


 Cite this: *RSC Adv.*, 2024, **14**, 29189

Optimizing workflow efficiency for analyzing low molecular weight endogenous peptides in colostrum†

Priti Panchal, ^a Reeju Rani, ^a Rohit Kumar, ^b Seema Malik, ^a Manishi Mukesh, ^c Jai Kumar Kaushik, ^a Monika Sodhi, ^c Ashok Kumar Mohanty, ^d and Sudarshan Kumar ^{*a}

Bovine milk and colostrum play pivotal roles in the nutritional support of both human and bovine infants. Colostrum, the initial milk secretion, is crucial for neonatal growth, providing essential nutrients, growth factors, immunity, and defense mechanisms through a diverse array of bioactive compounds, including bioactive proteins and peptides. Peptidomics, leveraging the potential health benefits of peptides derived from food and body fluids, has become prominent in contemporary research. Endogenous peptides (EPs) have gained notable scientific and commercial interest due to their potential biofunctional significance in areas such as immune health, antimicrobial, anti-inflammatory, antihypertensive, and antioxidant studies. In this investigation, we aimed to extract and analyze low molecular weight EPs from colostrum using four distinct peptide extraction methods, previously employed for EPs extraction from other bodily fluids. The efficiency of these methods was systematically compared and analysed to identify the most effective extraction technique for maximizing the identification of low molecular weight EPs from colostrum. This study represents a pioneering effort as no prior research has systematically compared different extraction methods for low molecular weight EPs from colostrum. Given the unique physical and chemical composition of colostrum compared to milk and other body fluids, a comprehensive analysis of EPs extraction methods was deemed essential. In the present study, we successfully extracted over 3200 EPs from colostrum using trichloroacetic acid (TCA) and a molecular weight cut off (MWCO) extraction method. The findings of this study revealed the extraction of EPs from colostrum, demonstrating potential inherent bioactivities as predicted by *in silico* tools.

Received 30th April 2024
 Accepted 1st August 2024

DOI: 10.1039/d4ra03199g
rsc.li/rsc-advances

Introduction

Peptidomics is the branch of proteomics that refers to the qualitative and quantitative analysis and characterization of the entire pool of peptides existing in the biological samples.¹⁻⁴ Peptidomics is a complex field that has successfully been applied in several areas of research for quantitation and identification of food bioactive peptides, characterization of food processing related proteolysis, profiling of human milk

peptides, profiling of peptides in gastric samples *etc*. Despite the thorough investigation of peptides produced by hydrolysis or fermentation, native peptides already existing in the mammary gland or other parts of the body are usually neglected. EPs are naturally existing peptides in the biological fluids. These EPs emanate by translational or proteolytic degradation through the action of proteases naturally present in the same biological system.¹ EPs are the low molecular weight components typically consisting of 2–50 amino acids (AA). EPs are extensively distributed throughout almost all biological fluids including blood plasma,^{1,5} urine,⁶ colostrum,^{7,8} milk⁹ and milk products¹⁰⁻¹⁸ and these peptides have several important biological functions. Peptidome profiling of 10 commercial dairy products identified 66 unique bioactive peptides with immunomodulatory, anti-hypertensive, anti-thrombotic, antimicrobial, anti-oxidative and opioid agonist functions.¹⁸ EPs get directly absorbed in the intestine and provide potential health benefits to new-borns. In a recent study, 8393 EPs were identified and quantified in the milk of two different breeds of camel. The identified peptides were functionally annotated to determine the bioactive functionality.¹⁰ Recently, a detailed

^aICAR-National Dairy Research Institute, Proteomics and Cell Biology Lab, Animal Biotechnology Division (ABTD), Karnal, Haryana, 132001, India. E-mail: kumarsudarshan@gmail.com

^bCollege of Public Health and Human Science, Oregon State University, Corvallis, OR, 97331, USA

^cICAR-National Bureau of Animal Genetic Resources, Karnal, Haryana, 132001, India

^dICAR-Central Institute for Research on Cattle, Meerut, India

† Electronic supplementary information (ESI) available: Table S1: different criteria used to evaluate the efficiency of TCA and MWCO extraction methods. Table S2: percentage peak area occupied by some representative EPs extracted via TCA precipitation and MWCO methods. See DOI: <https://doi.org/10.1039/d4ra03199g>

identification and characterization of short EPs in milk and its by-products including whole whey, skimmed whey, and whey permeate were carried out with the help of high-resolution mass spectrometry, and in total 79 short EPs were identified.¹¹ Several studies have been performed in recent years for the peptidome analysis of EPs in human, bovine, camel and other species milk.^{10,17,19-23} Though milk has been explored widely for the peptidome studies, comparatively less number of studies are available on the investigation of colostrum derived peptides. Several studies have proven the immunological and biological significance of colostrum^{24,25} especially in the neonate health development,²⁵ immunomodulation^{26,27}, regulation of cytokine production and improvement of gut health.²⁸ Colostrum contains a wide array of immune boosting factors including antimicrobial,^{26,29} antioxidative,³⁰ anti-inflammatory, antiviral and growth promoting proteins and peptides, immunoglobulins, lactoferrin, insulin like growth factor and other. A recent study reported isolation of bioactive peptides from buffalo colostrum whey produced by bacterial fermentation. It identified 40 potential immunomodulatory sequences and 16 antimicrobial sequences.²⁶ And most of the biologically important components are at higher concentration in colostrum than in milk.³¹ Nissen *et al.* performed a comparative study between bovine colostrum proteome and milk proteome to rank the proteins mutually and to generate a ratio between colostrum and milk proteome. They observed the highest concentration of most of proteins including osteopontin, total immunoglobulins, lactotransferrin, apolipoprotein, milk amyloid protein A, haptoglobin, transforming growth factor β -2 and transforming growth factor β -1 in the colostrum. The concentration of these proteins declined greatly in the milk.²⁴ Colostrum was found to have developmental and immunological importance since several proteins such as immunoglobulins, thyroglobulins, RNA binding proteins were at higher level in colostrum than in milk of Mediterranean and Murrah buffalo.²⁵ Colostrum is also found to promote the performance of athlete in post exercise inflammatory responses by reducing the hyper-permeability and maintaining the junctional integrity of the intestine.³² Colostrum derived peptides have recently been explored since these peptides are found to have significant bioactive functionality. Recently, a group of scientists studied the colostrum derived peptides generated by simulated digestion in oral, gastric, and duodenal phases. Some of the identified peptide sequences showed potential antimicrobial activity when predicted bioinformatically.³³ Ashok and coworkers investigated the antioxidative potency of peptic hydrolysate of buffalo colostrum whey and observed that the hydrolysates diminished the increased level of ROS, H_2O_2 and catalase.³⁰ While most of previous studies focused on the peptides generated by *in vitro* digestion of the proteins, there are very few studies on the profiling or identification of colostrum-derived EPs. In a previous study three naturally present peptides named isracidin, casecidin 15 and 17 were isolated and characterized from bovine colostrum and were found to have antimicrobial activity against *E. coli*.³⁴ Few other studies on colostrum-derived EPs prove the significance of these peptides.^{7,35,36} Given the different numbers of EPs reported in different studies, a comprehensive

research is required to achieve an effective extraction method of EPs from the colostrum in order to form an accurate and reproducible peptidome base. So, in this study, we aimed to explore and compare the efficiency of various methods which has been used previously for EPs extraction from various body fluids but colostrum. We introduced modifications in these methods in order to achieve the maximum identification of colostral peptides.

Material and method

Sample collection

Bovine colostrum samples were collected within 24 h of calving in sterile tubes and stored at -80°C until further use to minimize the milk protease activity, which has the potential to modify the peptide profile.

Ultracentrifugation

The fresh colostrum samples were pooled ($n = 8$) and defatted by centrifugation at $1000 \times g$ for 20 min at 4°C . Fat layer was removed and the skimmed colostrum infranate was collected in fresh tubes and stored at refrigerated condition for further use. Skimmed colostrum infranate was then subjected to ultracentrifugation in order to deplete casein which accounts for about 80% of the total proteins. Due to the abundance of casein in the colostrum, other less abundant proteins and peptides would be difficult to identify. For this skimmed colostrum was centrifuged at $70\,000 \times g$ for 60 min at 4°C and then the remaining fat layers was removed, supernatant was collected and again centrifuged at $100\,000 \times g$ for another 60 min at 4°C . Collected supernatant (casein-depleted colostrum) was stored at -80°C till further use (Fig. 1A).

