





Proteomic analysis of human milk reveals nutritional and immune benefits in the colostrum from mothers with COVID-19

Juanjuan Guo^{a,d,#}, Minjie Tan^{b,#}, Jing Zhu^{c,#}, Ye Tian^b, Huanyu Liu^{a,d}, Fan Luo^e, Jianbin Wang^f, Yanyi Huang^{b,g} , Yuanzhen Zhang^{a,d,h} , Yuexin Yangⁱ , Guanbo Wang^{b,g} 

^a Department of Gynaecology and Obstetrics, Zhongnan Hospital of Wuhan University, Wuhan, Hubei Province 430071, China

^b Institute for Cell Analysis, Shenzhen Bay Laboratory, Shenzhen, Guangdong Province 518132, China

^c Beijing Institute of Nutritional Sources, Beijing Academy of Science and Technology, Beijing 100069, China

^d Hubei Clinical Research Center for Prenatal Diagnosis and Birth Health, Wuhan 430071, China

^e State Key Laboratory of Virology/Institute of Medical Virology, School of Basic Medical Sciences, Wuhan University, Wuhan 430071, China

^f School of Life Sciences and Tsinghua-Peking Center for Life Sciences, Tsinghua University, Beijing 100084, China

^g Biomedical Pioneering Innovation Center, Peking University, Beijing 100871, China


^h Hubei Provincial Key Laboratory of Developmentally Originated Disease, Wuhan, Hubei Province 430071, China


ⁱ National Institute for Nutrition and Health, Chinese Center for Disease Control and Prevention, Beijing, 100050, China


Equal contribution to this work

Address correspondence to:

 Yanyi Huang, Email: yanyi@pku.edu.cn

 Yuanzhen Zhang, Email: zhangyuanzhen@whu.edu.cn

 Yuexin Yang, Email: yuexin_yang@sina.com

 Guanbo Wang, Email: guanbo.wang@pku.edu.cn

ABSTRACT

The range of benefits breastfeeding provides neonates and infants include nutrition, improved neonatal survival, and reduced morbidity from certain diseases. It also aids maternal health by speeding postpartum recovery. However, due to concern about the risk of SARS-CoV-2 transmission and the lack of evidence of breastmilk's protective effects against the virus, whether mothers with COVID-19 should be encouraged to breastfeed is under debate. Here, we present the results of proteomic and glycoproteomic studies of breast milk (colostrum and mature milk) from mothers with confirmed COVID-19. All colostrum samples exhibited significantly upregulated immune-related proteins, especially whey proteins with antiviral properties against SARS-CoV-2, and increased glycosylation levels and heterogeneity at those proteins. Such adaptive differences in milk from COVID-19 mothers tend to fade in mature milk from the same mothers one month postpartum. These results suggest the immune benefits of colostrum from mothers with COVID-19 and provide molecular-level insights that aid breastmilk feeding decisions in cases of active infection.

INTRODUCTION

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has caused over 171 million confirmed cases of coronavirus disease 2019 (COVID-19), which has resulted in over 3 million deaths globally as of June 2, 2021¹. Based on data from the United States, approximately 98,948 pregnant women had confirmed COVID-19 infections from January 22 to June 28, 2021, 0.14% of all COVID-19 cases. A systematic evidence review that followed the procedures in the Cochrane handbook for systematic reviews of interventions, identified studies that included mothers with suspected or confirmed COVID-19. As a result, in January 2021 the World Health Organization (WHO) updated their Interim Guidance on Clinical Management of COVID-19 to recommend that mothers with suspected or confirmed COVID-19 be encouraged to initiate and continue breastfeeding². However, concerns over the risk of viral transmission from mother to infant have inevitably hindered the choice to breastfeed. A three-month study in 2020 reported that the breastfeeding rate of infants under one week of age by mothers with COVID-19 was 8.8% in China³.

Breastmilk contains not only basic nutrients but also developmental immune system factors (e.g., antibacterial components, immunoglobulins, and immune cells) that help protect newborns from environmental pathogens. Tromp et al. (2017) found that breastfeeding was associated with a 30% reduction in lower respiratory tract infections in young children⁴. A member of the enveloped single-stranded RNA virus family, SARS-CoV-2 infects human bronchial epithelial cells, pneumocytes, and upper respiratory tract cells, and that can then lead to severe, life-threatening respiratory disease and lung injury⁵. Respiratory mucosal immunity with IgA, especially secretory IgA (sIgA), plays a major role in mounting an antiviral defense⁶. As the dominant antibody in milk, sIgA has a stronger protective effect than do IgGs because it can directly neutralize the virus while not activating the complement to cause inflammation⁷. The protection of breastfeeding against respiratory infections in infants is generally recognized⁸, and breast milk is an important sIgA source. However, the immune benefits of feeding breastmilk from mothers with COVID-19 remains elusive because of a lack of molecular-level characterization of that milk.

In this work, we used mass spectrometry (MS)-based proteomic analysis to characterize the immune-related proteins in breast milk collected from both six COVID-19 patients and ten healthy donors and then compared the results from those two groups. Both colostrum and mature milk were examined to monitor dynamic proteomic changes. Most human milk proteins are glycosylated and the glycosylation pattern importantly aids in morbidity reduction⁹ by both affecting proteolytic susceptibility¹⁰ and serving as immunomodulators¹¹ and competitive inhibitors of pathogen binding^{11,12}. Therefore, we expanded the scope of this study to the glycoproteomic level by thoroughly characterizing a panel of key immune-related proteins' glycosylation patterns to reveal the COVID-19-induced dynamic glycoproteome evolution. We show that combined proteomic and glycoproteomic strategies can provide straightforward nutrition and immunity information that can help professionals recommend appropriate infant feeding procedures for infected mothers.

RESULTS

Alteration of proteome profiles in milk from COVID patients

The basic clinical characteristics between the patient group and control group were not significantly different ($p > 0.05$, Table 1). Based on the results of nested RT-PCR methods¹³, all milk samples tested negative for SARS-CoV-2, which ruled out the risk of direct viral transmission through the milk. Also, milk is recognized to be mature after 16 days postpartum and completely mature by 4 to 6 weeks postpartum¹⁴. Meanwhile, the protein content in human milk gradually decreases throughout lactation, but the rate of decrease is much lower at the fully mature stage¹⁵ and the composition remains relatively consistent¹⁴. Therefore, the mature milk samples collected in this study represent a relative stable lactational period (Figure 1A). We conducted proteomic and glycoproteomic analyses of milk compositional changes in four sample groups: colostrum samples from COVID-19 patients (COVID-19 colostrum, $n = 6$) and from healthy donors (Ctrl colostrum, $n = 10$), and mature milk samples from COVID-19 patients (COVID-19 mature, $n = 6$) and from healthy donors (Ctrl mature, $n = 10$) (Figure 1B).

Table 1. Clinical characteristics of sample donors.

