

An Analysis of the proteomics approach to the glycosylated peptides of human milk

Seonghyeon Cho¹, Jong-Moon Park¹, Hookeun Lee¹, Jun Hwan Song², and Nam Mi Kang³ ★

¹College of Pharmacy, Gachon University, Hambakmoe-ro, Incheon 21936, Korea

²Department of Pediatrics, Soonchunhyang University, Suncheonhyang 6-gil, Cheonan 31151, Korea

³Department of Nursing, Konkuk University, Chungwondae-ro, Chungju 27478, Korea

(Received October 25, 2021; Revised December 30, 2021; Accepted January 13, 2022)

Abstract: Many studies have shown that advanced glycation end-products (AGEs) and glycation adducts are significantly linked to aging and disease. Particularly, the level of glycation in human milk is important because the AGE intake is closely related to AGE levels in infants. In this study, we used human milk samples obtained from four primiparae and four multiparae. We isolated proteins using acetone and trichloroacetic acid (TCA) precipitation. A total of 67 glycosylated proteins and 122 glycosylated peptides was quantified; among them, 19 glycosylated peptides were differentially expressed. We confirmed that the degree of glycation differed according to fertility. The study provides a foundation for using proteomics to evaluate the mother's milk quality and link between maternal health and human milk quality.

Key words: glycation, advanced glycation end-product, human milk, LC-MS/MS, fertility

1. Introduction

Milk is important for the health of infants because it contains nutrients, immune-related components, anti-infective factors, and metabolic enzymes, which greatly impact growth and development.¹ The composition of human milk is dynamic and depends on various environmental factors such as breastfeeding and the mother's diet.²

The protein concentration in human milk is 8.5 g/L, which is lower than the 34.5 g/L found in bovine milk.³ The protein ratio of whey: casein in human milk is between 50:50 and 80:20.⁴ Compared to in

bovine milk, the percentage of whey protein and ratio of carbohydrate in human milk are high.⁵ Thus, human milk is a better environment for glycation compared to bovine milk.

Glycation is a spontaneous post-translational modification that occurs at lysine, arginine, and the N-terminus of proteins.⁶ The active carbonyl group of the sugar and nucleophilic free amino group chemically react to form an unstable Schiff base. The Schiff base is rearranged to form Amadori compounds that form advanced glycation end-products (AGEs) through several reactions.⁷

Levels of serum AGEs between the mother and

★ Corresponding author

Phone : +82-(0)2-450-3033 Fax : +82-(0)450-4063

E-mail : nmkang03@kku.ac.kr

This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

infant are strongly correlated. In addition, increasing the intake of dietary AGEs also increases the level of serum AGEs in infants.⁸ In many studies, it has been confirmed that toxic substances such as dichlorodiphenyltrichloroethane, perfluorinated compounds, and polychlorinated biphenyls of human milk come from multiparae more than primiparae.⁹⁻¹¹

In a previous study, the automated boronate affinity enrichment method has been applied to bovine and human serum albumin using a liquid handling robotic system.¹² In this study, we removed *N*-glycan that interfered with glycosylated peptide enrichment by performing PNGase F digestion during the filter-aided sample preparation (FASP) method. Glycosylated peptides from human milk were enriched and analyzed using liquid chromatography-tandem mass spectrometry (LC-MS/MS)-based label-free quantification. We analyzed glycosylated peptides in human milk and their relationship with fertility.

2. Experimental

2.1. Chemicals and reagents

Iodoacetamide, ammonium bicarbonate (ABC), $MgCl_2$, urea, Tris(hydroxymethyl)aminomethane (Tris), trichloroacetic acid (TCA), and ammonium acetate (NH_4OAc) were purchased from Sigma-Aldrich (St. Louis, MO, USA). Tris (2-carboxyethyl) phosphine (TCEP) was obtained from Thermo Fisher Scientific (Waltham, MA, USA). Sequencing-grade modified trypsin was acquired from Promega (Madison, WI, USA). Acetic acid (HOAc), methanol, acetonitrile, and high-performance liquid chromatography-grade water were purchased from J.T. Baker (Phillipsburg, NJ, USA). Formic acid was from Fluka (Charlotte, NC, USA). PNGase F was purchased from New England Biolabs (Ipswich, MA, USA).

