Current Research in Food Science 3 (2020) 217–226

Contents lists available at ScienceDirect

Current Research in Food Science

journal homepage: www.editorialmanager.com/crfs/

Research Paper

Human milk and infant formulae: Peptide differences and the opportunity to address the functional gap

Cyril Lopez, Alessandro Adelfio, Audrey M. Wall *, Brendan Molloy, Thérèse A. Holton, Nora Khaldi

Nuritas Ltd, Joshua Dawson House, Dublin 2, D02 RY95, Ireland

ARTICLE INFO

Keywords:
Human milk (HM)
Formula milk (FM)
Peptide
Peptidomics
Bioactivity
Protein
Casein
LC-MS/MS

ABSTRACT

Bovine-derived formula milk (FM) is a common substitute to human milk (HM), but lacks key functional benefits associated with HM. Accordingly, there have been significant efforts to humanise FM. Recent research has demonstrated that HM-derived peptides convey an array of beneficial bioactivities. Given that peptides serve as important signalling molecules offering high specificity and potency, they represent a prime opportunity to humanise FM. To further understand how HM-derived peptides contribute to infant health, we used peptidomics and bioinformatics to compare the peptide profile of HM to commercially available FM. We found clear and substantial differences between HM and FM in terms of peptide physicochemical properties, protein coverage and abundance. We additionally identified 618 peptides specific to HM that represent an important untapped source to be explored for novel bioactivities. While further study is required, our findings highlight the potential of a peptide-based approach to address the functional gap in FM.

1. Introduction

In infant nutrition, breastfeeding for 6 months up to 2 years is considered the gold standard (WHO, 2003). Human Milk (HM) provides the optimal nutrients for healthy immune development and growth, and is associated with functional benefits which include reduced risk of infections, lower rates of asthma and diabetes mellitus, protection of preterm infants against necrotizing enterocolitis and improved cognitive development (Stevens et al., 2009; Lucas et al., 1998; Hays and Ra, 2005; Moossavi et al., 2018; Herrmann and Carroll, 2014). HM is comprised of macronutrients such as proteins, lipids, non-conjugated complex carbohydrates and an array of micronutrients such as vitamins A, C, B6, B12, and iodine. The protein content of HM consists of three major groups, namely mucins, caseins and whey, which are the main source of essential amino acids required for infant development (Khaldi et al., 2014). While the protein profile of HM is known to vary during lactation stages, it is generally considered to have a higher whey:casein ratio than bovine milk (Raikos and Dassios, 2014). Major HM proteins include α_{s1}-casein (α_{s1}-CN), β-casein (β-CN), κ-casein (κ-CN), α-Lactalbumin (α-LA), lactoferrin (LF), osteopontin (OPN) and serum albumin (SA), which convey a range of health benefits to infants such as immunomodulation, anti-inflammatory effects, antimicrobial effects, the development of a beneficial microbiome and increased absorption of iron and zinc (Lönnerdal, 2014). Moreover, in the past number of decades it has been established that functional peptides derived from these major proteins can additionally confer a wide range important benefits (Chatterton et al., 2013; Nielsen et al., 2017).

Advances in scientific research and the introduction of regulatory requirements have helped shape the content and functional benefits of FM. Bovine derived FM was first commercially developed as a substitute to HM in 1865 (Stevens et al., 2009). Early editions of infant formula consisted mainly of carbohydrates and fats but lacked essential vitamins and proteins (Fig. S1). Over time, changes were made to address the gaps within the nutritional composition of FM. Examples of this include the addition of micronutrients such as vitamin B6 and iron to alleviate observed deficiencies of these important factors in infants (Fomon, 2001). The 1960s saw FM milk adopt a casein:whey ratio similar to HM, while the 1980s saw the introduction of regulatory requirements such as quality control and standardisation of nutrient levels of FM (Fomon, 2001). To address allergies in infants, isolated soy-based formula was developed in the 1960s and hydrolysis of FM was adopted in the 1980s, as exposure to bovine milk is thought to trigger the onset of oral sensitivity to food antigens (Boyleet al., 2016). Interestingly, however, recent research has found no consistent evidence for the utility of

* Corresponding author.
E-mail address: info@nuritas.com (A.M. Wall).

https://doi.org/10.1016/j.crfs.2020.07.003

hydrolysed formulae in allergy prevention when compared to HM (Boyleet al., 2016).