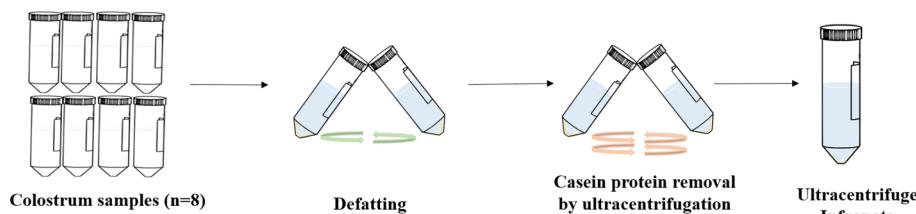
Peptide extraction

Peptide extraction from colostrum samples was performed by four different extraction methods. Equal volume of ultracentrifuged colostrum was taken for further processing (Fig. 1B).

Molecular weight cutoff

Low molecular weight EPs were isolated through centrifugation-based filtration utilizing a 9 kDa molecular weight cutoff (MWCO) device (Protein Concentrators®, Thermo Scientific, Pierce Biotechnologies, Rockford, USA). Initially, the filter underwent a washing step with ultrapure water, employing centrifugation at $3000 \times g$ for 15 minutes at 4°C . Subsequently, casein-depleted colostrum was mixed with acetonitrile (ACN) (Honeywell, France) at a 2 : 1 ratio in an ACN-compatible tube. After vortexing for 5 min and a 30 minute incubation at 4°C , the tube's contents were transferred into the filtration device and subjected to centrifugation under above mentioned conditions for 20 minutes. The retentate, harboring proteins of higher molecular weight, was discarded, while the filtrate containing peptides was collected, stored at -20°C and subjected to peptide enrichment and desalting using spin columns.

A



B

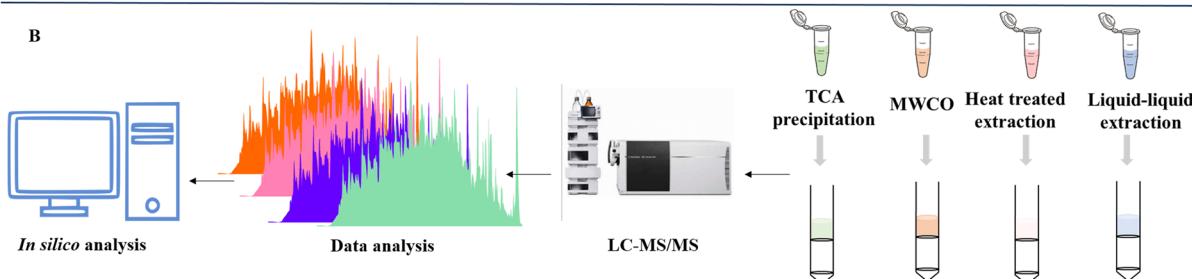


Fig. 1 Flow diagram depicting the collection and processing of colostrum samples. (A) After collection, colostrum was defatted and ultracentrifuged at high speed and infranate was collected. (B) For the extraction of EPs, four different peptide extraction methods were applied: TCA precipitation, MWCO, heat treated and liquid–liquid extraction. The samples were desalted and analysed using mass spectrometry.

TCA precipitation

EPs were extracted using the protocol established by Ferranti *et al.*² In summary, casein-depleted colostrum was treated with an equal volume of 20% trichloroacetic acid (TCA) (SRL, India) solution (w/v) in water, resulting in a final TCA concentration of 100 g L⁻¹. The mixture was vigorously vortexed for 5 minutes and subsequently subjected to centrifugation at 3000 × g for 20 minutes at 4 °C. The resulting supernatant was carefully collected and then processed through peptide enrichment and desalting.

Liquid-liquid extraction

This approach was adapted from established procedures with some modifications. To isolate EPs from colostrum, 2400 µL of methanol (Honeywell, France) was mixed with 800 µL of casein-depleted colostrum and thoroughly vortexed. Following this, 2135 µL of chloroform (SRL, India) was added and the mixture was vortexed once again. The mixture was continuously shaken for 1 hour at 4 °C. Subsequently, 1065 µL of cold water was added, and the sample underwent centrifugation at 10 000 × g for 20 minutes at 4 °C. The organic phase was discarded, and the remaining aqueous phase was collected followed by concentration in a speed vac concentrator (Eppendorf Concentrator Plus™, Eppendorf, Hamburg, Germany). The concentrate was then reconstituted in ultrapure water and subjected to peptide enrichment and desalting.

Heat-treated extraction

In this methodology, casein-depleted colostrum was heated in a microwave for a duration of 30 seconds. Subsequently, the resulting liquid fraction was collected and subjected to

centrifugation at 10 000 × g for a duration of 20 minutes at 4 °C. The supernatant, thus obtained, underwent collection and filtration through a molecular weight cutoff filter with a 10 kDa molecular weight cutoff. The resultant filtrate was further processed through peptide enrichment and desalting using spin desalting columns.

Peptide enrichment and desalting

Prior to the identification of peptides through LC-MS/MS, it is imperative to undertake the purification and enrichment of peptides. This is essential to enhance sensitivity in peptide identification and address the challenge of signal suppression arising from the concurrent presence of non-target molecules. The extracted peptides were subsequently subjected to desalting using solid-phase extraction. Samples were desalted using Pierce™ Peptide Desalting Spin Columns (Thermo Scientific, Rockford, USA). Column was first activated using 100% methanol followed by equilibration with 2% ACN, 0.1% TFA. The sample was loaded onto the column, followed by column wash with 2% ACN, 0.1% TFA and the bound peptides were subsequently eluted from column by 40% ACN, 0.1% TFA. The eluate was dried down by vacuum concentrator (Eppendorf Concentrator Plus™, Eppendorf, Hamburg, Germany) and identification of peptides was achieved by mass spectrometry.

Mass spectrometry

The identification of colostrum-derived EPs was done using high-resolution mass spectrometry (HRMS) (Agilent Technologies, Santa Clara, CA, USA) at the National Chemical Laboratory (Venture Center), Pune. All the samples were reconstituted in 50 µL 0.1% formic acid. The injection volume was 25 µL and



column used was Zorbax Eclipse Plus c18 column (150 mm \times 2.1 mm, 1.8 μ m). Sample was injected at a flow rate of 0.5 mL min $^{-1}$. MS analysis was performed with Agilent 6550 UHD QTOF MS equipped with Dual AJS ESI ion source. The chromatographic gradients were as shown below A: water with 0.1% formic acid; B: 0.1% formic acid in 90% ACN. The gradient employed was ramped from 0 to 8% B for 0 to 5 min, 8 to 27% B from 5 to 24 min, 27 to 99% B from 24 to 55 min, followed by 99% B for 5 min. The acquisition mode applied was auto MS/MS and precursor ion formed were acquired in the range of 100–1700 m/z in positive polarity. The scan rate for MS was 2 spectra per second and for MS/MS it was 3 spectra per second. Other instrument parameters used were: gas temp – 270 °C, gas flow – 11 L min $^{-1}$, Nebulizer – 25 psig, sheath gas temp – 295 °C, sheath gas flow 10.

Data analysis

Raw data file was analysed through Trans Proteomic Pipeline version 6.3.0 Arcus as described by Kumar *et al.*, 2021 (ref. 6) with some modifications. Briefly, .d file were converted to mzML format with the help of MSconvert (a proteowizard tool). This newly generated mzML file was then searched using comet search parameters for MS/MS search against UniProt *Bos taurus* proteome database (Proteome Id UP000009136). The search was performed using parameters for undigested peptides and for N and C-terminal unspecific cleavage. Up to four potential variable modifications were allowed on each peptide including phosphorylation at serine, threonine and tyrosine; methionine oxidation; asparagine and glutamine deamidation. Peptide mass tolerance was set 20 ppm and maximum missed cleavage were allowed up to 2. Remaining parameters were set as default. The .xml file generated after comet search was subjected to peptide prophet and iProphet. Peptide prophet and iProphet were run in Xpress in label free mode to calculate the peak area. Percentage peak area occupied by the sequences was calculated by dividing the sum of peak area to the peak area occupied by individual sequence. Finally, protein prophet was used to assign a protein probability score to each identified peptide. High confidence peptides were accepted based on the iProphet probability cut-off and error rate. An error rate of <0.05 was accepted for

identification of the peptides as correct hit. Peak area of the identified sequences was calculated by running the peptide prophet and iProphet in label free mode.

