Received: 7 September 2023 | Revised: 26 October 2023 | Accepted: 2 November 2023

DOI: 10.1002/jex2.128

RESEARCH ARTICLE

Gestational age at birth influences protein and RNA content in human milk extracellular vesicles

Brett Vahkal[1,2,3] | **Illimar Altosaar**[1] | **Eric Tremblay**[4] | **David Gagné**[4] | **Nico Hüttman**[5] | **Zoran Minic**[5] | **Marceline Côté**[1,2,3] | **Alexandre Blais**[1,2,3,6,7] | **Jean-François Beaulieu**[4] | **Emanuela Ferretti**[8]

[1]Department of Biochemistry, Microbiology and Immunology, University of Ottawa, Ottawa, Canada

[2]uOttawa Centre for Infection, Immunity, and Inflammation, Ottawa, Canada

[3]Ottawa Institute of Systems Biology, Ottawa, Canada

[4]Department of Immunology and Cell Biology, Université de Sherbrooke, Sherbrooke, Canada

[5]Faculty of Science, John L. Holmes Mass Spectrometry Facility, University of Ottawa, Ottawa, Canada

[6]Brain and Mind Institute, University of Ottawa, Ottawa, Canada

[7]Éric Poulin Centre for Neuromuscular Disease, Ottawa, Canada

[8]Department of Pediatrics, Division of Neonatology, Children's Hospital of Eastern Ontario, Ottawa, Canada

Correspondence
Emanuela Ferretti, Division of Neonatology, Children's Hospital of Eastern Ontario, 401 Smyth Road, Ottawa, ON, K1H 8L1, Canada.
Email: eferretti@toh.ca

Alexandre Blais, University of Ottawa, Faculty of Medicine, 451 Smyth Rd, Ottawa, ON, K1H 8M5, Canada.
Email: Alexandre.Blais@uottawa.ca

Jean-François Beaulieu, Université de Sherbrooke, Faculty of Medicine and Health Sciences, 3001, 12th Ave North, Sherbrooke, QC, J1H 5N4 Canada.
Email: Jean-Francois.Beaulieu@usherbrooke.ca

Present address
Illimar Altosaar, Proteins Easy Corp, 75 Campus Drive, Kemptville Agricultural College, North Grenville, Ontario K0G 1J0.

Funding information
Canadian Institutes of Health Research, Grant/Award Numbers: PJT 162423, PJT 183839; Ontario Ministry of Research and Innovation, Grant/Award Number: ER18-14-091; Western Canada Research Grid; Alliance de recherche numérique du Canada; Canada Research Chairs, Grant/Award Number: 950-232840

Abstract

Human milk extracellular vesicles (HM EVs) are proposed to protect against disease development in infants. This protection could in part be facilitated by the bioactive EV cargo of proteins and RNA. Notably, mothers birth infants of different gestational ages with unique needs, wherein the EV cargo of HM may diverge. We collected HM from lactating mothers within two weeks of a term or preterm birth. Following purification of EVs, proteins and mRNA were extracted for proteomics and sequencing analyses, respectively. Over 2000 protein groups were identified, and over 8000 genes were quantified. The total number of proteins and mRNA did not differ significantly between the two conditions, while functional bioinformatics of differentially expressed cargo indicated enrichment in immunoregulatory cargo for preterm HM EVs. In term HM EVs, significantly upregulated cargo was enriched in metabolism-related functions. Based on gene expression signatures from HM-contained single cell sequencing data, we proposed that a larger portion of preterm HM EVs are secreted by immune cells, whereas term HM EVs contain more signatures of lactocyte epithelial cells. Proposed differences in EV cargo could indicate variation in mother's milk based on infants' gestational age and provide basis for further functional characterisation.

KEYWORDS
EVs - extracellular vesicles, GA - gestational age, HM - Human milk, preterm, proteomics, RNA sequencing, term

1 | INTRODUCTION

Human milk (HM) is the gold standard in nutrition for infants (Ballard & Morrow, 2013). HM feeding can prevent the development of inflammatory gut diseases (Herrmann & Carroll, 2014; Underwood, 2013), promote antimicrobial protection, and immune modulation, via milk-contained oligosaccharides, RNA and bioactive proteins such as secretory antibodies, lactoferrins (LF) and cytokines (Ballard & Morrow, 2013; Floris et al., 2016; Kosaka et al., 2010; Rani et al., 2017; Ward et al., 2013). This biological regulation could extend to HM-contained extracellular vesicles (EVs), which carry proteins and nucleic acids (Galley & Besner, 2020). Uniquely, HM EVs can withstand digestion and are taken up by intestinal cells (Rani et al., 2017). Thus, EV-contained cargo is hypothesised to reach the infant gut lumen, setting the stage for a temporal biological effect—*ut digeratur*. Understanding the composition of proteins and nucleic acids carried by EVs may therefore inform on the downstream outcome of HM feeding.

As infants are born at different gestational ages (GA)—term (>37 weeks) or preterm (<37 weeks) (Spong et al., 2011), the composition of mother's milk varies. GA-specific differences in dozens of HM proteins have been measured (Cattaneo et al., 2019; Demers-Mathieu et al., 2017; Molinari et al., 2012; Trend et al., 2016; Wan et al., 2013). Preterm HM has increased proteolytic activity, which could aid in the digestion of milk nutrients in premature infants (Dallas et al., 2015). The increased concentration of epidermal growth factor, transforming growth factor alpha (TGF-α), soluble CD14, immunoglobulin A, LF and lysozyme in preterm HM (Dvorak et al., 2003; Montagne et al., 1999; Ronayne De Ferrer et al., 2021; Trend et al., 2016) could modulate the immune system and prevent development of inflammatory disease in the more vulnerable premature infants. The changes in HM composition may arise due to the unique needs of infants and could extend to the HM EV cargo.

Studies of the HM EV proteome have resulted in over 1000 proteins identified, with a total of 1963 proteins identified in mature HM EVs (>2 weeks postpartum) (Admyre et al., 2007; Larssen et al., 2017; Van Herwijnen et al., 2016; Wang et al., 2019; Yang et al., 2017). In GA-specific HM EVs, 46 proteins have been investigated for differential expression. LF and lactadherin-derived peptides were upregulated in preterm HM EVs and proposed to increase gut epithelial cell proliferation and migration (Wang et al., 2019). Consistent levels of LF have been previously identified in HM EVs, and the protein content was highly resistant to digestion (Liao et al., 2017).

For HM RNA, sequencing has been performed in the HM fat layer (Lemay et al., 2013), and cells (Gleeson et al., 2023; Martin Carli et al., 2020; Nyquist et al., 2022; Twigger et al., 2022), where the latter could also provide information about the origins of HM EVs (Hu et al., 2021). Several studies have focused on microRNA cargo (Alsaweed et al., 2015; Carrillo-Lozano et al., 2020; Do et al., 2017; Floris et al., 2016; Golan-Gerstl et al., 2017; Herwijnen et al., 2018; Kupsco et al., 2021; Lefèvre et al., 2019; Leiferman et al., 2019; Liao et al., 2017; Na et al., 2015; Shu et al., 2015; Zhou et al., 2011), mostly EV-contained, though with limited information on GA-specific changes (Floris et al., 2015; Kahn et al., 2018; Shiff et al., 2019). Two studies have previously characterised GA-specific differences in long-coding, circular RNAs, and mRNAs. For circular RNAs, both term and preterm HM EVs were enriched in VEGF-signalling pathway markers, wherein differentially expressed RNAs regulated inflammatory pathways via their predicted microRNA targets (Zhou et al., 2021). In long-coding RNA, differentially expressed transcripts were similarly enriched in regulation of the inflammatory response, specifically those regulated by cell surface receptors. In the same study, EV-contained mRNAs were sequenced. However, differentially expressed mRNA were not functionally characterised by term of birth, but rather described in the context of interactions with long-coding RNAs. These specific interactions focused on upregulated preterm EV mRNAs, which were determined to be enriched in immune, signalling and metabolic pathways (Yan et al., 2022).

Though differential composition of HM EV cargo of proteins has been described following both term and preterm birth (Chen et al., 2021; Kahn et al., 2018; Wang et al., 2019; Yan et al., 2022), an overall limited number of individual proteins with quantitative information have been identified. Additional low abundance proteins remain elusive, but their identification could expand functional interpretations for HM EVs. Prior studies have used label-based, or data-dependent mass spectrometry detection methods. Here, we applied a data-independent approach, which allows for increased proteome coverage (He et al., 2019; Meyer, 2019; Vahkal et al., 2021).

Similarly for HM EV mRNA, one study has previously examined EV-contained mRNA in term and preterm EVs, with a full characterisation focusing on long-coding RNAs (Yan et al., 2022). As such, quantitative, full-spectrum profiling of GA-specific HM EV cargo is lacking in the current field of HM EV biology.

In this study, we quantitatively measured GA-specific differences in HM EV cargo of proteins and mRNA. By employing mass spectrometry with data-independent acquisition we obtained a deep proteome coverage and quantification of over 2000 HM EV-contained proteins. Using RNA sequencing, total mRNA in GA-specific EVs was detected. With bioinformatic analyses, we postulate significant quantitative differences on individual protein and mRNA levels, wherein preterm HM EV contents have functional enrichments in immune-related pathways and processes. Term HM EVs contain diverse cargo with highly enriched metabolic-related functions, including pathways and processes implicated in signalling. To postulate the origin of vesicles, we compared the signatures of EV RNA cargo to that of previously published HM single cell sequencing. We show that preterm EVs cluster with immune cell signatures, whereas term EVs have signatures of lactocyte epithelial cells. Overall, our results indicate

differences in HM EV cargo based on duration of gestation, wherein the cellular origin and cargo of preterm HM EVs are more immune-related.

2 | MATERIALS AND METHODS

2.1 | Milk collection

HM from donors was collected according to ethics protocols at the University of Ottawa and Children's Hospital of Eastern Ontario (#H-03-20-5643). Once written consent was obtained, donors were asked to sterilise the breast with an antibacterial wipe and manually express their milk into a sterile collection tube. A total of 20 mL of milk was collected from each anonymous donor: 6 term and 6 preterm. Caesarean birth, antibiotic use and any active maternal genetic, immune and chronic inflammatory disease resulted in donor exclusion. Mother's age and parity, infant sex, age, gestational age and birth weight, when available, were recorded (Supplementary file—Donor information). HM was classified as term for GA of more than 37 weeks, and preterm for GA of less than 37 weeks (Spong et al., 2011). Following collection, HM was transported to the lab to commence EV isolation within 30 minutes.

2.2 | Extracellular vesicle isolation

Prior literature on milk EV isolation from diverse species has established that differential and ultracentrifugation is appropriate to obtain a sample enriched in EVs (Admyre et al., 2007; De La Torre Gomez et al., 2018; Pluchino & Smith, 2019; Yamada et al., 2012; Zhou et al., 2020). We have previously described an isolation method for HM EVs with an end goal of proteomics analysis (Vahkal et al., 2021). Briefly, collected milk was centrifuged at 4600 × *g* to remove fat, cells and cell debris. The skimmed milk was then centrifuged at 20,000 × *g* to remove the remaining fat, and pellet large apoptotic and mitochondrial vesicles. The supernatant was subjected to two rounds of ultracentrifugation at 100,000 × *g*. The pellet was resuspended in 400 μL of sterile PBS and stored in −80°C in aliquots until further analysis.

2.3 | Extracellular vesicle characterisation

HM EV surface markers and size were characterised using the Exoview R100 instrument (Nanoview BioSciences, USA), following protocols previously published (Jung et al., 2020). Briefly, EVs were incubated overnight on Exoview chips with immunocapture spots for CD81, CD63, CD9 and mIgG. Following incubation, the chips were washed three times with an incubation solution. The chips were stained with anti-CD9-AF488, CD63-AF647 and CD81-AF555 antibodies for 1 h. After another three washes, the chips were imaged using the Exoview R100 instrument. The acquired images were analysed for fluorescence and size utilising the ExoScan 2.5.5 acquisition software. Graphs were prepared using Prism 9.

HM EVs were visualised using scanning electron microscopy. EVs were prepared by re-suspending in PBS, then pelleted at 100,000 × *g* for 1.5 h, at 4°C. The EV pellet was fixed in 4% glutaraldehyde for 15 min, and subsequently washed twice with PBS for 10 min each. The pellet was resuspended in 10 μL ddH_2O and placed on an EM silicon chip to air-dry, followed by transfer to a 96-well plate. The sample was dehydrated in ascending ethanol (40%, 60%, 80%, 96%—98%) by pipetting ethanol down the side of the well thereby immersing the silicon chip. After ethanol was evaporated, the samples were left to dry at room temperature for 24 h. The following day, silicon chips were mounted on a carbon stage by removing the adhesive, and Au-coated. The samples were imaged on a JSM-7500F FESEM (JEOL) microscope (University of Ottawa Materials Characterization Core Facility).

2.4 | Proteomics

The liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) data acquisitions were performed at PhenoSwitch Bioscience Inc. (Sherbrooke, QC, Canada), using a Sciex TripleTOF 6600 instrument (Sciex, Foster City, CA). All LC-MS/MS equipment and settings were used as previously described (Gagné et al., 2022). In brief, six individual samples (30 µg each) were analysed in Sequential Window Acquisition of All Theoretical Mass Spectra (SWATH) acquisition mode, with a 60-min LC gradient. Analyst TF 1.8 software (Sciex) was used for instrument control, data processing and acquisition.

To analyse the MS data, we employed a hybrid strategy based on a previously described proteomics pipeline (Gagné et al., 2022). Firstly, we used MSConvert to convert the file format from WIFF to mzML (Adusumilli & Mallick, 2017). Then, we utilised proteomics analysis tools (MSFragger, Philosopher, EasyPQP) from the FragPipe GUI (v.17.1; https://fragpipe.nesvilab.org/) to search the human proteome (UP000005640, www.uniprot.org) and build preliminary libraries at 5% FDR (false discovery rate)

(Da Veiga Leprevost et al., 2020; Kong et al., 2017). Next, we retrieved protein IDs from these libraries and inputted them into the ID mapping tool from UniProt.org (www.uniprot.org/id-mapping) to generate a FASTA file of protein targets. Finally, we employed this FASTA file to perform a "library-free search" with methionine oxidation set as variable modification and MBR (match between runs) enabled in DIA-NN (GUI; v.1.8), applying a more stringent 1% FDR to ensure increased confidence in the identifications and quantifications of peptides and proteins (Demichev et al., 2020).

For differential expression analysis, the tsv file generated by DIA-NN was curated and directly imported into R. All analyses were performed using the pOmics R package (github.com/nicohuttmann/pOmics). Further analysis only considered proteins that were quantified in at least 50% of the samples and imputation was performed using a shifted Gaussian distribution (shift = 1.8 standard deviations (SD), width = 0.3 SD) to account for missing protein MaxLFQ intensities. The composition of the identified proteins was compared with all human proteins to identify over-represented Gene Ontology (GO) cellular components (Aleksander et al., 2023; Ashburner et al., 2000). To identify differentially abundant proteins between term and preterm groups, pairwise Student's *t*-tests were conducted. Due to the limited sample size, *p*-values were not corrected. Proteins were considered significantly changed based on a threshold of $p < 0.05$ and an absolute log2 fold-change >0.5.

Based on the sets of significantly up- and downregulated proteins, functional enrichment analysis (over-representation analysis) was performed using GO biological process parameters. Further functional annotations were done using STRING (https://www.string-db.org) and PANTHER (http://www.pantherdb.org/) databases according to respective presets.

2.5 | RNA sequencing and analysis

RNA from HM EVs was extracted using Qiagen miRNeasy kit (Qiagen, Germany) coupled with a phenol-chloroform precipitation step. RNA quality was assessed using LibQC (McGill Sequencing Facility). Total RNA from term and preterm HM EV samples ($n = 4$, each group) was sequenced at the McGill Sequencing Facility (Montreal, Canada). RNA libraries were prepared using NEBNextUltra II Directional RNA kit, and sequenced using Illumina NovaSeq 6000. Nf-core rnaseq nextflow pipeline was employed with default settings (Di Tommaso et al., 2017; Ewels et al., 2020). The reads were aligned to the GRCh38 reference genome using STAR (Dobin et al., 2013). Reads aligning to exons were counted using Subread *featureCounts*, and counts were summarised over genes (Liao et al., 2014).