2.2. Instrumental conditions

Experiments were conducted on a nano-LC-MS/MS system consisting of a Q Exactive™ Hybrid Quadrupole-Orbitrap MS (Thermo Fisher Scientific) coupled with a Dionex Ultimate 3000 HPLC (Sunnyvale, CA, USA). The solvent A as the mobile

phase was composed of 0.1 % formic acid in water and solvent B as the mobile phase was composed of 0.1 % formic acid in 80 % acetonitrile. The glycosylated peptides were loaded on an Acclaim™ PepMap™ 100 C18 nano-trap column (75 $\mu m \times 2$ cm, 3 μm particles, 100 °C pore size) by carrying out using a mobile phase at a flow rate of 3 $\mu L/min$ for 10min. The peptides were separated on a PepMap™ RSLC C18 nano-column (2 μm , 100 Å, 75 $\mu m \times 50$ cm) at a flow rate of 300 nL/min. Data-dependent acquisition was performed, and the top ten precursor peaks were fragmented with higher energy collisional dissociation and normalized collisional energy was 27. Ions were scanned at 70,000 in MS1 (the first level of mass analysis) and 17,500 in MS2 (the second level of mass analysis) over an MS scan range of 400-2000 m/z for both the MS1 and MS2 levels. The injection quantity was 300 ng. A NanoDrop™ 2000 Spectrophotometer (Thermo Fisher Scientific) was used.

2.3. Whey protein extraction

Human milk samples and clinical information were obtained with informed consent, and procedures were performed with the approval of the Institutional Review Board (IRB 7001355-202108-E-153). Information on the clinical samples is summarized in *Table 1*. The overall protocol is shown in *Fig. 1*. Human milk samples (3 mL) were treated with 3 mL of acetone, sonicated for 30 min at room temperature, and centrifuged at $3,600 \times g$ for 10 min. The protein pellet was separated from human milk and subjected to TCA precipitation. To remove casein protein, the protein pellet was treated with 60 mM $CaCl_2$, incubated for 1 h at 4 °C and centrifuged at $13,000 \times g$ for 30 min. The organic layer was removed, and the sample was centrifuged again. The bottom layer of the sample was extracted and filtered through a 0.45- μm mixed cellulose ester syringe filter from Advantec (Tokyo, Japan). The filtered sample was treated with a three-fold volume of cold acetone/10 % TCA (v/w) and incubated for 2 h at -20 °C. The sample was washed twice with cold acetone and dried using a HyperVAC-MAX VC2200 centrifugal vacuum concentrator (Hanil Scientific, Inc., Gochon-eup, South Korea).

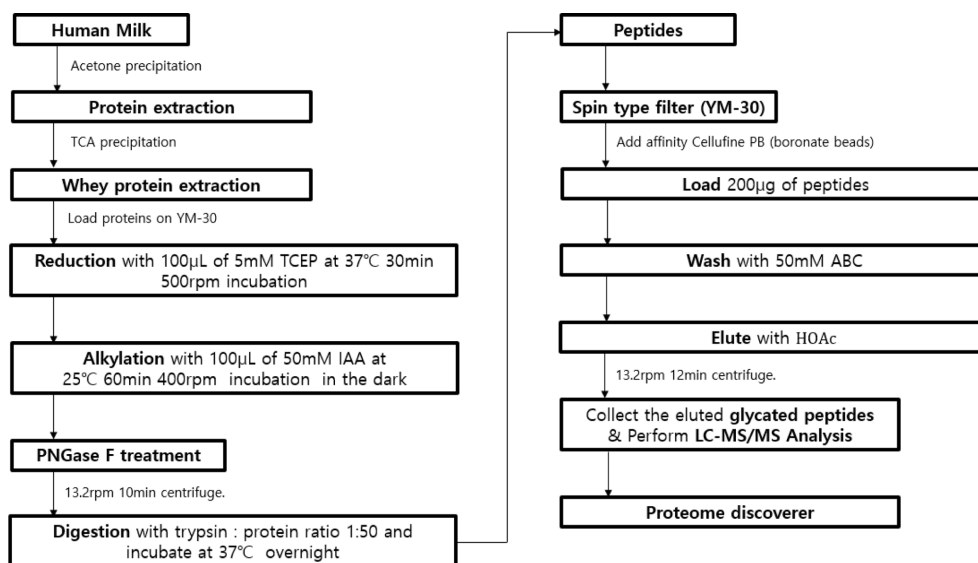


Fig. 1. Protocol of extraction, digestion, and enrichment of glycosylated peptides with boronate beads.