Statistical analysis

The qualitative analysis of identified peptides was conducted utilizing Microsoft Excel. Venn diagrams were constructed employing Venny 2.1.0.³⁷ Bar charts were generated using both MS Excel and GraphPad Prism version 8.0.1 (Graph Pad Software Inc., San Diego, CA, United States). Additional plots were crafted using SRplot,³⁸ and BoxplotR.³⁹ The molecular weight of identified peptides was computed through the utilization of TumorHPD.⁴⁰ To assess the hydrophobicity of peptide sequences, the GRAVY calculator (<https://www.gravy-calculator.de>) was employed.

Results

Evaluation of number of identified peptides in different methods

Total number of unique peptides identified in Yak colostrum are shown in Fig. 2A. The total number of unique peptides extracted were 3211, 2656, 3210 and 1716 in TCA precipitation, heat-treated, MWCO and Liquid–Liquid (LL) extraction, respectively. Peptide sequences were selected with an error rate of <0.05 (Fig. 2C). TCA precipitation and MWCO extraction methods were found to be more effective for extraction of low molecular weight EPs from Yak colostrum since both the methods identified equal number of peptides. TCA precipitation and MWCO methods yielded highest number of peptides among all the methods. Venn Diagram for the number of common sequences identified in the three methods shows that only 91 sequences were common through all the methods *i.e.* TCA precipitation, MWCO method, heat treated and LL extraction (Fig. 2B). A total of 171 sequences were common between TCA precipitation and MWCO method. And 197 sequences were common among TCA and heat treated extraction method while 138 sequences were common in TCA and LL extraction method. 2941, 2883, 2311, 1482 sequences were

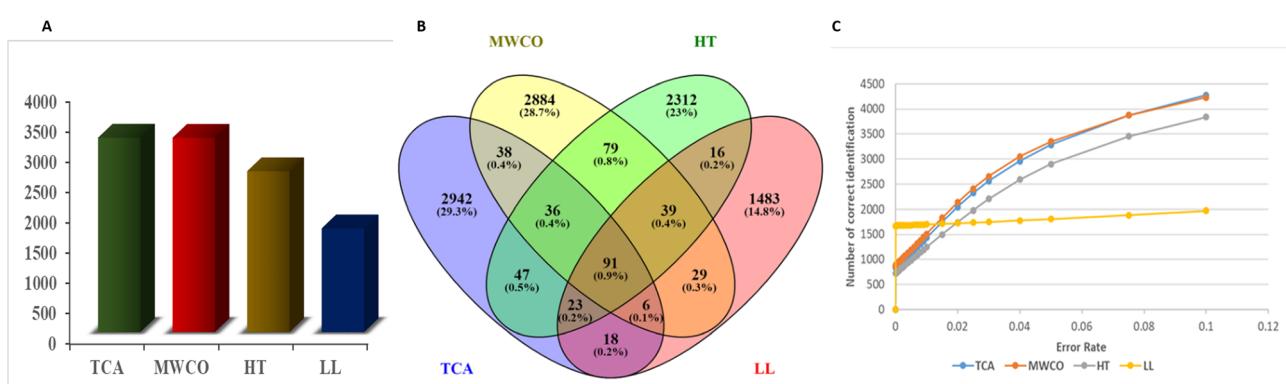


Fig. 2 Peptide sequences identified across different extraction methods. (A) Frequency of identified sequences in four methods: TCA precipitation, MWCO, HT and LL extraction methods. (B) Venn diagram displaying the number of common peptide sequences among TCA, MWCO, HT and LL methods. (C) Number of correct identification against the error rate in each method.



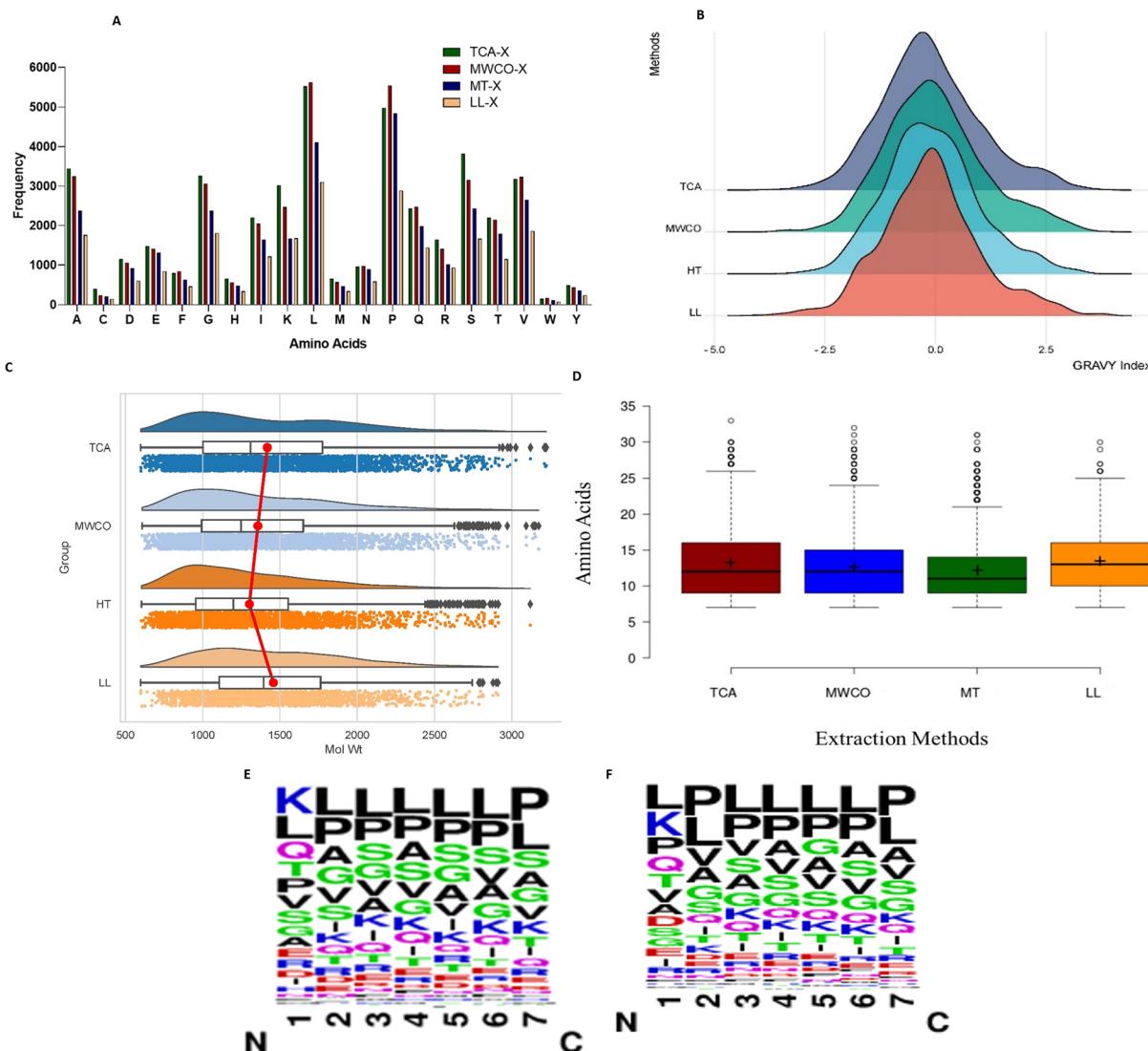


Fig. 3 Physico-chemical characterisation of identified sequences. (A) Amino acid frequency of identified sequences across four methods. (B) Gravy index of sequences. (C) Molecular mass of identified sequences in each method. (D) Peptide sequence length (in AA) of identified sequences across all four methods. (E) Sequence logos (7 amino acid from N-terminus) of sequences identified in TCA and (F) in MWCO.

exclusively identified in TCA, MWCO, HT and LL extraction methods respectively.

