70). Acetate and butyrate levels significantly and positively correlated with *C. mangenotii* (ρ = +0.72 and +0.80), and 5-aminopentanoate levels correlated positively with *B. longum* subsp. *infantis* and *C. mangenotii* (ρ = +0.72 and +0.70). Some metabolites, such as amino acids and other bacterial species, in infant feces were classified into cluster C, which had no constant trend throughout lactation.

4.5 Discussion

Breast milk is an ideal nutritional source for infants and contains various bioactive components such as HMOs and antibacterial proteins. However, relationships between milk bioactive components and the infant gut microbiome and fecal metabolites throughout lactation remain unclear. Our PCA findings revealed different clustering profiles in the infant gut microbiome and fecal metabolites between 1–2 and 3–5 months postpartum, and the diversity of the gut microbiome also increased 3 months onwards. Nevertheless, breast milk components did not fluctuate as lactation progressed. This indicated that the gut microbiome and fecal metabolites of the breastfed infant fluctuated after 3 months postpartum, even though milk compounds did not change. These changes in the infant intestinal tract might translate into optimization for the efficient digestion and absorption of breast milk. In fact, it has been reported that after the 3rd month of life, the functional maturation of the gastrointestinal tract progresses rapidly, and the fecal properties change dramatically [26].

Several studies have shown that an infant's gut microbiome could change significantly by the first year of life [4, 27], and this change might be an intestinal adaptation allowing the use of milk components such as undigested HMOs. As previously reported, 2′-FL in breast milk reflects maternal

secretor status [28, 29], because their production is a consequence of the fucosyltransferase encoded by the *FUT2* gene. Similarly, 3-FL in breast milk is contributed to the fucosyltransferase encoded by the *FUT3* gene. Since 2′-FL and 3-FL were detected in the breast milk in this study, this subject would be "secretory" type. It has been reported that the abundance of bifidobacteria in infants fed with secretory breast milk is higher than that in infants fed with non-secretory breast milk [14]. It has been reported that 2′-FL contained in breast milk promotes the growth of bifidobacteria in infants [30]; this may be the reason the infant in this study showed a high abundance of bifidobacteria.

Spearman correlation analyses of milk bioactive compounds in breast milk and infant feces revealed different tendencies between 1–2 and 3–5 months postpartum. The HMOs 3-FL, 3′-SL and 6′-SL positively correlated only at 3–5 months postpartum. This change in correlation suggested that the amounts of HMOs delivered into infant feces via breast milk changed between early and late lactation. The structural characteristics and prebiotic effects of HMOs assimilated by intestinal bacteria protect against pathogens [28]. The effects of HMOs might change as infants grow. A weak correlation for LF between breast milk and infant feces at 1–2 months, disappeared by 3–5 months postpartum. Digested LF exerts stronger antibacterial activities than the full-length intact forms [31]. The disappearance of the LF correlation between breast milk and infant feces might mean that the digestion of intact LF started at 3 months postpartum, as the digestive capacity of the infant developed. Furthermore, LF in infant feces significantly and negatively correlated with *B. breve* at 1–2 months postpartum. Some peptides isolated from human milk exert strong bifidogenic effects on *B. bifidum*, *B. breve*, and *B. longum* and are resistant to digestive enzymes [32]. Thus, digested LF in the gastrointestinal tract might promote the growth of *B. breve* at 1–2 months despite the immature digestive function of the infant. Neither IgA nor LYZ significantly correlated those between in breast milk and infant feces. Thus, our results suggest that antibacterial proteins in breast milk are not necessarily reflected in the content of the influx into the digestive tract of the infant.

Since IgA is detectable in the saliva and feces of infants aged 3–5 months [33], the fecal IgA herein might have been partially synthesized by the infant. Similarly to LF, the LYZ digested with pepsin exerts stronger antibacterial activities than the intact forms [34]. The weak correlation between LYZ in breast milk and infant feces might have been affected by degradation due to infant digestion. In addition, antibacterial proteins (IgA, LF, and LYZ) in infant feces did not significantly correlate with any bacterial species detected in the present study. However, antibacterial proteins are supplied via breast milk to the digestive tract, where they modify the infant gut microbiome [35–37]. Further studies are needed to clarify the effects of antibacterial proteins in breast milk on the infant gut microbiome.

Interestingly, the concentrations of HMOs such as 2′-FL, 3-FL, 3′-SL, and 6′-SL in breast milk were almost constant, while in infant feces these HMOs were no longer detected after the beginning of month 4. Therefore, our results suggest that the infant's gut microbiome began to metabolize these HMOs from month 4 onwards. Importantly, Spearman analysis of infant feces at 3–5 months postpartum revealed significant positive correlations between HMOs and *B. breve*. In contrast, HMOs and *B. longum* subsp. *infantis* were significantly and negatively correlated. Spearman analysis also revealed a significantly negative correlation between *B. breve* and *B. longum* subsp. *infantis* in infant feces in 3–5 months. Several *in vitro* studies have reported that 2′-FL, 3-FL, 3′-SL, and 6′-SL are not metabolized by *Bifidobacterium breve*, but are metabolized by *Bifidobacterium longum* subsp. *infantis* due to the high activity of glycosyl hydrolases and the structural differences in ATP-binding cassette transporters [38, 39]. Therefore, it is inferred that the dramatic decrease in HMOs in the infant feces was probably because of the "species switching" from *Bifidobacterium breve* to *Bifidobacterium longum* subsp. *infantis*. *Enterococcus faecalis*, categorized in cluster C, which did not evidence a constant trend throughout lactation, remained abundant during the switch from *B. breve* to *B. longum* subsp. *infantis*. Exopolysaccharides produced by *E. faecalis* promote the growth

of *B. longum* rather than *B. breve* [40]. Thus, the growth of *B. longum* could explain the switch from *B. breve* to *B. longum* subsp. *infantis* found herein, which might have been influenced by *E. faecalis*, which modulates inflammation in the gastrointestinal tract during the first few years of life, and affects the development of intestinal immunity associated with allergies [41, 42]. This indicates that the increase in *E. faecalis* at 2–4 months directly contributed to the development of the gastrointestinal tract of the infant studied herein.

Escherichia coli were relatively significantly and negatively correlated with HMOs in infant feces. Fucosylated and sialylated glycans in breast milk can bind to *E. coli* and suppress infection [43, 44]. Therefore, *E. coli* abundance in infant feces might have increased due to the depletion of HMO in infant feces. *Escherichia coli* produces acetate via acid fermentation under anaerobic conditions [45]. Thus, elevated acetate levels in the infant feces might have been partly due to *E, coli*, even though the correlation coefficient between acetate and *E. coli* was <0.70 at 3–5 months postpartum. Furthermore, supplementing preterm infants with bifidobacteria increases fecal acetate levels via metabolized HMOs, and consequently promotes the defense of epithelial and mucosal dendritic cells [46]. The metabolism of HMOs by *B. longum* subsp. *infantis* might have partly contributed to the elevated acetate in infant feces later during lactation in the present study. Fecal butyrate significantly and positively correlated with *C. mangenotii* at 3–5 months. The results of PICRUSt analysis and an anaerobic culture system have shown that *C. mangenotii* is a potential butyrate producer [47]. Therefore, the elevated butyrate levels in the infant feces are thought to be a result of increased *C. mangenotii* abundance during late lactation. Butyrate in the intestinal tract confers benefits, such as trophic and anti-inflammatory effects on epithelial cells, upon hosts [48, 49]. Butyrate contributes to maturation of the immune system by modulating the differentiation of regulatory T cells in the large intestine [50]. Therefore, elevated butyrate levels during late lactation could help immune development and inflammatory modulation in infants. Breast milk contains

acetate and butyrate, but they are rapidly absorbed in the upper gastrointestinal tract [51, 52]. These metabolites in infant feces are probably produced via fermentation by the gut microbiome.

Urea and hypoxanthine are metabolites that are respectively generated by amino and nucleic acid metabolism; these were categorized into cluster A, which was respectively abundant and scant during early and late-lactation. However, as the amounts of amino acids (or proteins) and nucleic acids required for a developing infant significantly change during lactation [53], the relationship between the gut microbiome and these nutritional components remains controversial. The genus *Clostridium* produces 5-aminopentanoate from proline [54]. Therefore, the increase in 5-aminopentanoate found herein might be due to an increase in the abundance of *C. neonatale* despite the absence of a significant correlation with 5-aminopentanoate. As far as we can ascertain, the function of 5-aminopentanoate in infants is unknown. However, 5-aminopentanoate can be a biomarker of cerebral ischemia [55], and it might therefore play an important role in the development of brain function in infants. Spearman analysis did not identify significant correlations between fecal fucose and any factors. Since fucose is a component of fucosylated HMOs [15], it might have been produced in infant feces via the metabolism of 2'-FL and 3-FL. However, the physiological significance of fucose requires further investigation.

The major limitation of this pilot study is that we analyzed only one infant and mother. Therefore, further investigation of a large sample is required to comprehensively determine the effects of milk compounds in the context of breastfeeding. We did not include colostrum and transitional milk secreted during the first month postpartum in our analysis, as they have different nutritional and functional components from mature breast milk [56]. This exclusion might explain the limited fluctuations in breast milk components throughout lactation in the present study.

In conclusion, microbiome analysis, NMR metabolomics, and Python's clustering and correlation analyses showed that the gut microbiome and fecal metabolites of the infant were

associated with milk bioactive components during a certain lactation period**.** In particular, HMOs in the breast milk did not fluctuate throughout the lactation period, but disappeared from the infant feces after the 4th lactation month. Moreover, at the same time a species switching from *B. breve* to *B. longum* subsp. *infantis*, *E. coli* and *Clostridioides mangenotii* occurred, accompanied by an increase in the levels of metabolites such as acetate and butyrate, in the infant feces. Importantly, in line with previous reports, our data suggest that the changes in metabolites in the infant feces might be linked to benefits such as maturation of immune function and protection against infection. Therefore, milk bioactive components, such as HMOs, might play different roles in the exclusively breastfed infants depending on the lactation period. In particular, the increase in metabolites such as acetate and butyrate in the late-lactation period might reflect the replacement from developmental benefits to the infant via breastfeeding to benefits by metabolism of the infant or its own gut microbiome.

To develop a mechanistic understanding of the complex interactions of microbial species, more detailed studies are required, such as co-culturing of the bacterial species that fluctuated in this study. However, the present findings could contribute to understanding of the relationship between milk bioactive components and the infant gut microbiome, and to the development of supplementary nutritional foods for breastfed infants during lactation.

4.6 Figures

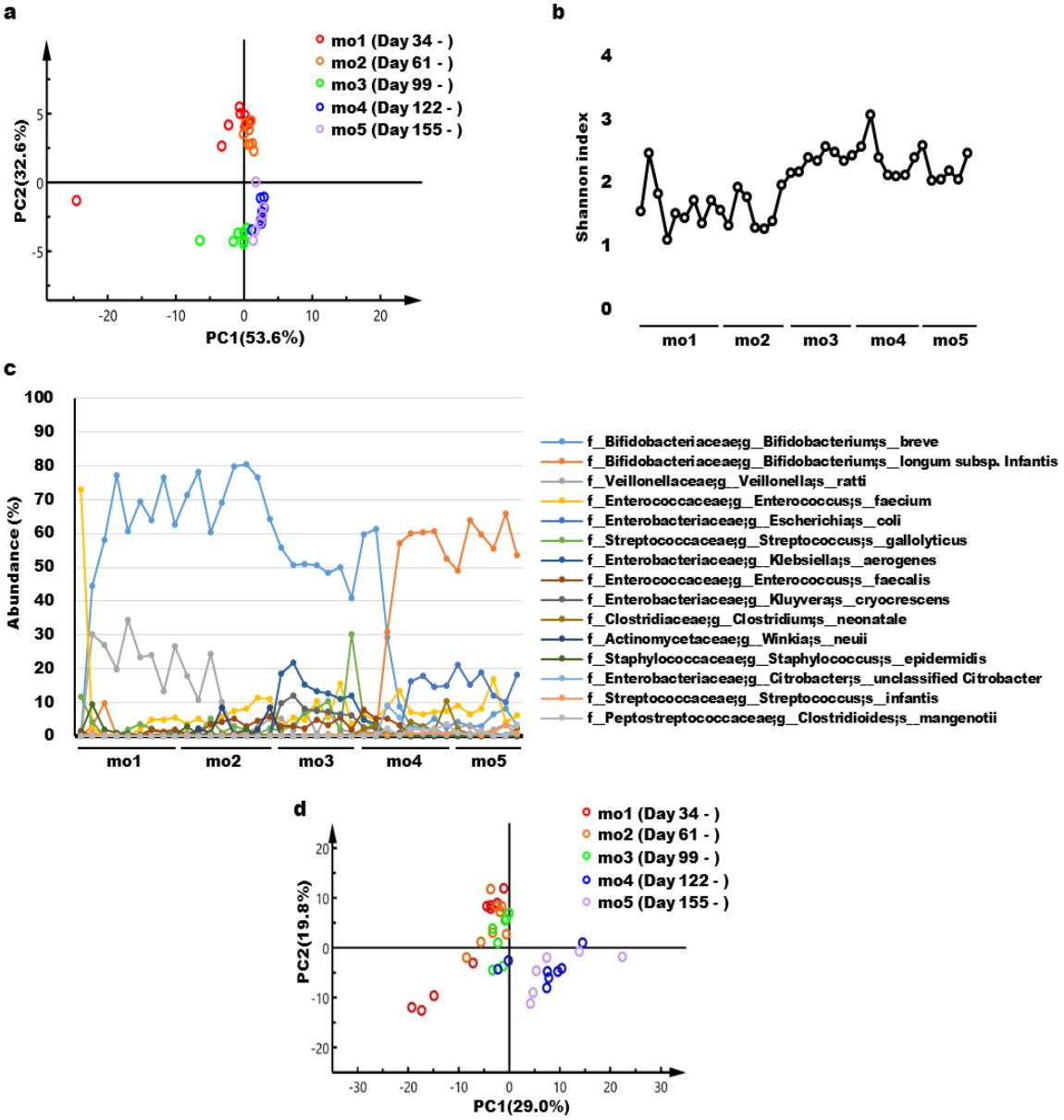


Figure 4-1. (a) Principal component analysis (PCA) score plots based on the weighted UniFrac distance metric of the microbiome in feces samples collected 1–5 months after delivery. Each symbol represents one individual feces sample. (b) Changes in the microbiome α-diversity, as assessed from the Shannon index, in feces samples collected 1–5 months after delivery. (c) Transitions of 15 bacterial species with an average abundance of 0.5% or more during the lactation period. (d) PCA score plots derived from the NMR spectra of feces samples collected 1–5 months after delivery. Each symbol represents one individual feces sample. mo, month.