Table 1. Characteristics of primiparae and multiparae

| Group | AGEs (years) | BMI (kg/cm ²) | Puerperal stress | Number of subjects | Sample label |
|------------|--------------|---------------------------|------------------|--------------------|--------------|
| Primiparae | 33.0 (30-39) | 22.62 (19.95-24.88) | 18.2 (14-21) | 4 | 1-4 |
| Multiparae | 36.3 (32-43) | 23.82 (20.08-26.40) | 20.5 (13-25) | 4 | 5-8 |

2.4. Sample preparation using FASP

Human milk whey protein was digested using FASP on a Microcon YM-30 filter from Millipore (Billerica, MA, USA). Each 200 µg of human milk whey protein was adjusted to a volume of 100 µL lysis buffer (8 M urea, 0.1 M Tris and HCl buffer, pH 8.5). The protein was reduced by incubation with TCEP at 37 °C for 30 min and alkylated with iodoacetamide at 25 °C for 1 h in the dark. After washing with lysis buffer and 50 mM ABC sequentially, the protein was treated with PNGase F to cleave N-linked-glycan chains from the protein and incubated at 37 °C for 3 h. After washing with 50 mM ABC, the protein was digested with trypsin (enzyme to protein ratio of 1:50; w/w) at 37 °C for 18 h. The digested protein was collected into new tubes, and trypsin was inactivated by acidification with 20 µL formic acid. Collected peptides were dried using a centrifugal vacuum concentrator.

2.5. Boronate affinity enrichment

Glycosylated peptides were enriched using Cellufine PB (500 µL, JNC Corporation, Tokyo, Japan), a phenyl borate affinity ligand. Loading buffer (50 mM MgCl₂, 250 mM NH₄OAc, 0.1 M Tris, pH 8.1), washing buffer (50 mM ABC and 0.1 M Tris, pH 8.1), and elution buffer (0.1 M HOAc, pH 3) were prepared in high-performance liquid chromatography grade water.¹² Cellufine PB was placed on the Microcon YM-30 filter and washed five times with 400 µL loading buffer for equilibration. The peptide sample was dissolved in 250 µL loading buffer and loaded in to YM-30 filter with Cellufine PB. The filter was washed thrice with 400 µL washing buffer. After washing, the retained glycosylated peptide was eluted twice using a 200 µL elution buffer. Eluted glycosylated peptides were desalted with ultra-micro spin C18 columns (Harvard Apparatus, Holliston, MA, USA). Desalted glycosylated peptides were dried using a centrifugal vacuum concentrator.

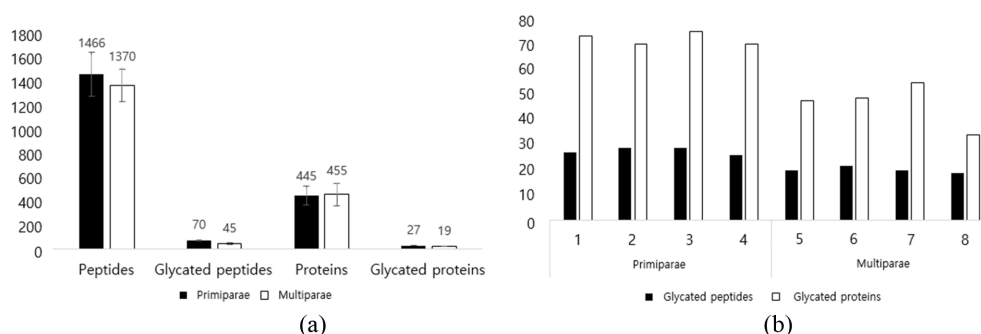


Fig. 2. The numbers of glycosylated peptides and proteins identified in the human milk sample. (a) Average number of identified peptides, glycosylated peptides, proteins and glycosylated proteins from human milk whey protein of four primiparae and multiparae, (b) Number of identified glycosylated proteins and peptides from each human milk sample.