Evaluation of amino acid profile of identified peptides in different methods

We summarized information regarding the prevalence of amino acids in the identified sequences as depicted in Fig. 3A. Leucine emerged as the most widely distributed amino acid in peptide sequences obtained through all extraction methods, except in LL where proline prevailed as the most abundant. Hydrophobic amino acids, such as P, G, L, V, and A, were consistently present in high frequency across all extraction methods. Additionally, hydrophilic amino acids like S, T, and Q were also enriched abundantly in all four methods.

Further, analysis of the sequence logos for the first seven amino acids from the N-terminus of sequences identified via

the TCA and MWCO methods was performed (Fig. 3E and F). It was observed that the first position was predominantly occupied by lysine, followed by leucine and glutamine in TCA method. While for the MWCO method, leucine, lysine, and proline were the most common in the first position. From the second position onwards, leucine and proline were the most abundant amino acids in both methods.

Evaluation of mass range and sequence length of identified peptides

The graphical representation in Fig. 3C illustrates that all employed methods primarily resolved low molecular weight peptides within the range of 700 Da to <3300 Da. Notably, LL peptide extraction method exhibited an absence of peptides exceeding 3000 Da. Across all methods, the prevalence of detected peptides fell within the range of 700 Da to 2000 Da.

Further analysis revealed that TCA, MWCO and heat treated extraction methods predominantly identified peptides in the range of 900 Da to <1000 Da (TCA – 10.64%, MWCO – 10.68%, HT – 12.01%), while LL extraction detected the highest number of peptides within the 1100 Da to <1200 Da range (10.31%) despite method-specific variations, the majority of peptides obtained through the different extraction methods displayed comparable patterns, both in terms of the number of amino acid residues and the distribution of molecular weights.

Further, it was observed that all the applied methods consistently retrieved sequences spanning from 7 amino acids (AA) to 30 amino acids (Fig. 3D). Significantly, the predominant sequence length ranged from 7AA to 18AA with 85.23%, 89.47%, 87.34%, and 91.41% for TCA, MWCO, LL extraction, and heat-treated extraction, respectively. It is noteworthy that the amino acid sequences obtained through MWCO and heat-treated extraction were significantly shorter, respectively accounting for 50.90% and 55.15%, within the range of 8AA to 12AA.

Hydropathicity of identified peptides

To determine the hydrophobicity of the sequences, we computed GRAVY index values for sequences identified through all four methods (Fig. 3B). As depicted in the figure, peptides identified across these methods predominantly exhibited GRAVY values ranging from –0.25 to 0.25. Nevertheless, 54–58% of the peptides identified through these methods displayed GRAVY index values below 0, signifying hydrophilic characteristics. Concurrently, there was also a presence of hydrophobic peptides, constituting 41–45% across different methods.

While each method resulted in identification of both hydrophilic and hydrophobic peptides, the prevalence of hydrophilic sequences was notably prominent. TCA extraction

method exhibited enrichment of hydrophilic peptides, constituting 58.17%, surpassing the other three methods (MWCO – 54.26%, HT – 54.94%, LL – 56.36%).

Parent protein of identified EPs

Parent proteins of identified EPs were identified and analysed as shown in Fig. 4. All the peptides identified through TCA, MWCO, HT, LL extraction methods were derived from 2647, 2587, 2155 and 1428 proteins respectively (Fig. 4A). Remarkably, 80 parent proteins were consistently identified across all the methods with TCA and MWCO exhibiting the highest commonality, sharing 581 proteins, accounting for 12.5%, as depicted in the Fig. 4A. This huge number of unique protein across the methods is due to different members of a family or groups have different Uniprot IDs.

Functional annotation of identified sequences

Identified peptide sequences were scrutinized for potential bioactive attributes utilizing bioinformatics tools. Peptide bioactivity prediction was executed employing PeptideRanker. PeptideRanker is neural network based tool which assigns a bioactive probability to each peptide sequence, ranging from 0 to 1, where scores nearing 1 indicate a high confidence for bioactive potential of given peptide sequence. Peptides attaining a score >0.50 were deemed to possess a significant probability to exhibit bioactivity. Remarkably, a substantial proportion of the identified sequences, exceeding 25%, exhibited scores greater than 0.5 in PeptideRanker, indicating a high-confidence bioactivity prediction (Fig. 5). Furthermore, over 60% of the sequences were within the intermediate range of PeptideRanker scores, ranging from 0.1 to 0.5, suggesting a moderate probability of bioactivity. These findings collectively

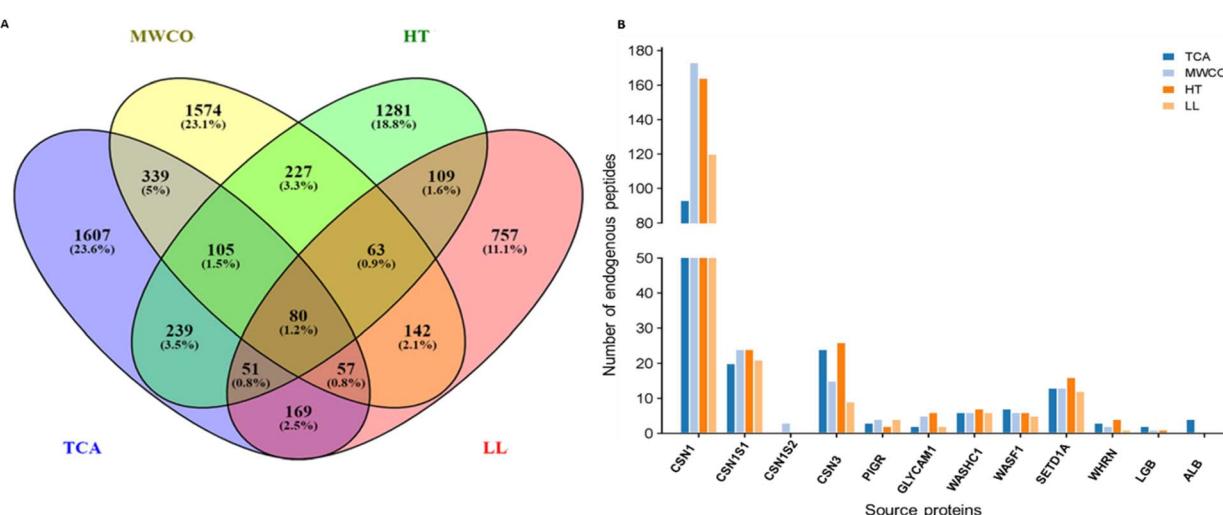


Fig. 4 Parent proteins analysis of identified endogenous peptides. (A) Venn diagram displaying the number of parent proteins from which the endogenous peptides were originated among four different methods; TCA, MWCO, HT, LL. (B) Graph showing number of identified peptides from major milk proteins. CSN1 – beta casein; CSN1S1 – alpha S1 casein; CSN1S2 – alpha S2 casein; CSN3 – kappa casein; PIGR – polymeric immunoglobulin receptor; GLYCAM1 – glycosylation dependent cell adhesion molecule 1; WASHC1 – WASH complex subunit 1; WASF1 – wiskott-aldrich syndrome protein family member 1; SETD1A – histone-lysine N-methyltransferase; WHRN – whirlin; LGB – beta lactoglobulin; ALB – serum albumin.



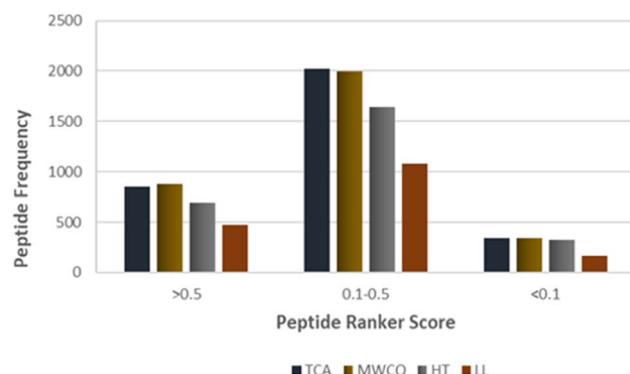


Fig. 5 PeptideRanker score of identified sequences against the frequency of peptides identified across TCA, MWCO, HT and LL extraction methods.

underscore the potential bioactivity inherent in the identified sequences across employed methods.