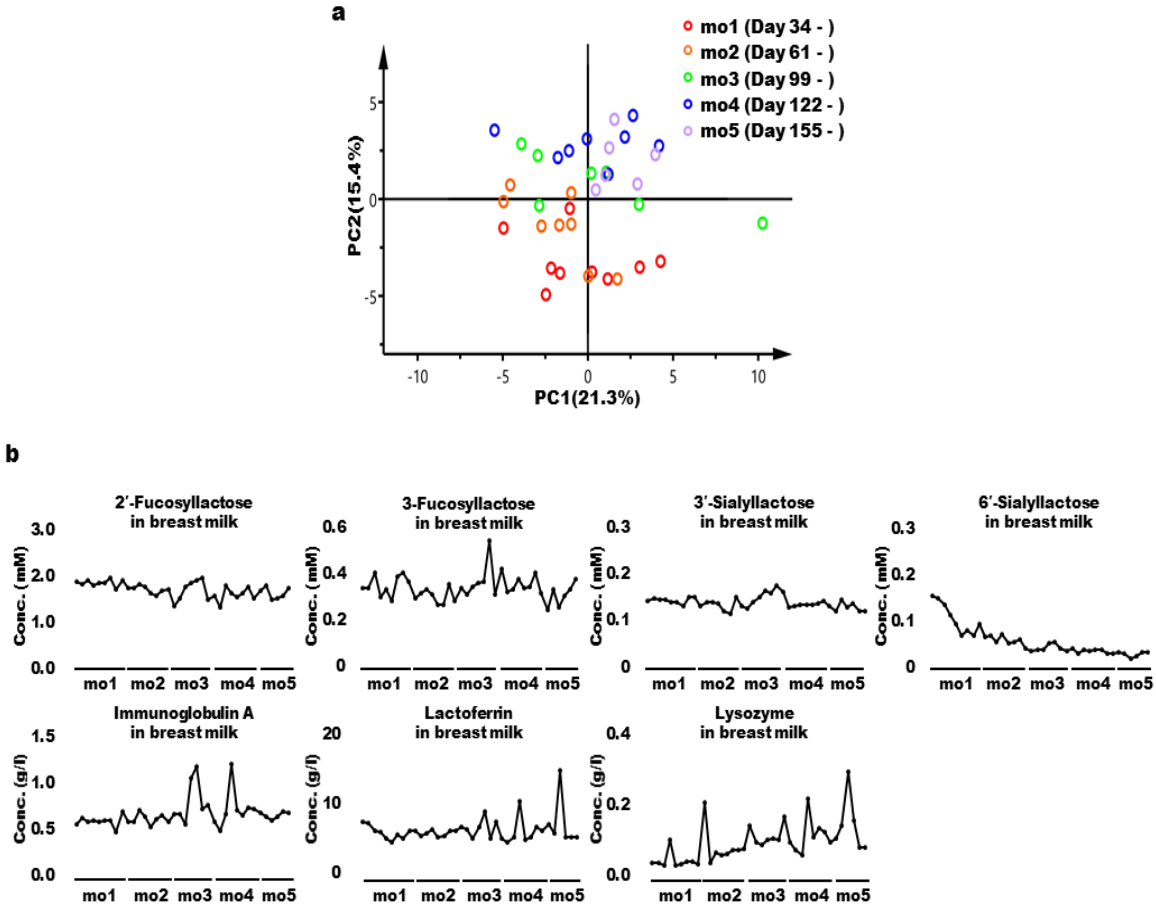


Figure 4-2. Principal component analysis and changes in the bioactive components of breast milk. (a) Principal component analysis (PCA) score plots derived from components such as low-molecular-weight metabolites, major nutrients, and antibacterial proteins in breast milk collected 1–5 months after delivery. Each symbol represents an individual breast milk sample. (b) Changes in levels of the HMOs IgA, LF, and LYZ in breast milk. mo, month.

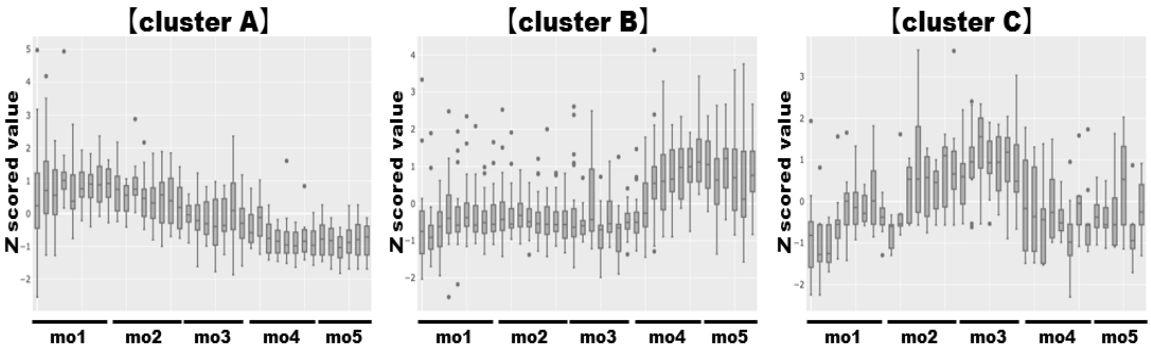


Figure 4-3. Clustering analysis of milk bioactive components, the microbiome at the species level, and the metabolites in infant feces. Those with similar characteristics (similar fluctuation patterns) were categorized into three clusters: cluster A, cluster B, and cluster C. "Cluster A" showed high values in the early-lactation period and low values in the late-lactation period; "cluster B" showed low values in the early-lactation period and high values in the late-lactation period; and "cluster C" did not show a constant trend throughout the lactation period. Categorized factors: (cluster A) 2'-Fucosyllactose, 3-Fucosyllactose, 3'-Sialyllactose, 6'-Sialyllactose, Formate, Hypoxanthine, Immunoglobulin A, Lactoferrin, Lactose, Lysozyme, Trimethylamine N-oxide, Tryptophan, Urea, *Bifidobacterium breve*, *Staphylococcus epidermidis*, *Veillonella ratti*. (cluster B) 5-Aminopentanoate, Acetate, Alanine, Aspartate, Butyrate, Cadaverine, Carnitine, Choline, Creatine, Fucose, Fumarate, Glucose, Lactate, o-Phosphocholine, Propionate, Pyruvate, Succinate, Taurine, Threonine, Tyrosine, *Bifidobacterium longum* subsp. *infantis*, *Citrobacter* unclassified *Citrobacter*, *Clostridioides mangenotii*, *Clostridium neonatale*, *Enterococcus faecium*, *Escherichia coli*, *Streptococcus infantis*. (cluster C) Glutamate, Glycine, Isoleucine, Leucine, Methionine, Phenylalanine, Valine, *Enterococcus faecalis*, *Klebsiella aerogenes*, *Kluyvera cryocrescens*, *Streptococcus gallolyticus*, *Winkia neuii*. mo, month.

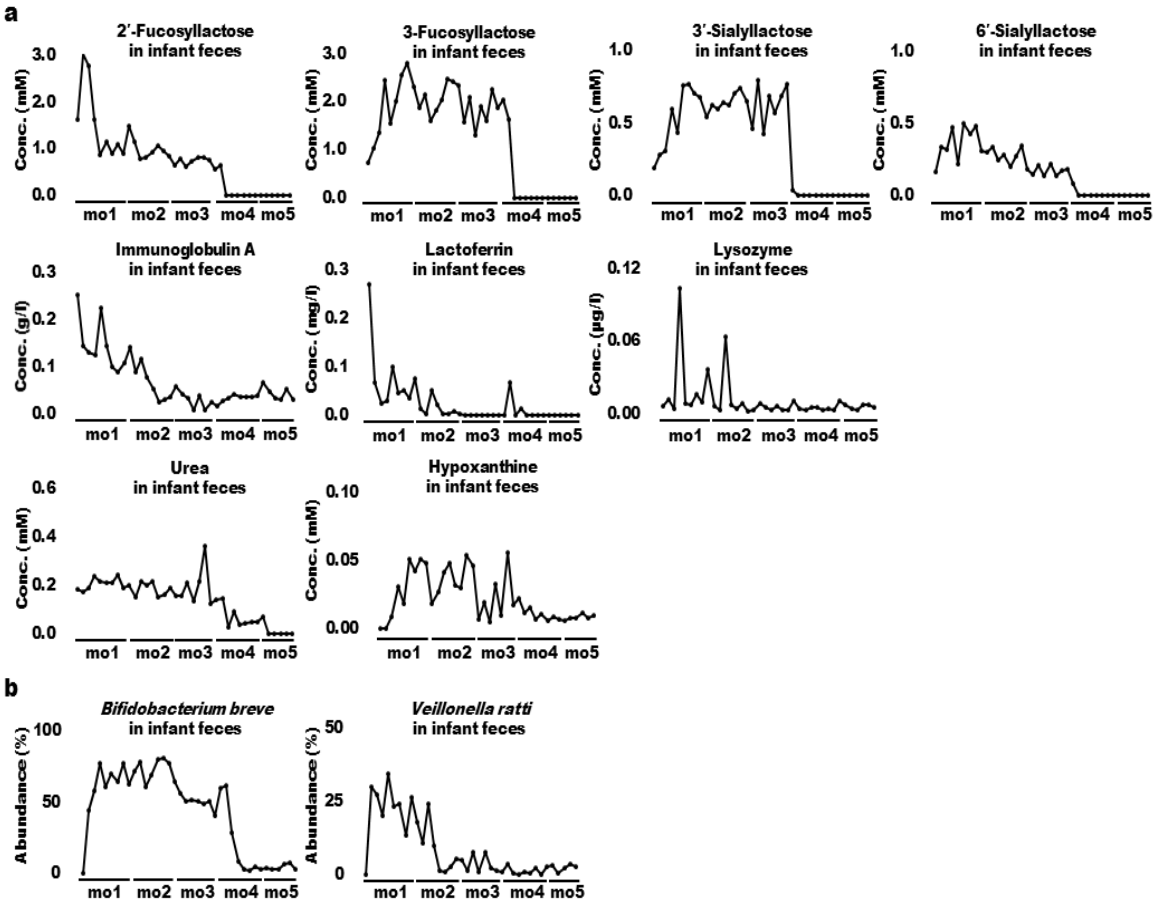


Figure 4-4. Changes in milk bioactive components, microbiome, and metabolites in infant feces with high similarity in cluster A. (a) Human milk oligosaccharides (2′-fucosyllactose, 3-fucosyllactose, 3′-sialyllactose, and 6′-sialyllactose), antibacterial proteins immunoglobulin A, lactoferrin, and lysozyme), urea, and hypoxanthine in infant feces. (b) *Bifidobacterium breve* and *Veillonella ratti*, in infant feces. mo, month.

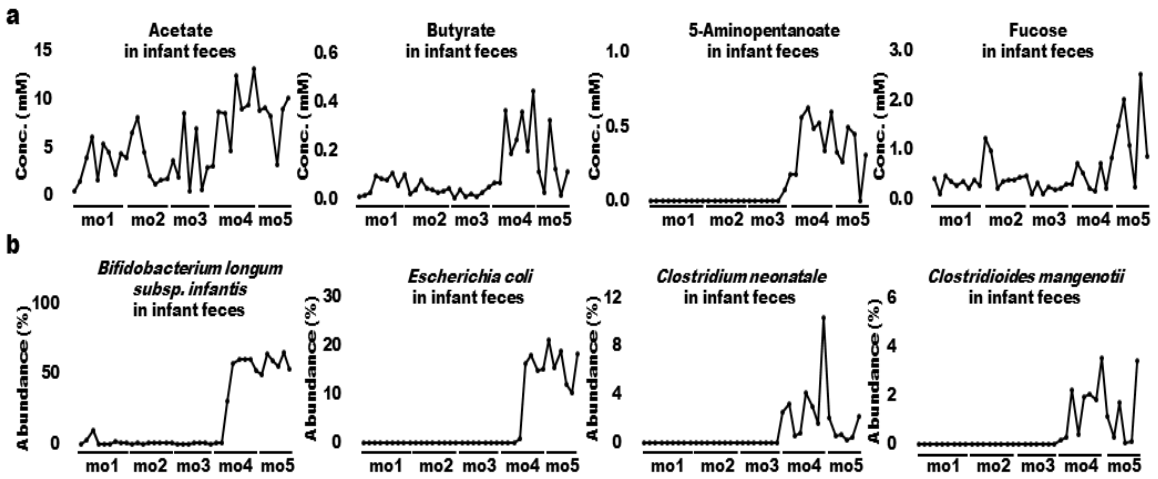


Figure 4-5. Changes in milk bioactive components, microbiome, and metabolites in infant feces with high similarity in cluster B. (a) Acetate, butyrate, 5-aminopentanoate, and fucose levels. (b) *Bifidobacterium longum* subsp. *infantis*, *Escherichia coli*, *Clostridium neonatale*, and *Clostridioides mangenotii*. mo, month.