	Characteristics	Case (n=6)	Control (n=10)	p-value
Maternal	age, median (IQR) ^a [years]	30.00 (27.75, 30.00)	33.50 (31.00, 34.75)	0.08
	pre-pregnancy BMI ^b , median (IQR) [kg/m ²]	19.00 (18.23, 22.03)	21.55 (20.35, 22.13)	0.22
	gestational weight gain, median (IQR) [kg]	18.50 (15.75, 19.00)	14.00 (12.25, 15.75)	0.06
	multiparous, n [%]	3 (50)	8 (80)	0.30
	undergraduate, n [%]	3 (50)	7 (70)	0.61
	postgraduate, n [%]	3 (50)	3 (30)	0.61
Diseases during pregnancy	anemia, n [%]	0 (0)	0 (0)	NA
	diabetes, n [%]	1 (17)	3 (30)	1.00
	hypertension, n [%]	0 (0)	0 (0)	NA
	pre-eclampsia, n [%]	1 (17)	0 (0)	0.38
Infant	male, n [%]	3 (50)	6 (60)	1.00
	gestational age on admission, median (IQR) [wk]	37.50 (37.00, 38.75)	39.00 (38.25, 39.00)	0.24
	birth weight, median (IQR) [kg]	2.99 (2.79, 3.14)	3.24 (3.10, 3.62)	0.07
	birth length, median (IQR) [cm]	49.50 (48.25, 50)	49.00 (48.00, 49.75)	0.82
	head circumference, median (IQR) [cm]	34.50 (34.00, 35.00)	35.00 (35.00, 35.25)	0.08
	Apgar 1-min, median (IQR)	9 (9, 9)	9 (9, 10)	0.08
	Apgar 5-min, median (IQR)	10 (10, 10)	10 (10, 10)	0.84

^aIQR: interquartile range

^bBMI: body masses index

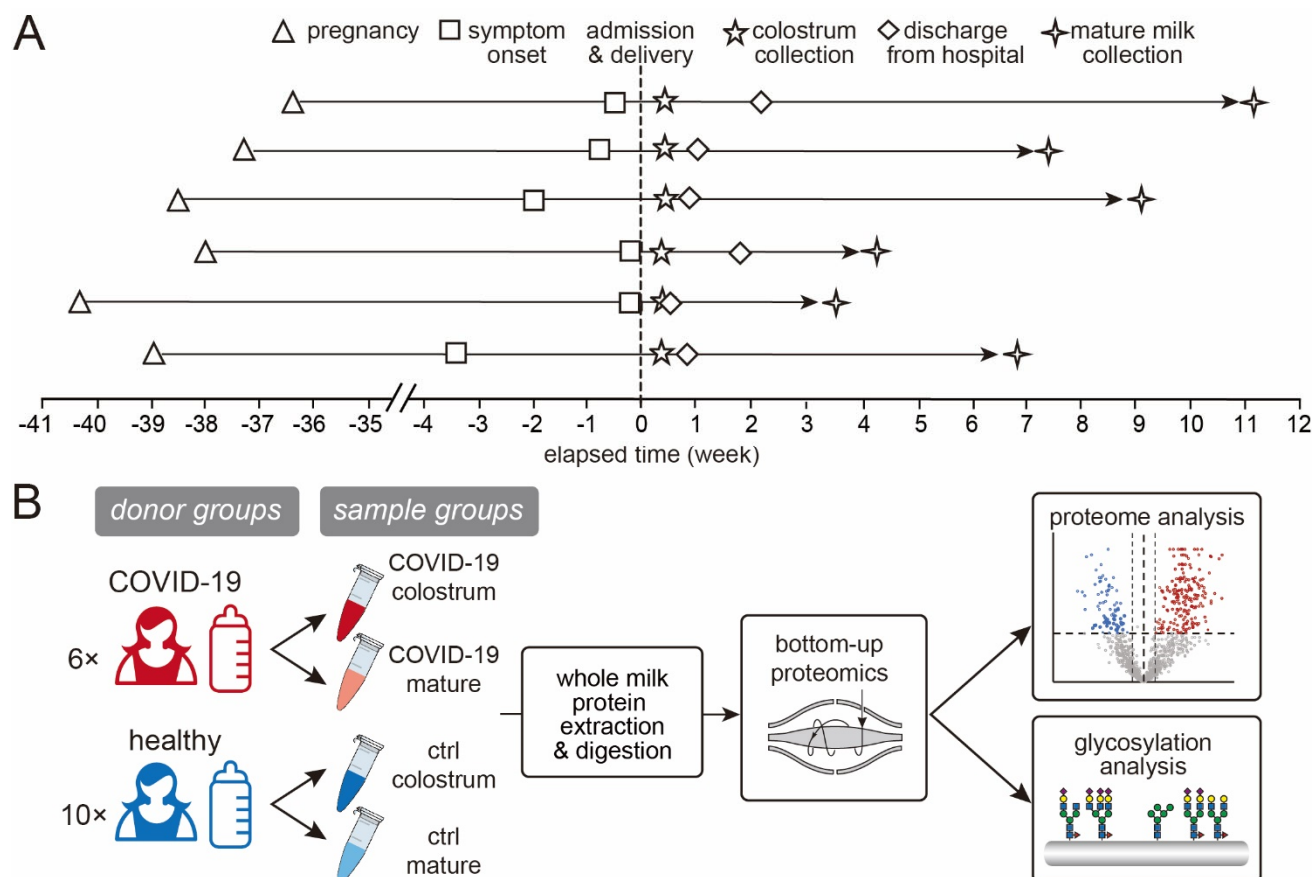


Figure 1. Study information and workflow. (A) The timeline of milk sample collection from the 6 COVID-19 patients and (B) a schematic workflow of this study.

Of the 1573 proteins identified in the four milk groups, 1439 were sufficiently abundant to be quantitatively monitored, and 1123 of those (78%) were shared by all groups (Figure 2A, Table S1). Principal components analysis (PCA) suggested a distinction between the COVID-19 colostrum group and the other three groups, indicating an alteration in milk proteome upon active SARS-CoV-2 infection (Figure 2B). The total protein content in the COVID-19 colostrum group were 4.1 times higher than that in the control colostrum group; particularly, casein proteins were 3.9 times lower and whey proteins were 7.2 times higher (Figure 2C). As the major components of milk, whey proteins and caseins (comprising about 45%–97% and 3%–55% of human milk proteins, respectively) figure importantly in nutrition and immunity¹⁶, including being sources of amino acids, enhancing micronutrient bioavailability, providing immunogenic training for innate and adaptive immunity, and promoting intestinal growth and maturation via interactions with the microbiome¹⁷. The proportion of casein proteins in the COVID-19 colostrum group was lower than that of the other groups ($p < 0.0001$; Figure 2C), likely because of the downregulation of all caseins (α -S1-, β -, κ -caseins) and an increase in the total abundance of whey proteins. In detail, Among the highly abundant milk proteins, beta-casein and immunoglobulins (predominated by sIgA1) were significantly down- and upregulated, respectively (Figure 2D). The differences between COVID-19 milk and Ctrl milk protein contents at the mature stage were not significant (Figure 2C and 2D).