Peptides sequences were screened in Milk Bioactive Peptide Database (MBPDB) to analyse bioactivities already reported

elsewhere. Table 1 below shows MBPDB database analysis results of peptide sequences extracted through all the applied methods.

Discussion

We realized that this is not the first study to compare the methods for EPs extraction. Recently, Dingess *et al.*⁴¹ compared different methods for extraction of EPs in human milk. In another study, four different protocols were compared for the extraction and purification of short endogenous plasma peptides.⁵ In a similar study EPs extraction method including ACN precipitation, ultrafiltration and size exclusion chromatography were evaluated from human plasma.¹ All of the previous studies investigating EPs extraction were conducted on either human or bovine milk or plasma or urine.^{1,42} Each biological fluid including milk, colostrum or serum is quite different in the composition as well as physical properties. A study conducted on the physical properties of colostrum and milk observed the decreased concentration of casein and IgG, change in colour from yellow to white, change in pH as the

Table 1 Table reporting the bioactive functionality of identified sequences retrieved from MBPDB and peptide extracted from the extraction method used in this study

Peptide sequence	Protein name	Protein ID	Species	Bioactivity	Method			
					TCA	MWCO	HT	LL
KVLPVPQK	P02666	Beta-casein	Bos taurus	Antioxidant, immunomodulatory, anti-inflammatory	✓	✓	✓	✓
EPVLGPVRGP	P02666	Beta-casein	Bos taurus	Cytomodulatory	✓	✓	✓	✓
VLGPVRGPF	P02666	Beta-casein	Bos taurus	ACE-inhibitory	✓	✓	✓	✓
EPVLGPVRGPFP	P02666	Beta-casein	Bos taurus	ACE-inhibitory	✓	✓	✓	✓
SQSKVLPVPQKAVPYPQ	P02666	Beta-casein	Bos taurus	Antioxidant	✓	✓	✓	✓
TPVVVPPFL	A0A344 × 7B9	Beta-casein	Bos mutus	Anticancer	✓	✓	✓	✓
KVLPVPQ	P02666	Beta-casein	Bos taurus	ACE-inhibitory, immunomodulatory, anti-inflammatory	✓	✓	✓	✓
RPKHPIKHQ	P02662	Alpha-S1-casein	Bos taurus	ACE-inhibitory	✓	✓	✓	
RPKHPIK	P02662	Alpha-S1-casein	Bos taurus	Antimicrobial	✓			✓
MPPPKYPVEP	P02666	Beta-casein	Bos taurus	ACE-inhibitory	✓			✓
ELNVPGEIVES	P02666	Beta-casein	Bos taurus	Antimicrobial	✓		✓	✓
SQSKVLPVPQ	P02666	Beta-casein	Bos taurus	ACE-inhibitory	✓		✓	✓
SKVLPVPQ	P02666	Beta-casein	Bos taurus	ACE-inhibitory	✓			✓
LPQNIPPLT	P02666	Beta-casein	Bos taurus	DPP-IV inhibitory			✓	✓
NLHLPLPLL	P02666	Beta-casein	Bos taurus	ACE-inhibitory	✓		✓	
VPYPQRDMPIQAFL	P02666	Beta-casein	Bos taurus	Antimicrobial	✓		✓	
YPFPGPPIPNS	P02666	Beta-casein	Bos taurus	Antianxiety	✓		✓	
LVYPFPGPPIPNSLPQN	P02666	Beta-casein	Bos taurus	ACE-inhibitory	✓		✓	
YQEPVLGPVRGPFIIV	P02666	Beta-casein	Bos taurus	ACE-inhibitory, anticancer, antimicrobial, antithrombotic, immunomodulatory	✓		✓	✓
LYQEPVLGPVRGPFIIV	P02666	Beta-casein	Bos taurus	Immunomodulatory	✓		✓	✓
LLYQEPVLGPVRGPFIIV	P02666	Beta-casein	Bos taurus	ACE-inhibitory	✓		✓	✓
TPVVVPPFLQP	P02666	Beta-casein	Bos taurus	ACE-inhibitory	✓			✓
QEPVLGPVRGPFIIV	P02666	Beta-casein	Bos taurus	ACE-inhibitory	✓			✓
YQEPVLGPVR	P02666	Beta-casein	Bos taurus	ACE-inhibitory, antioxidant, antithrombotic, Immunomodulatory	✓			
ENLHLPLPLL	P02666	Beta-casein	Bos taurus	ACE-inhibitory			✓	
EPVLGPVRGPFIIV	P02666	Beta-casein	Bos taurus	ACE-inhibitory				✓
VENLHLPLPLL	P02666	Beta-casein	Bos taurus	ACE-inhibitory				✓
DVENLHLPLPLL	P02666	Beta-casein	Bos taurus	Antimicrobial				✓



colostrum transits into milk.³¹ Furthermore, the peptidome profiling of colostrum poses significant challenges due to its high viscosity compared to mature milk. This viscosity, along with other factors, hinders the development of a reliable methodology for extracting low molecular weight EPs. The major technical hurdle in the extraction and analysis of low molecular weight EPs is that the majority of the colostrum proteome and peptidome is covered by the abundant proteins and the peptides present in the sample. These dominant proteins and peptides suppress the signal of low-abundant and low molecular weight peptides during MS analysis and consequently, hampers the sensitivity of detection. So, peptide extraction workflow adapted for milk, serum or urine cannot be directly applied to colostrum making it necessary to explore the efficiency of these methods on colostrum. In a previous study, Ning *et al.* extracted EPs from colostrum using TCA precipitation method and for this, the colostrum was diluted fourfold with ultrapure water. Although over 10 000 peptides were identified in this study, they originated from approximately 1100 unique parent proteins.⁴³ In the present study, although the number of peptides identified was 3211, the source proteins of these peptides were over 2500. Interestingly, 47.82% of EPs in Ning *et al.*'s study were from casein, while in the present study, casein accounted for only 4.26% in the TCA method and 6.69% in the MWCO method. This difference is likely due to the enhanced identification of peptides from low-abundance proteins. This suggests that while diluting colostrum effectively reduces its viscosity, it does not adequately address the challenge of detecting low-abundance proteins, which remain masked by highly-abundant proteins. Consequently, the sensitivity for detecting low-abundance proteins and peptides is compromised due to the presence of dominant proteins and peptides.

To achieve comprehensive peptidome and proteome coverage, combining ultracentrifugation with peptide extraction methods, is more effective for colostrum. This approach is supported by previous studies demonstrating that extensive proteome coverage, particularly for low-abundance proteins, was achieved in colostrum through ultracentrifugation. High-speed centrifugation was shown to significantly enhance the detection of low-abundance proteins, thereby contributing to a more exhaustive proteomic analysis.⁴⁴

Another bottleneck in the investigation of the low molecular weight EPs is that EPs are submerged in complex biological fluid which contains wide variety of compounds including lipids, carbohydrates, fats and salts. Also, EPs have huge heterogeneity in their size, abundance, charge and chemical composition. All these factors influence the extraction as well as detection of peptides. Therefore, it requires a unique workflow for the extraction of peptides.^{4,45}

Very few studies have examined the EPs profile of colostrum, and all of these studies are based on any particular method for the extraction of EPs. Laboratory and pilot scale peptide extraction was performed using ultrafiltration method for comprehensive peptide analysis from the whey permeate. They identified a total of 212 and 238 unique peptides in both the methods respectively and all of the peptides were

predominantly casein derived.⁸ In another study Jorgensen and co-workers isolated putative bioactive peptides from non-digested colostrum and their immune modulation response was assessed against some bacterial ligands such as lipopolysaccharides and peptidoglycans.⁷

The implementation of proficient and specialized sample preparation is imperative for the successful conduct of peptidomic studies. Unlike tryptic peptides, EPs lack specific C-terminal residues such as K and R, complicating their analysis. Therefore, developing an effective method for their efficient extraction is necessary. To the best of our knowledge there is no study conducted on the comparative analysis of EPs extraction methods from colostrum.