Differential expression was analysed using edgeR with a library size correction by the Trimmed Mean of M-values (TMM) algorithm. To remove unwanted variation, RUVseq method was employed, using two algorithms—RUVs and RUVr, with values of the "k" factor from 1 to 5 (Peixoto et al., 2015; Risso et al., 2014). Implementing RUVr or RUVs with increasing values of k improved the clustering of samples, and RUVr $k = 5$ was chosen for further downstream analysis based on sample clustering by PCA and *p*-value distribution histograms. Gene set enrichment analysis (GSEA) was further performed in R/Bioconductor using the *clusterProfiler* and *msigdb* packages (Liberzon et al., 2011; Subramanian et al., 2005; Wu et al., 2021). The gene sets used were limited to the MSigDB groups "Hallmark," C2, C3, C5 and C7. Only sets containing between 10 and 1000 genes were surveyed as a cut-off independent method to identify significantly enriched gene sets.

3 | RESULTS

3.1 | Validation of human milk EVs

To investigate small EV surface marker abundance, fluorescence based measurement was used (Jeppesen et al., 2019; Pluchino & Smith, 2019; Théry et al., 2018). HM EV surface markers CD9, CD63 and CD81 were identified on term and preterm EVs using Exoview R-100 (Nanoview Biosciences, USA). The Exoview platform utilises three channels of fluorescence for CD9, CD63 and CD81 detection using fluorescently tagged antibodies, and one channel for interferometric sizing. EVs from both groups of samples had presence of CD9, CD81 and CD63 signal (Figure 1a). The average size for EVs isolated from term or preterm HM was 60 and 59 nm, respectively (Figure 1b). To visualise EV integrity and size, scanning electron microscopy was performed. Spherical vesicles, up to 200 nm in size, were visible in the field of view (Figure 1c).

3.2 | Identification and differential expression of proteins from GA-specific human milk EVs

Label-free quantitative mass spectrometry of six HM samples ($n = 3$ each term and preterm) was performed using SWATH—a method of data independent acquisition for the detection peptides (Chutipongtanate & Greis, 2018). Over 19,000 unique peptides and 2000 protein groups were identified and quantified. Similar number of proteins were identified in term and preterm HM EVs—2094 and 1709, respectively (Supplementary data 1). In addition to 837 common proteins, term HM EVs contained 1257, while preterm milk samples contained 872 unique proteins (Figure 2).

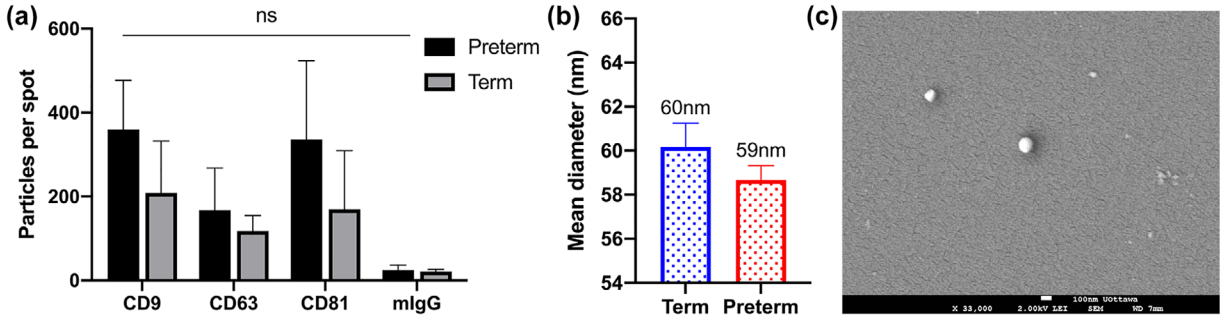

FIGURE 1 HM EV validation and surface marker characterization. (a) Surface markers CD9, CD63, CD81 and mIgG negative control, were identified on term ($n = 2$) and preterm ($n = 4$) EVs. ns, non-significant, two-way ANOVA. (b) Mean size of particles with a cut-off scale of 50–200 nm, bound to CD9, CD63 and CD81 capture probes, measured by interferometry. (c) Scanning electron micrograph of individual HM preterm vesicles, scale bar to 100 nm.

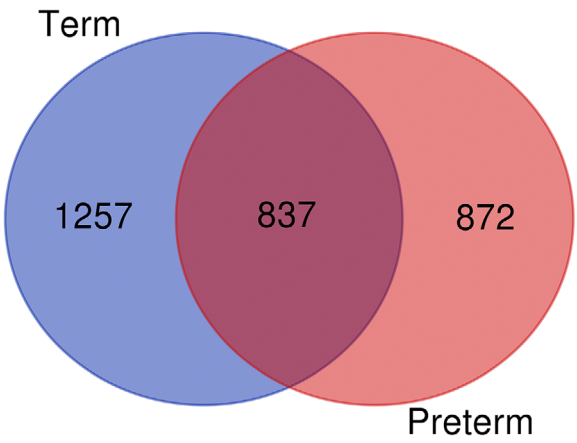

FIGURE 2 Number of proteins in term or preterm human milk EVs. (a) Total number of proteins identified using SWATH method – 2094 in term, 1709 in preterm; 837 proteins common to both; 1257 and 872 unique proteins in term and preterm cohorts respectively; $n = 3$. High FDR confidence proteins included in analysis.

Quantitative evaluation of protein abundances was compared between term and preterm HM EV using hierarchical cluster analysis. Large differences between samples and no distant clustering were observed among the two conditions (Figure 3a). Following differential expression analysis, few significantly ($p < 0.05$) up- or downregulated proteins were identified when compared to the total number of identified proteins (Figure 3b). The results revealed 15 and 25 significantly upregulated proteins in term and preterm HM EVs, respectively (Table 1). Cathepsin B (CTSB) and serpin family E member 2 (SERPINE2) were the most significantly upregulated proteins ($p < 0.04$) in term HM EVs, with SERPINE2 highest upregulated. In preterm HM EVs, endosulfine alpha (ENSA) was the most significantly upregulated protein, while elastin microfibril interfacer 3 (EMILIN3) and fibroblast growth factor binding protein 1 (FGFBP1) were among the highest upregulated proteins by fold-change.

3.3 | Functional annotation of GA-specific human milk EV protein cargo

Identified term and preterm proteins were subjected to biological process and pathway analysis. Proteins were queried in the PANTHER database for biological processes analysis, wherein both term and preterm samples revealed the presence of proteins related to immune system processes (Figure 4). The total number of proteins involved in each biological process was higher for preterm HM EVs (Figure 4 and Supplementary data 2). Reactome pathway analysis of total protein abundance using STRING database revealed that term EVs had a higher number of unique enriched pathways, including those related to developmental biology, nervous system development and infectious disease (Figure 5).

Differentially expressed proteins with significant up- and downregulation were submitted for combined analysis in STRING. In total, 11 GO biological processes, and neutrophil degranulation Reactome pathway were enriched in the dataset. Among the enriched biological processes, neutrophil degranulation and leukocyte activation involved in immune response were in the top three of largest enrichment effect in the dataset and included proteins that were both up- and downregulated in preterm HM EVs. Neutrophil degranulation and leukocyte activation processes included upregulated proteins HSPA8, PYGL, CTSB and

TABLE 1 Comparative analysis of differently abundant proteins in term and preterm HM EVs. Differential expression was assessed by pairwise Student's *t*-test and proteins were considered significantly changed based on $p < 0.05$ and absolute log 2 fold-change >0.5. Proteins with positive fold-change value were more highly abundant in term HM EVs, while those with negative fold-change were more highly abundant in preterm HM EVs.

Condition of birth	Protein name	Gene symbol	Fold-change (log2)	*p*-value
Term	Cathepsin B	*CTSB*	1.75	1.79e-03
	Serpin family E member 2	*SERPINE2*	2.7	9.31e-03
	Glycogen phosphorylase L	*PYGL*	1.53	1.60e-02
	Argininosuccinate synthase 1	*ASS1*	1.28	1.98e-02
	Guanine nucleotide-binding protein G(i) subunit alpha-2	*GNAI2*	0.91	2.11e-02
	Sorting nexin 3	*SNX3*	0.89	2.48e-02
	Vitronectin	*VTN*	1.27	2.53e-02
	Sorting nexin 2	*SNX2*	1.13	3.02e-02
	Cellular communication network factor 1	*CCN1*	2.34	3.16e-02
	Heat shock cognate 71 kDa protein	*HSPA8*	0.8	3.19e-02
	Solute carrier family 2 member 1	*SLC2A1*	1.37	3.65e-02
	CYFIP-related Rac1 interactor A	*LOC110522500*	0.52	3.66e-02
	Transmembrane p24 trafficking protein 7	*TMED7*	1.8	3.85e-02
	Tax1 binding protein 1	*TAX1BP1*	0.91	4.39e-02
	Claudin 8	*CLDN8*	1.15	4.67e-02
Preterm	Endosulfine alpha	*ENSA*	−1.29	1.91e-03
	Vesicle amine transport 1	*VAT1*	−0.54	4.51e-03
	Syntaxin binding protein 1	*STXBP1*	−0.82	6.59e-03
	Dehydrogenase/reductase 1	*DHRS1*	−1.81	1.13e-02
	Elastin microfibril interfacer 3	*EMILIN3*	−3.68	1.29e-02
	Platelet activating factor acetylhydrolase 1b catalytic subunit 2	*PAFAH1B2*	−0.66	1.56e-02
	Midkine	*MDK*	−1.98	1.72e-02
	Xanthine dehydrogenase	*XDH*	−0.95	1.94e-02
	Serine/threonine-protein kinase TAO1	*TAOK1*	−1.08	2.22e-02
	Albumin	*ALB*	−1.37	2.55e-02
	Cellular retinoic acid binding protein 2	*CRABP2*	−1.14	2.92e-02
	Unconventional myosin-VI	*MYO6*	−0.81	2.97e-02
	Thymosin beta 4 X-linked	*TMSB4X*	−1	3.13e-02
	Fibroblast growth factor binding protein 1	*FGFBP1*	−2.88	3.22e-02
	Growth regulating estrogen receptor binding 1	*GREB1*	−0.75	3.49e-02
	Phosphate cytidylyltransferase 2, ethanolamine	*PCYT2*	−1.37	3.64e-02
	Mucin-4	*MUC4*	−1.51	3.82e-02
	S100 calcium binding protein A7	*S100A7*	−1.26	4.11e-02
	Butyrophilin subfamily 1 member A1	*BTN1A1*	−1.02	4.19e-02
	Solute carrier organic anion transporter family member 4C1	*SLCO4C1*	−0.96	4.19e-02
	Caspase 14	*CASP14*	−0.68	4.33e-02
	Tandem C2 domains, nuclear	*TC2N*	−0.7	4.82e-02
	Calcium binding protein 39 like	*CAB39L*	−0.71	4.87e-02
	C1q and TNF related 1	*C1QTNF1*	−1.04	4.89e-02
	Secreted frizzled related protein 1	*SFRP1*	−1.57	4.95e-02

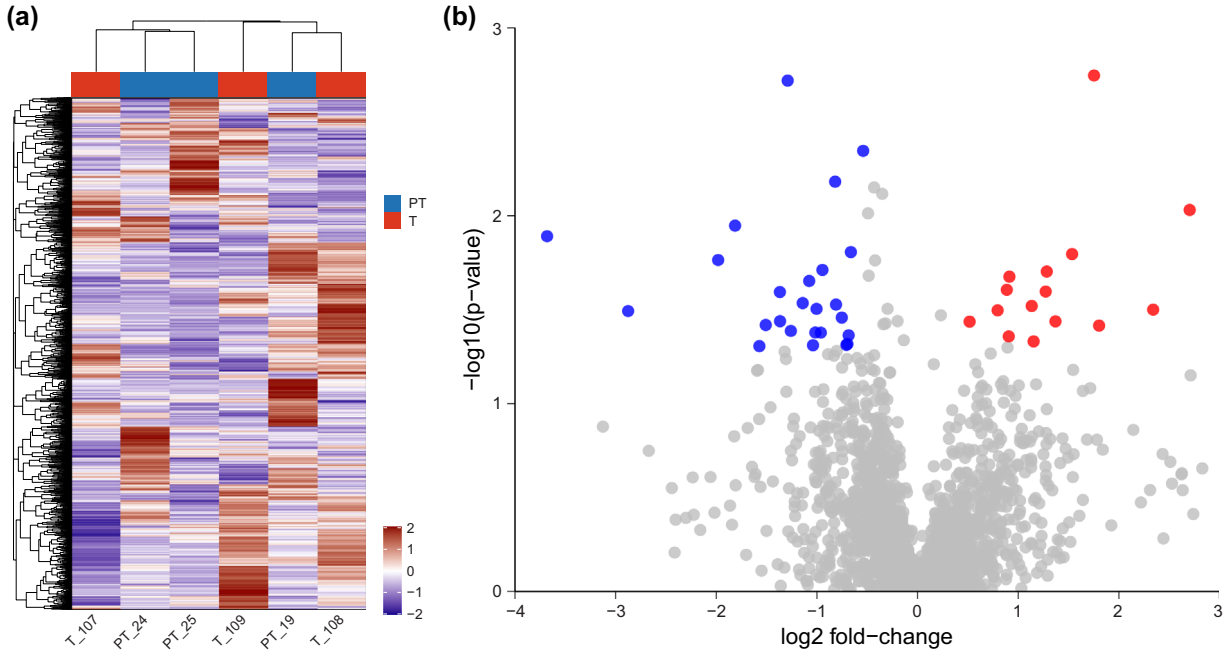

FIGURE 3 Differentially expressed proteins in term and preterm HM EVs. (a) Heatmap showing hierarchical clustering of Z-scored MaxLFQ intensities on protein and sample level. (b) Volcano plot of differentially expressed proteins. Proteins with positive log2 fold-change value were higher abundant term (red), proteins with a negative log2 fold-change were upregulated in preterm (blue), proteins not significantly changed are shown in grey. Significance was assigned to proteins with a $p < 0.05$, pairwise Student's *t*-test.

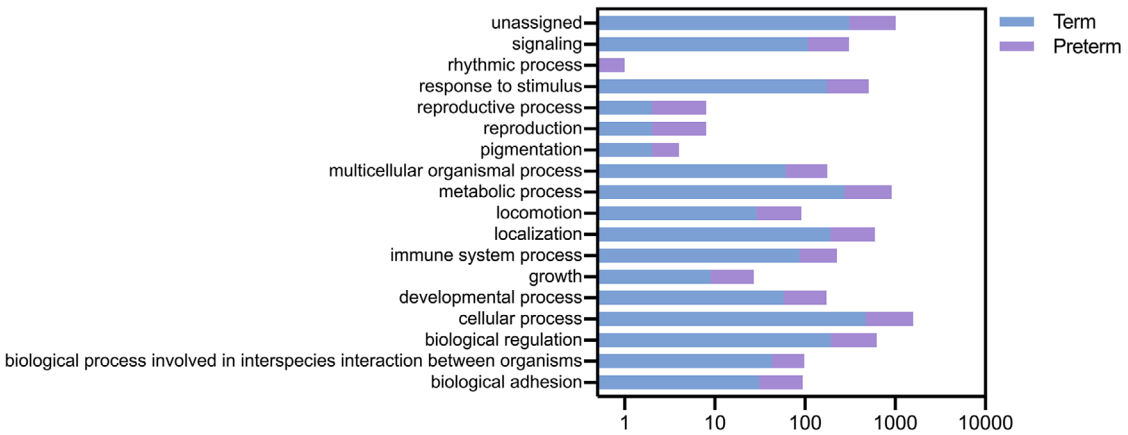

FIGURE 4 Biological processes identified with PANTHER database for all term and preterm HM EV proteins. X-axis represents the log of number of proteins assigned to each biological process that were detected in term (blue bars) or preterm (pink bars) HM EVs. Total number of proteins assigned to biological processes was larger in preterm HM EVs.

downregulated proteins VAT1, PAFAH1B2, RAB6A, GDI1 and MDK, latter of which was specific for leukocyte activation only. The identified biological processes for regulation of signalling and signal transduction contained proteins that were primarily downregulated in preterm HM EVs (Figure 6). The other enriched biological processes belonged to secretion, vesicle-mediated transport, cell communication and exocytosis functions, which included both up- and downregulated proteins (Supplementary data 3).