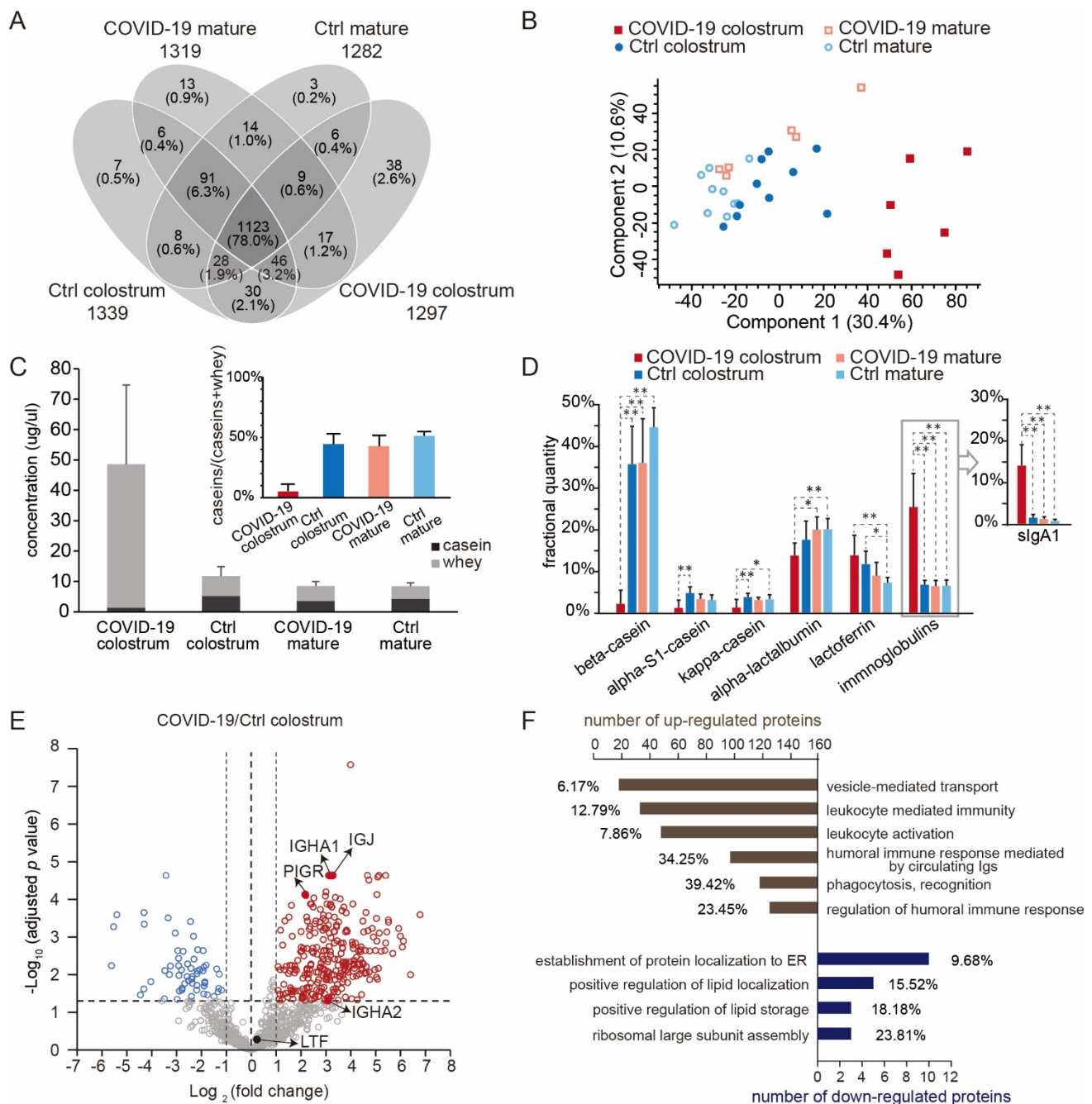


Figure 2. Comparison of human milk proteomes in different sample groups. **(A)** A Venn diagram showing the numbers of proteins that could be quantitatively monitored. The number of proteins for each group are shown below the group name. **(B)** Principal components analysis results of the milk proteome. **(C)** Bar graphs of the quantities of casein (black) and whey (gray) proteins. The inset shows the fractions of caseins. Bars indicate means and whiskers indicate SDs. **(D)** Comparisons of the fractional quantities of major milk proteins. For one-way ANOVA with Tukey HSD test, * $p < 0.05$, ** $p < 0.01$. **(E)** A volcano plot showing proteome changes between COVID-19 colostrum and Ctrl colostrum groups. **(F)** Significantly enriched biological processes of differently expressed proteins (p of each Gene Ontology [GO] term < 0.0002). The graph shows both numbers and percentages of differently expressed proteins in each term group.

Upregulation of immune-related proteins suggest the adaptive immune benefit of colostrum from COVID-19 patients

In comparing the COVID-19 colostrum and Ctrl colostrum groups (cut-off $p < 0.05$ and fold change >2), we found 282 upregulated proteins and 58 downregulated proteins (Figure 2E) in association with 34 and 7 enriched biological processes, respectively, among which we specify the most prominent ones in Figure 2F (see Table S2 for the list of enriched GO terms and their related proteins). The increased expressions of immunoglobulins in response to the viral infection served as the major contributors to those term groups. We did not directly measure sIgA because of the bottom-up approach used in this study. Consider that the intact multimeric form of sIgA incorporates IgA1/IgA2, a joining chain (JC), and a secretory component from the endoproteolytic cleavage of the polymeric immunoglobulin receptor (PIGR)¹⁸. So, since JC, PIGR, light chains, and variable regions may be used to assemble other forms of immunoglobulins, we instead used the detected quantities of immunoglobulin heavy constant A1 and A2 (IGHA1 and IGHA2) to represent the abundances of sIgA1 and sIgA2. Notably, ribosomal proteins were among the significantly downregulated proteins and they contributed to the GO term for both ribosomal large subunit assembly and protein localization establishment to the endoplasmic reticulum (ER, Figure 2F), both of which are activities associated with viral replication and transport.

Biological pathway analysis revealed 15 differentially expressed proteins that participate in the pathway related to coronavirus disease. Nine ribosomal proteins that might inhibit protein expression in the mammary gland were down-regulated, while some complement proteins (e.g., complement factor B, complement C3, and complement C4B) increased. Altogether, the proteome in COVID-19 colostrum featured a much higher proportion of immunoglobulins than did that of the Ctrl colostrum, thus implying that the former likely provides infants adaptive immune protection against coronavirus.

The increased microheterogeneity of overall protein glycosylation and the glycosylation levels of immune-related proteins in the COVID-19 colostrum proteome

Glycoprotein microheterogeneity originates from the various glycan structures that may attach to identical glycosylation sites on cell surfaces. Because microheterogeneity leads to differing preferences of glycan-based interactions and reorganization of dynamic epitopes¹⁹, it thus facilitates virus neutralization. Our glycoproteomic measurements identified 820 unique *N*-glycopeptides from 48 different glycoproteins (Figure 3A), results that mostly agreed with those of a previous study²⁰. Most of those glycoproteins were within the top 200 most abundant milk proteins (Figure S1A). Also, most of them had one to six identified glycosites, but tenascin (TNC), which is involved in protection against viral infections²¹, had 12 glycosites (Figure S1B). In the COVID-19 colostrum glycoproteome, we observed a high degree of microheterogeneity (Figure S1C) that was clearly distinct from those of the other three samples groups (Figure 3B). Among 103 glycosites, 66.0% had more than 1 glycoform and 18.4% harbored more than 10 glycoforms. Taking the glycosylation pattern into account, we detected only 178 (21.7%) of the 820 glycopeptides in all sample groups, while 33.2% (272) of the glycopeptides were detected exclusively in the COVID-19 colostrum group (Figure 3A). This was most likely

because of a dramatically improved glycosylation pattern diversity and slightly increased number of glycosites in that group (Figure S1D) rather than more protein identities (Figure S1E).