2.6. Data analysis and statistical interpretation

Thermo MS/MS raw files of analysis were searched using Proteome Discoverer™ software (ver. 2.4), and the *Human* database was downloaded from Uniprot. The appropriate consensus workflow included a peptide-spectrum match validation step and SEQUEST HT process for detection as a database search algorithm. Search parameters were set up as follows: 10 ppm of tolerances of precursor ion masses, 0.02 Da fragment ion mass, and maximum of five missed cleavages with trypsin enzyme. The modifications on the peptide sequence were as follows: static carbamidomethylation of cysteine (+57.012 Da), dynamic modifications of methionine oxidation (+15.995 Da), carbamylation of protein in N-terminus (+43.006 Da), and acetylation of protein in N-terminus (+42.011 Da). Glycation of lysine (+162.053 Da) for analyzing glycosylated proteins and deamidation of asparagine and glutamine (+0.984 Da) for evaluating the effect of PNGase F were added as the dynamic modifications in the Proteome Discoverer™ software search. The results below 1% of the false discovery rate were selected and filtered for at least more 6 amino acids in a peptide length.

3. Results and Discussion

In this study, human milk samples obtained from eight mothers were collected and whey protein extraction and glycosylated peptide enrichment were performed. Whey protein was extracted using acetone

precipitation and TCA precipitation and digested using FASP. PNGase F was used to remove *N*-glycan, and glycosylated peptides were enriched using boronate beads.

3.1. Qualitative analysis of human milk glycation enrichment

The numbers of glycosylated peptides and proteins identified in the human milk sample from primiparae were 102 and 52, respectively. The numbers of glycosylated peptides and proteins identified in the human milk sample from multiparae were 80 and 41, respectively. As shown in Fig. 2, average numbers of glycosylated peptides and proteins were greater in human milk samples from primiparae compared to those in multiparae.

3.2. Quantitative analysis of human milk peptides

We confirmed through quantitative analysis that the peptide level in human milk differs according to fertility. Fig. 3(a) shows the volcano plot obtained in the multiparae and primiparae groups, with 485 differentially expressed peptides (251 up and 234 down). Volcano plot was made of the ratio of normalized abundance between groups and P-value. The peptides identified and quantified according to fertility were visualized in a principal component analysis (PCA) plot in Fig. 3(b). PCA is based on signal intensity to compare samples. As shown in Fig. 3(c), a heatmap was generated for 3104 quantified peptides to compare