3.4 | GA-specific human milk EV cargo in previously identified human milk proteins

To further validate our proteomics findings, we compared our data to that of previously published proteomics analyses from HM, or HM EVs, whenever publicly available and clearly annotated.

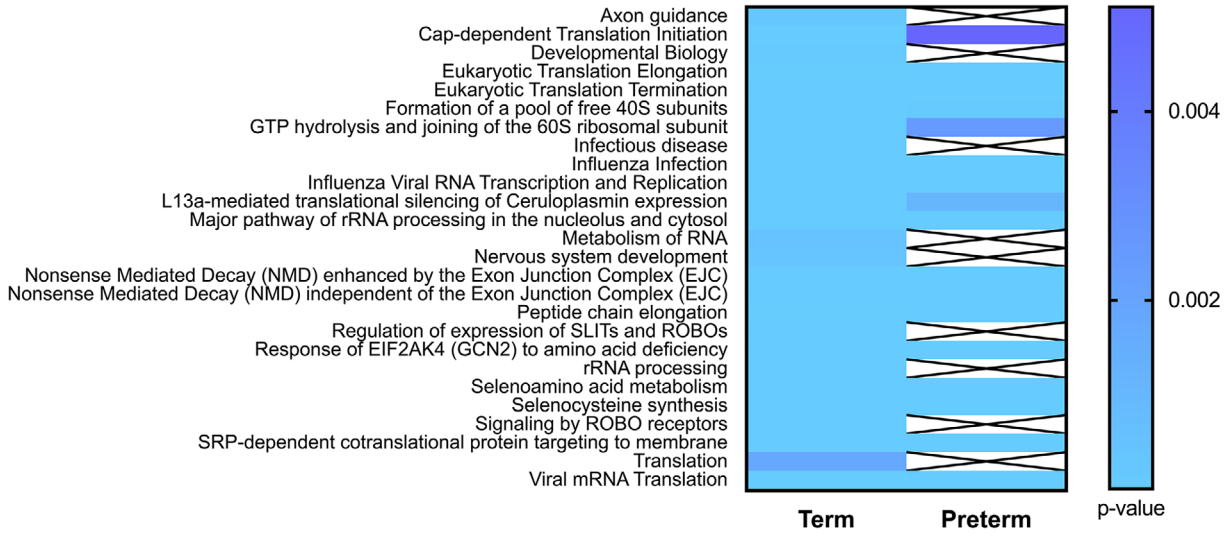

FIGURE 5 Reactome pathways identified with STRING database of all term and preterm HM EV proteins. Blue gradient cells represent the significance of pathway in the dataset. Term HM EV proteins were enriched in more diverse pathways with higher significance. Cells with black lines indicate lack of identification for corresponding pathway. Proteins with abundance values and FDR <1% were queried.

First, we compared our results against all the proteins curated in the Human Biofluid Repository (Shao et al., 2021) (human milk with a total number of proteins 2536, last accessed 14 March 2023). We visualised 568 overlapping proteins between all three groups, 686 between HM and preterm EVs, 75 between HM and term EVs (Figure 7a). Proteins unique to only term or preterm HM EVs were queried in STRING database to identify functional enrichments in the unique protein groups. A total of 1182 unique term and 186 unique preterm EVs proteins were found and analysed for biological process enrichment. As expected, term EVs had a larger number of enriched processes, 516, compared to the 96 in preterm EVs. A total of 74 processes were shared between the two groups, indicating similarity of preterm contents to term. Relevant to newborn immunity, shared processes included immune effector process, leukocyte activation and antigen processing and presentation (Supplementary data 4). Biological processes unique to term HM EVs included several related to exocytosis, and various immune related processes were enriched, including toll-like receptor signalling, response to bacterial lipopeptide and CDC42 protein signalling (Figure 7b), also present were regulation of T cell proliferation, IL-12 signalling, antigen processing, Fc-receptor signalling, processes related to leukocytes and neutrophils (Supplementary data 4). Biological processes unique to preterm were related to endocytosis and enzymatic processes (proteolysis and enzyme linked protein signalling). Further processes associated with immune system and signalling processes, neutrophil activation, as well as Ephrin and Tyrosine Kinase signalling pathways were identified (Figure 7b and Supplementary data 4).

Human milk composition is known to differ based on lactation period between colostrum (up to 4 days after birth), transitional (5 days to two weeks after birth) and mature (>1 month after birth) milk (Ballard & Morrow, 2013; Castellote et al., 2011). Herwijnen et al. have previously characterised the proteome from EVs of mature HM and identified 1964 unique proteins across seven donors (Van Herwijnen et al., 2016). To compare the EV proteome of mature HM to that of transitional term or preterm HM, we submitted the list of proteins identified in three donors of term and preterm HM EVs, alongside data obtained from Herwijnen et al. to Venn analysis (bioinformatics.psb.ugent.be/venn). In total, 1235 common proteins were identified among the three datasets. Mature HM EVs contained 728 unique proteins, transitional HM EVs had 615, of which 96 unique proteins were identified in term, and 19 in preterm HM EVs (Figure 8a). For functional characterisation, each set of unique proteins from mature and transitional HM EVs were submitted to STRING for enrichment analyses. In mature HM EVs, 107 GO biological processes, 6 KEGG pathways and 13 Reactome pathways were enriched (Supplementary data 5). A variety of biological processes were represented, including cell differentiation, metabolic regulation and immune mediation (Supplementary data 5). In transitional HM EVs, 315 GO biological processes, 35 KEGG pathways and 223 Reactome pathways were enriched (Supplementary data 6). The enrichments with the highest confidence interaction score from STRING analysis were plotted in BioRender and included several clusters related to proteasome, immune response, immune system process in GO biological process classification, and metabolic pathways, ribosome-related, and RNA transport in KEGG pathways (Figure 8b).

In addition to total protein identification, previous studies have also measured protein expression quantitatively. We first compared the differentially expressed proteins quantified in the present data to those previously identified in term and preterm skim HM by Molinari et al. (2012). Common significantly upregulated proteins in term HM EVs were cathepsin B and vitronectin. In preterm HM, both lists contained significantly upregulated albumin, and ras-related proteins—Rab-6 in HM EVs, and Rab-1A,

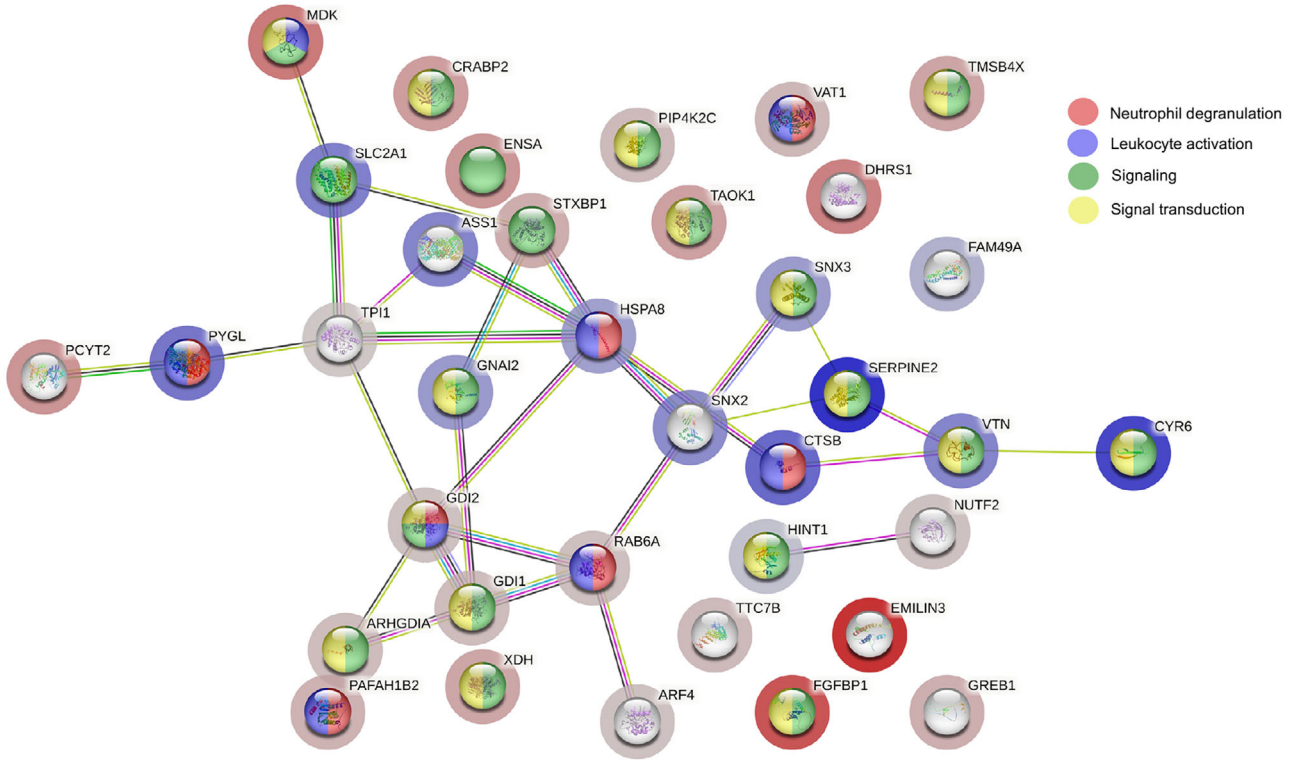

FIGURE 6 Biological processes enriched in significantly up- or downregulated preterm HM EV proteins. Upregulated proteins are annotated with a red halo, downregulated proteins have a blue halo. A total of 11 biological processes were assigned, with strongest evidence for neutrophil degranulation (red filled circles) and leukocyte activation in immune response (blue fill). Majority of proteins enriched in signaling (green fill) and signal transduction (yellow fill) were downregulated. $p < 0.04$, Student's t-test. Circles include 3D crystal structure of protein, if known. STRING analysis, confidence score >0.4–0.9 with ascending number of connecting lines.

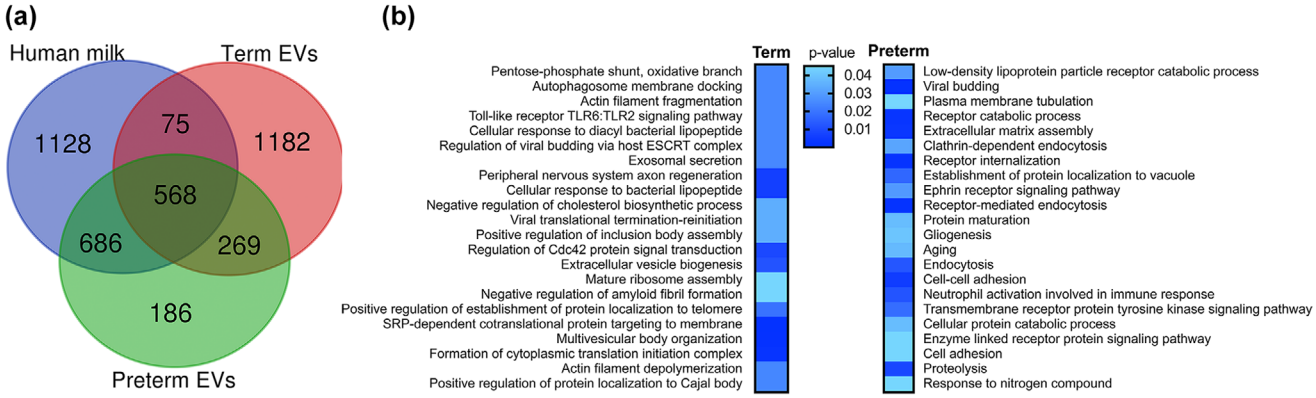

FIGURE 7 Analysis of proteins derived from HM and term and preterm HM EVs. (a) Venn diagram comparing proteins in Human Biofluid Repository (HBF) with term and preterm HM EVs. Out of 2536 proteins identified in human milk (HM), 568 were present in term and preterm EVs, while preterm EVs share 686, and term EVs share 75 proteins with HM. (b) Functional enrichments of proteins not present in HBF and unique to term and preterm EVs. Most significant processes in both groups included those involved in EV-related processes and immune system—extracellular vesicle biogenesis, multivesicular body organization, cellular response to bacterial lipopeptide, and toll-like receptor signalling pathway for term EVs; and endocytosis, neutrophil activation for preterm EVs. STRING analysis.

Rab-8A in skim preterm HM. The datasets differed for mucin-4, which was upregulated in preterm HM EVs, but downregulated in preterm skim milk (Supplementary data 7).

To investigate the protein content in relation to previous studies on enzymatic processes and infant digestion, we compared the present dataset to data reported by Nielsen et al. (2020) obtained from term and preterm human skim milk fractions. When compared with the respective conditions in our data, preterm EVs had 49 proteins in common with preterm skim milk, and term EVs had 22 proteins in common with term skim milk. Although preterm EVs demonstrated more proteins in common with skim

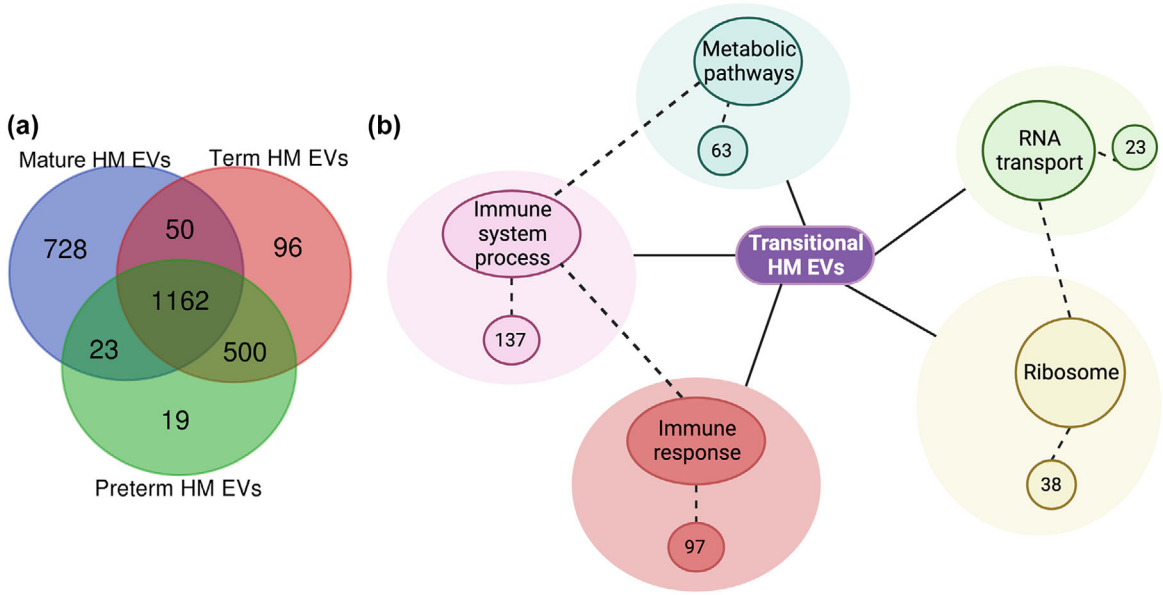

FIGURE 8 Analysis of proteins derived from mature HM EVs and transitional term and preterm HM EVs. (a) Venn diagram comparing proteins identified in mature HM EVs with transitional term and preterm HM EVs. Out of 1964 proteins identified in mature HM EVs, 1235 were present in term and preterm EVs, with a total of 615 unique proteins in transitional term and preterm HM EVs. (b) Functional enrichments of proteins unique to transitional HM EVs. Enriched clusters included pathways in metabolism (blue, KEGG pathway – 63 proteins), immune system process (pink, GO biological process – 137 proteins), immune response (red, GO biological process – 97 proteins), ribosome (yellow, KEGG pathway – 38 proteins), and RNA transport (green, KEGG pathway – 23 proteins), indicated by solid lines. Only high confidence interactions were included in STRING pathway analysis (>0.9 score), dashed lines indicate interactions between clusters.

milk, those proteins were still found in term HM EVs, but not in term skim milk. Thus, the proteins common to both EVs were analysed in STRING. As expected, many proteins had enzymatic features. The enriched biological processes were varied, but the processes with the highest abundance were related to the immune system and to metabolism (Figure 9 and Supplementary data 8).