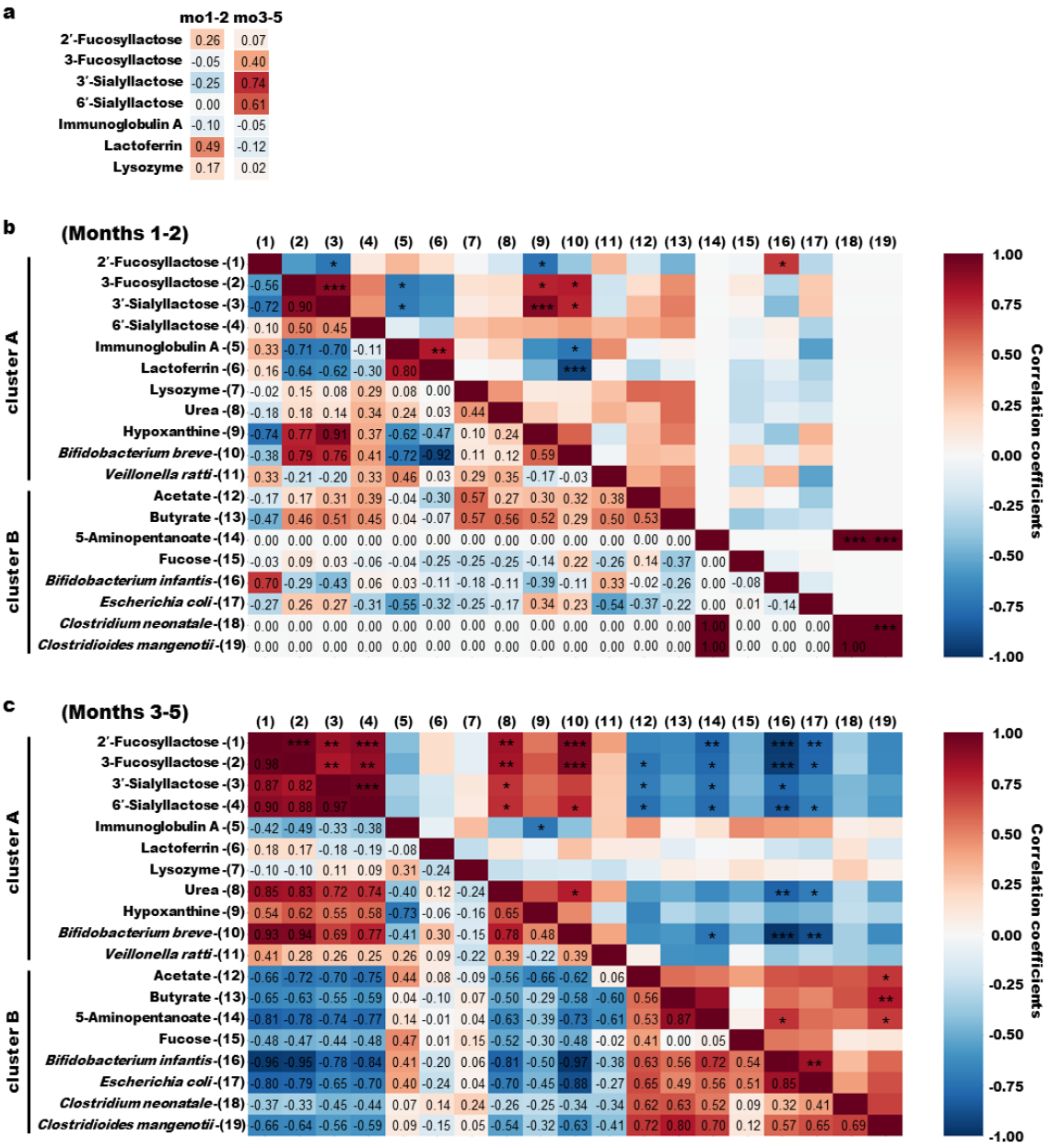


Figure 4-6. Spearman correlation heatmaps of milk bioactive components, infant gut microbiome, and fecal metabolites. (a) Spearman rank correlation coefficients between bioactive HMOs and antibacterial proteins in breast milk and those in infant feces at 1–2 and 3–5 months postpartum. (b) Spearman rank correlation coefficients of milk bioactive components, bacterial species, and metabolites in infant feces at 1–2 months postpartum with high similarity categorized into clusters A and B. (c) Spearman rank correlation coefficients of milk bioactive components, bacterial species, and metabolites in infant feces at 3–5 months postpartum with high similarity categorized into clusters A and B. Spearman ρ values are shown in different colors and |ρ| values are shown as *0.7–0.8, **0.8–0.9, and ***>0.9. mo, month.

4.7 Supplementary tables and figures

Supplementary table 4-S1. Relative abundance of the microbial species in infant feces samples collected 1–5 months after delivery

	Microbial species	Month 1									Month 2				
		Day 34	Day 36	Day 41	Day 46	Day 51	Day 53	Day 58-1	Day 58-2	Day 59	Day 61	Day 62	Day 66	Day 69	Day 72
						%							%		
1	f__Bifidobacteriaceae;g__Bifidobacterium;s__longum subsp. Infantis	0.00	3.13	9.94	0.00	0.06	0.00	1.73	1.02	0.89	0.21	0.60	0.15	0.74	0.61
2	f__Bifidobacteriaceae;g__Bifidobacterium;s__breve	0.13	44.61	58.09	77.39	60.69	69.31	63.97	76.65	62.53	71.32	78.16	60.51	69.04	79.87
3	f__Bifidobacteriaceae;g__Bifidobacterium;s__pseudocatenulatum	0.00	3.34	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4	f__Actinomycetaceae;g__Winkia;s__neuii	0.00	0.00	0.00	0.00	0.00	0.00	0.16	0.00	0.32	0.12	2.21	1.18	8.64	0.23
5	f__Enterobacteriaceae;g__Escherichia;s__coli	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.09
6	f__Enterobacteriaceae;g__Klebsiella;s__aerogenes	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
7	f__Enterobacteriaceae;g__Kluyvera;s__cryocrescens	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8	f__Enterobacteriaceae;g__Citrobacter;s__unclassified Citrobacter	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
9	f__Enterobacteriaceae;g__Citrobacter;s__farmeri	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10	f__Enterobacteriaceae;g__Raoultella;s__terrigena	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
11	f__Enterobacteriaceae;g__Citrobacter;s__freundii	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12	f__Enterobacteriaceae;g__Enterobacter;s__hormaechei	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
13	f__Enterobacteriaceae;g__Citrobacter;s__koseri	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
14	f__Pasteurellaceae;g__Haemophilus;s__parainfluenzae	10.28	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
15	f__Streptococcaceae;g__Streptococcus;s__gallolyticus	11.78	4.15	0.00	0.00	1.85	3.48	2.18	1.02	1.00	0.96	1.01	5.17	1.01	2.87
16	f__Streptococcaceae;g__Streptococcus;s__infantis	0.00	0.92	0.40	0.38	0.19	0.13	0.19	0.35	0.08	0.85	0.20	0.46	0.00	0.29
17	f__Streptococcaceae;g__Streptococcus;s__salivarius	0.38	1.20	0.89	0.99	1.48	0.64	0.30	0.35	0.49	0.75	0.37	0.33	0.20	0.60
18	f__Streptococcaceae;g__Streptococcus;s__peroris	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
19	f__Enterococcaceae;g__Enterococcus;s__faecium	72.90	2.43	0.70	0.13	0.44	1.98	4.89	4.91	5.66	3.79	5.19	2.88	5.80	7.56
20	f__Enterococcaceae;g__Enterococcus;s__faecalis	1.79	0.00	0.00	0.00	0.00	0.31	2.09	1.40	1.95	0.65	1.05	3.64	4.24	5.28
21	f__Enterococcaceae;g__Enterococcus;s__casseliflavus	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
22	f__Streptococcaceae;g__Streptococcus;s__thermophilus	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
23	f__Streptococcaceae;g__Streptococcus;s__sp.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
24	f__Staphylococcaceae;g__Staphylococcus;s__epidermidis	1.12	9.56	2.14	1.02	0.67	0.48	0.17	0.64	0.24	3.03	0.41	0.90	0.20	0.74
25	f__Staphylococcaceae;g__Staphylococcus;s__aureus	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
26	f__Staphylococcaceae;g__Staphylococcus;s__lugdunensis	0.14	0.30	0.00	0.00	0.00	0.15	0.20	0.39	0.20	0.37	0.00	0.23	0.14	0.42
27	f__Paenibacillaceae;g__Paenibacillus;s__sp.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
28	f__Bacillales Family XI. Incertae Sedis;g__Gemella;s__haemolysans	0.00	0.28	0.19	0.17	0.22	0.10	0.06	0.00	0.00	0.09	0.12	0.08	0.00	0.19
29	f__Clostridiaceae;g__Hungatella;s__effluvii	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
30	f__Clostridiaceae;g__Clostridium;s__sardiniense	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
31	f__Clostridiaceae;g__Clostridium;s__paraputrificum	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
32	f__Clostridiaceae;g__Clostridium;s__perfringens	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
33	f__Clostridiaceae;g__Clostridium;s__neonatale	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
34	f__Clostridiaceae;g__Clostridium;s__saccharobutylicum	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
35	f__Peptostreptococcaceae;g__Terrisporobacter;s__glycolicus	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
36	f__Peptostreptococcaceae;g__Clostridioides;s__mangenotii	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
37	f__Peptostreptococcaceae;g__Paraclostridium;s__benzoelyticum	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
38	f__Lachnospiraceae;g__Tyzzerella;s__nexilis	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
39	f__Veillonellaceae;g__Veillonella;s__ratti	0.00	30.08	27.13	19.93	34.40	23.41	24.06	13.28	26.64	17.85	10.67	24.32	9.98	1.25
40	f__Veillonellaceae;g__Veillonella;s__parvula	1.49	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.14	0.00	0.00

Supplementary table 4-S1, continued

	Microbial species	Month 2			Month 3							Month 4			
		Day 74	Day 76	Day 82	Day 99	Day 103	Day 104	Day 108	Day 110	Day 114	Day 118	Day 122	Day 124	Day 131	Day 134
			%					%						%	
1	f__Bifidobacteriaceae;g__Bifidobacterium;s__longum subsp. Infantis	0.48	0.43	0.50	0.00	0.18	0.09	0.33	0.30	0.44	0.12	1.01	1.14	30.84	57.13
2	f__Bifidobacteriaceae;g__Bifidobacterium;s__breve	80.50	76.69	64.37	55.71	50.69	51.08	50.80	48.45	50.16	40.80	59.64	61.26	29.16	8.74
3	f__Bifidobacteriaceae;g__Bifidobacterium;s__pseudocatenulatum	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.15	0.00	0.00	0.15	0.30	0.00	0.00
4	f__Actinomycetaceae;g__Winkia;s__neuii	0.00	2.05	8.53	0.00	0.00	0.00	0.00	0.12	0.44	0.00	0.00	0.00	0.23	0.57
5	f__Enterobacteriaceae;g__Escherichia;s__coli	0.00	0.12	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
6	f__Enterobacteriaceae;g__Klebsiella;s__aerogenes	0.00	0.00	0.00	18.67	21.92	15.44	13.34	12.71	11.09	11.96	5.03	3.66	0.28	0.28
7	f__Enterobacteriaceae;g__Kluyvera;s__cryocrescens	0.00	0.00	0.00	9.92	12.06	8.13	7.42	7.08	6.54	6.14	2.77	2.25	0.21	0.61
8	f__Enterobacteriaceae;g__Citrobacter;s__unclassified Citrobacter	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	9.27	5.73
9	f__Enterobacteriaceae;g__Citrobacter;s__farmeri	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.36	0.76
10	f__Enterobacteriaceae;g__Raoultella;s__terrigena	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.62	0.94	1.47	0.00
11	f__Enterobacteriaceae;g__Citrobacter;s__freundii	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.66	1.01	1.17	0.00
12	f__Enterobacteriaceae;g__Enterobacter;s__hormaechei	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.27	0.36	0.56	0.32
13	f__Enterobacteriaceae;g__Citrobacter;s__koseri	0.00	0.00	0.00	0.00	0.47	0.85	2.67	2.11	3.29	1.85	3.23	2.22	0.13	0.00
14	f__Pasteurellaceae;g__Haemophilus;s__parainfluenzae	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
15	f__Streptococcaceae;g__Streptococcus;s__gallolyticus	2.08	0.70	2.38	1.59	3.35	6.85	8.26	10.42	1.63	30.07	5.05	2.24	0.73	0.97
16	f__Streptococcaceae;g__Streptococcus;s__infantis	0.43	0.19	0.07	0.50	0.13	0.32	0.14	0.68	0.33	0.19	0.27	0.64	0.52	0.40
17	f__Streptococcaceae;g__Streptococcus;s__salivarius	1.43	0.51	0.39	0.42	0.25	0.25	0.12	0.23	0.31	0.27	0.23	0.58	0.16	0.26
18	f__Streptococcaceae;g__Streptococcus;s__peroris	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.05
19	f__Enterococcaceae;g__Enterococcus;s__faecium	8.22	11.29	11.17	2.83	5.52	5.02	10.41	5.70	15.75	3.52	6.38	4.88	8.86	13.29
20	f__Enterococcaceae;g__Enterococcus;s__faecalis	2.95	4.50	5.97	2.91	3.13	2.42	4.92	3.16	5.49	1.90	7.96	5.41	5.34	3.36
21	f__Enterococcaceae;g__Enterococcus;s__casseliflavus	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.22	0.00	0.85	1.11
22	f__Streptococcaceae;g__Streptococcus;s__thermophilus	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
23	f__Streptococcaceae;g__Streptococcus;s__sp.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
24	f__Staphylococcaceae;g__Staphylococcus;s__epidermidis	1.98	0.29	0.46	0.46	0.13	0.00	0.00	0.15	0.51	0.00	0.00	0.33	0.15	0.00
25	f__Staphylococcaceae;g__Staphylococcus;s__aureus	0.00	0.00	0.00	0.87	0.24	0.54	0.51	0.66	1.24	1.02	0.98	1.65	0.66	1.04
26	f__Staphylococcaceae;g__Staphylococcus;s__lugdunensis	0.70	0.49	0.71	0.97	0.40	1.32	0.32	0.42	0.68	0.82	0.52	0.89	0.21	0.00
27	f__Paenibacillaceae;g__Paenibacillus;s__sp.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.17	2.01	0.00	0.00
28	f__Bacillales Family XI. Incertae Sedis;g__Gemella;s__haemolysans	0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
29	f__Clostridiaceae;g__Hungatella;s__effluvii	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
30	f__Clostridiaceae;g__Clostridium;s__sardiniense	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.00	0.72	0.48
31	f__Clostridiaceae;g__Clostridium;s__paraputrificum	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
32	f__Clostridiaceae;g__Clostridium;s__perfringens	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.91	0.52	0.19
33	f__Clostridiaceae;g__Clostridium;s__neonatale	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.49	3.23	0.62	0.75
34	f__Clostridiaceae;g__Clostridium;s__saccharobutylicum	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
35	f__Peptostreptococcaceae;g__Terrisporobacter;s__glycolicus	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.08	2.42
36	f__Peptostreptococcaceae;g__Clostridioides;s__mangenotii	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.21	0.28	2.25	0.44
37	f__Peptostreptococcaceae;g__Paraclostridium;s__benzoelyticum	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
38	f__Lachnospiraceae;g__Tyzzerella;s__nexilis	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
39	f__Veillonellaceae;g__Veillonella;s__ratti	1.11	2.72	5.34	5.13	1.54	7.69	0.77	7.66	2.09	1.36	1.09	3.82	0.60	0.11
40	f__Veillonellaceae;g__Veillonella;s__parvula	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Supplementary table 4-S1, continued