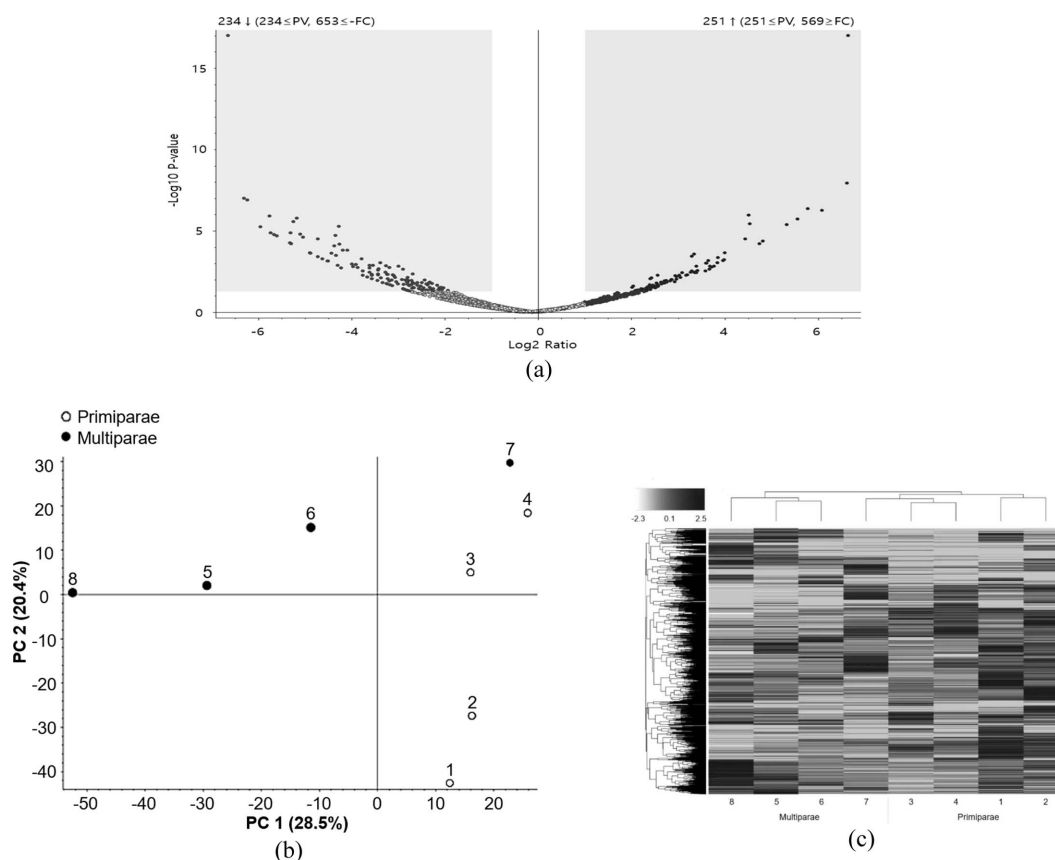


Fig. 3. Proteomics analysis results of human milk with peptide level. (a) Multiparae vs primiparae on volcano plot. Volcano plot shows the 485 differentially expressed peptides between the groups, (b) PCA plot shows an overall understanding of differentially expressed peptides and illustrates how samples differ according to fertility group, (c) Heatmap visualize the comparison of normalized abundance of peptides between groups.

the differences in the normalized abundance of peptides between groups.

3.3. Qualitative analysis of human milk glycation enrichment

As shown in Fig. 4, we quantitatively analyzed human milk according to fertility and found that 19 of 94 glycosylated peptides were differentially expressed. The 100% stacked bar chart shows the average values of the 19 differentially expressed peptides (DEPs) in the group divided by fertility. A list of the 19 glycosylated DEPs with primiparae and multiparae is shown in Table 2. The 14 glycosylated DEPs down-regulated and 5 glycosylated DEPs up-regulated in

multiparae. The 19 glycosylated DEPs belong to 10 albumin proteins. Glycosylated albumin was studied as potential factor for glycemic control.¹³ So, we confirmed that glycemic control was better in human milk of multiparae. The degree of glycation in patients with diabetes is higher than that in normal subjects, as reported in several previous studies.^{12,14} As the nutritional levels derived the human milk are related to the maternal health, the glycation degree is a potential factor that determining the quality of human milk.¹⁵ Studies have shown that the intake of AGEs increases the level of AGEs in the body.¹⁶ Based on these results, primiparae should consume food that includes less AGEs to positively impact the baby's nutrition.