A study has previously compared the proteomes of term and preterm HM EVs (Wang et al., 2019). Since the total protein list was not publicly available, we compared the reported significantly up- and downregulated peptides to those differentially expressed in our data. In both datasets vitronectin was significantly upregulated in term HM EVs, and albumin was significantly upregulated in preterm HM EVs. Out of the 16 reported proteins that were differentially expressed in Wang et al., all but RNASE1 and VIMENTIN were found in the present dataset, though not significantly differentially expressed (Supplementary data 9).

3.5 | Quantification of mRNA from GA-specific human milk EVs

RNA of HM EVs from term and preterm samples was sequenced to quantify the mRNA expression. In total, over 12,000 unique transcripts were detected in both term and preterm HM EVs (Figure 10a). On the gene level, term and preterm HM EVs shared 5522 genes, with an additional 1067 and 2912 unique genes in each group, respectively (Figure 10b).

Preterm and term HM EV mRNA was then queried in the STRING database to identify enriched biological processes and pathways. Term HM EVs contained double the total number of enrichments in biological processes – 163 versus 80 in preterm HM EVs, but same number of KEGG pathway enrichments (Supplementary data 1b). Pathway analysis in Reactome database revealed a similar number of pathways between term (191) and preterm HM EVs (189). Both had an enrichment in IL-1 and IL-12-related pathways (Supplementary data 1b). For enriched biological processes, antimicrobial humoral response, IL-1 response, oxygen level response, and Fc receptor, IL-12, NF-kappaB, and wnt signalling were unique to term HM EVs, whereas cellular response to IL-7 and innate immune response in mucosa were unique to preterm HM EVs (Figure 11 and Supplementary data 2b).

Furthermore, the lists of unique genes from term and preterm HM EVs were queried in the STRING database. In term HM EVs, genes with the highest abundance included translation and ribosomal genes (*EIF3CL, RPL34*) as well as cell death related genes (*PDCD5 AND CASP4*). Enriched biological processes included energy related functions only (Supplementary data 3b). In the unique preterm gene set, one Reactome pathway for translation was enriched (Supplementary data 3b).

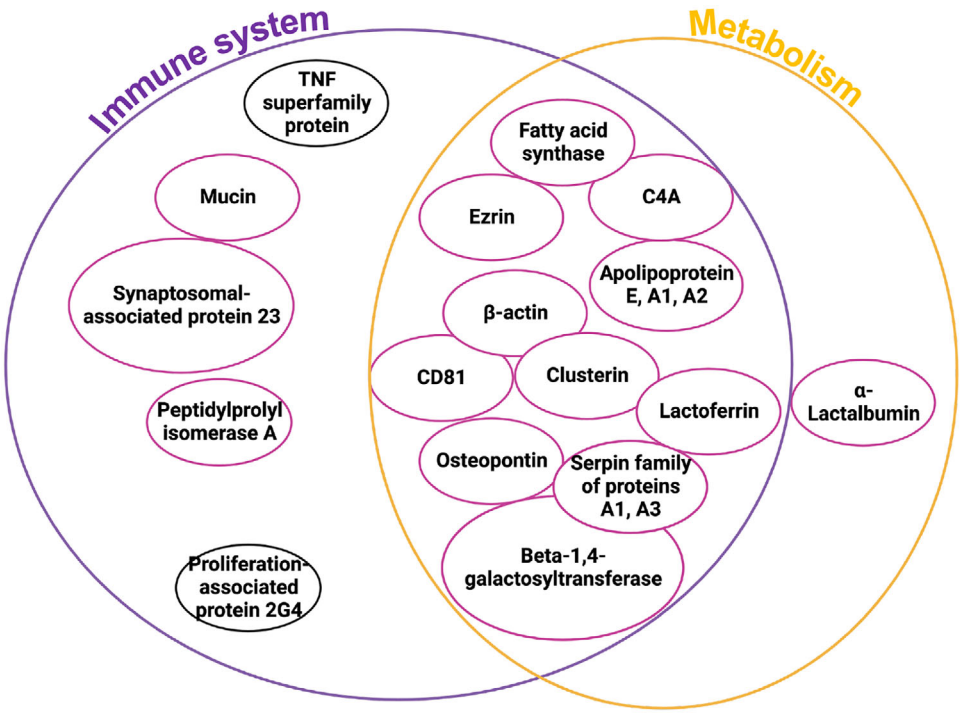

FIGURE 9 Select HM EV proteins common to skim HM with biological process enrichment analysis. Majority of the proteins formed a network of interactions (pink circles) and categorised to both immune system and metabolism-related biological processes. STRING analysis, medium confidence interactions (>0.400 score), total input of 31 proteins.

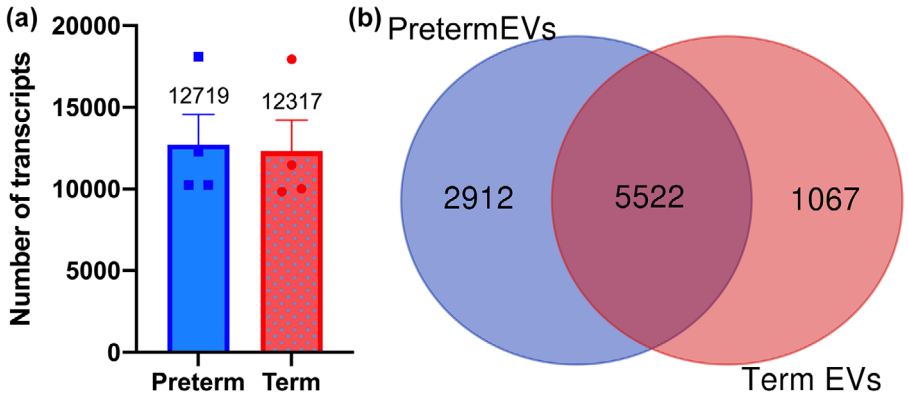

FIGURE 10 RNA transcripts identified in term and preterm HM EVs following RNA sequencing. (a) Number of total transcripts identified for preterm – 12,719 and term – 12,317, average $n = 4$ for each preterm and term. (b) 5522 genes are shared between term and preterm HM EVs, while preterm has 2912 and term 1067 unique genes.

3.6 | Differential expression of GA-specific human milk EV mRNA cargo

Next, mRNA raw counts were subjected to differential expression analysis. Term and preterm HM EV biological replicates clustered within their respective groups (Figure 12a). Overall, high fold-change differences were measured between the two groups, and a modest number of significantly expressed genes were identified, when compared to the size of the total gene pool (Figure 12b). A total of 153 and 285 genes were significantly (adjusted p-value <0.05) upregulated in term and preterm HM EVs, respectively (Supplementary data 4b). The top 20 most differentially expressed genes in term and preterm HM EVs are listed in Table 2.

All significantly differentially expressed genes were analysed in STRING for networks interactions and functional associations. Genes more highly expressed in term HM EVs were enriched in 11 GO biological processes, 24 KEGG and 4 Reactome pathways, with the majority of genes related to metabolic processes only, with no immune enrichment in any of the queried databases (Figure 13 and Supplementary data 5b). Differentially expressed genes in preterm HM EVs contained genes enriched in 361 GO

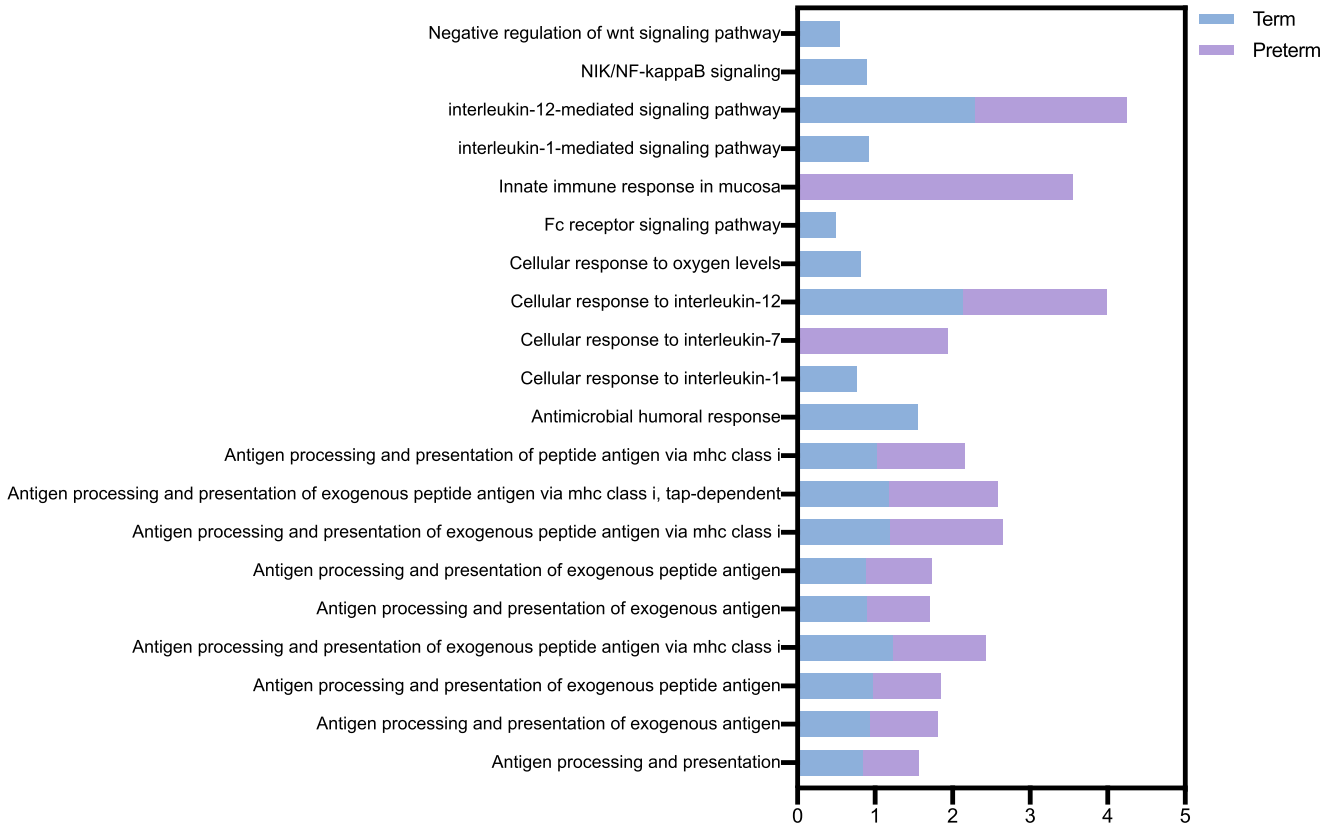

FIGURE 11 Enriched biological processes identified with STRING database of all term and preterm HM EV mRNA. Immune system and signalling related processes were selected for plotting. X-axis represents the enrichment score for term (blue bars) or preterm (pink bars). Enrichment was similar between shared identified biological processes for term and preterm HM EVs. Unique processes for antimicrobial humoral response, IL-1 response, oxygen level response, and Fc receptor, IL-12, NF-kappaB and wnt signalling were identified for term HM EVs, whereas IL-7 and innate immune response in mucosa were enriched in preterm HM EVs only. Transcripts with a TPM value >0 were queried.

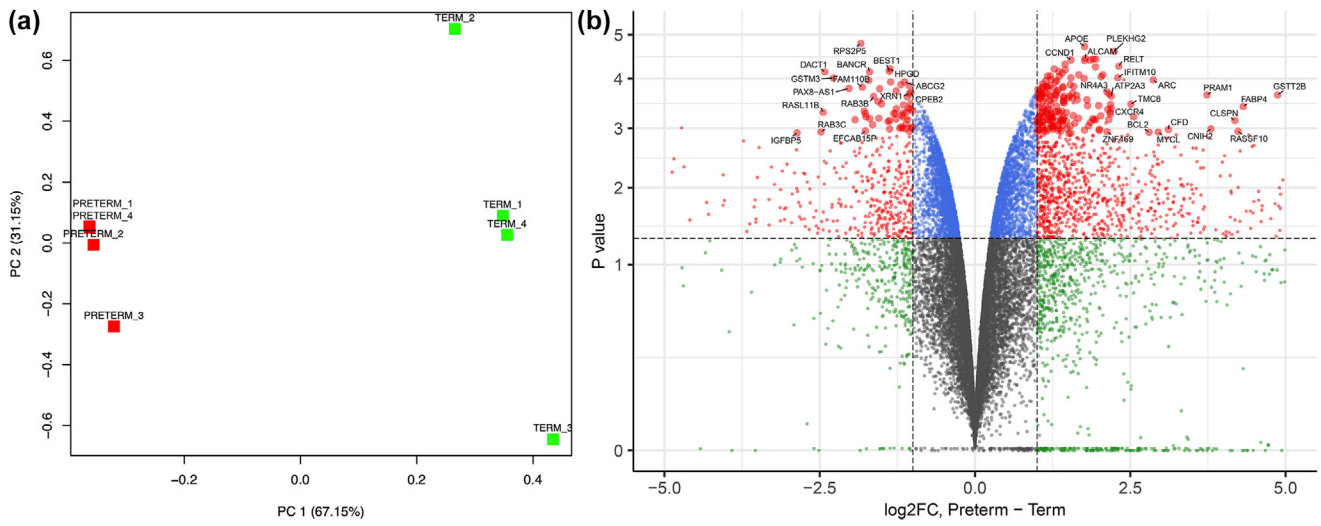

FIGURE 12 Differentially expressed mRNA in term and preterm HM EVs. (a) Clustering analysis of term (green) and preterm (red) HM EVs following TMM RUVr $k = 5$ normalisation. Both clustered with their respective conditions. (b) Volcano plot of differentially expressed genes. Genes with positive log2 fold-change (FC) value were higher abundant preterm, genes with a negative log2 FC were upregulated in term, genes not significantly changed are annotated in green and grey. Significance was assigned to mRNA with an adjusted *p*-value (FDR) of <0.05 (dashed line). Plot is constrained to log2 FC <5, genes with a log2 FC >1 and FDR <0.05 are annotated.

TABLE 2 Differential mRNA expression in term and preterm HM EVs. Genes were considered significantly changed based on padj <0.05 and log 2 fold-change >1. Genes with positive fold-change value were more highly expressed in preterm, those with negative fold-change were more highly expressed in term samples. The 20 most differentially expressed genes are displayed.