	Microbial species	Month 4				Month 5					
		Day 137	Day 142	Day 145	Day 148	Day 155	Day 157	Day 158	Day 159	Day 161	Day 163
		%				%					
1	f__Bifidobacteriaceae;g__Bifidobacterium;s__longum subsp. Infantis	60.08	60.23	60.78	52.53	49.01	64.02	59.77	55.48	65.78	53.57
2	f__Bifidobacteriaceae;g__Bifidobacterium;s__breve	3.09	2.50	5.31	2.81	4.44	3.00	2.99	6.49	8.17	3.33
3	f__Bifidobacteriaceae;g__Bifidobacterium;s__pseudocatenulatum	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4	f__Actinomycetaceae;g__Winkia;s__neuii	0.23	0.35	0.24	0.00	0.21	0.29	0.42	0.00	0.00	0.06
5	f__Enterobacteriaceae;g__Escherichia;s__coli	16.15	17.81	14.74	15.14	21.02	15.28	18.79	12.11	10.24	18.20
6	f__Enterobacteriaceae;g__Klebsiella;s__aerogenes	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
7	f__Enterobacteriaceae;g__Kluyvera;s__cryocrescens	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8	f__Enterobacteriaceae;g__Citrobacter;s__unclassified Citrobacter	0.56	0.75	0.72	0.48	1.15	0.76	0.77	0.92	0.30	0.66
9	f__Enterobacteriaceae;g__Citrobacter;s__farmeri	0.09	0.07	0.08	0.10	0.15	0.10	0.08	0.19	0.00	0.08
10	f__Enterobacteriaceae;g__Raoultella;s__terrigena	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
11	f__Enterobacteriaceae;g__Citrobacter;s__freundii	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12	f__Enterobacteriaceae;g__Enterobacter;s__hormaechei	0.48	0.45	0.29	0.54	0.66	0.34	0.43	0.22	0.00	0.42
13	f__Enterobacteriaceae;g__Citrobacter;s__koseri	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
14	f__Pasteurellaceae;g__Haemophilus;s__parainfluenzae	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
15	f__Streptococcaceae;g__Streptococcus;s__gallolyticus	0.25	0.10	0.19	0.00	0.39	0.55	0.33	0.17	0.12	0.19
16	f__Streptococcaceae;g__Streptococcus;s__infantis	0.64	1.11	1.13	0.76	1.14	0.81	0.30	1.22	3.17	0.73
17	f__Streptococcaceae;g__Streptococcus;s__salivarius	0.10	0.00	0.21	0.13	0.28	0.12	0.10	0.11	0.51	0.11
18	f__Streptococcaceae;g__Streptococcus;s__peroris	0.07	0.15	0.11	0.08	1.21	0.96	0.26	0.56	1.90	0.54
19	f__Enterococcaceae;g__Enterococcus;s__faecium	7.15	6.44	7.06	7.53	9.26	6.39	8.24	16.79	4.20	6.21
20	f__Enterococcaceae;g__Enterococcus;s__faecalis	0.52	0.37	0.18	0.67	0.20	0.13	0.16	0.12	0.07	0.00
21	f__Enterococcaceae;g__Enterococcus;s__casseliflavus	0.42	0.89	0.00	0.91	2.61	2.05	1.46	1.76	0.00	0.51
22	f__Streptococcaceae;g__Streptococcus;s__thermophilus	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
23	f__Streptococcaceae;g__Streptococcus;s__sp.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
24	f__Staphylococcaceae;g__Staphylococcus;s__epidermidis	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
25	f__Staphylococcaceae;g__Staphylococcus;s__aureus	0.90	0.73	1.48	0.47	0.50	0.23	0.21	0.30	0.69	0.36
26	f__Staphylococcaceae;g__Staphylococcus;s__lugdunensis	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
27	f__Paenibacillaceae;g__Paenibacillus;s__sp.	0.00	0.00	0.00	0.06	0.09	0.00	0.00	0.00	0.00	0.26
28	f__Bacillales Family XI. Incertae Sedis;g__Gemella;s__haemolysans	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
29	f__Clostridiaceae;g__Hungatella;s__effluvii	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
30	f__Clostridiaceae;g__Clostridium;s__sardiniense	0.18	0.27	0.41	0.14	0.37	0.33	0.33	0.51	0.28	3.39
31	f__Clostridiaceae;g__Clostridium;s__paraputrificum	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
32	f__Clostridiaceae;g__Clostridium;s__perfringens	0.35	0.18	0.18	0.18	0.17	0.35	0.20	0.12	0.32	1.95
33	f__Clostridiaceae;g__Clostridium;s__neonatale	4.17	3.00	1.58	10.36	2.07	0.60	0.66	0.19	0.48	2.20
34	f__Clostridiaceae;g__Clostridium;s__saccharobutylicum	0.00	0.00	0.00	1.07	0.00	0.00	0.00	0.00	0.00	0.00
35	f__Peptostreptococcaceae;g__Terrisporobacter;s__glycolicus	1.58	1.76	1.13	2.25	1.39	0.14	2.19	0.48	0.00	1.14
36	f__Peptostreptococcaceae;g__Clostridioides;s__mangenotii	1.94	2.09	1.86	3.55	1.14	0.31	1.71	0.10	0.15	3.43
37	f__Peptostreptococcaceae;g__Paraclostridium;s__benzoelyticum	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
38	f__Lachnospiraceae;g__Tyzzerella;s__nexilis	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
39	f__Veillonellaceae;g__Veillonella;s__ratti	1.03	0.51	2.35	0.25	2.54	3.25	0.61	2.16	3.61	2.66
40	f__Veillonellaceae;g__Veillonella;s__parvula	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Supplementary table 4-S2. Metabolites and their chemical shifts from ^{1}H NMR data in infant feces samples*

	Metabolite	δ¹H (ppm)					
1	2′-Fucosyllactose	1.23(d)	4.28(dd)	5.22(d)			
2	2-Hydroxyisovalerate	0.82(d)	0.95(d)				
3	3-Fucosyllactose	1.17(d)	5.44(d)				
4	3′-Sialyllactose	1.79(t)	3.28(t)	4.52(d)			
5	5-Aminopentanoate	1.61(m)	1.67(m)	2.22(t)	3.00(t)		
6	6′-Sialyllactose	1.74(t)	4.42(d)				
7	Acetate	1.91(s)					
8	Alanine	1.47(d)	3.78(q)				
9	Aspartate	2.66(dd)	2.80(dd)	3.89(dd)			
10	Butyrate	0.88(t)	1.54(m)	2.15(t)			
11	Cadaverine	1.46(m)	1.70(m)	3.01(t)			
12	Carnitine	3.21(s)	3.40(d)	3.43(dd)			
13	Choline	3.19(s)	3.51(m)	4.06(m)			
14	Creatine	3.02(s)	3.92(s)				
15	Formate	8.44(s)					
16	Fucose	1.19(d)	1.23(d)	3.44(dd)	4.54(d)	5.19(d)	
17	Fumarate	6.51(s)					
18	Glucose	3.23(dd)	3.39(t)	3.48(t)	4.64(d)	5.22(d)	
19	Glutamate	2.12(m)	2.33(m)	2.36(m)	3.75(dd)		
20	Glycine	3.55(s)					
21	Hypoxanthine	8.18(s)	8.20(s)				
22	Isoleucine	0.93(t)	1.00(d)	1.46(m)	3.66(d)		
23	Lactate	1.32(d)	4.11(q)				
24	Lactose	3.28(t)	3.93(d)	4.44(d)	4.45(d)	4.66(d)	5.22(d)
25	Leucine	0.94(d)	0.95(d)	1.67(m)	1.70(m)	3.73(dd)	
26	Methionine	2.11(m)	2.13(s)	2.63(t)	3.85(dd)		
27	o-Phosphocholine	3.20(s)	4.15(m)				
28	Phenylalanine	7.31(d)	7.36(t)	7.42(t)			
29	Propionate	1.04(t)	2.17(q)				
30	Pyruvate	2.36(s)					
31	Succinate	2.39(s)					
32	Taurine	3.25(t)	3.41(t)				
33	Threonine	1.31(d)	3.58(d)	4.24(m)			
34	Trimethylamine N-oxide	3.25(s)					
35	Tryptophan	7.19(m)	7.53(m)	7.74(m)			
36	Tyrosine	6.88(m)	7.18(m)				
37	Urea	5.77(s)					
38	Valine	0.98(d)	1.03(d)	2.26(m)	3.60(d)		

*Key: s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; dd, doublet of doublet.

Supplementary table 4-S3. Concentration of the metabolites in infant feces samples collected 1–5 months after delivery*

	Metabolite	Month 1									Month 2								Month 3	
		Day 34	Day 36	Day 41	Day 46	Day 51	Day 53	Day 58-1	Day 58-2	Day 59	Day 61	Day 62	Day 66	Day 69	Day 72	Day 74	Day 76	Day 82	Day 99	Day 103
1	2′-Fucosyllactose (mM)	1.634	3.016	2.776	1.626	0.876	1.151	0.900	1.092	0.913	1.483	1.149	0.796	0.831	0.934	1.073	0.957	0.854	0.638	0.791
2	2-Hydroxyisovalerate (mM)	0.004	0.001	0.002	0.001	0.003	0.002	0.007	0.004	0.004	0.002	0.003	0.005	0.007	0.014	0.013	0.010	0.016	0.003	0.018
3	3-Fucosyllactose (mM)	0.718	1.039	1.354	2.456	1.539	2.022	2.562	2.797	2.323	1.891	2.157	1.593	1.834	2.031	2.471	2.416	2.352	1.573	2.109
4	3′-Sialyllactose (mM)	0.195	0.283	0.307	0.601	0.436	0.758	0.767	0.708	0.677	0.541	0.623	0.594	0.639	0.622	0.708	0.744	0.650	0.460	0.791
5	5-Aminopentanoate (mM)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
6	6′-Sialyllactose (mM)	0.163	0.336	0.319	0.474	0.216	0.505	0.433	0.482	0.314	0.297	0.342	0.249	0.284	0.202	0.276	0.346	0.183	0.152	0.209
7	Acetate (mM)	0.507	1.489	3.959	6.129	1.730	5.424	4.568	2.243	4.367	3.944	6.552	8.202	4.561	2.163	1.201	1.665	1.873	3.655	1.967
8	Alanine (mM)	0.081	0.066	0.100	0.201	0.252	0.185	0.155	0.153	0.150	0.122	0.165	0.194	0.173	0.212	0.155	0.177	0.196	0.222	0.219
9	Aspartate (mM)	0.057	0.097	0.101	0.359	0.163	0.361	0.388	0.478	0.390	0.169	0.219	0.249	0.324	0.328	0.340	0.434	0.292	0.129	0.177
10	Butyrate (mM)	0.011	0.014	0.025	0.098	0.084	0.079	0.107	0.056	0.099	0.022	0.041	0.076	0.046	0.040	0.028	0.034	0.041	0.006	0.039
11	Cadaverine (mM)	0.048	0.036	0.031	0.069	0.060	0.117	0.213	0.147	0.190	0.070	0.077	0.105	0.135	0.093	0.095	0.128	0.128	0.053	0.083
12	Carnitine (mM)	0.006	0.003	0.005	0.008	0.005	0.012	0.005	0.008	0.004	0.006	0.005	0.007	0.004	0.009	0.013	0.006	0.004	0.015	0.004
13	Choline (mM)	0.015	0.009	0.004	0.009	0.037	0.009	0.009	0.007	0.008	0.018	0.013	0.027	0.014	0.018	0.008	0.007	0.010	0.042	0.009
14	Creatine (mM)	0.070	0.074	0.010	0.035	0.027	0.031	0.026	0.043	0.030	0.018	0.015	0.018	0.025	0.013	0.037	0.036	0.029	0.002	0.019
15	Formate (mM)	1.216	1.335	1.483	1.374	1.309	1.010	1.147	0.905	0.925	0.935	1.159	1.168	0.749	0.894	0.626	0.806	0.714	0.951	0.489
16	Fucose (mM)	0.414	0.103	0.460	0.358	0.280	0.351	0.207	0.388	0.272	1.229	0.979	0.209	0.355	0.381	0.380	0.444	0.474	0.102	0.320
17	Fumarate (mM)	0.001	0.000	0.002	0.002	0.000	0.005	0.002	0.002	0.002	0.002	0.003	0.003	0.004	0.003	0.003	0.002	0.004	0.001	0.002
18	Glucose (mM)	0.017	0.128	0.477	0.391	0.269	0.328	0.144	0.128	0.193	1.065	0.854	0.223	0.161	0.120	0.134	0.153	0.146	0.181	0.214
19	Glutamate (mM)	0.152	0.141	0.204	0.457	0.254	0.452	0.487	0.597	0.458	0.279	0.426	0.496	0.586	0.755	0.670	0.959	0.839	0.456	0.841
20	Glycine (mM)	0.000	0.000	0.035	0.055	0.168	0.130	0.109	0.174	0.100	0.092	0.129	0.116	0.153	0.231	0.224	0.238	0.191	0.144	0.221
21	Hypoxanthine (mM)	0.000	0.000	0.008	0.030	0.018	0.051	0.042	0.051	0.048	0.018	0.027	0.041	0.048	0.031	0.030	0.053	0.046	0.006	0.019
22	Isoleucine (mM)	0.029	0.037	0.038	0.059	0.058	0.063	0.066	0.071	0.055	0.049	0.049	0.071	0.084	0.078	0.065	0.085	0.066	0.069	0.074
23	Lactate (mM)	0.097	0.003	2.942	1.159	2.065	1.233	0.971	0.086	0.727	2.871	2.809	3.233	1.580	0.229	0.051	0.317	0.115	6.081	0.411
24	Lactose (mM)	5.275	6.211	1.000	1.115	0.707	1.038	0.948	1.039	1.006	2.025	1.541	0.730	0.839	0.706	0.911	0.796	0.867	0.334	0.408
25	Leucine (mM)	0.073	0.063	0.060	0.096	0.134	0.116	0.101	0.085	0.069	0.070	0.086	0.139	0.187	0.147	0.148	0.157	0.149	0.153	0.165
26	Methionine (mM)	0.013	0.012	0.016	0.024	0.024	0.023	0.017	0.025	0.016	0.015	0.016	0.015	0.016	0.014	0.021	0.018	0.018	0.021	0.015
27	O-Phosphocholine (mM)	0.001	0.001	0.000	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.000	0.001	0.001
28	Phenylalanine (mM)	0.033	0.025	0.022	0.033	0.042	0.041	0.035	0.051	0.039	0.023	0.032	0.047	0.067	0.044	0.050	0.059	0.063	0.043	0.048
29	Propionate (mM)	0.121	0.080	0.074	0.255	0.254	0.251	0.386	0.285	0.439	0.052	0.169	0.306	0.325	0.459	0.276	0.226	0.313	0.013	0.369
30	Pyruvate (mM)	0.157	0.022	0.081	0.247	0.088	0.139	0.065	0.046	0.085	0.154	0.169	0.231	0.096	0.068	0.032	0.069	0.074	0.078	0.082
31	Succinate (mM)	0.065	0.008	0.114	0.083	0.045	0.136	0.049	0.036	0.049	0.182	0.226	0.230	0.125	0.438	0.153	0.139	0.245	0.121	0.231
32	Taurine (mM)	0.017	0.064	0.056	0.035	0.036	0.049	0.042	0.050	0.046	0.082	0.051	0.035	0.057	0.053	0.064	0.080	0.069	0.073	0.068
33	Threonine (mM)	0.117	0.032	0.102	0.264	0.198	0.256	0.139	0.099	0.141	0.117	0.122	0.117	0.059	0.044	0.049	0.071	0.043	0.192	0.115
34	TMAO (mM)	0.008	0.015	0.016	0.022	0.015	0.020	0.028	0.033	0.026	0.019	0.013	0.022	0.009	0.028	0.028	0.021	0.005	0.009	0.002
35	Tryptophan (mM)	0.027	0.019	0.011	0.026	0.015	0.024	0.029	0.036	0.030	0.020	0.019	0.026	0.036	0.021	0.026	0.031	0.030	0.018	0.021
36	Tyrosine (mM)	0.025	0.032	0.031	0.048	0.047	0.044	0.035	0.038	0.033	0.033	0.028	0.032	0.036	0.029	0.036	0.031	0.026	0.046	0.025
37	Urea (mM)	0.182	0.172	0.191	0.234	0.216	0.211	0.212	0.244	0.188	0.199	0.152	0.218	0.198	0.217	0.150	0.163	0.187	0.155	0.155
38	Valine (mM)	0.065	0.059	0.060	0.101	0.122	0.130	0.123	0.163	0.102	0.072	0.097	0.155	0.205	0.168	0.144	0.170	0.154	0.147	0.180
39	Immunoglobulin A (g/l)	0.252	0.144	0.130	0.124	0.226	0.144	0.098	0.089	0.108	0.143	0.089	0.117	0.078	0.052	0.025	0.030	0.035	0.056	0.040
40	Lactoferrin (mg/l)	0.271	0.068	0.025	0.031	0.099	0.046	0.052	0.037	0.076	0.014	0.005	0.052	0.024	0.004	0.004	0.008	0.003	0.000	0.002
41	Lysozyme (μg/l)	0.006	0.012	0.005	0.104	0.009	0.008	0.017	0.010	0.037	0.007	0.004	0.064	0.008	0.005	0.009	0.003	0.003	0.009	0.006

*The relative concentration of each metabolite is its integral value normalized to TSP, which was added to the NMR buffer as an internal standard.