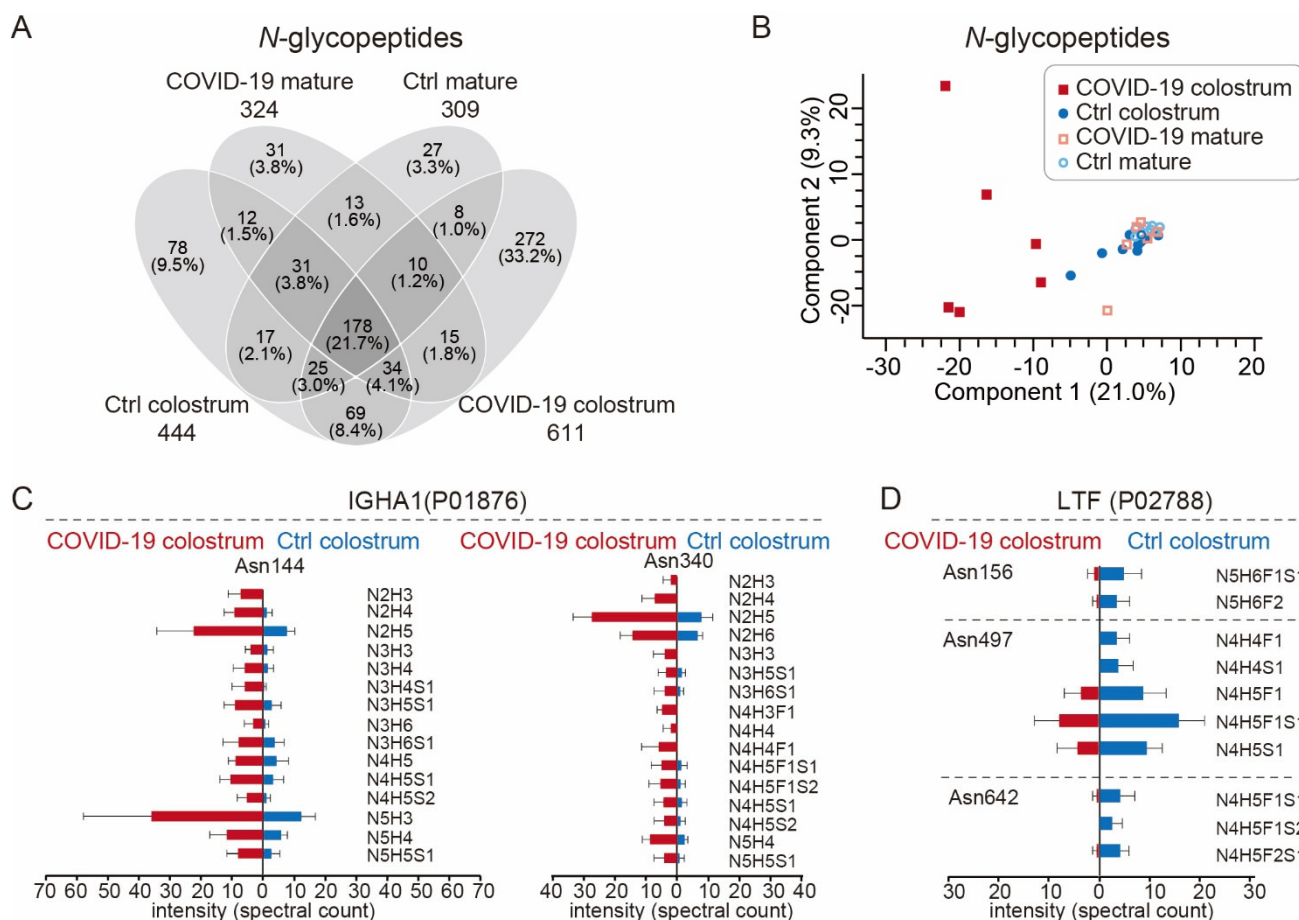


Figure 3. The glycoproteomes of each sample group. **(A)** A Venn diagram showing the numbers of *N*-glycopeptides that could be quantitatively monitored. The number of *N*-glycopeptides for each group are shown below the group name. **(B)** Principal components analysis plot of *N*-glycopeptides. **(C and D)** Site-specific changes in *N*-glycopeptides of immunoglobulin heavy constant alpha 1 (IGHA1, **C**) and lactoferrin (LTF, **D**) between COVID-19 colostrum and Ctrl colostrum groups. Bars indicate means and whiskers indicate SDs.

We observed 83 upregulated and 17 downregulated *N*-glycopeptides, as well as site-specific *N*-glycosylation changes, in the COVID-19 colostrum (Figure S2). As the major components in milk, the two most abundant glycoproteins, lactoferrin (LTF) and IgA1, exhibited different glycosylation pattern alteration trends in response to COVID-19. The elevated level of IgA, a key component in adaptive immunity, was accompanied by increases in both microheterogeneity at IGHA1's Asn144, and Asn340 *N*-glycosylation sites and those glycoforms' expression levels (Figure 3C). In contrast, LTF, which is associated with innate immunity²², showed reduced microheterogeneity (Figure 3D).

COVID-19 colostrum's immune-related proteome profile changes fade with milk maturation

Compared to levels in the COVID-19 colostrum, the proportions of whey proteins, especially IgA1, decreased

significantly in the COVID-19 mature group (Figure 2C and 2D). Besides the overlapping proteins that were differentially expressed at different lactational stages in the Ctrl groups, 201 proteins were downregulated while 83 proteins were upregulated in COVID-19 milk during its maturation (Figure S3A). Fourteen and eight biological processes changed significantly for down- and up-regulated proteins, respectively (Table S3, Figure S3B). Leukocyte mediated immunity, adaptive immune response, and complement activations involved the largest number of down-regulated proteins, and phagocytosis recognition and regulation of complement activation involved the highest proportion of all associated genes. Up-regulated proteins during maturation were clustered largely in the SRP-dependent co-translational protein targeting membrane and ubiquitin ligase inhibitor activity GO terms. Also, multiple ribosomal proteins contributed to those clusters. The downregulated proteins clearly show that immune-related proteins had obviously decreased during the mature milk phase and with the COVID-19 patients' recoveries.

The COVID-19 colostrum's unique glycosylation features also appeared to fade with milk maturation. Seven *N*-glycopeptides were back-upregulated and 55 were back-downregulated in the COVID-19 mature sample group (Figure S4A). The glycosite-specific changes at sIgA and LTF in COVID-19 colostrum, with respect to Ctrl groups, underwent reverse evolvement in the COVID-19 mature group, producing glycosylation patterns that were similar to those of the Ctrl mature milk (Figure S4B). Using the changes between Ctrl colostrum and Ctrl mature as the reference, sIgA's microheterogeneity and glycosylation levels at Asn144 and Asn340 changed more prominently than other IGHA1 *N*-glycopeptides did (Figure S4C), and so did LTF at Asn156, 497, and 642 (Figure S4D). See Figure S5 for an extended list of *N*-glycoproteome site-specific changes between Ctrl colostrum and Ctrl mature groups. In summary, the proteome and glycoproteome profiles exhibited in the mature milk of the COVID-19 mature and Ctrl mature groups (Figure 2B, 3B, S6A and S6B) were not substantially different.

DISCUSSION

In healthy mother-infant dyads, human milk provides nutritional and protective components for newborns facing pathogenic challenges¹⁷. During inflammation or infection in either mother²³ or infant²⁴, the changes in breastmilk components not only reflects the maternal-infant health status^{25,26}, but also provides the neonate with protection from diseases ranging from diarrhea to acute respiratory tract infections²⁷. However, since SARS-CoV-2 RNA has been detected in human breastmilk, women with COVID-19 are greatly concerned about feeding their infants breastmilk²⁸⁻³¹. Chinese experts have recommended that COVID-19 infected mothers delay breastfeeding for periods of up to two weeks after delivery³², and many patients chose to begin breastfeeding after recovery. Also, some hospitals are implementing practices that conflict with WHO's policy on the subject³³. All infected mothers of newborns must understand the nutritional and immune benefits of breastmilk so they may knowledgeably consider the option of feeding it to their infants.

COVID-19 infection clearly leads to significant changes in profiles of both proteome and glycosylation patterns in breastmilk, and they both undergo dynamic evolution as milk maturation proceeds. Two major components

of human breastmilk, casein and whey proteins, are down- and up-regulated, respectively, and several studies have found that the composition of human breastmilk changes according to the infant's demands^{34,35}. Fetus oxygen is supplied completely through placental transport, and the most striking features of COVID-19 placentas are decidual arteriopathy and other maternal vascular malperfusion features, either one of which can cause fetal hypoxia³⁶. Hypoxia in neonatal rats leads to delayed enzyme maturation in the small intestine, which is the main site for protein absorption. The activities of those enzymes returned to control levels 3 weeks after recovery³⁷. Maternal stressors in late pregnancy can also impair the fetal gut, further limiting the ability of the newborn to absorb nutrients, especially proteins from colostrum³⁸. Absorption of high-MW casein proteins is more difficult than absorption of relatively low-MW whey proteins. Given that COVID-19 neonates likely suffered fetal hypoxia and maternal stresses due to the disease, the decreased levels of casein proteins and increased whey protein abundance in COVID-19 colostrum benefits the neonatal gut's ability to effectively absorb milk proteins during the first few days postpartum. As the neonatal gut recovers from hypoxia damage, both casein and whey protein abundances return to normal levels in mature milk (Figure 2C). Since caseins in human milk provide essential amino acids³⁹ and transport divalent cations (such as Ca^{2+} and Zn^{2+})^{40,41}, the decline of caseins in COVID-19 colostrum might affect absorption and transportation of some essential minerals in the gut of the nursing infant.