Gestational age at birth	Gene symbol	Fold-change (log2)	*p*-value
Term	*GSTM1*	−12.05	0.045
	RPSI4_1_156	−5.89	0.045
	RP4-765C7.2	−4.72	0.047
	IGFBP5	−2.87	0.049
	RAB3C	−2.48	0.047
	RASL11B	−2.45	0.047
	DACT1	−2.42	0.045
	GSTM3	−2.28	0.045
	PAX8-AS1	−2.03	0.045
	PDLIM1	−1.93	0.047
	RPS2P5	−1.84	0.045
	FAM110B	−1.81	0.045
	BEST3	−1.78	0.047
	BTG3	−1.77	0.047
	EFCAB15P	−1.77	0.047
	GLYATL2	−1.76	0.047
	TMEM200A	−1.71	0.045
	BANCR	−1.70	0.045
	GAPT	−1.70	0.047
	RAB6B	−1.65	0.047
Preterm	*GSTT2B*	4.87	0.045
	FABP4	4.32	0.047
	RASSF10	4.23	0.047
	CLSPN	4.19	0.047
	CNIH2	3.80	0.047
	PRAM1	3.73	0.045
	CFD	3.11	0.047
	MYCL	2.95	0.048
	ARC	2.87	0.045
	BCL2	2.80	0.048
	CXCR4	2.55	0.047
	TMC8	2.51	0.047
	RP11-295K3	2.40	0.045
	RELT	2.32	0.045
	IFITM10	2.30	0.045
	PLEKHG2	2.24	0.045
	ATP2A3	2.19	0.045
	EGR2	2.19	0.047
	WAS	2.18	0.047
	FERMT3	2.18	0.047

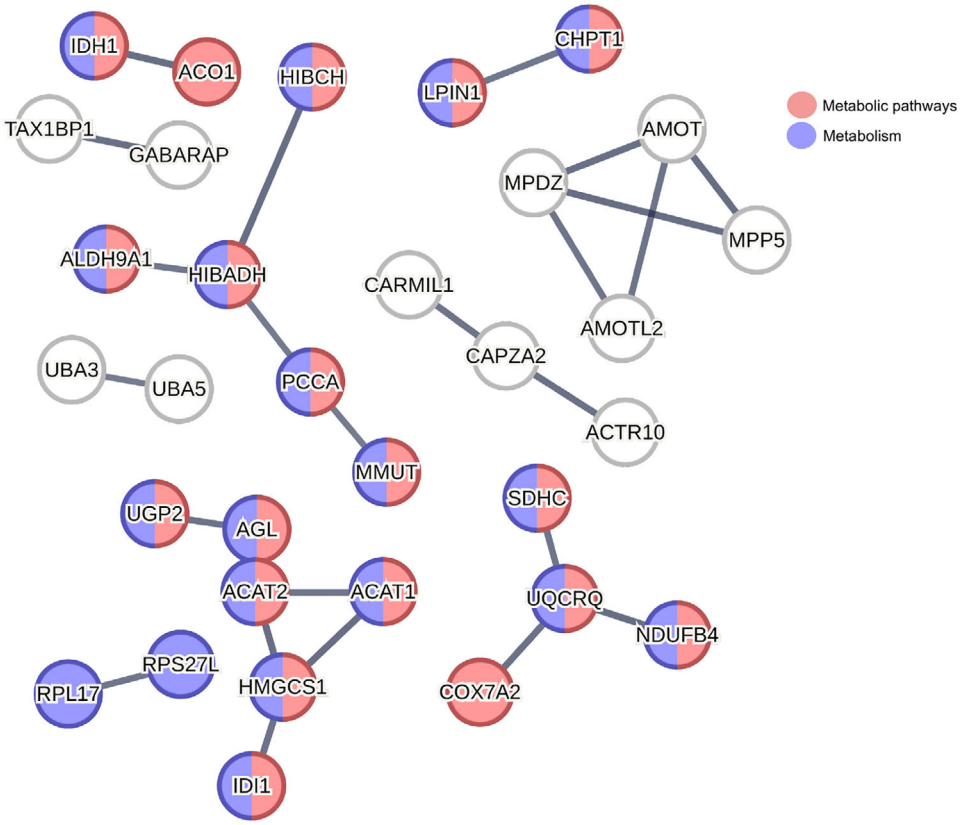

FIGURE 13 Biological processes enriched in significantly ($p < 0.05$) upregulated term HM EV mRNA. A total of 11 biological processes, 24 KEGG and 4 Reactome pathways were enriched. Metabolism-related enrichment was seen for many differentially expressed genes. KEGG metabolic pathway (red fill) included 33 mRNA in the total dataset, and Reactome metabolism (blue fill) included 42 mRNA in the total dataset. Other significantly increased interacting mRNA are annotated with a white fill. STRING analysis, confidence score 0.9, unconnected nodes removed.

biological processes, 6 KEGG and 8 Reactome pathways, with high confidence associations including developmental and immune system processes, and TNF signalling pathway (Figure 14 and Supplementary data 6b). Both conditions resulted in enrichment of sub-cellular compartments—extracellular-exosome for term, and endosome system for preterm.

To identify meaningful processes and pathways without the limits imposed by arbitrarily chosen differential expression cutoffs, GSEA was performed. In term HM EVs, upregulated genes were constrained to five clusters (Figure 15a), with an enrichment in metabolic processes, including amino acid metabolism, proteasome regulation and metabolic cycles. Genes upregulated in preterm HM EVs were constrained to a formation of maximum six clusters (Figure 15b). Many enrichments were immune related, including biological processes in neutrophil activation, antigen processing and presentation, T cell selection, and genomic spatial events in several immune cells (monocytes, CD4+ T cells, monocytes).

3.7 | GA-specific human milk EV cargo in previously identified human milk mRNA

Since cellular origin of EVs can determine the cargo and downstream biological effects, we sought to compare the gene signatures of term and preterm EVs to human milk single cells. The marker genes listed by Nyquist et al. were clustered with term or preterm HM EVs. The gene signatures corresponding to T cells, neutrophils, dendritic cells, fibroblasts, lactocyte epithelial cells (clusters LC1 and LC2) and macrophages (GPNMB+ and CSN1S1) were included (Nyquist et al., 2022). Briefly, gene lists with fold-change >1 and percent distance >0.5 were retained. Any gene not detected in our dataset was removed from the gene sets. Following GSEA analysis of retained gene sets, any gene set smaller than 10 genes or larger than 500 genes was eliminated.

Macrophages (GPNMB+) and neutrophils were the most upregulated in the preterm group, when compared to term samples (Figure 16a). In addition to the highest upregulation seen for signatures of GPNMB+ macrophages and neutrophils in preterm HM EVs, they too contained the highest count of differentially expressed genes (Figure 16). Also significantly increased were T cell, CSN1S1p+ macrophages, LC1 lactocyte epithelial cells and fibroblasts. Term HM EVs were enriched for LC2 lactocyte epithelial cell gene signatures (Figure 16a). GA-specific gene enrichment data was also processed with TMM only, and TMM RUV ($k = 5$), wherein signatures for LC1, LC2 and fibroblasts were more enriched for term HM EVs (Supplementary data 7b).

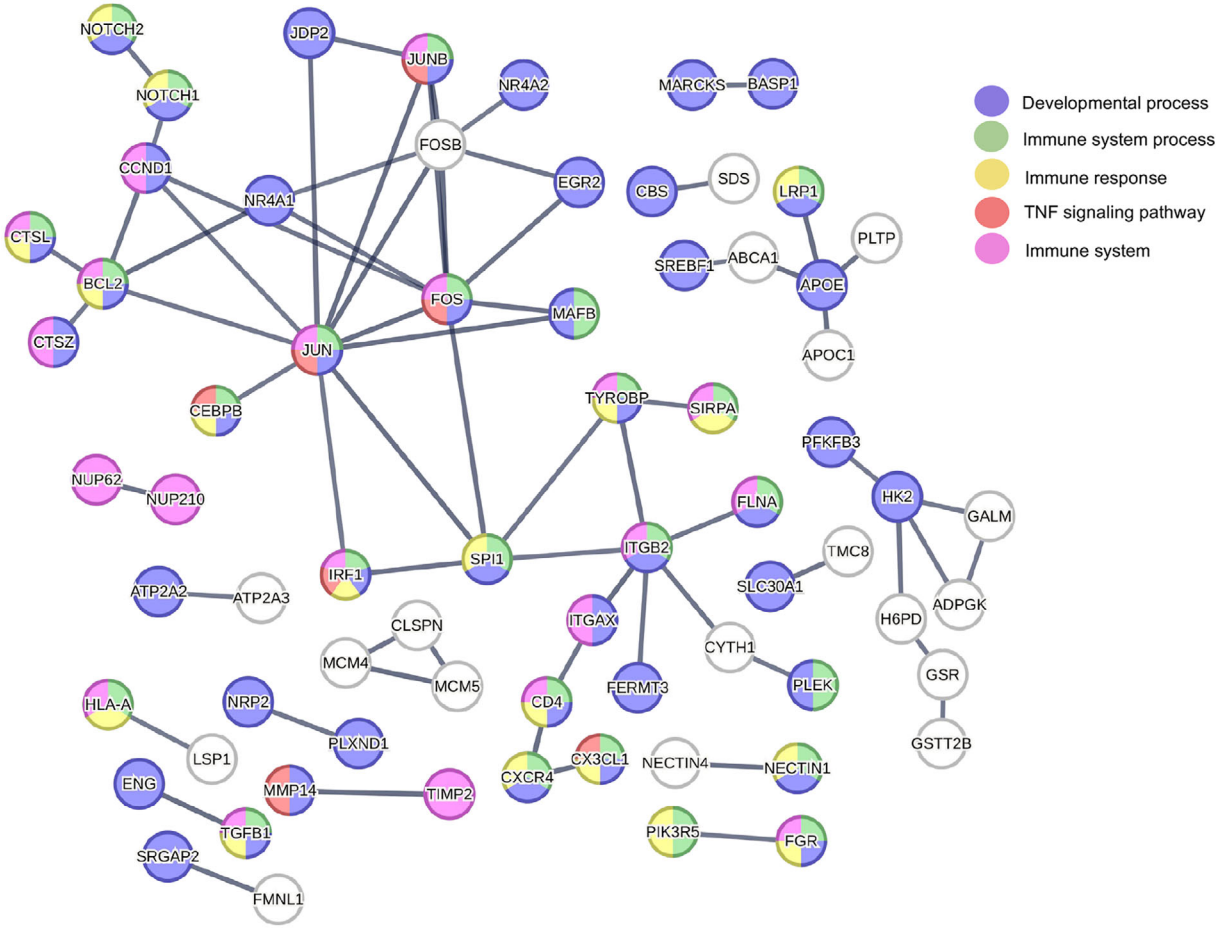

FIGURE 14 Biological processes enriched in significantly ($p < 0.05$) upregulated preterm HM EV mRNA. A total of 361 biological processes, 6 KEGG and 8 Reactome pathways were enriched. Immune system enrichment was seen for many differentially expressed genes. Total mRNA in the dataset included 118 for developmental process (blue fill), 60 for immune system (green fill) and 38 for immune response (yellow fill, GO biological process). For KEGG database, 8 mRNA in the total dataset were enriched for TNF signaling pathway (red fill). Immune system in Reactome database included 56 mRNA in the total dataset (pink fill). Other significantly increased interacting mRNA are annotated with a white fill. STRING analysis, confidence score >0.9, unconnected nodes removed.

However, in all analyses, preterm HM EVs consistently displayed high enrichment for immune cell signatures (Supplementary data 7b and Figure 16).

4 | DISCUSSION

In this study, we have quantitatively characterised EV cargo from term and preterm HM. We observed that preterm HM EVs had a higher number of proteins and mRNA related to immune system processes. Whereas we propose that term HM EVs upregulate additional metabolic pathways. Analysis of biological processes from total protein and RNA abundance suggested a more diverse biological regulation by term HM EVs.

A recent study by Wang et al. characterised proteins in term and preterm HM EVs using iTRAQ labelling (Wang et al., 2019). Our study complements their findings by using LC-MS/MS protein identification with label-free analysis allowing for the identification of an increased number of peptides. We reported over 19,000 total peptides, which greatly exceeds the 719 peptides identified previously (Wang et al., 2019). In this study, the SWATH technique further facilitated the identification of lower abundance proteins (Whitman & Lynch, 2019; Zhang et al., 2020). As such, the current findings provide a foundational basis for further functional studies into the unique proteins identified in term and preterm HM (Table 1 and Supplementary data). Freiría-Martínez et al. also reported proteomics analysis of term and preterm whole HM and EVs (Freiría-Martínez et al., 2023). Between the two groups reporting differentially expressed proteins, both describe upregulation of lysozyme, and Wang et al. also report increased levels of LF, in preterm HM EVs (Freiría-Martínez et al., 2023; Wang et al., 2019). Although these proteins were present in our dataset, we did not detect significantly different expression levels. It is important to note that the previous studies had

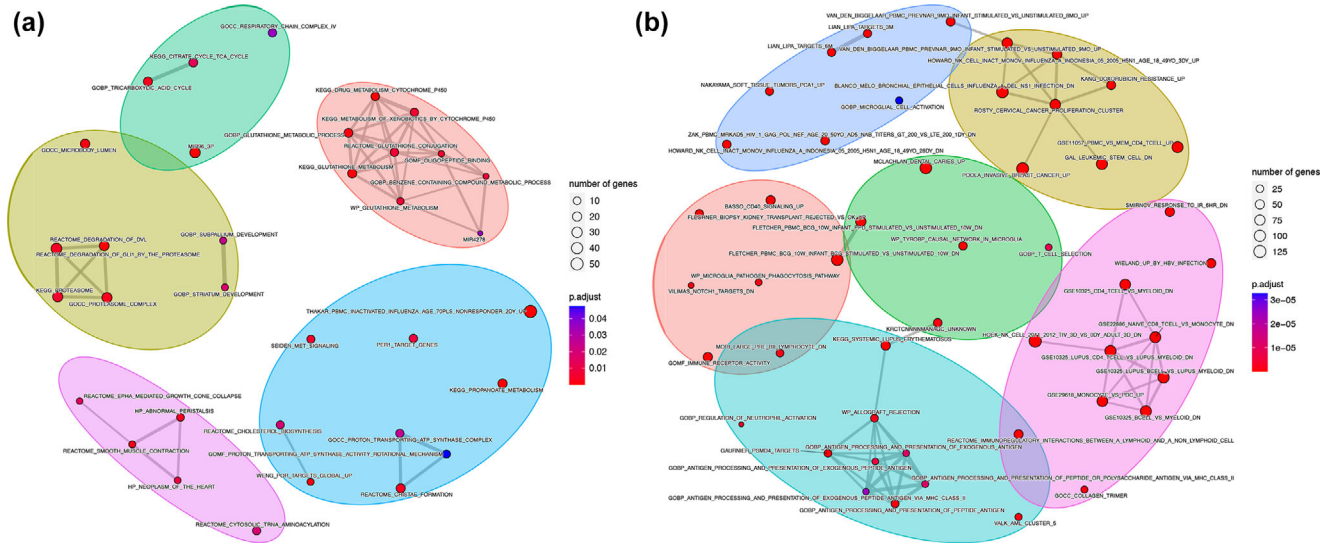

FIGURE 15 Gene set enrichment analysis of enriched processes, pathways, and functions for differentially expressed term and preterm HM EV mRNA. (a) Term HM EV-cargo had the highest number of gene clusters for metabolism (red shading). (b) Preterm HM EV-cargo was diverse and shared between several enriched gene sets, indicated by overlapping borders. Immune-related functions (light blue and pink shading) were highly enriched. Gene sets with an adjusted *p*-value of <0.05 were plotted.