Supplementary table 4-S3, continued*

	Metabolite	Month 3					Month 4								Month 5					
		Day 104	Day 108	Day 110	Day 114	Day 118	Day 122	Day 124	Day 131	Day 134	Day 137	Day 142	Day 145	Day 148	Day 155	Day 157	Day 158	Day 159	Day 161	Day 163
1	2′-Fucosyllactose (mM)	0.628	0.736	0.817	0.805	0.755	0.568	0.644	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	2-Hydroxyisovalerate (mM)	0.007	0.018	0.011	0.025	0.024	0.013	0.009	0.010	0.009	0.009	0.009	0.005	0.013	0.010	0.009	0.003	0.026	0.002	0.004
3	3-Fucosyllactose (mM)	1.292	1.910	1.611	2.247	1.866	2.048	1.645	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4	3′-Sialyllactose (mM)	0.425	0.687	0.568	0.683	0.770	0.042	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
5	5-Aminopentanoate (mM)	0.000	0.000	0.000	0.000	0.071	0.177	0.176	0.559	0.626	0.492	0.523	0.341	0.603	0.328	0.265	0.498	0.450	0.000	0.310
6	6′-Sialyllactose (mM)	0.134	0.221	0.138	0.176	0.184	0.084	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
7	Acetate (mM)	8.534	0.553	7.004	0.643	2.965	3.121	8.736	8.607	4.717	12.41	8.994	9.469	13.23	8.900	9.109	8.301	3.307	9.036	10.18
8	Alanine (mM)	0.248	0.174	0.233	0.231	0.253	0.195	0.260	0.291	0.341	0.405	0.229	0.339	0.432	0.473	0.419	0.268	0.648	0.239	0.353
9	Aspartate (mM)	0.107	0.113	0.098	0.082	0.130	0.218	0.128	0.450	0.750	0.501	0.646	0.471	0.670	0.294	0.170	0.583	0.591	0.124	0.349
10	Butyrate (mM)	0.012	0.020	0.012	0.025	0.047	0.069	0.064	0.363	0.184	0.242	0.356	0.201	0.443	0.112	0.025	0.325	0.123	0.013	0.112
11	Cadaverine (mM)	0.069	0.078	0.070	0.090	0.091	0.068	0.064	0.214	0.181	0.123	0.167	0.140	0.158	0.117	0.083	0.181	0.160	0.045	0.133
12	Carnitine (mM)	0.005	0.003	0.004	0.003	0.007	0.003	0.005	0.009	0.006	0.005	0.007	0.007	0.008	0.014	0.007	0.010	0.005	0.011	0.011
13	Choline (mM)	0.027	0.006	0.026	0.005	0.015	0.011	0.016	0.016	0.013	0.021	0.019	0.030	0.027	0.028	0.030	0.032	0.021	0.033	0.031
14	Creatine (mM)	0.028	0.018	0.017	0.023	0.014	0.009	0.006	0.030	0.058	0.045	0.047	0.067	0.060	0.084	0.045	0.065	0.059	0.044	0.049
15	Formate (mM)	0.708	0.457	0.732	0.438	0.655	0.650	0.949	0.876	0.712	1.351	0.746	1.108	0.880	0.793	0.635	0.446	0.938	0.747	0.837
16	Fucose (mM)	0.099	0.245	0.195	0.208	0.288	0.290	0.720	0.534	0.209	0.159	0.729	0.208	0.824	1.480	2.036	1.084	0.255	2.523	0.863
17	Fumarate (mM)	0.001	0.005	0.001	0.001	0.001	0.002	0.001	0.016	0.025	0.005	0.012	0.005	0.012	0.010	0.002	0.023	0.026	0.004	0.006
18	Glucose (mM)	0.163	0.134	0.306	0.099	0.220	0.162	0.715	0.480	0.197	0.243	0.606	0.327	0.796	0.750	0.510	0.684	0.283	0.739	0.419
19	Glutamate (mM)	0.443	0.741	0.542	0.791	0.807	0.721	0.524	0.960	0.995	0.623	0.829	0.581	1.072	0.543	0.342	1.038	0.721	0.197	0.512
20	Glycine (mM)	0.270	0.229	0.243	0.249	0.175	0.109	0.215	0.055	0.084	0.133	0.073	0.160	0.170	0.119	0.186	0.134	0.245	0.214	0.217
21	Hypoxanthine (mM)	0.005	0.032	0.009	0.055	0.017	0.022	0.011	0.016	0.007	0.011	0.005	0.008	0.006	0.005	0.008	0.007	0.011	0.007	0.009
22	Isoleucine (mM)	0.098	0.072	0.068	0.074	0.063	0.039	0.036	0.036	0.044	0.052	0.038	0.061	0.050	0.058	0.055	0.042	0.081	0.045	0.073
23	Lactate (mM)	7.617	0.039	6.313	0.035	1.555	1.940	6.159	3.226	0.591	5.560	4.339	8.311	3.837	5.942	9.531	4.307	2.380	7.981	8.756
24	Lactose (mM)	0.320	0.352	0.486	0.410	0.510	0.183	0.770	0.156	0.078	0.084	0.134	0.085	0.117	0.902	0.618	0.180	0.080	0.903	0.537
25	Leucine (mM)	0.202	0.169	0.167	0.164	0.155	0.076	0.067	0.108	0.109	0.099	0.096	0.118	0.079	0.104	0.095	0.118	0.135	0.078	0.106
26	Methionine (mM)	0.024	0.017	0.017	0.013	0.014	0.018	0.014	0.011	0.022	0.019	0.007	0.024	0.012	0.014	0.015	0.021	0.021	0.012	0.019
27	O-Phosphocholine (mM)	0.003	0.001	0.002	0.001	0.001	0.000	0.001	0.002	0.003	0.004	0.004	0.003	0.004	0.003	0.002	0.004	0.002	0.001	0.002
28	Phenylalanine (mM)	0.070	0.051	0.053	0.071	0.045	0.018	0.020	0.017	0.025	0.030	0.021	0.042	0.024	0.049	0.036	0.024	0.062	0.025	0.045
29	Propionate (mM)	0.080	0.250	0.092	0.217	0.322	0.372	0.128	0.382	0.616	0.458	0.492	0.336	0.739	0.223	0.061	0.231	0.287	0.033	0.132
30	Pyruvate (mM)	0.244	0.007	0.207	0.006	0.161	0.043	0.285	0.202	0.059	0.184	0.144	0.236	0.169	0.370	0.350	0.266	0.063	0.236	0.410
31	Succinate (mM)	0.214	0.273	0.248	0.022	0.130	0.404	0.322	0.486	0.506	0.327	0.549	0.552	0.589	0.668	0.478	0.228	1.304	0.301	0.224
32	Taurine (mM)	0.193	0.079	0.107	0.062	0.098	0.051	0.182	0.093	0.149	0.138	0.086	0.201	0.108	0.113	0.127	0.168	0.125	0.208	0.247
33	Threonine (mM)	0.120	0.040	0.081	0.016	0.071	0.088	0.081	0.044	0.120	0.152	0.112	0.155	0.155	0.134	0.173	0.160	0.191	0.105	0.125
34	TMAO (mM)	0.006	0.002	0.008	0.011	0.005	0.006	0.006	0.004	0.002	0.002	0.004	0.003	0.004	0.006	0.005	0.003	0.002	0.008	0.007
35	Tryptophan (mM)	0.024	0.014	0.011	0.031	0.009	0.014	0.013	0.022	0.018	0.019	0.014	0.017	0.020	0.010	0.014	0.015	0.017	0.009	0.017
36	Tyrosine (mM)	0.062	0.023	0.032	0.018	0.021	0.022	0.026	0.023	0.036	0.042	0.034	0.051	0.041	0.040	0.042	0.034	0.064	0.032	0.048
37	Urea (mM)	0.211	0.134	0.214	0.361	0.121	0.140	0.147	0.024	0.092	0.040	0.044	0.047	0.048	0.068	0.000	0.000	0.000	0.000	0.000
38	Valine (mM)	0.173	0.152	0.167	0.171	0.148	0.063	0.074	0.056	0.085	0.077	0.057	0.105	0.090	0.141	0.126	0.073	0.213	0.090	0.113
39	Immunoglobulin A (g/l)	0.033	0.008	0.037	0.007	0.024	0.016	0.027	0.032	0.041	0.036	0.034	0.037	0.037	0.067	0.045	0.032	0.029	0.052	0.030
40	Lactoferrin (mg/l)	0.000	0.000	0.000	0.001	0.001	0.000	0.070	0.000	0.015	0.000	0.000	0.000	0.000	0.001	0.001	0.001	0.000	0.001	0.000
41	Lysozyme (μg/l)	0.004	0.007	0.003	0.004	0.011	0.004	0.003	0.006	0.006	0.004	0.005	0.004	0.011	0.008	0.005	0.004	0.008	0.008	0.005

*The relative concentration of each metabolite is its integral value normalized to TSP, which was added to the NMR buffer as an internal standard.

Supplementary table 4-S4. Metabolites and their chemical shifts from 1H NMR data in breast milk samples*

	Metabolite	δ¹H (ppm)						
1	2′-Fucosyllactose	1.22(d)	4.24(dd)	5.22(d)	5.30(d)			
2	3-Fucosyllactose	1.17(d)	5.43(d)					
3	3′-Sialyllactose	1.79(t)	3.28(t)	4.53(d)				
4	6′-Sialyllactose	1.74(t)	4.42(d)					
5	Acetate	1.91(s)						
6	Acetone	2.22(s)						
7	Alanine	1.47(d)	3.78(q)					
8	Aspartate	2.67(dd)	2.80(dd)	3.89(dd)				
9	Betaine	3.25(s)	3.89(s)					
10	Butyrate	0.88(t)	1.54(m)	2.14(t)				
11	Caprate	0.86(t)	1.27(m)	1.53(m)	2.16(t)			
12	Caprylate	0.85(t)	1.28(m)	1.53(m)	2.16(t)			
13	Choline	3.19(s)	3.51(m)	4.06(m)				
14	Citrate	2.52(d)	2.66(d)					
15	Creatine	3.02(s)	3.92(s)					
16	Creatinine	3.03(s)	4.04(s)					
17	Formate	8.44(s)						
18	Fucose	1.21(d)	1.23(d)	4.54(d)	5.20(d)			
19	Fumarate	6.51(s)						
20	Galactose	3.48(dd)	3.98(dd)	4.57(d)	5.27(d)			
21	Glucose	3.23(dd)	3.39(t)	3.40(t)	3.48(t)	3.89(dd)	4.63(d)	5.23(d)
22	Glutamate	2.04(m)	2.12(m)	2.32(m)	2.36(m)			
23	Glutamine	2.11(m)	2.14(m)	2.42(m)	2.46(m)			
24	Hippurate	7.54(m)	7.62(m)	7.81(m)				
25	Histidine	7.06(s)	7.81(s)					
26	Hypoxanthine	8.17(s)	8.20(s)					
27	Isoleucine	0.93(t)	1.00(d)					
28	Lactate	1.32(d)	4.11(q)					
29	Lactose	3.28(t)	3.54(dd)	3.55(dd)	3.59(m)	3.95(dd)	5.22(d)	
30	Leucine	0.94(d)	0.95(d)	1.67(m)	1.70(m)			
31	Methionine	2.19(m)	2.63(t)					
32	2-Oxoglutarate	2.43(t)	3.00(t)					
33	o-Phosphocholine	3.21(s)	4.15(m)					
34	Phenylalanine	7.32(d)	7.36(t)	7.42(t)				
35	Succinate	2.39(s)						
36	Taurine	3.25(t)	3.41(t)					
37	Threonine	1.31(d)	4.25(m)					
38	Tyrosine	6.89(m)	7.18(m)					
39	Urea	5.78(s)						
40	Valine	0.97(d)	1.03(d)	2.26(m)				

*Key: s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; dd, doublet of doublet.