Whey proteins show antiviral activity against SARS-CoV-2 and its related pangolin coronavirus (GX_P2V) by blocking viral attachment and viral replication⁴², and immunoglobulins are major elevators of overall whey protein levels (Figure 2D). Immunoglobulins in breast milk trigger antigenic responses to infectious pathogens of maternal mucosa-associated lymphoid tissue, or the bronchomammary axis, and is most likely encountered by the infant^{43,44}. Immunoglobulin expression in blood is dramatically increased in the serum proteome of COVID-19 patients who had seroconversion, and most of the immunoglobulin regions are significantly changed in response to SARS-CoV-2 infection⁴⁵. SARS-CoV-2-specific antibodies detected in milk samples collected from COVID-19 patients were capable of neutralizing SARS-CoV-2 in vitro^{46,47}. Breastmilk immunoglobulins complement the neonate's immature adaptive immune system⁴⁸ and protect it by preventing infectious agents on the mucosal membranes from entering tissues, rather than by functioning in the circulation. This contrasts with IgG's transplacental protection that is most efficient against microbes in the blood and tissues and that functions through a complementary system that induces energy consumption and destructive tissue inflammation⁴⁹. The major immunoglobulin in human milk, sIgA, serves as a first line defense in mucosal areas, using mechanisms that include intracellular neutralization, virus excretion, and immune exclusion^{50,51}. When a pathogen enters the mother's upper airways, the Peyer's patch acquires the pathogen and its antigens are presented by M cells to circulating B cells that then migrate to the serosal (basolateral) side of the mammary epithelial cell, and it produces IgAs. As the IgAs move from the serosal to the luminal (apical) side of the mammary epithelial cell, they are glycosylated and complexed to form sIgA, which is secreted into milk^{17,52}. Since sIgA is more resistant to protein hydrolysis than IgG is, it can function on the mucosal membranes primarily in the gastrointestinal tract and to some extent in the respiratory tract as well⁴⁹. Because the volume

of milk ingested increases as a neonate grows, sIgA, which usually occurs in higher concentrations in colostrum than in mature milk, is nevertheless maintained at a relatively constant total amount throughout lactation. On average, the proportion of sIgA1 in the milk proteins of COVID-19 colostrum was 8.6 times greater than in the Ctrl colostrum (Figure 2D), thus indicating that the mothers' responses to the infection and additional adaptive immunological protection for the infants from viruses may be introduced through the gastrointestinal or respiratory tract. Indeed, it has been reported that a portion of sIgA present in milk from COVID-19 patients were RBD-specific^{53,54}. Such protection may be further strengthened by other adaptive immunity-related proteins that appear in higher abundances in COVID-19 colostrum than in Ctrl colostrum (Figure 2F).

Ribosomal proteins are the major class of downregulated whey proteins in COVID-19 colostrum (Figure 2F), even with the increased total amount of whey proteins. Ribosomal proteins are essential for translating genetic information from mRNA into proteins⁵⁵, and they dictate ribosomal large subunit assembly and the establishment of protein localization to the ER. Since SARS-CoV-2 has no cellular structure of its own, it can only replicate by utilizing ribosomes in infected cells to make viral polyprotein and then processing and transporting that polyprotein formed in ER ribosomes⁵⁶. Downregulation of those ribosomal proteins in COVID-19 colostrum may be associated with the milk protein synthesis status that helps inhibit viral replication and viral polyprotein transport in the mammary glands of infected, lactating mothers.

The glycoproteins, especially immunoglobulins in COVID-19 colostrum, generally exhibit much higher degrees of glycan microheterogeneity (Figure 3), a feature that may help recognize more dynamic epitopes and favor different glycan-based interactions¹⁹. Consider that SARS-CoV-2 assembly and its attachment and entry to a host cell is closely related to its heavily glycosylated spike protein⁵⁷, which has 22 occupied *N*-glycosylated sites⁵⁸ and diverse glycosylation⁵⁹. Glycans at the surface of SARS-CoV-2 are essential for viral entry, as shown in vitro when blocked *N*-glycan biosynthesis at the oligomannose stage dramatically reduced viral entry into HEK293T cells⁶⁰. When SARS-CoV-2 enters the human body, it first recognizes sialic acid and searches for its receptor. Next, very likely the viral envelope glycans interact with sialic acid-binding lectins that are expressed in a host cell-dependent manner⁶¹. Accordingly, the more that human milk glycoproteins have glycan microheterogeneity, the better their ability might be to interact with SARS-CoV-2 via multiple glycan-based molecular events in the infants' gastrointestinal tract. That then helps block SARS-CoV-2 binding to its receptor, ACE2, on the cell surface.

The WHO's breastfeeding recommendations for COVID-19 mothers are largely based on both the significant consequences of not breastfeeding and the separation between mother and child, as well as on the low risk of vertical transmission of COVID-19 through breastfeeding. Here, the proteomic and glycoproteomic data suggest the nutritional and immune benefits of breastmilk feeding at the molecular level. Since all prominent differences between COVID-19 and Ctrl colostrum proteomes, in terms of the casein/whey ratio, differently expressed proteins and the related significantly enriched biological processes, and the microheterogeneity of glycosylation all fade as milk maturation proceeds, those benefits may be attenuated by delaying breastmilk

feeding. Based on the premise that contact transmission between COVID-19 mothers and their newborns is prevented, our results provide helpful information for selection of approaches to lactation.

METHODS

Sample collection

Human milk samples were collected from six patients with confirmed COVID-19 and from ten healthy donors. First, colostrum samples were collected within 3-4 days post-partum, and then mature milk samples were collected 30-45 days postpartum, upon the COVID-19 patients' recoveries and release from quarantine or during the healthy donors' postpartum checkups. The sample sizes and groups were (1) 6 in the group of colostrum from COVID-19 patients (COVID-19 colostrum), (2) 6 in the group of mature milk from COVID-19 patients (COVID-19 mature), (3) 10 in the group of colostrum from healthy donors (Ctrl colostrum), and (4) 10 in the group of mature milk from healthy donors (Ctrl mature) (Figure 1B). This study was approved by the Medical Ethics Committee of Zhongnan Hospital of Wuhan University (approval no. 2020031). Written consent was obtained from each patient.

Milk sample preparation for proteomic analysis

The in-solution digestion of whole milk was adapted from previous methods⁶². Briefly, a 20 µl milk sample was added to 80 µl of protein lysis buffer (ThermoFisher, USA) with a protease inhibitor (ThermoFisher, USA). The whole milk buffer was sonicated 20 mins in a water bath with cold water and put on ice for another 30 mins. Impurities were removed by methanol/chloroform protein precipitation and the precipitate was air dried. The pellet was further dissolved in an 8 M urea (Sigma, Germany)/50 mM ammonia bicarbonate (Alfa Aesar, Germany) solution, incubated with 10 mM dithiothreitol (Sigma, Canada) at 37 °C for 1 h and followed by alkylation with 20 mM iodoacetamide (Sigma, USA) at 37°C in the dark for 30 mins. For digestion, the samples were incubated with a trypsin/lys-C mixture (w:w = 1:50; Promega, USA) at 37°C for 16 h and then the digest reaction was quenched with 1% formic acid (Fisher Scientific, Czech Republic). The resulting peptides were desalted using a C18 column (ThermoFisher, USA) prior to lyophilization. The same tryptic peptides for proteomic analysis were applied directly to glycoproteomic analysis without further glycopeptide enrichment.