In our study we thoroughly studied various extraction methods for the efficient extraction of EPs from colostrum. Although all the methods identified more than thousands peptide sequences but TCA and MWCO identified highest number of peptide sequences. Comparatively less number of identification in LL extraction can be the co-precipitation of peptides (Fig. 2A). Dingess and coworkers analysed three different methods to identify endogenous peptides from human milk which were TCA precipitation, MWCO and LL extraction and were able to identify maximum of 3237 peptides using TCA precipitation. Our study aims to develop an efficient method for extraction of EPs from bovine colostrum which has different chemical composition than human milk and in the present study we identified 3211 peptides using TCA precipitation and 3210 unique peptides using MWCO method. Peptides identified using MWCO methods were higher in number than previous studies where this same method was used to identify EPs. The less number of identification using the molecular weight cutoff method in previous studies could be due to clogging of the filter membrane which occurs due to the high molecular weight proteins present in the colostrum which hinders the subsequent filtration of peptides through the filter membrane. Also, a study speculated the hindrance in peptide extraction using 10 kDa filters can be due to adsorption of peptides on filter membrane *via* electrostatic interaction.⁴⁶ Therefore, we used ACN in combination with MWCO for extraction of EPs. ACN helps to dissociate the interaction of low abundant proteins and peptides which remain bound to the carrier proteins so these small and low abundant proteins and peptides become available to be detected by LC-MS/MS.³ Additionally, this huge number of exclusive identification may stem from a single amino acid residue difference at both the N and C-termini. Another reason for this could be the diversification of peptide sequences across methods which may contribute to this large number. Notably, heat treatment is implicated in inducing chemical reactions such as Maillard reactions and oxidation, leading to the emergence of post-translationally modified sequences.¹²

Our objective was to isolate low molecular weight peptides, a goal successfully realized through each employed EPs extraction method. All methods demonstrated a predominant presence of peptides within the mass range of 700–2000 Da and amino acid residues in sequences between 7 and 20 amino acids mostly (Fig. 3C and D). Furthermore, our study aligns with



prior research focused on extracting low molecular weight peptides, confirming the consistency of our findings with existing literature.^{23,41}

Parent proteins analysis of identified sequences revealed that some major proteins were consistently enriched through all the methods including beta casein, alpha S1 casein, kappa casein, GlyCAM, WASH complex, WASL, polymeric immunoglobulin receptor, whirlin protein, histone lysine-N-methyltransferase, proline rich coiled coil and some uncharacterized proteins (Fig. 4). It was noteworthy that no high confidence peptide was obtained from alpha S2 in TCA, heat treated and LL extraction methods, although peptides were obtained from Alpha S2 through MWCO method. Beta lactoglobulin protein contributed peptides in TCA, MWCO and heat treated method but no peptide was observed in LL extracted method. In line with the previous studies, no EPs were observed from alpha lactalbumin.¹⁷ Other major proteins such as perilipin and butyrophilin, reported in prior research, yielded few high-confidence peptides in our study. Peptides from butyrophilin were exclusively observed in the heat-treated method and peptide originating from perilipin were identified in MWCO and HT methods only. It is noteworthy that EPs from proteins such as whirlin, WASH complex, and WASL were not previously reported in the milk or colostrum of any species, expanding our understanding of the peptide composition in these protein complexes.

An examination of the GRAVY index, molecular weight, amino acid length, and other physico-chemical properties of the identified peptides revealed that all methods enriched comparable peptide types. This indicates no discernible distinctions in the chemical-physical attributes of the identified peptides (Fig. 3A–D).

Functional annotation of the sequences indicated their potential bioactivity. All methods successfully extracted peptides previously obtained from cow or yak milk or colostrum including those with demonstrated antihypertensive, anti-oxidative, anti-inflammatory, and immunomodulatory effects (Table 1). Although, there were slight variations in the N or C-terminus or sequence length of these peptides across different extraction methods as shown in Table 1, all similarly extracted peptides exhibited proven bioactivity.

Some key endogenous peptides have been previously isolated from colostrum. For instance, casecidin 15 (YQEPVLGPVRGPFPPI) and casecidin 17 (YQEPVLGPVRGPFPPIIV) from bovine colostrum have demonstrated antimicrobial properties. Another peptide, isracidine (RPKHPIKHQGLP-QEVLNENLLRF), also shows antimicrobial activity and has been isolated from bovine colostrum.³⁴ In the present study, either the full-length sequence or a fragment within these sequences (casecidin 15 or casecidin 17 or isracidine) was isolated using all methods. Specifically, casecidin 17 was isolated through MWCO, HT, and liquid–liquid extraction, but not through TCA. However, fragments of this peptide was present in the TCA based extraction method, which also has proven biological activities. Additionally, the fragment YQEPVLGPVR, found within the sequences of casecidin 15 and 17, was observed only through the MWCO method. Previous studies have

demonstrated its anti-inflammatory, antioxidative, and immunomodulatory potential.⁴⁷ Furthermore, another peptide, KVLPVPQ, with proven ACE-inhibitory, immunomodulatory, and anti-inflammatory properties,^{48,49} was consistently isolated through all extraction methods. Overall, most methods were successful in isolating peptides with bioactive potential, with the MWCO method being particularly effective in extracting the exact sequence of casecidin 17 and its other fragments having proven biological activities. Therefore, this study confirms the bioactive potential of peptides extracted from colostrum, with consistent results across different extraction methods despite slight variations in peptide sequences.

These findings collectively underscore the potential bioactivity inherent in the identified sequences across employed methods.

Given the similarity in the physicochemical characteristics—such as amino acid profile, sequence length, and molecular weight—of the peptides identified *via* different extraction methods, it is challenging to determine the most effective extraction technique for extraction of EPs from colostrum. The bioactive potential of the peptides, assessed through PeptideRanker and validated using MBPDB, further complicates direct comparison of the efficacy of these methods since all the method successfully extracted peptides with bioactive peptides.

To evaluate the efficiency of the applied extraction techniques, we considered two primary metrics: the number of identified sequences and the number of proteins covered by each method for extracting EPs. Our findings revealed that TCA precipitation and MWCO methods yielded the highest number of peptides, derived from a broad range of parent proteins, particularly from low abundant proteins. Conversely, the other two methods evaluated demonstrated lower peptide yields. These results underscore the importance of employing TCA precipitation or MWCO methods. Thus, we focused our comparative analysis on TCA and MWCO methods. Several measures were taken into consideration to assess efficiency of these methods (Table S1†): 1. Number of peptides identified and number of unique peptides in both methods, 2. Percentage peak area occupied by total peptides, shared peptides, and key peptides reported in literature, 3. Number of parent proteins, 4. Percentage of peptides scoring more than 0.5 or more than 0.1 in PeptideRanker, 5. Number of key endogenous peptides identified, 6. Amino acid composition and 7. The major proteins from which the EPs have already been isolated from milk and colostrum.

The number of peptides identified was nearly equal for both methods. However, TCA identified a higher number of unique peptides, as shown in Fig. 2B. Additionally, TCA precipitation method resulted in a higher percentage peak area for total identified peptides, peptides shared by both methods, and the number of parent proteins contributing to the extraction of EPs.

Sequence logos analysis of the first seven amino acids from the N-terminus revealed that leucine and proline were the most dominant amino acids in nearly all positions for both methods. Despite these strengths, the bioactive functionality analysis showed that MWCO identified more bioactive sequences compared to TCA. Specifically, 27.25% of sequences identified



by MWCO had a PeptideRanker score of 0.5 or higher, compared to 26.41% for TCA. Furthermore, 20 peptides identified through MWCO displayed bioactive potential when screened through the MBPDB, compared to 13 peptides for TCA.

Moreover, the percentage peak area occupied by key EPs identified by both methods was higher in the MWCO method. Another interesting metric advocating the use of the MWCO method is the observation that major proteins, such as perilipin and alpha S2 casein, yielded EPs only in MWCO and not in TCA.

Therefore, while the TCA precipitation method excels in identifying a large number of sequences and parent proteins, the MWCO method is superior for isolating bioactive EPs from almost all the major proteins and EPs suitable for therapeutic screening.