Supplementary table 4-S5. Concentration of the components in breast milk samples collected 1–5 months after delivery*

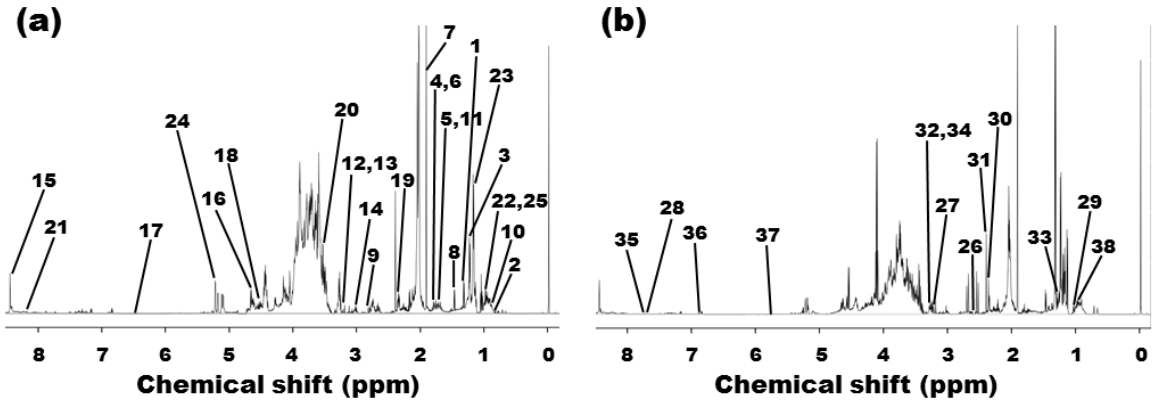
	Component	Month 1									Month 2								Month 3	
		Day 34	Day 36	Day 41	Day 46	Day 51	Day 53	Day 58-1	Day 58-2	Day 59	Day 61	Day 62	Day 66	Day 69	Day 72	Day 74	Day 76	Day 82	Day 99	Day 103
1	2′-Fucosyllactose (mM)	1.874	1.817	1.914	1.808	1.845	1.844	1.966	1.711	1.915	1.732	1.750	1.818	1.781	1.643	1.589	1.701	1.710	1.361	1.526
2	3-Fucosyllactose (mM)	0.343	0.341	0.410	0.305	0.335	0.286	0.390	0.407	0.368	0.296	0.317	0.335	0.312	0.267	0.267	0.355	0.285	0.339	0.313
3	3′-Sialyllactose (mM)	0.142	0.149	0.145	0.144	0.139	0.138	0.131	0.151	0.150	0.131	0.139	0.139	0.138	0.121	0.114	0.149	0.132	0.126	0.141
4	6′-Sialyllactose (mM)	0.155	0.149	0.136	0.113	0.094	0.070	0.080	0.069	0.095	0.066	0.070	0.056	0.071	0.054	0.055	0.062	0.042	0.035	0.040
5	Acetate (mM)	0.010	0.016	0.008	0.013	0.007	0.019	0.019	0.006	0.017	0.009	0.014	0.009	0.008	0.007	0.010	0.016	0.007	0.009	0.012
6	Acetone (mM)	0.006	0.005	0.004	0.009	0.006	0.005	0.008	0.005	0.005	0.006	0.005	0.007	0.008	0.006	0.006	0.006	0.006	0.005	0.006
7	Alanine (mM)	0.155	0.229	0.208	0.200	0.141	0.171	0.157	0.193	0.254	0.212	0.226	0.187	0.175	0.205	0.193	0.200	0.165	0.246	0.163
8	Aspartate (mM)	0.036	0.049	0.054	0.058	0.037	0.042	0.026	0.030	0.104	0.075	0.049	0.055	0.040	0.042	0.069	0.025	0.043	0.068	0.051
9	Betaine (mM)	0.091	0.090	0.089	0.090	0.090	0.091	0.094	0.088	0.095	0.092	0.089	0.087	0.086	0.090	0.078	0.095	0.089	0.084	0.083
10	Butyrate (mM)	0.053	0.098	0.120	0.125	0.107	0.037	0.368	0.065	0.279	0.105	0.126	0.056	0.131	0.077	0.085	0.259	0.115	0.134	0.181
11	Caprate (mM)	0.087	0.202	0.287	0.260	0.186	0.073	0.527	0.123	0.568	0.195	0.212	0.101	0.160	0.112	0.142	0.410	0.204	0.221	0.308
12	Caprylate (mM)	0.158	0.329	0.353	0.374	0.212	0.100	0.475	0.144	0.439	0.249	0.241	0.087	0.161	0.121	0.157	0.375	0.236	0.182	0.280
13	Choline (mM)	0.034	0.059	0.025	0.038	0.020	0.048	0.030	0.020	0.038	0.028	0.038	0.035	0.031	0.022	0.033	0.043	0.027	0.026	0.027
14	Citrate (mM)	2.294	1.380	2.698	1.648	1.298	1.079	2.568	2.115	2.376	1.430	1.427	1.380	2.358	2.118	1.793	2.292	1.301	1.748	1.779
15	Creatine (mM)	0.035	0.050	0.033	0.034	0.037	0.041	0.035	0.032	0.041	0.041	0.039	0.040	0.039	0.033	0.042	0.042	0.033	0.029	0.044
16	Creatinine (mM)	0.064	0.043	0.059	0.066	0.056	0.058	0.053	0.047	0.065	0.069	0.062	0.058	0.056	0.056	0.055	0.062	0.047	0.042	0.055
17	Formate (mM)	0.017	0.013	0.016	0.014	0.013	0.012	0.021	0.011	0.013	0.014	0.013	0.011	0.012	0.012	0.011	0.012	0.012	0.011	0.017
18	Fucose (mM)	0.323	0.269	0.315	0.296	0.362	0.299	0.347	0.422	0.384	0.384	0.318	0.453	0.370	0.413	0.335	0.336	0.428	0.447	0.470
19	Fumarate (mM)	0.002	0.004	0.002	0.002	0.001	0.002	0.002	0.001	0.006	0.002	0.002	0.002	0.001	0.002	0.001	0.001	0.002	0.002	0.001
20	Galactose (mM)	0.761	0.721	0.737	0.735	0.752	0.770	0.797	0.684	0.817	0.797	0.743	0.753	0.774	0.762	0.642	0.809	0.787	0.700	0.675
21	Glucose (mM)	1.613	1.991	1.691	2.038	1.745	2.570	1.261	1.732	2.475	1.692	2.233	1.770	1.872	1.479	1.678	1.402	1.044	1.551	1.306
22	Glutamate (mM)	1.052	1.174	1.134	1.330	1.096	1.393	0.993	1.136	1.682	1.173	1.281	1.443	1.122	1.190	1.480	1.105	0.927	1.146	1.159
23	Glutamine (mM)	0.345	0.370	0.374	0.387	0.539	0.447	0.376	0.614	0.677	0.514	0.383	0.613	0.617	0.654	0.632	0.535	0.445	0.524	0.422
24	Hippurate (mM)	0.007	0.011	0.016	0.004	0.007	0.009	0.017	0.009	0.006	0.007	0.011	0.006	0.009	0.006	0.006	0.006	0.006	0.004	0.003
25	Histidine (mM)	0.038	0.038	0.040	0.040	0.039	0.034	0.032	0.035	0.047	0.036	0.033	0.046	0.040	0.027	0.033	0.026	0.023	0.024	0.021
26	Hypoxanthine (mM)	0.008	0.005	0.002	0.007	0.003	0.007	0.004	0.003	0.006	0.004	0.005	0.002	0.003	0.003	0.004	0.005	0.003	0.002	0.003
27	Isoleucine (mM)	0.007	0.021	0.007	0.010	0.012	0.010	0.008	0.007	0.010	0.012	0.006	0.016	0.007	0.004	0.008	0.008	0.005	0.011	0.010
28	Lactate (mM)	0.047	0.151	0.049	0.057	0.039	0.049	0.044	0.042	0.103	0.048	0.057	0.118	0.036	0.041	0.030	0.079	0.046	0.084	0.044
29	Lactose (mM)	104.2	186.7	186.6	187.8	186.2	191.0	184.0	112.2	196.1	192.8	189.2	183.3	187.0	117.7	169.7	193.2	188.3	182.3	179.5
30	Leucine (mM)	0.018	0.041	0.020	0.023	0.029	0.026	0.018	0.019	0.029	0.032	0.017	0.041	0.024	0.015	0.027	0.020	0.016	0.018	0.020
31	Methionine (mM)	0.014	0.017	0.012	0.014	0.016	0.016	0.011	0.008	0.015	0.019	0.007	0.020	0.011	0.011	0.011	0.012	0.011	0.011	0.011
32	2-Oxoglutarate (mM)	0.035	0.058	0.034	0.058	0.025	0.065	0.040	0.021	0.059	0.032	0.055	0.043	0.027	0.025	0.037	0.039	0.029	0.036	0.032
33	O-Phosphocholine (mM)	0.889	0.750	0.863	0.895	0.766	0.785	0.717	0.737	0.839	0.827	0.705	0.682	0.734	0.594	0.601	0.716	0.672	0.534	0.556
34	Phenylalanine (mM)	0.012	0.019	0.012	0.011	0.015	0.018	0.011	0.013	0.015	0.017	0.012	0.019	0.013	0.008	0.012	0.009	0.008	0.007	0.008
35	Succinate (mM)	0.009	0.009	0.007	0.009	0.006	0.010	0.007	0.006	0.009	0.010	0.009	0.009	0.007	0.007	0.008	0.007	0.007	0.006	0.006
36	Taurine (mM)	0.352	0.440	0.430	0.451	0.465	0.503	0.487	0.374	0.633	0.493	0.533	0.574	0.431	0.387	0.433	0.494	0.377	0.371	0.437
37	Threonine (mM)	0.062	0.074	0.073	0.060	0.068	0.052	0.055	0.075	0.080	0.066	0.039	0.081	0.066	0.063	0.088	0.053	0.032	0.047	0.065
38	Tyrosine (mM)	0.010	0.011	0.021	0.009	0.012	0.013	0.006	0.012	0.016	0.020	0.008	0.020	0.008	0.006	0.015	0.008	0.006	0.008	0.010
39	Urea (mM)	2.428	2.183	2.118	2.442	2.754	2.376	2.318	2.245	2.319	3.428	1.801	2.241	2.232	2.025	2.238	1.584	2.415	2.223	2.307
40	Valine (mM)	0.040	0.051	0.047	0.042	0.058	0.044	0.044	0.051	0.053	0.056	0.037	0.069	0.050	0.037	0.048	0.041	0.030	0.037	0.040
41	Fat (g/100ml)	4.1	3.9	3.4	5.2	2.3	6.7	4.2	0.9	4.1	2.4	5.3	3.2	2.5	2.9	4.4	3.6	3.0	1.7	1.9
42	Protein (g/100ml)	0.9	1.0	0.9	0.9	0.8	0.6	0.9	0.9	1.0	0.9	0.7	0.7	0.4	0.9	0.8	0.9	0.7	0.7	0.3
43	Carbohydrate (g/100ml)	8.2	8.1	8.2	8.0	8.4	8.2	8.2	8.6	8.3	8.4	8.2	8.3	7.9	8.5	8.2	8.3	8.2	8.2	8.3
44	Total Solid (g/100ml)	13.5	13.3	12.9	14.5	11.8	15.8	13.6	10.8	13.7	12.0	14.5	12.7	11.1	12.5	13.8	13.2	12.2	10.9	10.7
45	Immunoglobulin A (g/l)	0.558	0.621	0.575	0.589	0.576	0.597	0.596	0.472	0.692	0.585	0.578	0.704	0.631	0.524	0.608	0.645	0.588	0.669	0.671
46	Lactoferrin (g/l)	7.495	7.222	6.208	6.095	5.115	4.562	5.686	5.010	6.246	6.242	5.468	5.897	6.361	5.330	5.499	6.122	6.146	6.672	6.445
47	Lysozyme (g/l)	0.037	0.036	0.030	0.104	0.029	0.035	0.041	0.042	0.035	0.208	0.039	0.065	0.060	0.063	0.072	0.074	0.077	0.143	0.095

*The relative concentration of each metabolite is its integral value normalized to TSP, which was added to the NMR buffer as an internal standard.