Liquid chromatography–tandem mass spectrometry

In preparation for the liquid chromatography (LC)–tandem MS (MS/MS) analysis the lyophilized tryptic peptides were reconstituted with 10% formic acid. The peptides were then separated by an UltiMate 3000 RSLCnano LC System (Thermo Fisher Scientific, USA) equipped with an Acclaim PepMap 100-C18 trap column (ThermoFisher) and an Acclaim PepMap RSLC-C18 analytical column (ThermoFisher). Mobile-phase solvents A and B consisted of 0.1% formic acid in water and acetonitrile, respectively. Trapping was performed at a flow rate of 10 µl/min for 5 min with 3% B and separation was performed at a passive split flow of 300 nl/min for 95 min with 8% to 50% B over 51 min, 50% to 99% B over 14 min, 99% B for 15 min, and 99% to 3% B over 10 min. The eluted peptides were electrosprayed at a 2.0 kV spray and analyzed by an Orbitrap Eclipse Tribrid

mass spectrometer (Thermo Fisher Scientific, USA) coupled to the LC system. The MS analyses were operated in data-dependent acquisition (DDA) mode using higher-energy collision dissociation (HCD) for MS/MS fragmentation.

MS settings for proteomic analysis

The term “proteomic analysis” in this work refers to analysis of the proteome regardless of protein modifications. The MS1 scan range was 300–1500 m/z with a 60,000 resolution at 200 m/z, and the automated gain control (AGC) set at 4e5 with a maximum injection time of 50 ms. With charge-states screening enabled, precursor ions in 2+ to 7+ charge states with intensities >5e4 were mass-selected for MS/MS (isolation window: 1.6 m/z). A 45 s dynamic exclusion was set with a 10-ppm exclusion window. The mass spectrometer was operated in DDA mode with a 2 s cycle time. The MS/MS scans were acquired with a 1.4 m/z isolation window, followed by HCD dissociation with a normalized collision energy of 30% and Orbitrap detection with a 100–2000 m/z scan range, a 30,000 resolution (at 200 m/z), and an AGC of 4e5 with a 54 ms maximum injection time.

MS settings for glycoproteomic analysis

The term “glycoproteomic analysis” in this work refers to the determination of the positions and identities of the repertoire of glycans attached to the identified peptides. The MS1 scan range was 375–2000 m/z with a 60,000 resolution at 200 m/z, and the AGC was set at 4e5 with a maximum injection time of 50 ms. With charge-states screening enabled, precursor ions in 2+ to 7+ charge states with intensities >5e4 were mass-selected for MS/MS (isolation window: 1.6 m/z). A 60 s dynamic exclusion was set with a 10-ppm exclusion window and a 3 s cycle time. MS/MS spectra were acquired with Orbitrap across a 120–4000 m/z range at 30,000 resolution with the AGC target set to 5e4 and a 75 ms maximum injection time. An HCD was performed with 30% normalized collision energy. Detection of at least three out of 15 oxonium ions of glycopeptides (127.0390+, 145.0495+, 163.0601+, 243.0264+, 405.0793+, 138.055+, 168.0655+, 186.0761+, 204.0867+, 274.0921+, 292.1027+, 366.1395+, 407.1660+, 512.1974+, 657.2349+) triggers electron-transfer/higher-energy collision dissociation (ETHCD) for the same precursor (AGC target: 2e5; maximum injection time: 250 ms; supplemental activation for electron-transfer dissociation: 30% normalized collision energy). In addition, the same samples were analyzed in parallel using the same settings except that stepped HCD (i.e., collision energy of 20%, 30%, 40% and a 1.6 Th isolation window) instead of ETHCD was used to fragment the glycans and the maximum injection time was set to 75 ms.

Data processing

Following Zhu et al. (2020)¹⁹, the raw shotgun LC-MS/MS data were searched with Proteome Discoverer version 2.3 (Thermo Scientific) using the Sequest HT search engine against a UniProt Swiss-Prot database⁶³: *Homo sapiens* (canonical and isoform) (October 2020; 26,566 entries). We used fixed Cys carbamidomethylation, variable methionine (Met) oxidation of peptides, and variable Met loss and acetylation of protein N-terminuses as search variables. Cleavage specificity was set for trypsin with an allowed maximum of two missed cleavages.

For database searching, we used a 10-ppm precursor mass tolerance and a fragment mass tolerance of 0.02 Da followed by data filtering using Percolator, thus resulting in a 1% false discovery rate (FDR). We accepted only peptide to spectrum matches (PSMs) with XCorr >2.2, and we then used the full proteome search result as a focused database for glycopeptide identification (1573 entries). For label-free quantification (LFQ), we used the minora feature detector node with high PSM confidence, a minimum of five non-zero points in a chromatographic trace, a minimum of two isotopes, and a 0.2 min maximum retention time difference for isotope peaks. Then Proteome Discoverer consensus workflow was used to both open the search results and enable retention time (RT) alignment with a maximum 5 min RT shift and 10 ppm mass tolerance to match the precursor between runs. We used the extracted ion chromatogram intensities for the LFQ of peptides. The abundance of each identified protein was determined by averaging ion signals representing each protein's three best ionized unique peptides. Finally, we estimated each protein's relative abundance as the proportion of its abundance to the sum of protein abundances ⁶⁴.

Byonic version 3.10.2 (Protein Metrics Inc.) was used to search the glycoproteomic data against the previously determined targeted milk protein database (1573 entries). The search was based on the above-mentioned shotgun proteome analysis strategy and used the following search parameters: trypsin digestion with a maximum of three missed cleavages; 10 ppm precursor ion mass tolerance; fragmentation type, HCD for stepped HCD files and EThcD for EThcD files; 20 ppm fragment mass tolerance; cysteine carbamidomethylation as a fixed modification; Met oxidation and acetylation at the peptide N-terminus as variable modifications. For glycan analysis, we used a Byonic database of 132 human glycan entries. The maximum number of precursors per scan was set to one and the FDR to 1%. Only PSMs to the EThcD spectra or more than two PSMs to the HCD spectra, both with non-negligible error probabilities $|\log \text{Prob}| > 2.0$ and Byonic scores ≥ 200 , were accepted. The total number of spectra counts for each glycopeptide was used as the quantified value for further analysis.

Statistical Analysis

All statistical analyses were performed using SPSS 21.0 (SPSS Inc., Chicago, IL, USA), Perseus v.1.6.14.0⁶⁵, or under the R Statistical Computing Environment (v. 4.0.2). The Mann-Whitney *U*-test was used to assess differences in skewed data. The Fisher's exact probability method was used to test the categorical variables, and one-way analysis of variance or Student's *t*-test were used to compare the differences in normally distributed data. Principal component analysis and plots were used to conduct data reduction and to visualize the differences in all samples. Statistical significance for all tests was $P < 0.05$. Proteins with an adjusted *p*-value < 0.05 were considered to have undergone significant protein enrichment. For identification of up and down-regulated proteins, a Cytoscape (v3.6.0) App ClueGO was⁶⁶ used based on the terms "biological processes" in the human database. Statistical significance was set for a K score of 0.8. The function "GO Term fusion" was selected and GO term restriction levels were set at 1–20, with a minimum of two proteins or 4% proteins in each GO term.

Data and Software Availabilities

The proteome data (dataset identifier PXD024776) have been deposited to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the iProX partner repository PRIDE⁶⁷. Raw and fasta files for glycopeptome data have been deposited to MassIVE (<https://massive.ucsd.edu/ProteoSAFe/static/massive.jsp>) with the identifier MSV000087229.

ACKNOWLEDGEMENTS

This study was supported by grants from National Natural Science Foundation of China (NSFC 21974069), Hubei Provincial Science and Technology Department Novel Coronavirus Pneumonia Emergency Science and Technology Project (Grant No. 2020FCA011), Wuhan Novel Coronavirus Pneumonia Emergency Science and Technology Tackling Key Project (Grant No. 2020020201010011), China Postdoctoral Science Foundation (No. 2021T140067), and the Fundamental Research Funds for the Central Universities (Grant No. 2042020kfxg17).