Conclusions

This research addresses the development of an efficient extraction procedure from the colostrum matrix, evaluating and comparing the efficacy of various methods for extracting EPs from colostrum. While previous studies have demonstrated EPs extraction from milk, serum, and other biological fluids, to our knowledge, no study has compared the extraction efficiency of methods specifically on colostrum. The novelty of this study lies in the extraction of bioactive EPs from colostrum using peptide extraction methods combined with ultracentrifugation to deplete the most abundant proteins (casein). Using ultracentrifugation in combination with the existing methods with some modifications enables the identification of less abundant proteins and peptides. In conclusion, we recommend the use of MWCO for the identification of peptides with bioactive properties due to its comparatively wider coverage of bioactive peptides in available database. However, the TCA method remains valuable for quantitative purposes and for identifying a high number of peptides from diverse parent proteins. Both methods have their unique strengths and can be chosen based on the specific goals of the research, ensuring a comprehensive approach to peptide extraction and analysis. Through our detailed evaluation and selection process, we have achieved the purpose of our research by identifying the most suitable technique for peptide extraction based on the desired outcomes. This work significantly contributes to bridging the gap in developing efficient extraction procedures for the colostrum matrix.

Data availability

The mass spectrometry proteomics data have been deposited with the ProteomeXchange consortium *via* the PRIDE⁵⁰ partner repository with the dataset identifier PXD051733 (PRIDE reference: 1-20240425-095929-3277975).

Author contributions

P. P.: conceptualization, design and conduct of experiments, data analysis, interpretation of results, investigation, data

visualization, writing – original draft preparation, review and editing; R. R.: writing – original draft preparation, review and editing; R. K.: design of experiments, writing – review and editing; S. M.: writing – review and editing; S. K.: sample collection, conceptualization, supervision, data analysis, writing – review and editing; M. M.: sample collection, suggestion in design of experiment, writing – review and editing, J. K. K., M. S. and A. K. M.: suggestion in design of experiment, writing – review and editing. All authors read and approved the final manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The author acknowledges the director of the National Dairy Research Institute for funding this work. The author also acknowledges the Department of Biotechnology for providing fellowship. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium *via* the PRIDE⁵⁰ partner repository with the dataset identifier PXD051733 (PRIDE reference: 1-20240425-095929-3277975).

References

- 1 L. P. Aristoteli, M. P. Molloy and M. S. Baker, Evaluation of endogenous plasma peptide extraction methods for mass spectrometric biomarker discovery, *J. Proteome Res.*, 2007, **6**(2), 571–581.
- 2 P. Ferranti, M. V. Traisci, G. Picariello, A. Nasi, V. Boschi, M. Siervo, C. Falconi, L. Chianese and F. Addeo, Casein proteolysis in human milk: tracing the pattern of casein breakdown and the formation of potential bioactive peptides, *J. Dairy Res.*, 2004, **71**(1), 74–87.
- 3 K. Merrell, K. Southwick, S. W. Graves, M. S. Esplin, N. E. Lewis and C. D. Thulin, Analysis of low-abundance, low-molecular-weight serum proteins using mass spectrometry, *J. Biomol. Tech.*, 2004, **15**(4), 238.
- 4 B. J. Kim and D. C. Dallas, Systematic examination of protein extraction, proteolytic glycopeptide enrichment and MS/MS fragmentation techniques for site-specific profiling of human milk N-glycoproteins, *Talanta*, 2021, **224**, 121811.
- 5 S. Piovesana, A. Cerrato, M. Antonelli, B. Benedetti, A. L. Capriotti, C. Cavaliere, C. M. Montone and A. Lagana, A clean-up strategy for identification of circulating endogenous short peptides in human plasma by zwitterionic hydrophilic liquid chromatography and untargeted peptidomics identification, *J. Chromatogr. A*, 2020, **1613**, 460699.
- 6 R. Kumar, S. A. Ali, S. K. Singh, V. Bhushan, J. K. Kaushik, A. K. Mohanty and S. Kumar, Peptide profiling in cow urine reveals molecular signature of physiology-driven pathways and in-silico predicted bioactive properties, *Sci. Rep.*, 2021, **11**(1), 12427.



7 A. L. Jørgensen, H. R. Juul-Madsen and J. Stagsted, Colostrum and bioactive, colostral peptides differentially modulate the innate immune response of intestinal epithelial cells, *J. Pept. Sci.*, 2010, **16**(1), 21–30.

8 D. C. Dallas, V. Weinborn, J. M. de Moura Bell, M. Wang, E. A. Parker, A. Guerrero, K. A. Hettinga, C. B. Lebrilla, J. B. German and D. Barile, Comprehensive peptidomic and glycomic evaluation reveals that sweet whey permeate from colostrum is a source of milk protein-derived peptides and oligosaccharides, *Food Res. Int.*, 2014, **63**, 203–209.

9 C. M. Montone, A. L. Capriotti, A. Cerrato, M. Antonelli, G. La Barbera, S. Piovesana, A. Laganà and C. Cavaliere, Identification of bioactive short peptides in cow milk by high-performance liquid chromatography on C18 and porous graphitic carbon coupled to high-resolution mass spectrometry, *Anal. Bioanal. Chem.*, 2019, **411**, 3395–3404.

10 L. Zhang, B. Han, B. Luo, Y. Ni, N. Bansal and P. Zhou, Characterization of endogenous peptides from dromedary and bactrian camel milk, *Eur. Food Res. Technol.*, 2022, **248**(4), 1149–1160.

11 C. M. Montone, S. E. Aita, C. Cavaliere, A. Cerrato, A. Laganà, S. Piovesana and A. L. Capriotti, High-resolution mass spectrometry and chemometrics for the detailed characterization of short endogenous peptides in milk by-products, *Molecules*, 2021, **26**(21), 6472.

12 M. Wölk, S. Milkovska-Stamenova and R. Hoffmann, Comprehensive profiling of the native and modified peptidomes of raw bovine milk and processed milk products, *Foods*, 2020, **9**(12), 1841.

13 J. Zhu, K. A. Dingess, M. Mank, B. Stahl and A. J. Heck, Personalized profiling reveals donor-and lactation-specific trends in the human milk proteome and peptidome, *J. Nutr.*, 2021, **151**(4), 826–839.

14 X. Wang, Y. Sun, F. Wang, L. You, Y. Cao, R. Tang, J. Wen and X. Cui, A novel endogenous antimicrobial peptide CAMP 211-225 derived from casein in human milk, *Food Funct.*, 2020, **11**(3), 2291–2298.

15 A. K. Enjapoori, S. Kukuljan, K. M. Dwyer and J. A. Sharp, In vivo endogenous proteolysis yielding beta-casein derived bioactive beta-casomorphin peptides in human breast milk for infant nutrition, *Nutrition*, 2019, **57**, 259–267.

16 K. A. Dingess, I. Gazi, H. W. van den Toorn, M. Mank, B. Stahl, K. R. Reiding and A. J. Heck, Monitoring human milk β -casein phosphorylation and O-glycosylation over lactation reveals distinct differences between the proteome and endogenous Peptidome, *Int. J. Mol. Sci.*, 2021, **22**(15), 8140.

17 D. C. Dallas, A. Guerrero, E. A. Parker, L. A. Garay, A. Bhandari, C. B. Lebrilla, D. Barile and J. B. German, Peptidomic profile of milk of Holstein cows at peak lactation, *J. Agric. Food Chem.*, 2014, **62**(1), 58–65.

18 M. Bhattacharya, J. Salcedo, R. C. Robinson, B. M. Henrick and D. Barile, Peptidomic and glycomic profiling of commercial dairy products: identification, quantification and potential bioactivities, *NPJ Sci. Food*, 2019, **3**(1), 4.

19 D. C. Dallas, A. Guerrero, N. Khaldi, P. A. Castillo, W. F. Martin, J. T. Smilowitz, C. L. Bevins, D. Barile, J. B. German and C. B. Lebrilla, Extensive *in vivo* human milk peptidomics reveals specific proteolysis yielding protective antimicrobial peptides, *J. Proteome Res.*, 2013, **12**(5), 2295–2304.

20 S. D. Nielsen, R. L. Beverly and D. C. Dallas, Peptides released from foremilk and hindmilk proteins by breast milk proteases are highly similar, *Front. Nutr.*, 2017, **4**, 54.

21 A. Guerrero, D. C. Dallas, S. Contreras, A. Bhandari, A. Cánovas, A. Islas-Trejo, J. F. Medrano, E. A. Parker, M. Wang, K. Hettinga and S. Chee, Peptidomic analysis of healthy and subclinically mastitic bovine milk, *Int. Dairy J.*, 2015, **46**, 46–52.