Supplementary table 4-S5, continued*

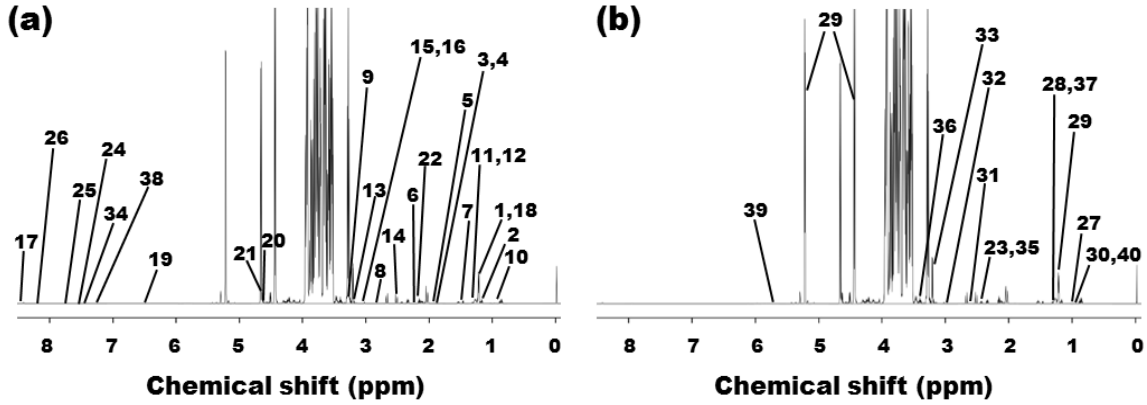
	Component	Month 3					Month 4								Month 5					
		Day 104	Day 108	Day 110	Day 114	Day 118	Day 122	Day 124	Day 131	Day 134	Day 137	Day 142	Day 145	Day 148	Day 155	Day 157	Day 158	Day 159	Day 161	Day 163
1	2′-Fucosyllactose (mM)	1.777	1.854	1.899	1.958	1.499	1.591	1.344	1.801	1.648	1.544	1.622	1.761	1.527	1.696	1.802	1.487	1.534	1.576	1.744
2	3-Fucosyllactose (mM)	0.344	0.366	0.372	0.547	0.314	0.425	0.322	0.333	0.378	0.342	0.347	0.408	0.317	0.249	0.334	0.257	0.306	0.334	0.379
3	3′-Sialyllactose (mM)	0.150	0.165	0.159	0.175	0.163	0.130	0.131	0.135	0.135	0.133	0.136	0.143	0.131	0.121	0.144	0.129	0.137	0.121	0.122
4	6′-Sialyllactose (mM)	0.039	0.052	0.056	0.040	0.035	0.041	0.029	0.038	0.037	0.040	0.039	0.031	0.031	0.032	0.030	0.020	0.025	0.033	0.033
5	Acetate (mM)	0.008	0.013	0.024	0.020	0.022	0.012	0.009	0.021	0.021	0.014	0.016	0.013	0.024	0.018	0.022	0.023	0.029	0.018	0.017
6	Acetone (mM)	0.007	0.004	0.005	0.004	0.011	0.005	0.005	0.005	0.004	0.006	0.006	0.003	0.005	0.008	0.007	0.012	0.007	0.004	0.004
7	Alanine (mM)	0.136	0.199	0.242	0.221	0.156	0.218	0.216	0.266	0.166	0.244	0.183	0.220	0.278	0.231	0.159	0.105	0.181	0.407	0.196
8	Aspartate (mM)	0.032	0.069	0.064	0.060	0.027	0.097	0.049	0.153	0.025	0.055	0.048	0.060	0.045	0.075	0.022	0.021	0.074	0.091	0.047
9	Betaine (mM)	0.093	0.090	0.083	0.098	0.088	0.086	0.086	0.091	0.083	0.093	0.092	0.094	0.091	0.090	0.096	0.090	0.090	0.088	0.099
10	Butyrate (mM)	0.107	0.237	0.397	0.392	0.326	0.183	0.097	0.309	0.343	0.203	0.300	0.144	0.430	0.293	0.357	0.293	0.459	0.270	0.306
11	Caprate (mM)	0.187	0.426	0.729	0.611	0.405	0.371	0.211	0.456	0.632	0.354	0.433	0.311	0.803	0.611	0.575	0.452	0.761	0.672	0.572
12	Caprylate (mM)	0.170	0.357	0.499	0.415	0.277	0.285	0.217	0.583	0.419	0.294	0.292	0.309	0.406	0.359	0.414	0.374	0.319	0.361	0.404
13	Choline (mM)	0.027	0.045	0.059	0.042	0.039	0.033	0.027	0.053	0.052	0.040	0.071	0.047	0.070	0.037	0.067	0.057	0.067	0.055	0.056
14	Citrate (mM)	1.543	1.459	1.360	1.835	1.906	1.626	1.569	1.377	1.725	1.917	2.138	1.849	1.717	1.495	1.536	1.884	1.841	1.931	1.771
15	Creatine (mM)	0.039	0.055	0.064	0.043	0.043	0.048	0.031	0.046	0.041	0.042	0.048	0.062	0.054	0.046	0.041	0.041	0.059	0.052	0.049
16	Creatinine (mM)	0.050	0.054	0.059	0.047	0.057	0.050	0.039	0.058	0.045	0.046	0.041	0.055	0.049	0.062	0.047	0.048	0.051	0.048	0.044
17	Formate (mM)	0.016	0.016	0.018	0.015	0.017	0.019	0.018	0.017	0.020	0.014	0.016	0.015	0.014	0.016	0.011	0.017	0.014	0.012	0.013
18	Fucose (mM)	0.545	0.527	0.479	0.437	0.469	0.422	0.441	0.300	0.388	0.444	0.432	0.536	0.393	0.395	0.378	0.388	0.407	0.344	0.381
19	Fumarate (mM)	0.001	0.003	0.004	0.002	0.002	0.002	0.002	0.007	0.003	0.002	0.001	0.003	0.005	0.002	0.004	0.003	0.002	0.006	0.003
20	Galactose (mM)	0.807	0.734	0.723	0.848	0.774	0.690	0.715	0.895	0.660	0.756	0.751	0.734	0.755	0.821	0.780	0.722	0.762	0.703	0.778
21	Glucose (mM)	1.322	1.561	1.500	1.230	1.454	1.955	1.346	2.433	1.253	1.826	1.551	1.604	1.948	1.825	1.284	1.465	1.412	2.512	1.546
22	Glutamate (mM)	1.204	1.508	1.508	1.225	1.328	1.279	1.135	1.625	1.184	1.219	1.130	1.311	1.205	1.299	1.248	1.193	1.133	1.186	1.194
23	Glutamine (mM)	0.546	0.488	0.322	0.468	0.508	0.533	0.568	0.669	0.558	0.708	0.468	0.577	0.500	0.511	0.674	0.636	0.673	0.695	0.770
24	Hippurate (mM)	0.010	0.013	0.006	0.006	0.017	0.010	0.008	0.003	0.012	0.003	0.005	0.011	0.005	0.004	0.006	0.008	0.004	0.009	0.012
25	Histidine (mM)	0.024	0.032	0.032	0.026	0.027	0.021	0.021	0.029	0.030	0.028	0.021	0.028	0.025	0.036	0.034	0.030	0.031	0.021	0.029
26	Hypoxanthine (mM)	0.003	0.004	0.002	0.003	0.004	0.003	0.002	0.004	0.004	0.003	0.003	0.004	0.005	0.006	0.006	0.006	0.004	0.005	0.006
27	Isoleucine (mM)	0.013	0.010	0.032	0.009	0.010	0.008	0.006	0.012	0.010	0.007	0.012	0.015	0.009	0.007	0.011	0.009	0.014	0.006	0.006
28	Lactate (mM)	0.025	0.048	0.157	0.049	0.066	0.046	0.054	0.068	0.072	0.052	0.055	0.072	0.171	0.066	0.080	0.062	0.074	0.130	0.048
29	Lactose (mM)	190.2	187.9	172.5	203.9	188.3	185.3	179.7	171.1	184.4	194.5	195.1	89.0	198.6	189.0	197.4	186.5	198.1	205.0	122.1
30	Leucine (mM)	0.030	0.033	0.059	0.023	0.027	0.020	0.016	0.028	0.029	0.024	0.024	0.031	0.020	0.020	0.030	0.027	0.032	0.015	0.024
31	Methionine (mM)	0.013	0.016	0.015	0.012	0.013	0.012	0.010	0.012	0.010	0.012	0.013	0.014	0.009	0.011	0.015	0.013	0.014	0.009	0.014
32	2-Oxoglutarate (mM)	0.026	0.039	0.074	0.040	0.052	0.031	0.036	0.022	0.061	0.041	0.038	0.046	0.075	0.045	0.062	0.081	0.056	0.079	0.049
33	O-Phosphocholine (mM)	0.668	0.629	0.615	0.614	0.623	0.609	0.517	0.530	0.538	0.614	0.572	0.523	0.615	0.756	0.645	0.569	0.606	0.561	0.614
34	Phenylalanine (mM)	0.011	0.015	0.023	0.014	0.011	0.011	0.008	0.015	0.015	0.012	0.012	0.014	0.009	0.012	0.015	0.011	0.020	0.010	0.014
35	Succinate (mM)	0.006	0.007	0.008	0.006	0.009	0.005	0.006	0.004	0.006	0.005	0.005	0.005	0.006	0.007	0.009	0.010	0.008	0.009	0.009
36	Taurine (mM)	0.468	0.685	0.676	0.479	0.430	0.497	0.379	0.648	0.469	0.585	0.533	0.618	0.575	0.509	0.564	0.620	0.637	0.613	0.589
37	Threonine (mM)	0.076	0.084	0.091	0.049	0.069	0.076	0.042	0.074	0.072	0.073	0.064	0.072	0.045	0.069	0.090	0.064	0.080	0.066	0.082
38	Tyrosine (mM)	0.008	0.015	0.024	0.012	0.008	0.005	0.005	0.010	0.011	0.009	0.009	0.013	0.006	0.010	0.011	0.006	0.015	0.005	0.011
39	Urea (mM)	2.753	2.439	2.677	2.498	1.637	2.163	1.936	1.514	1.922	1.921	1.900	2.259	2.405	2.387	2.094	2.498	2.000	1.737	1.947
40	Valine (mM)	0.055	0.051	0.074	0.034	0.047	0.042	0.028	0.042	0.045	0.040	0.041	0.051	0.034	0.035	0.043	0.039	0.044	0.025	0.045
41	Fat (g/100ml)	1.2	3.6	6.7	4.0	3.2	2.0	1.0	3.4	4.1	1.8	3.3	1.9	3.7	2.9	3.9	3.0	4.1	2.9	3.4
42	Protein (g/100ml)	0.7	0.7	0.6	0.6	0.7	0.7	0.7	0.5	0.7	0.7	0.6	0.6	0.7	1.0	0.8	0.9	0.7	1.0	0.9
43	Carbohydrate (g/100ml)	8.4	8.2	7.7	8.3	8.2	8.4	8.5	6.9	8.0	8.1	8.0	8.3	7.9	8.1	8.0	8.2	8.0	8.3	8.3
44	Total Solid (g/100ml)	10.6	12.9	15.3	13.2	12.4	11.4	10.5	11.1	13.2	11.0	12.2	11.1	12.7	12.2	13.0	12.2	13.0	12.3	12.9
45	Immunoglobulin A (g/l)	0.553	1.057	1.181	0.719	0.759	0.585	0.484	0.669	1.210	0.703	0.656	0.733	0.719	0.684	0.640	0.596	0.636	0.691	0.683
46	Lactoferrin (g/l)	5.123	6.829	9.000	5.042	7.545	5.165	4.523	5.328	10.38	4.950	5.196	6.685	6.253	7.122	5.862	14.81	5.332	5.344	5.247
47	Lysozyme (g/l)	0.090	0.104	0.105	0.102	0.169	0.094	0.072	0.060	0.218	0.108	0.136	0.125	0.094	0.106	0.141	0.295	0.157	0.082	0.081

*The relative concentration of each metabolite is its integral value normalized to TSP, which was added to the NMR buffer as an internal standard.



Supplementary figure 4-S1. ^{1}H NMR spectra of infant feces samples at 1 and 5 months after delivery.

Typical 600 MHz ^{1}H NMR spectra of infant feces samples at months (a) 1 and (b) 5. Identified metabolites: 1: 2'-Fucosyllactose; 2: 2-Hydroxyisovalerate; 3: 3-Fucosyllactose; 4: 3'-Sialyllactose; 5: 5-Aminopentanoate; 6: 6'-Sialyllactose; 7: Acetate; 8: Alanine; 9: Aspartate; 10: Butyrate; 11: Cadaverine; 12: Carnitine; 13: Choline; 14: Creatine; 15: Formate; 16: Fucose; 17: Fumarate; 18: Glucose; 19: Glutamate; 20: Glycine; 21: Hypoxanthine; 22: Isoleucine; 23: Lactate; 24: Lactose; 25: Leucine; 26: Methionine; 27: o-Phosphocholine; 28: Phenylalanine; 29: Propionate; 30: Pyruvate; 31: Succinate; 32: Taurine; 33: Threonine; 34: Trimethylamine N-oxide; 35: Tryptophan; 36: Tyrosine; 37: Urea; 38: Valine.



Supplementary figure 4-S2. ^{1}H NMR spectra of breast milk samples at 1 and 5 months after delivery.

Typical 600 MHz ^{1}H NMR spectra of breast milk samples at months (**a**) 1 and (**b**) 5. Identified metabolites: 1: 2′-Fucosyllactose; 2: 3-Fucosyllactose; 3: 3′-Sialyllactose; 4: 6′-Sialyllactose; 5: Acetate; 6: Acetone; 7: Alanine; 8: Aspartate; 9: Betaine; 10: Butyrate; 11: Caprate; 12: Caprylate; 13: Choline; 14: Citrate; 15: Creatine; 16: Creatinine; 17: Formate; 18: Fucose; 19: Fumarate; 20: Galactose; 21: Glucose; 22: Glutamate; 23: Glutamine; 24: Hippurate; 25: Histidine; 26: Hypoxanthine; 27: Isoleucine; 28: Lactate; 29: Lactose; 30: Leucine; 31: Methionine; 32: 2-Oxoglutarate; 33: o-Phosphocholine; 34: Phenylalanine; 35: Succinate; 36: Taurine; 37: Threonine; 38: Tyrosine; 39: Urea; 40: Valine.

4.8 References

[1] Carlisle EM, Morowitz MJ. The intestinal microbiome and necrotizing enterocolitis. Curr Opin Pediatr. 2013;25:382–7.

[2] Torrazza RM, Neu J. The altered gut microbiome and necrotizing enterocolitis. Clin Perinatol. 2013;40:93–108.

[3] Lawson MAE, O´Neill IJ, Kujawska M, Javvadi SG, Wijeyesekera A, Flegg F, et al. Breast milk-derived human milk oligosaccharides promote *Bifidobacterium* interactions within a single ecosystem. ISME J. 2020;14:635–48.

[4] Nagpal R, Tsuji H, Takahashi T, Nomoto K, Kawashima K, Nagata S, et al. Ontogenesis of the gut microbiota composition in healthy, full-term, vaginally born and breast-fed infants over the first 3 years of life: A quantitative bird's-eye view. Front Microbiol. 2017;8:1–9.

[5] Kramer MS, Kakuma R. Optimal duration of exclusive breastfeeding (Review). Cochrane Database Syst Rev. 2002;1:CD003517.

[6] Garofalo R. Cytokines in human milk. J Pediatr. 2010;156:S36–40.

[7] Donovan SM, Odle J. Growth factors in milk as mediators of infant development. Annu Rev Nutr. 1994;14: 147–67.

[8] Hennart PF, Brasseur DJ, Delogne-Desnoeck JB, Dramaix MM, Robyn CE. Lysozyme, lactoferrin, and secretory immunoglobulin A content in breast milk: Influence of duration of lactation, nutrition status, prolactin status, and parity of mother. Am J Clin Nutr. 1991;53:32–9.

[9] Miralles O, Sánchez J, Palou A, Picó CA. Physiological role of breast milk leptin in body weight control in developing infants. Obesity. 2006;14:1371–7.

[10] Feldman R, Eidelman AI. Direct and indirect effects of breast milk on the neurobehavioral and cognitive development of premature infants. Dev Psychobiol. 2003;43:109–19.

[11] Ten-Doménech I, Ramos-Garcia V, Piñeiro-Ramos JD, Gormaz M, Parra-Llorca A, Vento M, et

al. Current practice in untargeted human milk metabolomics. Metabolites. 2020;22:43.

[12] Bravi F, Wiens F, Decarli A, Dal Pont A, Agostoni C, Ferraroni M. Impact of maternal nutrition on breast-milk composition: A systematic review. Am J Clin Nutr. 2016;104:646–62.

[13] Verduci E, Banderali G, Barberi S, Radaelli G, Lops A, Betti F, et al. Epigenetic effects of human breast milk. Nutrients. 2014;6:1711–24.

[14] Lewis ZT, Totten SM, Smilowitz JT, Popovic M, Parker E, Lemay DG, et al. Maternal fucosyltransferase 2 status affects the gut bifidobacterial communities of breastfed infants. Microbiome. 2015;3:15–7.

[15] Doare K, Le Holder B, Bassett A, Pannaraj PS. Mother's Milk: A purposeful contribution to the development of the infant microbiota and immunity. Front Immunol. 2018;9:361.

[16] Cai X, Duan Y, Li Y, Wang J, Mao Y, Yang Z, et al. Lactoferrin level in breast milk: a study of 248 samples from eight regions in China. Food Funct. 2018;9:4216–22.

[17] Lönnerdal B. Nutritional and physiologic significance of human milk proteins. Am J Clin Nutr. 2003;77:1537S–43S.

[18] Hennart PF, Brasseur DJ, Delogne-Desnoeck JB, Dramaix MM, Robyn CE. Lysozyme, lactoferrin, and secretory immunoglobulin A content in breast milk: influence of duration of lactation, nutrition status, prolactin status, and parity of mother. Am J Clin Nutr. 1991;53:32–9.

[19] Fukuda S, Toh H, Hase K, Oshima K, Nakanishi Y, Yoshimura K, et al. Bifidobacteria can protect from enteropathogenic infection through production of acetate. Nature. 2011;469:543–9.

[20] Maslowski KM, Vieira AT, Ng A, Kranich J, Sierro F, Di Yu, et al. Regulation of inflammatory responses by gut microbiota and chemoattractant receptor GPR43. Nature. 2009;461:1282–6.