AUTHOR CONTRIBUTIONS

J.Z., G.W., Y.Z., Y.Y., and Y.H. designed the research. J.G. and H.L. collected samples and informed patient consents. F.L. stored the specimens and performed virus detection. J.Z., M.T., and Y.T. performed proteomic research and data analysis. M.T., J.Z., G.W., and J.G. wrote the manuscript. J.W. and Y.H. revised the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

REFERENCES

- 1 Li, N. *et al.* Memory CD4(+) T cells are generated in the human fetal intestine. *Nature immunology* **20**, 301-312, doi:10.1038/s41590-018-0294-9 (2019).
- 2 Organization, W. H. COVID-19 clinical management: living guidance, 25 January 2021. (World Health Organization, 2021).
- 3 Wang, Y. *et al.* Impact of Covid-19 in pregnancy on mother's psychological status and infant's neurobehavioral development: a longitudinal cohort study in China. *BMC medicine* **18**, 347, doi:10.1186/s12916-020-01825-1 (2020).
- 4 Tromp, I. *et al.* Breastfeeding and the risk of respiratory tract infections after infancy: The Generation R Study. *PloS one* **12**, e0172763, doi:10.1371/journal.pone.0172763 (2017).
- 5 V'kovski, P., Kratzel, A., Steiner, S., Stalder, H. & Thiel, V. Coronavirus biology and replication: implications for SARS-CoV-2. *Nature reviews. Microbiology* **19**, 155-170, doi:10.1038/s41579-020-00468-6 (2021).
- 6 Woof, J. M. & Russell, M. W. Structure and function relationships in IgA. *Mucosal immunology* **4**, 590-597, doi:10.1038/mi.2011.39 (2011).
- 7 Shehata, L. *et al.* Systematic comparison of respiratory syncytial virus-induced memory B cell responses in two anatomical compartments. *Nature communications* **10**, 1126, doi:10.1038/s41467-019-09085-1 (2019).
- 8 Chantry, C. J., Howard, C. R. & Auinger, P. Full breastfeeding duration and associated decrease in respiratory tract infection in US children. *Pediatrics* **117**, 425-432, doi:10.1542/peds.2004-2283 (2006).
- 9 Georgi, G., Bartke, N., Wiens, F. & Stahl, B. Functional glycans and glycoconjugates in human milk. *Am J Clin Nutr*

- 98**, 578S-585S, doi:10.3945/ajcn.112.039065 (2013).
- 10 Hayashi, H. & Yamashita, Y. Role of N-glycosylation in cell surface expression and protection against proteolysis of the intestinal anion exchanger SLC26A3. *Am J Physiol Cell Physiol* **302**, C781-795, doi:10.1152/ajpcell.00165.2011 (2012).
- 11 Liu, B. & Newburg, D. S. Human milk glycoproteins protect infants against human pathogens. *Breastfeed Med* **8**, 354-362, doi:10.1089/bfm.2013.0016 (2013).
- 12 Dixon, D. L. The Role of Human Milk Immunomodulators in Protecting Against Viral Bronchiolitis and Development of Chronic Wheezing Illness. *Children (Basel)* **2**, 289-304, doi:10.3390/children2030289 (2015).
- 13 Chen, H. *et al.* Clinical characteristics and intrauterine vertical transmission potential of COVID-19 infection in nine pregnant women: a retrospective review of medical records. *Lancet* **395**, 809-815, doi:10.1016/S0140-6736(20)30360-3 (2020).
- 14 Ballard, O. & Morrow, A. L. Human milk composition: nutrients and bioactive factors. *Pediatric clinics of North America* **60**, 49-74, doi:10.1016/j.pcl.2012.10.002 (2013).
- 15 Lönnerdal, B., Erdmann, P., Thakkar, S. K., Sauser, J. & Destailats, F. Longitudinal evolution of true protein, amino acids and bioactive proteins in breast milk: a developmental perspective. *The Journal of nutritional biochemistry* **41**, 1-11, doi:10.1016/j.jnutbio.2016.06.001 (2017).
- 16 Liao, Y. *et al.* Absolute Quantification of Human Milk Caseins and the Whey/Casein Ratio during the First Year of Lactation. *Journal of proteome research* **16**, 4113-4121, doi:10.1021/acs.jproteome.7b00486 (2017).
- 17 Zhu, J. & Dingess, K. A. The Functional Power of the Human Milk Proteome. *Nutrients* **11**, doi:10.3390/nu11081834 (2019).
- 18 Johansen, F. E., Braathen, R. & Brandtzaeg, P. Role of J chain in secretory immunoglobulin formation. *Scand J Immunol* **52**, 240-248, doi:10.1046/j.1365-3083.2000.00790.x (2000).
- 19 Pancera, M. *et al.* Structural basis for diverse N-glycan recognition by HIV-1-neutralizing V1-V2-directed antibody PG16. *Nature structural & molecular biology* **20**, 804 (2013).
- 20 Zhu, J. *et al.* Quantitative Longitudinal Inventory of the N-Glycoproteome of Human Milk from a Single Donor Reveals the Highly Variable Repertoire and Dynamic Site-Specific Changes. *J Proteome Res* **19**, 1941-1952, doi:10.1021/acs.jproteome.9b00753 (2020).
- 21 Fouda, G. G. *et al.* Tenascin-C is an innate broad-spectrum, HIV-1-neutralizing protein in breast milk. *Proceedings of the National Academy of Sciences* **110**, 18220-18225 (2013).
- 22 Dierick, M., Vanrompay, D., Devriendt, B. & Cox, E. Lactoferrin, a versatile natural antimicrobial glycoprotein that modulates the host's innate immunity. *Biochemistry and cell biology = Biochimie et biologie cellulaire* **99**, 61-65, doi:10.1139/bcb-2020-0080 (2021).
- 23 Wockel, A., Abou-Dakn, M., Beggel, A. & Arck, P. Inflammatory breast diseases during lactation: health effects on the newborn-a literature review. *Mediators Inflamm* **2008**, 298760, doi:10.1155/2008/298760 (2008).
- 24 Riskin, A. *et al.* Changes in immunomodulatory constituents of human milk in response to active infection in the nursing infant. *Pediatr Res* **71**, 220-225, doi:10.1038/pr.2011.34 (2012).
- 25 Zhang, L. *et al.* Changes over lactation in breast milk serum proteins involved in the maturation of immune and digestive system of the infant. *J Proteomics* **147**, 40-47, doi:10.1016/j.jprot.2016.02.005 (2016).
- 26 Gao, X. *et al.* Temporal changes in milk proteomes reveal developing milk functions. *Journal of proteome research* **11**, 3897-3907, doi:10.1021/pr3004002 (2012).
- 27 Raheem, R. A., Binns, C. W. & Chih, H. J. Protective effects of breastfeeding against acute respiratory tract infections and diarrhoea: Findings of a cohort study. *Journal of paediatrics and child health* **53**, 271-276, doi:10.1111/jpc.13480 (2017).
- 28 Groß, R. *et al.* Detection of SARS-CoV-2 in human breastmilk. *Lancet (London, England)* **395**, 1757-1758,

doi:10.1016/s0140-6736(20)31181-8 (2020).