22 W. Yu, Y. Yu, W. Wang, Y. Li, I. M. Szeto and Y. Jin, Effect of sample preparation on analysis of human milk endogenous peptides using liquid chromatography-tandem mass spectrometry, *Sepu*, 2021, **39**(5), 463–471.

23 S. Piovesana, A. L. Capriotti, C. Cavaliere, G. La Barbera, R. Samperi, R. Z. Chiozzi and A. Laganà, Peptidome characterization and bioactivity analysis of donkey milk, *J. Proteom.*, 2015, **119**, 21–29.

24 A. Nissen, P. H. Andersen, E. Bendixen, K. L. Ingvarstsen and C. M. Røntved, Colostrum and milk protein rankings and ratios of importance to neonatal calf health using a proteomics approach, *J. Dairy Sci.*, 2017, **100**(4), 2711–2728.

25 M. Abdel-Hamid, P. Yang, I. Mostafa, A. Osman, E. Romeih, Y. Yang, Z. Huang, A. A. Awad and L. Li, Changes in whey proteome between Mediterranean and Murrah buffalo colostrum and mature milk reflect their pharmaceutical and medicinal value, *Molecules*, 2022, **27**(5), 1575.

26 R. Kashyap, K. S. Narayan and S. Vij, Identification of antibacterial and immunomodulatory bioactive peptides generated from buffalo (*Bubalus bubalis*) colostrum whey fermented by *Lactobacillus rhamnosus* C25: LC-MS/MS-based analysis, *J. Funct. Foods*, 2022, **95**, 105158.

27 F. S. Fajardo-Espinoza, C. Ordaz-Pichardo, U. Sankar, A. Romero-Rojas, M. A. Moreno-Eutimio and H. Hernández-Sánchez, In vitro cytomodulatory and immunomodulatory effects of bovine colostrum whey protein hydrolysates, *Int. J. Food Sci. Technol.*, 2021, **56**(5), 2109–2121.

28 A. Chae, A. Aitchison, A. S. Day and J. I. Keenan, Bovine colostrum demonstrates anti-inflammatory and antibacterial activity in *in vitro* models of intestinal inflammation and infection, *J. Funct. Foods*, 2017, **28**, 293–298.

29 T. K. Kh, Peptides of trypsin hydrolyzate in bovine colostrum, *Food Process. Tech. Technol.*, 2023, **53**(1), 150–158.

30 N. R. Ashok, K. H. Vivek and H. S. Aparna, Antioxidative role of buffalo (*Bubalus bubalis*) colostrum whey derived peptides during oxidative damage, *Int. J. Pept. Res. Ther.*, 2019, **25**, 1501–1508.

31 B. D. Madsen, M. D. Rasmussen, M. O. Nielsen, L. Wiking and L. B. Larsen, Physical properties of mammary secretions in relation to chemical changes during



transition from colostrum to milk, *J. Dairy Res.*, 2004, **71**(3), 263–272.

32 H. Dziewiecka, H. S. Buttar, A. Kasperska, J. Ostapiuk-Karolczuk, M. Domagalska, J. Cichoń and A. Skarpańska-Stejnborn, A systematic review of the influence of bovine colostrum supplementation on leaky gut syndrome in athletes: diagnostic biomarkers and future directions, *Nutrients*, 2022, **14**(12), 2512.

33 I. B. Campanhon, P. F. de Aguiar, F. F. Bezerra, M. R. Soares and A. G. Torres, Human colostrum *in vitro* protein digestion: peptidomics by LC-Orbitrap-HRMS and prospection for bioactive peptides *via* bioinformatics, *Br. J. Nutr.*, 2023, **24**, 1–25.

34 G. A. Birkemo, O. O'sullivan, R. P. Ross and C. Hill, Antimicrobial activity of two peptides casecidin 15 and 17, found naturally in bovine colostrum, *J. Appl. Microbiol.*, 2009, **106**(1), 233–240.

35 Y. Zhou, L. Zhang, Z. Yu, A. Zhang, W. Wu, W. Chen, X. Yan, H. Liu, Y. Hu, C. Jiang and Y. Xu, Peptidomic analysis reveals multiple protection of human breast milk on infants during different stages, *J. Cell. Physiol.*, 2019, **234**(9), 15510–15526.

36 X. Cui, Y. Li, L. Yang, L. You, X. Wang, C. Shi, C. Ji and X. Guo, Peptidome analysis of human milk from women delivering macrosomic fetuses reveals multiple means of protection for infants, *Oncotarget*, 2016, **7**(39), 63514.

37 J. C. Oliveros, *VENNY. An interactive tool for comparing lists with Venn Diagrams*. 2007. <https://bioinfogp.cnb.csic.es/tools/venny/index.html>.

38 D. Tang, M. Chen, X. Huang, G. Zhang, L. Zeng, G. Zhang, S. Wu and Y. Wang, SRplot: a free online platform for data visualization and graphing, *PLoS ONE*, 2023, **18**(11), e0294236.

39 M. Spitzer, J. Wildenhain, J. Rappaport and M. Tyers, BoxPlotR: a web tool for generation of box plots, *Nat. Methods*, 2014, **11**(2), 121–122.

40 A. Sharma, P. Kapoor, A. Gautam, K. Chaudhary, R. Kumar, J. S. Chauhan, A. Tyagi and G. P. Raghava, Computational approach for designing tumor homing peptides, *Sci. Rep.*, 2013, **3**(1), 1607.

41 K. A. Dingess, H. W. van den Toorn, M. Mank, B. Stahl and A. J. Heck, Toward an efficient workflow for the analysis of the human milk peptidome, *Anal. Bioanal. Chem.*, 2019, **411**, 1351–1363.

42 A. Guerrero, D. C. Dallas, S. Contreras, S. Chee, E. A. Parker, X. Sun, L. Dimapasoc, D. Barile, J. B. German and C. B. Lebrilla, Mechanistic peptidomics: factors that dictate specificity in the formation of endogenous peptides in human milk, *Mol. Cell. Proteom.*, 2014, **13**(12), 3343–3351.

43 J. Ning, M. Yang, Q. Zhu, X. Luo and X. Yue, Peptidomics comparison of endogenous peptides derived from parent proteins in bovine colostrum and mature milk, *LWT*, 2024, **191**, 115648.

44 A. Nissen, E. Bendixen, K. L. Ingvarsson and C. M. Røntved, In-depth analysis of low abundant proteins in bovine colostrum using different fractionation techniques, *Proteomics*, 2012, **12**(18), 2866–2878.

45 J. Peng, H. Zhang and H. Niu, Peptidomic analyses: the progress in enrichment and identification of endogenous peptides, *TrAC, Trends Anal. Chem.*, 2020, **125**, 115835.

46 Y. Du, D. Wu, Q. Wu and Y. Guan, Quantitative evaluation of peptide-extraction methods by HPLC-triple-quad MS-MS, *Anal. Bioanal. Chem.*, 2015, **407**, 1595–1605.

47 K. Sowmya, M. I. Bhat, R. K. Bajaj, S. Kapila and R. Kapila, Buffalo milk casein derived decapeptide (YQEPVLPVPR) having bifunctional anti-inflammatory and antioxidative features under cellular milieu, *Int. J. Pept. Res. Ther.*, 2019, **25**, 623–633.

48 C. Adams, F. Sawh, J. M. Green-Johnson, H. J. Taggart and J. L. Strap, Characterization of casein-derived peptide bioactivity: differential effects on angiotensin-converting enzyme inhibition and cytokine and nitric oxide production, *J. Dairy Sci.*, 2020, **103**(7), 5805–5815.

49 M. Maeno, N. Yamamoto and T. Takano, Identification of an antihypertensive peptide from casein hydrolysate produced by a proteinase from *Lactobacillus helveticus* CP790, *J. Dairy Sci.*, 1996, **79**(8), 1316–1321.

50 Y. Perez-Riverol, J. Bai, C. Bandla, D. García-Seisdedos, S. Hewapathirana, S. Kamatchinathan, D. J. Kundu, A. Prakash, A. Frericks-Zipper, M. Eisenacher and M. Walzer, The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences, *Nucleic Acids Res.*, 2022, **50**(D1), D543–D552.