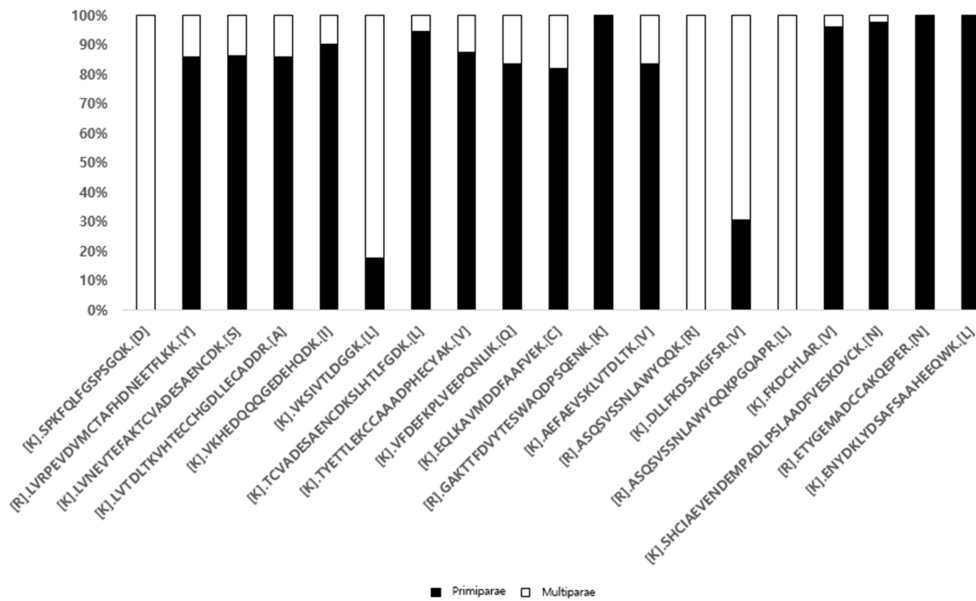


Fig. 4. The 100 % stacked bar chart of the 19 differently expressed glycosylated peptides was calculated as the average abundance in each group.

Table 2. List of 19 significantly different glycosylated peptides from human milk between multiparae and primiparae group

| Uniprot accession ID | protein name | Annotated Sequence | Glycation sites | # PSMs | Theo. MH+ [Da] | # Missed Cleavages | Abundance Ratio P-Value: Multiparae / Primiparae | Peptide level change Multiparae/ Primiparae |
|----------------------|------------------------------------|--|-----------------|--------|----------------|--------------------|--|---|
| E7EQB2 | Lactotransferrin | [K].SPKFQFLFGSPSGQK.[D] | K291 | 1 | 1669.84318 | 1 | 1E-17 | UP |
| P02768 | Albumin | [R].LVRPEVDVMCTAFHDNEETFLKK.[Y] | K160 K161 | 21 | 3102.46459 | 1 | 0.001749215 | DOWN |
| P02768 | Albumin | [K].LVNEVTEFAKTCVADESAENCDK.[S] | K75 | 35 | 2791.22845 | 1 | 5.49926E-05 | DOWN |
| P02768 | Albumin | [K].LVTDLTKVHTECCHGDLLECAADDR.[A] | K264 | 10 | 3019.34416 | 1 | 0.007569634 | DOWN |
| P05814 | Beta-casein | [K].VKHEDQQQGEDEHQDK.[I] | K40 | 2 | 2111.91121 | 1 | 0.040743058 | DOWN |
| P05413 | Fatty acid-binding protein, heart | [K].VKSIVTLDGGK.[L] | K82 | 6 | 1278.71512 | 1 | 0.036965002 | UP |
| P02768 | Albumin | [K].TCVADESAENCDKSLHTLFGDK.[L] | K88 | 3 | 2660.13382 | 1 | 0.019741575 | DOWN |
| P02768 | Albumin | [K].TYETTLKCCAAADPHECYAK.[V] | K383 | 14 | 2680.12115 | 1 | 0.013090098 | DOWN |
| P02768 | Albumin | [K].VFDEFKPLVEEPQNLIK.[Q] | K402 | 15 | 2207.14819 | 0 | 2.02564E-05 | DOWN |
| P02768 | Albumin | [K].EQLKAVMDDFAAFVEK.[C] | K569 | 22 | 2002.96779 | 1 | 0.037578216 | DOWN |
| X6R868 | Bile salt-activated lipase | [R].GAKTTFDVYTESWAQDPSQENK.[K] | K384 | 1 | 2664.19476 | 1 | 1E-17 | DOWN |
| P02768 | Albumin | [K].AEFAEVSKLVTDLTK.[V] | K257 | 50 | 1812.9477 | 1 | 5.92131E-05 | DOWN |
| P06311 | Immunoglobulin kappa variable 3-20 | [R].ASQSVSSNLAWYQQK.[R] | K59 | 9 | 1859.86576 | 0 | 1E-17 | UP |
| E7EQB2 | Lactotransferrin | [K].DLLFKDSAIGFSR.[V] | K307 | 4 | 1630.83228 | 1 | 0.036703587 | UP |
| A0A087X130 | Immunoglobulin kappa constant | [R].ASQSVSSNLAWYQQKPGQAPR.[L] | K59 | 2 | 2466.18956 | 0 | 1E-17 | UP |
| E7EQB2 | Lactotransferrin | [K].FKDCHLAR.[V] | K249 | 1 | 1208.57283 | 1 | 8.1829E-05 | DOWN |
| P02768 | Albumin | [K].SHCIAEVENDEMPADPLSLAAD-FVESKDVCK.[N] | K337 | 1 | 3638.61827 | 1 | 0.021431817 | DOWN |
| P02768 | Albumin | [R].ETYGEMADCCAKQEPER.[N] | K117 | 2 | 2235.88387 | 1 | 1E-17 | DOWN |
| H3BV03 | Protein fantom | [K].ENYDKLYDSAFSAAHEEQWK.[L] | K364 K379 | 1 | 2755.18934 | 1 | 1E-17 | DOWN |