- 29 Tam, P. C. K. *et al.* Detectable severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in human breast milk of a mildly symptomatic patient with coronavirus disease 2019 (COVID-19). *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, doi:10.1093/cid/ciaa673 (2020).
- 30 Costa, S. *et al.* Excretion of SARS-CoV-2 in human breast milk. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases*, doi:10.1016/j.cmi.2020.05.027 (2020).
- 31 Zhu, C. *et al.* Breastfeeding Risk from Detectable Severe Acute Respiratory Syndrome Coronavirus 2 in Breastmilk. *The Journal of infection* **81**, 452-482, doi:10.1016/j.jinf.2020.06.001 (2020).
- 32 Wang, L. *et al.* Chinese expert consensus on the perinatal and neonatal management for the prevention and control of the 2019 novel coronavirus infection (First edition). *Annals of translational medicine* **8**, 47, doi:10.21037/atm.2020.02.20 (2020).
- 33 Perrine, C. G. *et al.* Implementation of Hospital Practices Supportive of Breastfeeding in the Context of COVID-19 - United States, July 15-August 20, 2020. *MMWR. Morbidity and mortality weekly report* **69**, 1767-1770, doi:10.15585/mmwr.mm6947a3 (2020).
- 34 Fujita, M. *et al.* In poor families, mothers' milk is richer for daughters than sons: a test of Trivers-Willard hypothesis in agropastoral settlements in Northern Kenya. *American journal of physical anthropology* **149**, 52-59, doi:10.1002/ajpa.22092 (2012).
- 35 Michaelsen, K. F., Skaft, L., Badsberg, J. H. & Jørgensen, M. Variation in macronutrients in human bank milk: influencing factors and implications for human milk banking. *Journal of pediatric gastroenterology and nutrition* **11**, 229-239, doi:10.1097/00005176-199008000-00013 (1990).
- 36 Shanes, E. D. *et al.* Placental Pathology in COVID-19. *American journal of clinical pathology* **154**, 23-32, doi:10.1093/ajcp/aqaa089 (2020).
- 37 Lee, P. C., Struve, M. & Raff, H. Effects of hypoxia on the development of intestinal enzymes in neonatal and juvenile rats. *Experimental biology and medicine (Maywood, N.J.)* **228**, 717-723, doi:10.1177/153537020322800611 (2003).
- 38 Osorio, J. S. Gut health, stress, and immunity in neonatal dairy calves: the host side of host-pathogen interactions. *Journal of animal science and biotechnology* **11**, 105, doi:10.1186/s40104-020-00509-3 (2020).
- 39 Lönnerdal, B. Human milk proteins: key components for the biological activity of human milk. *Advances in experimental medicine and biology* **554**, 11-25 (2004).
- 40 Lönnerdal, B. Bioactive Proteins in Human Milk: Health, Nutrition, and Implications for Infant Formulas. *The Journal of pediatrics* **173 Suppl**, S4-9, doi:10.1016/j.jpeds.2016.02.070 (2016).
- 41 Dingess, K. A. *et al.* Monitoring Human Milk β -Casein Phosphorylation and O-Glycosylation Over Lactation Reveals Distinct Differences between the Proteome and Endogenous Peptidome. *International journal of molecular sciences* **22**, doi:10.3390/ijms22158140 (2021).
- 42 Fan, H. *et al.* The effect of whey protein on viral infection and replication of SARS-CoV-2 and pangolin coronavirus in vitro. *Signal Transduct Target Ther* **5**, 275, doi:10.1038/s41392-020-00408-z (2020).
- 43 Brandtzaeg, P. The mucosal immune system and its integration with the mammary glands. *The Journal of pediatrics* **156**, S8-15, doi:10.1016/j.jpeds.2009.11.014 (2010).
- 44 Goldman, A. S. Modulation of the gastrointestinal tract of infants by human milk. Interfaces and interactions. An evolutionary perspective. *The Journal of nutrition* **130**, 426s-431s, doi:10.1093/jn/130.2.426S (2000).
- 45 Geyer, P. E. *et al.* High-resolution longitudinal serum proteome trajectories in COVID-19 reveal patients-specific seroconversion. *BMJ MedRxiv*, 2021.2002.2022.21252236, doi:10.1101/2021.02.22.21252236 %J medRxiv (2021).

- 46 van Keulen, B. J. *et al.* Human Milk from Previously COVID-19-Infected Mothers: The Effect of Pasteurization on Specific Antibodies and Neutralization Capacity. *Nutrients* **13**, doi:10.3390/nu13051645 (2021).
- 47 Pace, R. M. *et al.* Characterization of SARS-CoV-2 RNA, Antibodies, and Neutralizing Capacity in Milk Produced by Women with COVID-19. *mBio* **12**, doi:10.1128/mBio.03192-20 (2021).
- 48 Hurley, W. L. & Theil, P. K. Perspectives on immunoglobulins in colostrum and milk. *Nutrients* **3**, 442-474, doi:10.3390/nu3040442 (2011).
- 49 Hanson, L. A. & Korotkova, M. The role of breastfeeding in prevention of neonatal infection. *Seminars in neonatology : SN* **7**, 275-281, doi:10.1016/s1084-2756(02)90124-7 (2002).
- 50 Mantis, N. J., Rol, N. & Corthesy, B. Secretory IgA's complex roles in immunity and mucosal homeostasis in the gut. *Mucosal Immunol* **4**, 603-611, doi:10.1038/mi.2011.41 (2011).
- 51 Van de Perre, P. Transfer of antibody via mother's milk. *Vaccine* **21**, 3374-3376 (2003).
- 52 Newburg, D. S. Innate immunity and human milk. *J Nutr* **135**, 1308-1312, doi:10.1093/jn/135.5.1308 (2005).
- 53 Fox, A. *et al.* Robust and Specific Secretory IgA Against SARS-CoV-2 Detected in Human Milk. *iScience* **23**, 101735, doi:<https://doi.org/10.1016/j.isci.2020.101735> (2020).
- 54 Narayanaswamy, V. *et al.* Humoral and Cell-Mediated Immune Response in Colostrum from Women Diagnosed Positive for SARS-CoV-2. *Breastfeeding Medicine* **16**, 987-994, doi:10.1089/bfm.2021.0082 (2021).
- 55 Ryu, J. & Lee, C. Regulatory Nucleotide Sequence Signals for Expression of the Genes Encoding Ribosomal Proteins. *Frontiers in genetics* **11**, 501, doi:10.3389/fgene.2020.00501 (2020).
- 56 Bhatt, P. R. *et al.* Structural basis of ribosomal frameshifting during translation of the SARS-CoV-2 RNA genome. *Science (New York, N.Y.)*, doi:10.1126/science.abf3546 (2021).
- 57 Li, Y. *et al.* The Importance of Glycans of Viral and Host Proteins in Enveloped Virus Infection. *Front Immunol* **12**, 638573, doi:10.3389/fimmu.2021.638573 (2021).
- 58 Watanabe, Y., Allen, J. D., Wrapp, D., McLellan, J. S. & Crispin, M. Site-specific glycan analysis of the SARS-CoV-2 spike. *Science* **369**, 330-333, doi:10.1126/science.abb9983 (2020).
- 59 Grant, O. C., Montgomery, D., Ito, K. & Woods, R. J. Analysis of the SARS-CoV-2 spike protein glycan shield reveals implications for immune recognition. *Sci Rep* **10**, 14991, doi:10.1038/s41598-020-71748-7 (2020).
- 60 Yang, Q. *et al.* Inhibition of SARS-CoV-2 viral entry upon blocking N- and O-glycan elaboration. *Elife* **9**, doi:10.7554/eLife.61552 (2020).
- 61 Wielgat, P., Rogowski, K., Godlewska, K. & Car, H. Coronaviruses: Is Sialic Acid a Gate to the Eye of Cytokine Storm? From the Entry to the Effects. *Cells* **9**, doi:10.3390/cells9091963 (2020).
- 62 Roberts, D. S. *et al.* Structural O-Glycoform Heterogeneity of the SARS-CoV-2 Spike Protein Receptor-Binding Domain Revealed by Top-Down Mass Spectrometry. *Journal of the American Chemical Society*, doi:10.1021/jacs.1c02713 (2021).
- 63 UniProt, C. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* **47**, D506-D515, doi:10.1093/nar/gky1049 (2019).
- 64 Ahrne, E., Molzahn, L., Glatter, T. & Schmidt, A. Critical assessment of proteome-wide label-free absolute abundance estimation strategies. *Proteomics* **13**, 2567-2578, doi:10.1002/pmic.201300135 (2013).
- 65 Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods* **13**, 731-740, doi:10.1038/nmeth.3901 (2016).
- 66 Mlecnik, B., Galon, J. & Bindea, G. Comprehensive functional analysis of large lists of genes and proteins. *Journal of Proteomics* **171**, 2-10, doi:<https://doi.org/10.1016/j.jprote.2017.03.016> (2018).
- 67 Ma, J. *et al.* iProX: an integrated proteome resource. *Nucleic Acids Res* **47**, D1211-D1217, doi:10.1093/nar/gky869 (2019).