[21] Kimura I, Inoue D, Maeda T, Hara T, Ichimura A, Miyauchi S, et al. Short-chain fatty acids and ketones directly regulate sympathetic nervous system via G protein-coupled receptor 41 (GPR41). Proc Natl Acad Sci U S A. 2011;108:8030–5.

[22] Komatsu Y, Shimizu Y, Yamano M, Kikuchi M, Nakamura K, Ayabe T, et al. Disease progression-associated alterations in fecal metabolites in SAMP1/YitFc mice, a Crohn's disease model. Metabolomics. 2020;16:48.

[23] Horigome A, Okubo R, Hamazaki K, Kinoshita T, Katsumata N, Uezono Y, et al. Association between blood omega-3 polyunsaturated fatty acids and the gut microbiota among breast cancer survivors. Benef Microbes. 2019;10:751–8.

[24] Komatsu Y, Wada Y, Izumi H, Shimizu T, Takeda Y, Aizawa T. ^{1}H NMR metabolomic and transcriptomic analyses reveal urinary metabolites as biomarker candidates in response to protein undernutrition in adult rats. Br J Nutr. 2020;in press https://doi.org/10.1017/S0007114520003281.

[25] Narasimhan R, Coras R, Rosenthal SB, Sweeney SR, Lodi A, Tiziani S, et al. Serum metabolomic profiling predicts synovial gene expression in rheumatoid arthritis. Arthritis Res Ther. 2018;20:1–11.

[26] Commare CE, Tappenden KA. Development of the infant intestine: Implications for nutrition support. Nutr Clin Pract. 2007;22:159–73.

[27] Fan W, Huo G, Li X, Yang L, Duan C. Impact of diet in shaping gut microbiota revealed by a comparative study in infants during the first six months of life. J Microbiol Biotechnol. 2014;24:133–43.

[28] Bode L. The functional biology of human milk oligosaccharides. Early Hum Dev. 2015;91:619–22.

[29] Smilowitz JT, O'Sullivan A, Barile D, German JB, Lönnerdal B, Slupsky CM. The human milk metabolome reveals diverse oligosaccharide profiles. J Nutr. 2013;143:1709–18.

[30] Salli K, Anglenius H, Hirvonen J, Hibberd AA, Ahonen I, Saarinen MT, et al. The effect of 2′-fucosyllactose on simulated infant gut microbiome and metabolites; a pilot study in comparison to GOS and lactose. Sci Rep. 2019;9:1–15.

[31] Bellamy W, Takase M, Yamauchi K, Wakabayashi H, Kawase K, Tomita M. Identification of the bactericidal domain of lactoferrin. Biochim Biophys Acta. 1992;1121:130–6.

[32] Liepke C, Adermann K, Raida M, Mägert H, Forssmann W, Zucht H. Human milk provides peptides highly stimulating the growth of bifidobacteria. Eur J Biochem. 2002;269:712–8.

[33] Maruyama K, Hida M, Kohgo T, Fukunaga Y. Changes in salivary and fecal secretory IgA in infants under different feeding regimens. Pediatr Int. 2009;51:342–5.

[34] Ibrahim HR, Inazaki D, Abdou A, Aoki T, Kim M. Processing of lysozyme at distinct loops by pepsin: a novel action for generating multiple antimicrobial peptide motifs in the newborn stomach. Biochim Biophys Acta. 2005;1726:102–14.

[35] Mastromarino P, Capobianco D, Campagna G, Laforgia N, Drimaco P, Dileone A, et al. Correlation between lactoferrin and beneficial microbiota in breast milk and infant's feces. Biometals. 2014;27:1077–86.

[36] Minami J, Odamaki T, Hashikura N, Abe F, Xiao JZ. Lysozyme in breast milk is a selection factor for bifidobacterial colonisation in the infant intestine. Benef Microbes. 2016;7:53–60.

[37] Planer JD, Peng Y, Kau AL, Blanton LV, Ndao IM, Tarr PI, et al. Development of the gut microbiota and mucosal IgA responses in twins and gnotobiotic mice. Nature. 2016;534:263–6.

[38] Sela DA, Chapman J, Adeuya A, Kim JH, Chen F, Whitehead TR, et al. The genome sequence of Bifidobacterium longum subsp. infantis reveals adaptations for milk utilization within the infant microbiome. Proc Natl Acad Sci U S A. 2008;105:18964–9.

[39] Sakanaka M, Hansen ME, Gotoh A, Katoh T, Yoshida K, Odamaki T, et al. Evolutionary adaptation in fucosyllactose uptake systems supports bifidobacteria-infant symbiosis. Sci Adv. 2019;5:eaaw7696.

[40] Kansandee W, Moonmangmee D, Moonmangmee S, Itsaranuwat P. Characterization and Bifidobacterium sp. growth stimulation of exopolysaccharide produced by Enterococcus faecalis

EJRM152 isolated from human breast milk. Carbohydr Polym. 2019;206:102–9.

[41] Wang S, Hibberd ML, Pettersson S, Lee YK. Enterococcus faecalis from healthy infants modulates inflammation through MAPK signaling pathways. PLoS One. 2014;9:e97523.

[42] Björkstén B, Sepp E, Julge K, Voor T, Mikelsaar M. Allergy development and the intestinal microflora during the first year of life. J Allergy Clin Immunol. 2001;108:516–20.

[43] Morrow AL, Ruiz-Palacios GM, Jiang X, Newburg DS. Human-milk glycans that inhibit pathogen binding protect breast-feeding infants against infectious diarrhea. J Nutr. 2005;135:1304–7.

[44] Martín-Sosa S, Martín MJ, Hueso P. The sialylated fraction of milk oligosaccharides is partially responsible for binding to enterotoxigenic and uropathogenic Escherichia coli human strains. J Nutr. 2002;132:3067–72.

[45] Clark DP. The fermentation pathways of Escherichia coli. FEMS Microbiol Lett. 1989;63:223–34.

[46] Alcon-Giner C, Dalby MJ, Caim S, Ketskemety J, Shaw A, Sim K, et al. Microbiota Supplementation with Bifidobacterium and Lactobacillus Modifies the Preterm Infant Gut Microbiota and Metabolome: An Observational Study. Cell Reports Med. 2020;1:100077.

[47] Chai L, Lu Z, Zhang X, Ma J, Xu P, Qian W, et al. Zooming in on butyrate-producing Clostridial consortia in the fermented grains of baijiu via gene sequence-guided microbial isolation. Front Microbiol. 2019;10:1397.

[48] Inan MS, Rasoulpour RJ, Yin L, Hubbard AK, Rosenberg DW, Giardina C. The luminal short-chain fatty acid butyrate modulates NF-κB activity in a human colonic epithelial cell line. Gastroenterology. 2000;118:724–34.

[49] Säemann MD, Böhmig GA, Österreicher CH, Burtscher H, Parolini O, Diakos C, et al. Anti-inflammatory effects of sodium butyrate on human monocytes: potent inhibition of IL-12 and up-

regulation of IL-10 production. FASEB J. 2000;14:2380–2.

[50] Furusawa Y, Obata Y, Fukuda S, Endo TA, Nakato G, Takahashi D, et al. Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. Nature. 2013;504:446–50.

[51] Waller AP, Geor RJ, Spriet LL, Heigenhauser GJF, Lindinger MI. Oral acetate supplementation after prolonged moderate intensity exercise enhances early muscle glycogen resynthesis in horses. Exp Physiol. 2009;94:888–98.

[52] Ferreira TM, Leonel AJ, Melo MA, Santos RRG, Cara DC, Cardoso VN, et al. Oral supplementation of butyrate reduces mucositis and intestinal permeability associated with 5-fluorouracil administration. Lipids. 2012;47:669–78.

[53] Butte NF. Energy requirements of infants. Public Health Nutr. 2005;8:953–67.

[54] Stickland LH. Studies in the Metabolism of the Strict Anaerobes (Genus Clostridium). II The Reduction of Proline by Cl. sporogenes. Biochem J. 1935;29:288–90.

[55] Lee Y, Khan A, Hong S, Jee, S. H. & Park, Y. H. A metabolomic study on high-risk stroke patients determines low levels of serum lysine metabolites: A retrospective cohort study. Mol Biosyst. 2017;13:1109–20.

[56] Castellote C, Casillas R, Ramírez-Santana C, Pérez-Cano FJ, Castell M, Moretones MG, et al. Premature delivery influences the immunological composition of colostrum and transitional and mature human milk. J Nutr. 2011;141:1181–7.

Concluding remarks

NMR metabolomics was applied to the fields of pathology and nutrition, search for reliable biomarker candidates that reflect pathological and nutritional conditions, and further elucidate the mechanism to fluctuate these metabolites. The results of each chapter are summarized below.

1. Fecal metabolites, such as SCFAs, glucose, xylose, taurine, and choline, dramatically fluctuated with disease progression, and that the fluctuations in SAMP1/YitFc mice could be partially influenced by alterations in the gut microbiome. Several fecal metabolites of SAMP1/YitFc mice had already fluctuated at the very early stage of CD before the histological changes associated with disease progression became apparent. Those fecal metabolites showing characteristic fluctuations may be potential candidate biomarkers for predicating disease progression.

2. The fluctuation of several metabolites identified in urine, plasma, and liver samples were caused by the LP diet ingestion in adult rats. Among the identified metabolites, taurine and TMAO, the final metabolites that are not further metabolized, fluctuated not only in urine but also in the liver. Notably, changes in taurine and TMAO levels were detectable in the urine by week 2 after LP diet ingestion, which may reflect protein undernutrition at an early stage. Hepatic gene expressions of taurine and TMAO biosynthesis enzymes also corresponded to these fluctuations. Thus, fluctuations in urinary taurine and TMAO are considered to have occurred as a result of substantially responses to protein undernutrition in the liver. Furthermore, taurine and TMAO might also be potential mediators of the development of skeletal muscle disorders and lipid abnormalities, which could be associated with LP nutritional status. It follows that taurine and TMAO may serve as potential biomarker candidates of protein undernutrition and targets for

prevention of secondary diseases associated with protein undernutrition.

3. HMOs in the breast milk did not fluctuate throughout the lactation period, but disappeared from the infant feces after the fourth lactation month. Moreover, at the same time a species switching from *Bifidobacterium breve* to *Bifidobacterium longum* subsp. *infantis*, *Escherichia coli*, *Clostridium neonatale*, and *Clostridioides mangenotii* occurred, accompanied by an increase in the levels of metabolites such as acetate, and butyrate, in the infant feces. Importantly, in line with previous reports, our data suggest that the changes in metabolites in the infant feces might be linked to benefits such as maturation of immune function and protection against infection.

The expansion of NMR metabolomics into the pathological and nutritional fields could contribute to the treatment and prevention of CD and protein malnutrition. However, clinical trials will be required to utilize the metabolites detected in this study as reliable biomarkers. Furthermore, we have shown that metabolites in breast milk can be associated with gut microbiome in infants. This study could help understand the relationship between breast milk metabolites and infant gut microbiome, and may contribute to the development of nutritional and functional foods for infants. I will be very pleased if this research could lead to further industrial application of metabolomics and contribute to the development of science.

List of papers related to this study

[1] **Komatsu Y**, Shimizu Y, Yamano M, Kikuchi M, Nakamura K, Ayabe T, Aizawa T. Disease progression-associated alterations in fecal metabolites in SAMP1/YitFc mice, a Crohn's disease model. Metabolomics. 2020;16:48.

[2] **Komatsu Y**, Wada Y, Izumi H, Shimizu T, Takeda Y, Aizawa T. ^{1}H NMR metabolomic and transcriptomic analyses reveal urinary metabolites as biomarker candidates in response to protein undernutrition in adult rats. Br J Nutr. 2021;125:633–43.

[3] **Komatsu Y**, Kumakura D, Seto N, Izumi H, Takeda Y, Nakaoka S, Aizawa T. Association of milk components with fecal microbiome and metabolites in a mother-infant dyad. *Pediatr Res*. 2021; *Submitted*.

Acknowledgements

Firstly, I would like to thank Professor Tomoyasu Aizawa (Graduate School of Life Science, Hokkaido University) for giving me the opportunity of starting NMR metabolomics research, for his supervision throughout this study and for judging of this dissertation.

I would also like to express my deepest appreciation to Professor Makoto Demura and Professor Hisashi Haga for accepting the judging of this dissertation despite their busy schedules.

I would also like to express my deep gratitude to Professor Tokiyoshi Ayabe, Dr. Kiminori Nakamura, and Dr. Shinji Nakaoka who provided valuable suggestions and discussion in this study.

I would also like to thank Dr. Hiroyuki Kumeta, Dr. Yasuhiro Kumaki, and Dr. Yuki Ohnishi (Open Facility Division, Global Facility Center, Creative Research Institution, Hokkaido University), for performing the NMR analysis using an NMR spectrometer and for providing insight and expertise that greatly assisted this research. I would also like to thank laboratory members in Hokkaido University, namely, Dr. Yu Shimizu, Ms. Megumi Yamano, Dr. Mani Kikuchi, and Mr. Daiki Kumakura for supporting the animal experiment and NMR analysis.

I would like to express my sincere gratitude to Michio Miyahara (President and Representative Director, Morinaga Milk Industry Co., Ltd.), Teiichiro Okawa (Executive Vice President and Representative Director, Morinaga Milk Industry Co., Ltd.), Fumiaki Abe (Managing Executive Officer and Chief Director of Research Division and Director of Food Ingredients & Technology Institute, Morinaga Milk Industry Co., Ltd.), Yasuhiro Takeda (Director of Food Research & Development Institute, Morinaga Milk Industry Co., Ltd.), and Kazuhiro Miyaji (Director of Wellness & Nutrition Science Institute, Morinaga Milk Industry Co., Ltd.) for their deep understanding and support for this research.

I would also like to express my deep gratitude to Takashi Shimizu (Deputy Director of Next Generation Science Institute, Morinaga Milk Industry Co., Ltd.), Hirohiko Nakamura (Head of Health

& Nutrition Research Group, Morinaga Milk Industry Co., Ltd.), Dr. Hirohisa Izumi (Health & Nutrition Research Group, Morinaga Milk Industry Co., Ltd.), Dr. Tatsuya Ehara (Health & Nutrition Research Group, Morinaga Milk Industry Co., Ltd.), and Dr. Yasuaki Wada (Health & Nutrition Research Group, Morinaga Milk Industry Co., Ltd.) for their great understanding and cooperation in conducting this research and preparing this paper.

I also wish to express thanks to all of my colleagues at Hokkaido University and Morinaga Milk Industry Co., Ltd. for their supports and encouragements.

Finally, I really appreciated my parents, Tadamitsu and Naoko and my family members, Yutaro, Shinjiro, and Yumi encouraging me to go through this PhD process and providing their supports at any step.

Yosuke Komatsu

March, 2021