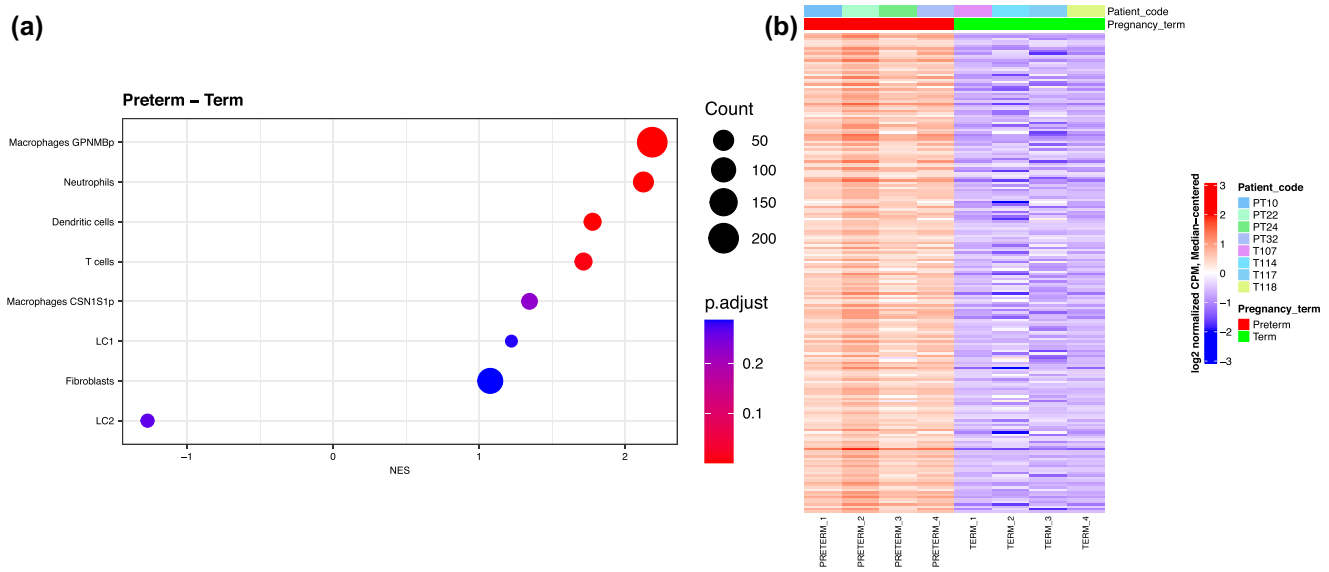

FIGURE 16 Single cell gene signatures in differentially expressed term and preterm HM EV mRNA. (a) Single cell sequencing signatures for GPNMB+ macrophages, neutrophils, dendritic cells, T cells, CSN1S1+ macrophages, lactocyte cluster 1 (LC1), fibroblasts and lactocyte cluster 2 (LC2) in differentially expressed term and preterm HM EV mRNA. Higher expression in all but one single cell signatures is present for preterm HM EVs (NES >0), term HM EVs have a higher expression for LC2 signatures only. (b) Heatmap of core set of most upregulated genes in preterm HM EVs for GPNMB+ macrophage signatures. Genes with log foldchange >1 and percent distance of >0.5 were retained, TMM with RUVr $k = 5$ normalisation.

isolated HM EVs from colostrum, mature milk or both, while we obtained transitional HM. In HM, most abundant proteins include LF and lysozyme (Donovan, 2019), which may also oversaturate the mass spectra and interfere in the identification of lower abundance proteins (Boehmer et al., 2010).

A previous study showed upregulation of cathepsin D in the milk of mothers that delivered prematurely (Demers-Mathieu et al., 2017). Proteases in human milk are suggested to aid digestion in infants with an added role of influencing infant immune system (Dallas et al., 2013; Demers-Mathieu et al., 2017). To our knowledge, presence of cathepsin B has not been previously reported in proteomic studies of HM EVs, though it has been detected among the top 20 most abundant transcripts in colostrum HM fat layer (Lemay et al., 2013). Using proteomic analysis, we were able to detect a significantly higher level of cathepsin B in term HM EVs, when compared to preterm samples. Additionally, in our mRNA data we were able to detect cathepsin B in both

term and preterm HM EV. However, there was no statistically significant differences in expression (Supplementary data). The availability of cathepsin B protein could support digestion in infants during feeding.

The most significantly upregulated protein in term HM EVs was SERPINE2, which is a member of the serpin family of proteins that inhibit serine proteases (Silverman et al., 2001). A previous study investigated anti-inflammatory effects in a premature infant intestine model, where epidermal growth factor treatment upregulated several *serpine* genes, including SERPINE2 (Ménard et al., 2012). Antiproteases from the serpin family are important in infant digestion (Dallas et al., 2012), and have been shown to regulate inflammatory processes (Santoro et al., 2015). Heat shock protein 70 was also upregulated in term HM EVs and may provide intestinal protection in infants by supporting the intestinal barrier, as previously demonstrated in a rat pup model (Liedel et al., 2011).

In preterm HM EVs, caspase 14 protease was significantly upregulated. Its main proposed function is to regulate epidermal barrier formation (Denecker et al., 2008; Markiewicz et al., 2021; Pistritto et al., 2002). In the context of a premature birth, the support of barrier formation and immune regulation could be, in part, attributed to caspase 14. Additional cell proliferation and differentiation could be promoted by fibroblast growth factor binding protein 1 (Abuharbeid et al., 2006), one of the most upregulated proteins we found in preterm HM EVs (Table 1).

Another highly upregulated protein was elastin microfibril interfacer 3 or EMILIN-3, which can act as a regulator of TGF β (Schiavinato et al., 2012). Moreover, this protein is implicated in various infant inflammatory diseases and is an important protein for immune response regulation (Cho et al., 2016; Torres-Castro et al., 2020). Under conditions of cellular stress, promoting gut epithelial cell survival is another route of protection, which is increased following HM EV treatment (Martin et al., 2018). Epithelial-to-mesenchymal transition (EMT) may occur as a result of microbe-induced inflammation to facilitate gut repair (Hofman & Vouret-Craviari, 2012; Kalluri & Weinberg, 2009) and HM EVs have been shown to affect EMT, via their TGFβ2 cargo, in normal and cancerous breast epithelial cells (Qin et al., 2016). Whether these TGFβ-dependent responses are regulated by additional EV carrier proteins, such as EMILIN-3, remains to be determined.

In HM EV mRNA, we noted further GA-specific differences. The gestational age-specific differences were similar to those seen on the protein level, with more immune-related enrichments for preterm when functional enrichments were analysed in STRING. It is notable that while the FDR value was similar between the top 200 significantly differentially expressed mRNA in preterm HM EVs, the fold-change varied 1- to 5-fold. When comparing the top 20 differentially expressed genes, many of the significantly upregulated genes in preterm EVs were implicated in immune signalling—*BCL2, IFITM10, CXCR4, CFD, WAS, FERMT3*. The anti-apoptotic signalling of BCL2 is prevalent during lactation and mammary tissue remodelling (Colitti, 2012). The significant upregulation of BCL2 in preterm HM EVs may represent the stage of gland development at the time of premature birth. For infants, increased presence of complement factor D (CFD) in preterm milk may help support the incomplete complement functions presenting in premature infants (Cacho & Lawrence, 2017). Similarly, CXCR4 is an important cell receptor expressed in immune cells to mitigate haematopoiesis (Zou et al., 1998) and may further support development of infant immune system, specifically innate (Shalekoff et al., 2004). Among the genes with a 2-fold increase in expression was *HLA-H*. Transfer of antibodies to infants via HM is well established and essential for immune system development (Blais et al., 2006), however the role for HLA molecules is less clear. It has been suggested that HLA molecules may be important in regulating or establishing tolerance, but more work is needed (Savulescu et al., 2020).

In term HM EVs, significant increases in expression of several metabolic related mRNAs were observed. *GSTM1* had a 12-fold increase, the highest of any other gene measured by several folds, while *GSTM3* had an over 2-fold increase. GSTMs are part of the glutathione transferase family of enzymes and implicated in detoxification processes, thus being potent antioxidants (Bhattacharjee et al., 2013; Sharma et al., 2004). Preterm HM EVs also contained significant increases in members of glutathione antioxidant defence system—*GSR* and *GSTT2B* with 1.4- and 4.8-fold increases, respectively. Since breastfed infants are more protected against development of many diseases (Le Doare et al., 2018; O'reilly et al., 2021), the abundant presence of antioxidant promoting factors in HM may contribute to the overall health of infant. Term HM EVs also had increased expression in mRNA implicated in cellular signalling, such as ras and rab-family of proteins (*RASL11B, RAB3C, RAB6B*), insulin-like growth factor-binding protein 5 (*IGFBP5*), and *DACT1*. Significant increases seen in *PDLIM1* in term EVs may promote infant gut epithelial cell migration and survival (Zhou et al., 2021), an effect previously seen in vitro following treatment with term and preterm HM EVs (Wang et al., 2019).

HM EVs are purified from the skim milk fraction, while the separation of HM fat and cells occurs in the top layer and pellet, respectively. In previous human milk mRNA studies, the fat and cellular portions have been subjected to comprehensive RNA sequencing (Gleeson et al., 2023; Lemay et al., 2013; Nyquist et al., 2022; Twigger et al., 2022). Whenever specified in the studies, the sequencing data showed enrichment of casein and LA transcripts, which we did not see in our data and could further emphasise a more cellular signalling-related function for EVs versus nutrient for other milk fractions. The HM EV sequencing data also allows for the comparison between human milk cells and EVs. Whether or not the cells found in milk are secretors of EVs remains to be determined but clustering the corresponding transcriptomes provides insight.

Single cell sequencing of HM has previously revealed leukocyte epithelial cells throughout lactation, as well as immune cells such as macrophages, T cells and neutrophils, and stem cells (Gleeson et al., 2023; Martin Carli et al., 2020; Sharp et al., 2016; Twigger et al., 2022). It has been hypothesised that stem cells from HM could survive digestion and have regenerative and

developmental roles in the infant (Ninkina et al., 2019). In mother-infant dyads, gene expression patterns in HM cells, specifically the JAK/STAT pathway genes, associated to infant microbiome outcomes (Johnson et al., 2023).

When analysing cellular signatures of HM EVs, both term and preterm EVs had immune cell signatures, though they were significantly higher in preterm EVs. The origin of HM EVs may explain the functional effects seen prior studies, especially when inflammatory regulation was examined (Zonneveld et al., 2021). In our transcriptomic data, preterm mRNA was most significantly upregulated for gene signatures of GPNMB+ macrophages, generally attributed to M2 macrophages and an anti-inflammatory phenotype (Zhang et al., 2022; Zhou et al., 2017). Given the vulnerability of infants, especially premature infants, to gut inflammatory diseases such as necrotising enterocolitis, the EV cargo may provide significant protection. It is especially notable that EVs themselves have been shown to survive digestion in simulated conditions (Kahn et al., 2018; Liao et al., 2017) and have tissue bioavailability following oral and intravenous supplementation in mice (Manca et al., 2018). In a functional study, term HM EVs were shown to regulate T cell activation and were proposed to allow development of a more immune tolerogenic phenotype (Zonneveld et al., 2021). Prior analyses from mature HM EV proteins and surface markers have also indicated an immune, epithelial and mesenchymal stem cell origin for the EVs (Giovanazzi et al., 2023; Van Herwijnen et al., 2016). Given the overall enrichment of preterm HM EVs in immune-related protein, mRNA and cell signatures, it remains to be seen whether preterm EVs promote immunoregulatory effects in vitro and in vivo.

Limitations of this work include the lower biological sample number, with milk from 12 mothers analysed in total. However, collecting fresh milk was imperative for isolation of EVs, thus experimental analyses proceeded with fresh samples available at the time. While every effort was made to collect milk at the same lactation stage post birth, some variation between term and preterm was present, and could have influenced the composition of HM EV cargo (Supplementary file—Donor information). The interpretation of data is based on in silico analyses, which require in vitro and in vivo validation of the cargo differences, especially to allow confirmation of the more immunoregulatory profile of preterm HM EVs proposed here.

Characterising the term and preterm HM EV proteome and mRNA could provide insight for future studies investigating possible milk deficiencies and potential supplementation of donor milk or formula. Whether premature infants should be preferentially supplemented with premature HM, or vice versa in term infants, might be an important consideration to ensure the best possible outcomes for all infants, especially fragile premature infants. Overall, our data has demonstrated consistent diversity between term and preterm HM EVs. Temporal effects of the cargo may be of importance, allowing HM EVs to have an immediate effect upon uptake via proteins, and delayed effect via translation of mRNA. Exact functional effects and clinical relevance of the differences in GA-specific EV cargo remain to be elucidated.

AUTHOR CONTRIBUTIONS

Brett Vahkal: Conceptualization; data curation; formal analysis; investigation; methodology; validation; visualization; writing—original draft; writing—review and editing. **Illimar Altosaar**: Conceptualization; funding acquisition; investigation; methodology; project administration; supervision; writing—review and editing. **Eric Tremblay**: Formal analysis; investigation; methodology; resources; writing—review and editing. **David Gagné**: Data curation; formal analysis; methodology; resources; software; validation; writing—review and editing. **Nico Hüttman**: Data curation; formal analysis; methodology; resources; software; validation; visualization; writing—review and editing. **Zoran Minic**: Formal analysis; methodology; resources; software; validation; writing—review and editing. **Marceline Côté**: Investigation; methodology; project administration; supervision; validation; writing—original draft; writing—review and editing. **Alexandre Blais**: Data curation; formal analysis; methodology; resources; software; validation; visualization; writing—review and editing. **Jean-François Beaulieu**: Conceptualization; funding acquisition; investigation; methodology; project administration; writing—review and editing. **Emanuela Ferretti**: Conceptualization; investigation; methodology; project administration; supervision; writing—review and editing.

ACKNOWLEDGEMENTS

We would like to thank Dr. Yun Liu from the uOttawa Materials Characterization Core Facility for technical support; Dr. Ray Eby, Dr. Leif Anderson and Nanoview Biosciences for EV surface marker characterization; Prof. Maxim Berezovski at the uOttawa John L. Holmes Mass Spectrometry Facility for guidance on proteomics analysis; and Dr. Senthilkumar Kailasam at the Canadian Centre for Computational Genomics for assistance with RNA sequencing. Additional thanks to the support staff at CHEO Research Institute, especially NICU project lead coordinator Chantal Horth, and research assistants Rebecca Grimwood, Denise Campuzano for patient recruitment, sample collection, and Samira Chamaa for administrative support. Figures 8b and 9 were created with BioRender.com. This research was funded by the Canadian Institutes of Health Research: grant number PJT 162423 to J-F. B., E.F., and I.A; CHIR Grant number PJT 183839 to A.B. B.V. was supported by the Ontario Graduate Scholarship and a stipend from Estonian University of Life Sciences. M.C. is a Canada Research Chair in Molecular Virology and Antiviral Therapeutics (950-232840) and recipient of Ontario Ministry of Research, Innovation and Science Early Researcher Awards (ER18-14-091). This research was enabled in part by support provided by WestGrid and the Digital Research Alliance of Canada (https://alliancecan.ca).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

ORCID

Brett Vahkal https://orcid.org/0009-0003-8327-7557

REFERENCES

Abuharbeid, S., Czubayko, F., & Aigner, A. (2006). The fibroblast growth factor-binding protein FGF-BP. *International Journal of Biochemistry & Cell Biology*, *38*, 1463–1468.

Admyre, C., Johansson, S. M., Qazi, K. R., Filén, J.-J., Lahesmaa, R., Norman, M., Neve, E. P. A., Scheynius, A., & Gabrielsson, S. (2007). Exosomes with immune modulatory features are present in human breast milk. *Journal of Immunology*, *179*, 1969–1978.

Adusumilli, R., & Mallick, P. (2017). Data conversion with ProteoWizard msConvert. *Methods in Molecular Biology*, *1550*, 339–368.

Aleksander, S. A., Balhoff, J., Carbon, S., Cherry, J. M., Drabkin, H. J., Ebert, D., Feuermann, M., Gaudet, P., Harris, N. L., Hill, D. P., Lee, R., Mi, H., Moxon, S., Mungall, C. J., Muruganugan, A., Mushayahama, T., Sternberg, P. W., Thomas, P. D., & Westerfield, M., Gene Ontology Consortium. (2023). The Gene Ontology knowledgebase in 2023. *Genetics*, *224*, iyad031.

Alsaweed, M., Hartmann, P., Geddes, D., & Kakulas, F. (2015). Micrornas in breastmilk and the lactating breast: Potential immunoprotectors and developmental regulators for the infant and the mother. *International Journal of Environmental Research and Public Health*, *12*, 13981–14020.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., & Sherlock, G. (2000). Gene ontology: Tool for the unification of biology. *Nature Genetics*, *25*, 25–29.

Ballard, O., & Morrow, A. L. (2013). Human milk composition. *Pediatric Clinics of North America*, *60*, 49–74.

Bhattacharjee, P., Paul, S., Banerjee, M., Patra, D., Banerjee, P., Ghoshal, N., Bandyopadhyay, A., & Giri, A. K. (2013). Functional compensation of glutathione S-transferase M1 (GSTM1) null by another GST superfamily member, GSTM2. *Scientific Reports*, *3*, 2704.