4. Conclusions

We isolated whey proteins from human milk using acetone and TCA precipitation. We quantified 67 glycosylated proteins and 122 glycosylated peptides using a modified FASP with PNGase F. Nineteen glycosylated peptides were differentially expressed in human milk samples by classifying according to fertility. We confirmed difference between the glycosylated peptide level in human milk and fertility. We hope to analyze more human milk samples to study of glycosylation for validation. Other omics studies should also be conducted to identify relationship between human milk-derived components associated with fertility.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT). (No. 2020R1A2C1005082).

References

1. Y. Liao, R. Alvarado, B. Phinney and B. Lönnerdal, *J. Proteome Res.*, **10**(4), 1746-1754 (2011).
2. O. Ballard and A. L. Morrow, *Pediatr. Clin. North Am.*, **60**(1), 49-74 (2013).
3. W. E. Heine, P. D. Klein and P. J. Reeds, *J. Nutr.*, **121**(3), 277-283 (1991).
4. B. Lönnerdal, *Am. J. Clin. Nutr.*, **77**(6), 1537S-1543S (2003).
5. A. M. Prentice and A. Prentice, *Proc. Nutr. Soc.*, **54**(2), 391-400 (1995).
6. L. Maillard, *J. Neonatol.*, **54**, 66-68 (1912).
7. C. S. Kim, S. Park and J. Kim, *J. Exerc. Nutrition Biochem.*, **21**(3), 55-61 (2017).
8. A. Gupta and J. Uribarri, *Horm. Res. Paediatr.*, **85**(5), 291-300 (2016).
9. J. G. Dorea, A. C. Cruz Granja and M. L. Lacayo Romero, *Ann. Nutr. Metab.*, **41**(4), 250-254 (1997).
10. M. Molas Guzmán, C. Clementini, M. D. Pérez-Cárceles, S. Jiménez Rejón, A. Cascone and T. Martellini, *Science of The Total Environment*, **544**, 595-600 (2016).
11. J. Mes, D. J. Davies, J. Doucet, D. Weber and E. McMullen, *Food Additives and Contaminants*, **10**(4), 429-441 (1993).
12. J. Jeon, J. Yang, J. M. Park, N. Y. Han, Y. B. Lee and H. Lee, *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.*, **1092**, 88-94 (2018).
13. I. Shimizu, T. Kohzuma, M. Koga, *J. Lab. Precis. Med.*, **4**, 23 (2019).
14. L. Zhang, Q. Zhang, *Proteomics Clin. Appl.*, **14**(2), 1900104 (2019).
15. E. Isganaitis, S. Venditti, T. J. Matthews, C. Lerin, E. W. Demerath and D. A. Fields, *Am. J. Clin. Nutr.*, **110**(1), 111-120 (2019).
16. C.-S. Kim, S. Park and J. Kim, *J. Exerc. Nutrition Biochem.*, **21**(3), 55-61 (2017).

Authors' Positions

| | |
|----------------|--------------------|
| Seonghyeon Cho | : Graduate Student |
| Jong-Moon Park | : Professor |
| Hookeun Lee | : Professor |
| Jun Hwan Song | : Professor |
| Nam Mi Kang | : Professor |