Blais, D. R., Harrold, J., & Altosaar, I. (2006). Killing the messenger in the nick of time: Persistence of breast milk scd14 in the neonatal gastrointestinal tract. *Pediatric Research*, *59*, 371–376.

Boehmer, J. L., Degrasse, J. A., Mcfarland, M. A., Tall, E. A., Shefcheck, K. J., Ward, J. L., & Bannerman, D. D. (2010). The proteomic advantage: Label-free quantification of proteins expressed in bovine milk during experimentally induced coliform mastitis. *Veterinary Immunology and Immunopathology*, *138*, 252–266.

Cacho, N. T., & Lawrence, R. M. (2017). Innate immunity and breast milk. *Frontiers in Immunology*, *8*, 584.

Carrillo-Lozano, E., Sebastián-Valles, F., & Knott-Torcal, C. (2020). Circulating microRNAs in breast milk and their potential impact on the infant. *Nutrients*, *12*, 3066.

Castellote, C., Casillas, R., Ramírez-Santana, C., Pérez-Cano, F. J., Castell, M., Moretones, M. G., López-Sabater, M. C., & Franch, Í. (2011). Premature delivery influences the immunological composition of colostrum and transitional and mature human milk. *Journal of Nutrition*, *141*, 1181–1187. bioinformatics.psb.ugent.be/venn

Cattaneo, C., Caramaschi, A., Uga, E., Braghin, M., Cosi, G., Peila, C., Strozzi, M. C., Sabatini, M., Gazzolo, D., Manfredi, M., & Cavaletto, M. (2019). Analysis of toll-like receptors in human milk: Detection of membrane-bound and soluble forms. *Journal of Immunology Research*, *2019*, 4078671.

Chen, W., Chen, X., Qian, Y., Wang, X., Zhou, Y., Yan, X., Yu, B., Yao, S., Yu, Z., Zhu, J., & Han, S. (2021). Lipidomic profiling of human milk derived exosomes and their emerging roles in the prevention of necrotizing enterocolitis. *Molecular Nutrition & Food Research*, *65*, 2000845.

Cho, S. X., Berger, P. J., Nold-Petry, C. A., & Nold, M. F. (2016). The immunological landscape in necrotising enterocolitis. *Expert Reviews in Molecular Medicine*, *18*, e12.

Chutipongtanate, S., & Greis, K. D. (2018). Multiplex biomarker screening assay for urinary extracellular vesicles study: A targeted label-free proteomic approach. *Scientific Reports*, *8*, 15039.

Colitti, M. (2012). BCL-2 family of proteins and mammary cellular fate. *Anatomia, Histologia, Embryologia*, *41*, 237–247.

Dallas, D. C., Guerrero, A., Khaldi, N., Castillo, P. A., Martin, W. F., Smilowitz, J. T., Bevins, C. L., Barile, D., German, J. B., & Lebrilla, C. B. (2013). Extensive in vivo human milk peptidomics reveals specific proteolysis yielding protective antimicrobial peptides. *Journal of Proteome Research*, *12*, 2295–2304.

Dallas, D. C., Murray, N. M., & Gan, J. (2015). Proteolytic systems in milk: Perspectives on the evolutionary function within the mammary gland and the infant. *Journal of Mammary Gland Biology and Neoplasia*, *20*, 133–147.

Dallas, D. C., Underwood, M. A., Zivkovic, A. M., & German, J. B. (2012). Digestion of protein in premature and term infants. *Journal of Nutritional Disorders & Therapy*, *02*, 112.

Da Veiga Leprevost, F., Haynes, S. E., Avtonomov, D. M., Chang, H.-Y., Shanmugam, A. K., Mellacheruvu, D., Kong, A. T., & Nesvizhskii, A. I. (2020). Philosopher: A versatile toolkit for shotgun proteomics data analysis. *Nature Methods*, *17*, 869–870.

De La Torre Gomez, C., Goreham, R. V., Bech Serra, J. J., Nann, T., & Kussmann, M. (2018). Exosomics—A review of biophysics, biology and biochemistry of exosomes with a focus on human breast milk. *Frontiers in Genetics*, *9*, 92.

Demers-Mathieu, V., Nielsen, S. D., Underwood, M. A., Borghese, R., & Dallas, D. C. (2017). Analysis of milk from mothers who delivered prematurely reveals few changes in proteases and protease inhibitors across gestational age at birth and infant postnatal age. *Journal of Nutrition*, *147*, 1152–1159.

Demichev, V., Messner, C. B., Vernardis, S. I., Lilley, K. S., & Ralser, M. (2020). DIA-NN: Neural networks and interference correction enable deep proteome coverage in high throughput. *Nature Methods*, *17*, 41–44.

Denecker, G., Ovaere, P., Vandenabeele, P., & Declercq, W. (2008). Caspase-14 reveals its secrets. *Journal of Cell Biology*, *180*, 451–458.

Di Tommaso, P., Chatzou, M., Floden, E. W., Barja, P. P., Palumbo, E., & Notredame, C. (2017). Nextflow enables reproducible computational workflows. *Nature Biotechnology*, *35*, 316–319.

Do, D. N., Li, R., Dudemaine, P.-L., & Ibeagha-Awemu, E. M. (2017). MicroRNA roles in signalling during lactation: An insight from differential expression, time course and pathway analyses of deep sequence data. *Scientific Reports*, *7*, 44605.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*, 15–21.

Donovan, S. M. (2019). Human milk proteins: Composition and physiological significance. *Nestlé Nutrition Institute Workshop Series*, *90*, 93–101.

Dvorak, B., Fituch, C. C., Williams, C. S., Hurst, N. M., & Schanler, R. J. (2003). Increased epidermal growth factor levels in human milk of mothers with extremely premature infants. *Pediatric Research*, *54*, 15–19.

Ewels, P. A., Peltzer, A., Fillinger, S., Patel, H., Alneberg, J., Wilm, A., Garcia, M. U., Di Tommaso, P., & Nahnsen, S. (2020). The nf-core framework for community-curated bioinformatics pipelines. *Nature Biotechnology*, *38*, 276–278.

Floris, I., Billard, H., Boquien, C.-Y., Joram-Gauvard, E., Simon, L., Legrand, A., Boscher, C., Rozé, J.-C., Bolaños-Jiménez, F., & Kaeffer, B. (2015). MiRNA analysis by quantitative PCR in preterm human breast milk reveals daily fluctuations of hsa-miR-16-5p. *PLoS ONE*, *10*, e0140488.

Floris, I., Kraft, J., & Altosaar, I. (2016). Roles of microRNA across prenatal and postnatal periods. *International Journal of Molecular Sciences*, *17*, 1994.

Freiría-Martínez, L., Iglesias-Martínez-Almeida, M., Rodríguez-Jamardo, C., Rivera-Baltanás, T., Comís-Tuche, M., Rodrígues-Amorím, D., Fernández-Palleiro, P., Blanco-Formoso, M., Álvarez-Chaver, P., Diz-Chaves, Y., Gonzalez-Freiria, N., Martín-Forero-Maestre, M., Fernández-Feijoo, C. D., Suárez-Albo, M., Fernández-Lorenzo, J. R., Guisán, A. C., Olivares, J. M., & Spuch, C. (2023). Proteomic analysis of exosomes derived from human mature milk and colostrum of mothers with term, late preterm, or very preterm delivery. *Analytical Methods*, *15*, 4905–4917.

Gagné, D., Shajari, E., Thibault, M.-P., Noël, J.-F., Boisvert, F.-M., Babakissa, C., Levy, E., Gagnon, H., Brunet, M. A., Grynspan, D., Ferretti, E., Bertelle, V., & Beaulieu, J.-F. (2022). Proteomics profiling of stool samples from preterm neonates with SWATH/DIA mass spectrometry for predicting necrotizing enterocolitis. *International Journal of Molecular Sciences*, *23*, 11601.

Galley, J. D., & Besner, G. E. (2020). The therapeutic potential of breast milk-derived extracellular vesicles. *Nutrients*, *12*, 745.

Giovanazzi, A., Van Herwijnen, M. J. C., Kleinjan, M., Van Der Meulen, G. N., & Wauben, M. H. M. (2023). Surface protein profiling of milk and serum extracellular vesicles unveils body fluid-specific signatures. *Scientific Reports*, *13*, 8758.

Gleeson, J. P., Chaudhary, N., Fein, K. C., Doerfler, R., Hredzak-Showalter, P., & Whitehead, K. A. (2023). Profiling of mature-stage human breast milk cells identifies six unique lactocyte subpopulations. *Science Advances*, *8*, eabm6865.

Golan-Gerstl, R., Elbaum Shiff, Y., Moshayoff, V., Schecter, D., Leshkowitz, D., & Reif, S. (2017). Characterization and biological function of milk-derived miRNAs. *Molecular Nutrition & Food Research*, *61*, 1700009.

He, B., Shi, J., Wang, X., Jiang, H., & Zhu, H.-J. (2019). Label-free absolute protein quantification with data-independent acquisition. *Journal of Proteomics*, *200*, 51–59.

Herrmann, K., & Carroll, K. (2014). An exclusively human milk diet reduces necrotizing enterocolitis. *Breastfeeding and Medications*, *9*, 184–190.

Herwijnen, M. J. C. V., Driedonks, T. A. P., Snoek, B. L., Kroon, A. M. T., Kleinjan, M., Jorritsma, R., Pieterse, C. M. J., Hoen, E. N. M. N.-'T., & Wauben, M. H. M. (2018). Abundantly present mirnas in milk-derived extracellular vesicles are conserved between mammals. *Frontiers in Nutrition*, *5*, 81.

Hofman, P., & Vouret-Craviari, V. (2012). Microbes-induced EMT at the crossroad of inflammation and cancer. *Gut Microbes*, *3*, 176–185.

Hu, Y., Thaler, J., & Nieuwland, R. (2021). Extracellular vesicles in human milk. *Pharmaceuticals (Basel)*, *14*, 1050.

Jeppesen, D. K., Fenix, A. M., Franklin, J. L., Higginbotham, J. N., Zhang, Q., Zimmerman, L. J., Liebler, D. C., Ping, J., Liu, Q. I., Evans, R., Fissell, W. H., Patton, J. G., Rome, L. H., Burnette, D. T., & Coffey, R. J. (2019). Reassessment of exosome composition. *Cell*, *177*, 428–445. e18.

Johnson, K. E., Heisel, T., Allert, M., Fürst, A., Yerabandi, N., Knights, D., Jacobs, K. M., Lock, E. F., Bode, L., Fields, D. A., Rudolph, M. C., Gale, C. A., Albert, F. W., Demerath, E. W., & Blekhman, R. (2023). Human milk variation is shaped by maternal genetics and impacts the infant gut microbiome. *BioRxiv*, 2023.01.24.525211 https://doi.org/10.1101/2023.01.24.525211

Jung, H. H., Kim, J. I.- Y., Lim, J. I. E., & Im, Y.-H. (2020). Cytokine profiling in serum-derived exosomes isolated by different methods. *Scientific Reports*, *10*, 14069.

Kahn, S., Liao, Y., Du, X., Xu, W., Li, J., & Lönnerdal, B. O. (2018). Exosomal microRNAs in milk from mothers delivering preterm infants survive in vitro digestion and are taken up by human intestinal cells. *Molecular Nutrition & Food Research*, *62*, 1701050.

Kalluri, R., & Weinberg, R. A. (2009). The basics of epithelial-mesenchymal transition. *Journal of Clinical Investigation*, *119*, 1420–1428.

Kong, A. T., Leprevost, F. V., Avtonomov, D. M., Mellacheruvu, D., & Nesvizhskii, A. I. (2017). MSFragger: Ultrafast and comprehensive peptide identification in mass spectrometry–based proteomics. *Nature Methods*, *14*, 513–520.

Kosaka, N., Izumi, H., Sekine, K., & Ochiya, T. (2010). microRNA as a new immune-regulatory agent in breast milk. *Silence*, *1*, 7.

Kupsco, A., Prada, D., Valvi, D., Hu, L., Petersen, M. S., Coull, B., Grandjean, P., Weihe, P., & Baccarelli, A. A. (2021). Human milk extracellular vesicle miRNA expression and associations with maternal characteristics in a population-based cohort from the Faroe Islands. *Scientific Reports*, *11*, 5840.

Larssen, P., Wik, L., Czarnewski, P., Eldh, M., Löf, L., Ronquist, K. G., Dubois, L., Freyhult, E., Gallant, C. J., Oelrich, J., Larsson, A., Ronquist, G., Villablanca, E. J., Landegren, U., Gabrielsson, S., & Kamali-Moghaddam, M. (2017). Tracing cellular origin of human exosomes using multiplex proximity extension assays. *Molecular & Cellular Proteomics*, *16*, 502–511.

Le Doare, K., Holder, B., Bassett, A., & Pannaraj, P. S. (2018). Mother's milk: A purposeful contribution to the development of the infant microbiota and immunity. *Frontiers in Immunology*, *9*, 361. https://doi.org/10.3389/fimmu.2018.00361

Lefèvre, C., Venkat, P., Kumar, A., Modepalli, V., & Nicholas, K. R. (2019). Comparative analysis of milk microRNA in the therian lineage highlights the evolution of lactation. *Reproduction, Fertility, and Development*, *31*, 1266.

Leiferman, A., Shu, J., Upadhyaya, B., Cui, J., & Zempleni, J. (2019). Storage of extracellular vesicles in human milk, and microRNA profiles in human milk exosomes and infant formulas. *Journal of Pediatric Gastroenterology and Nutrition*, *69*, 235–238.

Lemay, D. G., Ballard, O. A., Hughes, M. A., Morrow, A. L., Horseman, N. D., & Nommsen-Rivers, L. A. (2013). RNA sequencing of the human milk fat layer transcriptome reveals distinct gene expression profiles at three stages of lactation. *PLoS ONE*, *8*, e67531.

Liao, Y., Du, X., Li, J., & Lönnerdal, B. O. (2017). Human milk exosomes and their microRNAs survive digestion in vitro and are taken up by human intestinal cells. *Molecular Nutrition & Food Research*, *61*, 1700082.

Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, *30*, 923–930.

Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., & Mesirov, J. P. (2011). Molecular signatures database (MSigDB) 3.0. *Bioinformatics*, *27*, 1739–1740.

Liedel, J. L., Guo, Y., Yu, Y., Shiou, S.-R. U., Chen, S., Petrof, E. O., Hu, S., Musch, M. W., & Claud, E. C. (2011). Mother's milk-induced hsp70 expression preserves intestinal epithelial barrier function in an immature rat pup model. *Pediatric Research*, *69*, 395–400.

Manca, S., Upadhyaya, B., Mutai, E., Desaulniers, A. T., Cederberg, R. A., White, B. R., & Zempleni, J. (2018). Milk exosomes are bioavailable and distinct microRNA cargos have unique tissue distribution patterns. *Scientific Reports*, *8*, 11321.

Markiewicz, A., Sigorski, D., Markiewicz, M., Owczarczyk-Saczonek, A., & Placek, W. (2021). Caspase-14-From biomolecular basics to clinical approach. A review of available data. *International Journal of Molecular Sciences*, *22*, 5575.

Martin, C., Patel, M., Williams, S., Arora, H., & Sims, B. (2018). Human breast milk-derived exosomes attenuate cell death in intestinal epithelial cells. *Innate Immunity*, *24*, 278–284.

Martin Carli, J. F., Trahan, G. D., Jones, K. L., Hirsch, N., Rolloff, K. P., Dunn, E. Z., Friedman, J. E., Barbour, L. A., Hernandez, T. L., Maclean, P. S., Monks, J., Mcmanaman, J. L., & Rudolph, M. C. (2020). Single cell RNA sequencing of human milk-derived cells reveals sub-populations of mammary epithelial cells with molecular signatures of progenitor and mature states: A novel, non-invasive framework for investigating human lactation physiology. *Journal of Mammary Gland Biology and Neoplasia*, *25*, 367–387.

Ménard, D., Tremblay, E., Ferretti, E., Babakissa, C., Perron, N., Seidman, E. G., Levy, E., & Beaulieu, J.-F. (2012). Anti-inflammatory effects of epidermal growth factor on the immature human intestine. *Physiological Genomics*, *44*, 268–280.

Meyer, J. (2019). Fast proteome identification and quantification from data-dependent acquisition–tandem mass spectrometry (DDA MS/MS) using free software tools. *Methods and Protocols*, *2*, 8.

Molinari, C. E., Casadio, Y. S., Hartmann, B. T., Livk, A., Bringans, S., Arthur, P. G., & Hartmann, P. E. (2012). Proteome mapping of human skim milk proteins in term and preterm milk. *Journal of Proteome Research*, *11*, 1696–1714.

Montagne, P., Cuillière, M. L., Molé, C., Béné, M. C., & Faure, G. (1999). Immunological and nutritional composition of human milk in relation to prematurity and mothers' parity during the first 2 weeks of lactation. *Journal of Pediatric Gastroenterology and Nutrition*, *29*, 75–80.

Na, R. S., Ea, G. X., Sun, W., Sun, X. W., Qiu, X. Y., Chen, L. P., & Huang, Y. F. (2015). Expressional analysis of immune-related miRNAs in breast milk. *Genetics and Molecular Research [Electronic Resource]*, *14*, 11371–11376.

Nielsen, S. D., Beverly, R. L., Underwood, M. A., & Dallas, D. C. (2020). Differences and similarities in the peptide profile of preterm and term mother's milk, and preterm and term infant gastric samples. *Nutrients*, *12*, 2825.

Ninkina, N., Kukharsky, M. S., Hewitt, M. V., Lysikova, E. A., Skuratovska, L. N., Deykin, A. V., & Buchman, V. L. (2019). Stem cells in human breast milk. *Human Cell*, *32*, 223–230.

Nyquist, S. K., Gao, P., Haining, T. K. J., Retchin, M. R., Golan, Y., Drake, R. S., Kolb, K., Mead, B. E., Ahituv, N., Martinez, M. E., Shalek, A. K., Berger, B., & Goods, B. A. (2022). Cellular and transcriptional diversity over the course of human lactation. *Proceedings of the National Academy of Sciences*, *119*, e2121720119.

O'reilly, D., Dorodnykh, D., Avdeenko, N. V., Nekliudov, N. A., Garssen, J., Elolimy, A. A., Petrou, L., Simpson, M. R., Yeruva, L., & Munblit, D. (2021). Perspective: The role of human breast-milk extracellular vesicles in child health and disease. *Advances in Nucleic Acids Research*, *12*, 59–70.

Peixoto, L., Risso, D., Poplawski, S. G., Wimmer, M. E., Speed, T. P., Wood, M. A., & Abel, T. (2015). How data analysis affects power, reproducibility and biological insight of RNA-seq studies in complex datasets. *Nucleic Acids Research*, *43*, 7664–7674.

Pistritto, G., Jost, M., Srinivasula, S. M., Baffa, R., Poyet, J.-L., Kari, C., Lazebnik, Y., Rodeck, U., & Alnemri, E. S. (2002). Expression and transcriptional regulation of caspase-14 in simple and complex epithelia. *Cell Death and Differentiation*, *9*, 995–1006.

Pluchino, S., & Smith, J. A. (2019). Explicating exosomes: Reclassifying the rising stars of intercellular communication. *Cell*, *177*, 225–227.

Qin, W., Tsukasaki, Y., Dasgupta, S., Mukhopadhyay, N., Ikebe, M., & Sauter, E. R. (2016). Biology of human tumors exosomes in human breast milk promote EMT. *Clinical Cancer Research*, *22*, 4517–4524.

Rani, P., Vashisht, M., Golla, N., Shandilya, S., Onteru, S. K., & Singh, D. (2017). Milk miRNAs encapsulated in exosomes are stable to human digestion and permeable to intestinal barrier in vitro. *Journal of Functional Foods*, *34*, 431–439.

Rani, P., Yenuganti, V. R., Shandilya, S., Onteru, S. K., & Singh, D. (2017). The hidden bioactive component of milk. *Trends in Food Science & Technology*, *65*, 94–102.

Risso, D., Ngai, J., Speed, T. P., & Dudoit, S. (2014). Normalization of RNA-seq data using factor analysis of control genes or samples. *Nature Biotechnology*, *32*, 896–902.

Ronayne De Ferrer, P. A., Baroni, A., Sambucetti, M. E., López, N. E., & Ceriani Cernadas, J. M. (2021). Lactoferrin levels in term and preterm milk. *Journal of the American College of Nutrition*, *19*, 370–373.

Santoro, A., Conde, J., Scotece, M., Abella, V., Lois, A., Lopez, V., Pino, J., Gomez, R., Gomez-Reino, J. J., & Gualillo, O. (2015). SERPINE2 inhibits IL-1α-induced MMP-13 expression in human chondrocytes: Involvement of ERK/NF-κB/AP-1 pathways. *PLoS ONE*, *10*, e0135979.

Savulescu, D. M., Groome, M., Malfeld, S. C. K., Madhi, S., Koen, A., Jones, S., Duxbury, V., Scheuermaier, K., De Assis Rosa, D., & Suchard, M. (2020). HLA antibody repertoire in infants suggests selectivity in transplacental crossing. *American Journal of Reproductive Immunology*, *84*, e13264.

Schiavinato, A., Becker, A.-K. A., Zanetti, M., Corallo, D., Milanetto, M., Bizzotto, D., Bressan, G., Guljelmovic, M., Paulsson, M., Wagener, R., Braghetta, P., & Bonaldo, P. (2012). EMILIN-3, peculiar member of elastin microfibril interface-located protein (EMILIN) family, has distinct expression pattern, forms oligomeric assemblies, and serves as transforming growth factor β (TGF-β) antagonist. *Journal of Biological Chemistry*, *287*, 11498–11515.

Shalekoff, S., Gray, G. E., & Tiemessen, C. T. (2004). Age-related changes in expression of CXCR4 and CCR5 on peripheral blood leukocytes from uninfected infants born to human immunodeficiency virus type 1-infected mothers. *Clinical and Diagnostic Laboratory Immunology*, *11*, 229–234.

Shao, D., Huang, L., Wang, Y., Cui, X., Li, Y., Wang, Y., Ma, Q., Du, W., & Cui, J. (2021). HBFP: A new repository for human body fluid proteome. *Database*, *2021*, baab065.

Sharma, R., Yang, Y., Sharma, A., Awasthi, S., & Awasthi, Y. C. (2004). Antioxidant role of glutathione S-transferases: Protection against oxidant toxicity and regulation of stress-mediated apoptosis. *Antioxidants & Redox Signaling*, *6*, 289–300.

Sharp, J. A., Lefèvre, C., Watt, A., & Nicholas, K. R. (2016). Analysis of human breast milk cells: Gene expression profiles during pregnancy, lactation, involution, and mastitic infection. *Functional & Integrative Genomics*, *16*, 297–321.

Shiff, Y. E., Reif, S., Marom, R., Shiff, K., Reifen, R., & Golan-Gerstl, R. (2019). MiRNA-320a is less expressed and miRNA-148a more expressed in preterm human milk compared to term human milk. *Journal of Functional Foods*, *57*, 68–74.

Shu, J., Chiang, K., Zempleni, J., & Cui, J. (2015). Computational characterization of exogenous microRNAs that can be transferred into human circulation. *PLoS ONE*, *10*, e0140587.

Silverman, G. A., Bird, P. I., Carrell, R. W., Church, F. C., Coughlin, P. B., Gettins, P. G. W., Irving, J. A., Lomas, D. A., Luke, C. J., Moyer, R. W., Pemberton, P. A., Remold-O'donnell, E., Salvesen, G. S., Travis, J., & Whisstock, J. C. (2001). The serpins are an expanding superfamily of structurally similar but functionally diverse proteins. Evolution, mechanism of inhibition, novel functions, and a revised nomenclature. *Journal of Biological Chemistry*, *276*, 33293–33296.

Spong, C. Y., Mercer, B. M., D'alton, M., Kilpatrick, S., Blackwell, S., & Saade, G. (2011). Timing of indicated late-preterm and early-term birth. *Obstetrics and Gynecology*, *118*, 323–333.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., & Mesirov, J. P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *PNAS*, *102*, 15545–15550.

Théry, C., Witwer, K. W., Aikawa, E., Alcaraz, M. J., Anderson, J. D., Andriantsitohaina, R., Antoniou, A., Arab, T., Archer, F., Atkin-Smith, G. K., Ayre, D. C., Bach, J.-M., Bachurski, D., Baharvand, H., Balaj, L., Baldacchino, S., Bauer, N. N., Baxter, A. A., Bebawy, M., … Zuba-Surma, E. K. (2018). Minimal information for studies of extracellular vesicles 2018 (MISEV2018): A position statement of the International Society for Extracellular Vesicles and update of the MISEV2014 guidelines. *Journal of Extracellular Vesicles*, *7*, 1535750.

Torres-Castro, P., Grases-Pintó, B., Abril-Gil, M., Castell, M., Rodríguez-Lagunas, M. J., Pérez-Cano, F. J., & Franch, À. (2020). Modulation of the systemic immune response in suckling rats by breast milk TGF-β2, EGF and FGF21 supplementation. *Nutrients*, *12*, 1888.

Trend, S., Strunk, T., Lloyd, M. L., Kok, C. H., Metcalfe, J., Geddes, D. T., Lai, C. T., Richmond, P., Doherty, D. A., Simmer, K., & Currie, A. (2016). Levels of innate immune factors in preterm and term mothers' breast milk during the 1st month postpartum. *British Journal of Nutrition*, *115*, 1178–1193.

Twigger, A.-J., Engelbrecht, L. K., Bach, K., Schultz-Pernice, I., Pensa, S., Stenning, J., Petricca, S., Scheel, C. H., & Khaled, W. T. (2022). Transcriptional changes in the mammary gland during lactation revealed by single cell sequencing of cells from human milk. *Nature Communications*, *13*, 562.

Underwood, M. A. (2013). Human milk for the premature infant. *Pediatric Clinics of North America*, *60*, 189–207.

Vahkal, B., Kraft, J., Ferretti, E., Chung, M., Beaulieu, J.-F., & Altosaar, I. (2021). Review of methodological approaches to human milk small extracellular vesicle proteomics. *Biomolecules*, *11*, 833.

Van Herwijnen, M. J. C., Zonneveld, M. I., Goerdayal, S., Nolte-'T Hoen, E. N. M., Garssen, J., Stahl, B., Maarten Altelaar, A. F., Redegeld, F. A., & Wauben, M. H. M. (2016). Comprehensive proteomic analysis of human milk-derived extracellular vesicles unveils a novel functional proteome distinct from other milk components. *Molecular & Cellular Proteomics*, *15*, 3412–3423.

Wan, J., Cui, X.-W., Zhang, J., Fu, Z. I.-Y. I., Guo, X. I.-R., Sun, L. I.-Z., & Ji, C.-B. O. (2013). Peptidome analysis of human skim milk in term and preterm milk. *Biochemical and Biophysical Research Communications*, *438*, 236–241.

Wang, X., Yan, X., Zhang, L. E., Cai, J., Zhou, Y., Liu, H., Hu, Y., Chen, W., Xu, S., Liu, P., Chen, T., Zhang, J., Cao, Y., Yu, Z., & Han, S. (2019). Identification and peptidomic profiling of exosomes in preterm human milk: Insights into necrotizing enterocolitis prevention. *Molecular Nutrition & Food Research*, *63*, 1801247. https://doi.org/10.1002/mnfr.201801247

Ward, T. L., Hosid, S., Ioshikhes, I., & Altosaar, I. (2013). Human milk metagenome: A functional capacity analysis. *BMC Microbiology*, *13*, 1–12.

Whitman, J. D., & Lynch, K. L. (2019). Optimization and comparison of information-dependent acquisition (IDA) to sequential window acquisition of all theoretical fragment ion spectra (SWATH) for high-resolution mass spectrometry in clinical toxicology. *Clinical Chemistry*, *65*(7), 862–870. https://doi.org/10.1373/clinchem.2018.300756

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., Fu, X., Liu, S., Bo, X., & Yu, G. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation*, *2*, 100141.

Yamada, T., Inoshima, Y., Matsuda, T., & Ishiguro, N. (2012). Comparison of methods for isolating exosomes from bovine milk. *Journal of Veterinary Medical Science*, *74*, 1523–1525.

Yan, X., Liu, L., Yao, S., Chen, Y., Yu, Q., Jiang, C., Chen, W., Chen, X., & Han, S. (2022). LncRNA and mRNA profiles of human milk-derived exosomes and their possible roles in protecting against necrotizing enterocolitis. *Food & Function*, *13*, 12953–12965.

Yang, M., Song, D., Cao, X., Wu, R., Liu, B., Ye, W., Wu, J., & Yue, X. (2017). Comparative proteomic analysis of milk-derived exosomes in human and bovine colostrum and mature milk samples by iTRAQ-coupled LC-MS/MS. *Food Research International*, *92*, 17–25.

Zhang, F., Ge, W., Ruan, G., Cai, X., & Guo, T. (2020). Data-independent acquisition mass spectrometry-based proteomics and software tools: A glimpse in 2020. *Proteomics*, *20*, 1900276.

Zhang, H., Zhang, S., Dang, X., Lin, L., Ren, L., & Song, R. (2022). GPNMB plays an active role in the M1/M2 balance. *Tissue & Cell*, *74*, 101683.

Zhou, J.-K., Fan, X., Cheng, J., Liu, W., & Peng, Y. (2021). PDLIM1: Structure, function and implication in cancer. *Cell Stress*, *5*, 119–127.

Zhou, L., Zhuo, H., Ouyang, H., Liu, Y., Yuan, F., Sun, L., Liu, F., & Liu, H. (2017). Glycoprotein non-metastatic melanoma protein b (Gpnmb) is highly expressed in macrophages of acute injured kidney and promotes M2 macrophages polarization. *Cellular Immunology*, *316*, 53–60.

Zhou, M., Weber, S. R., Zhao, Y., Chen, H., & Sundstrom, J. M. (2020). Methods for exosome isolation and characterization. in *Exosomes* (pp. 23–38). Elsevier. https://doi.org/10.1016/B978-0-12-816053-4.00002-X

Zhou, Q. I., Li, M., Wang, X., Li, Q., Wang, T., Zhu, Q. I., Zhou, X., Wang, X., Gao, X., & Li, X. (2011). Immune-related microRNAs are abundant in breast milk exosomes. *International Journal of Biological Sciences*, *8*, 118–123.

Zhou, Y., Yu, Z., Wang, X., Chen, W., Liu, Y., Zhang, Y., Yin, J., & Han, S. (2021). Exosomal circRNAs contribute to intestinal development via the VEGF signalling pathway in human term and preterm colostrum. *Aging (Albany, NY)*, *13*, 11218–11233.

Zonneveld, M. I., Van Herwijnen, M. J. C., Fernandez-Gutierrez, M. M., Giovanazzi, A., De Groot, A. M., Kleinjan, M., Van Capel, T. M. M., Sijts, A. J. A. M., Taams, L. S., Garssen, J., De Jong, E. C., Kleerebezem, M., Nolte-'T Hoen, E. N. M., Redegeld, F. A., & Wauben, M. H. M. (2021). Human milk extracellular vesicles target nodes in interconnected signalling pathways that enhance oral epithelial barrier function and dampen immune responses. *Journal of Extracellular Vesicles*, *10*, e12071.

Zou, Y.-R., Kottmann, A. H., Kuroda, M., Taniuchi, I., & Littman, D. R. (1998). Function of the chemokine receptor CXCR4 in haematopoiesis and in cerebellar development. *Nature*, *393*, 595–599.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Vahkal, B., Altosaar, I., Tremblay, E., Gagné, D., Hüttman, N., Minic, Z., Côté, M., Blais, A., Beaulieu, J.-F., & Ferretti, E. (2024). Gestational age at birth influences protein and RNA content in human milk extracellular vesicles. *Journal of Extracellular Biology*, *3,* e128. https://doi.org/10.1002/jex2.128