Striving to close the functional gap between HM and FM, supplementation has been the main innovation to FM. For instance, strategies to improve immune development have included FM supplementation with oligosaccharides, such as galacto-oligosaccharides (GOS), fructo-oligosaccharides (FOS) and more recently HM oligosaccharides (HMO) (Vandenplas et al., 2014). While GOS and FOS have been shown to decrease the incidence of infection and boost immune development, they fail to attain the level of protection provided by HM (Vandenplas et al., 2014). HMO's are an important and diverse group of indigestible sugars, both structurally and biologically, which are the third most abundant component of HM (Wiciński et al., 2020). HMO's within HM confer numerous activities such as anti-viral/bacterial and the development of the immune system and intestinal microflora (Wiciński et al., 2020). The addition of either 2′-fucosyllactose (2′FL) and lacto-N-neotetraose (LNnT) HMOs to FM has demonstrated a promising prebiotic effect (Chichlowski et al., 2011), and whilst encouraging, more research is required to determine if HMOs provide greater benefits than those of GOS and FOS (Vandenplas, 2017). Additional supplementation strategies include long-chain polyunsaturated fatty acids (LCPUFAs), such as docosahexaenoic acid (DHA), which have been added to FM due to their reported roles in allergy prevention, promoting cognitive function and visual acuity (McCann and Ames, 2005; Qawasmi et al., 2013). Despite their inclusion in FM since 2002, the benefits of LCPUFA supplementation in full term and preterm infants are still being investigated (McCann and Ames, 2005; Qawasmi et al., 2013).

There is increasing evidence highlighting the importance of the peptide component in HM (Lönnerdal, 2014; Dallas et al., 2014; Nielsen et al., 2017). Peptides are an important, and indeed the largest, class of signalling molecules in animals and plants, contributing to a vast array of functions. Peptides exhibit high specificity resulting in greater potency (Craik et al., 2013). Functional peptides derived from protein complexes of HM convey a number of important health benefits including anti-microbial, immunomodulatory, anti-inflammatory, anti-oxidant, dipeptidyl peptidase IV inhibitory, opioid, anti-hypertensive and nutrient uptake effects (Chatterton et al., 2013; Nielsen et al., 2017a). With advances in mass spectrometry and bioinformatics, it is now feasible to investigate distinct peptides in HM and FM (Lambers et al., 2015; Holton et al., 2016; Dallas, 2013; Khaldi et al., 2014; Manguy et al., 2017). Indeed recent investigations into the peptide profiles and associated bioactivity of HM at different lactation stages are providing important insights into the peptide components of HM (Nielsen et al., 2017; Dingess et al., 2019). However, there remains a need for side by side comparisons of the endogenous peptide profiles within HM and FM.

Accordingly, in this study we analysed undigested HM samples to profile endogenous HM peptides and compared them to the peptide profiles found in commercially available FM. Using LC-MS/MS in combination with bioinformatic tools, we characterised and compared the distinct properties of peptides found in HM and FM. By increasing our understanding of FM peptide profiles relative to those in HM, we may be able to unlock the vast array of health benefits conferred by HM peptides. In particular, the identification of peptides specific to HM that are conserved across mothers are an important source of potential functional activities that are not captured in FM.

2. Methods

2.1. Samples

HM samples were donated by 3 healthy, volunteer mothers (Dublin, Ireland) who delivered at term. All samples were from a period of 8–25 weeks postpartum, with each mother providing at least 3 independent samples from within this time range. Following the approach of Dallas et al. (2014), samples were collected from milk expressed by volunteers at home via breast milk pump. Samples were stored immediately in home freezers before transport on dry ice for storage at −80 °C until experimental preparation. All mothers provided informed written consent, in accordance with ethical standards.

Two commercially available brands of FM were purchased locally (Dublin, Ireland). Three samples of each brand were sourced, each originating from independent production batches, resulting in a total of 6 FM samples. Both brands are tailored for full-term infants aged 0–26 weeks. The nutritional information supplied by the respective manufacturers of each brand are displayed in supplementary material (Table S1).

2.2. Sample Preparation and mass spectrometry analysis

Samples (100 μL) were desalted and concentrated using 30-kDa Spin-X UF centrifugal concentrators (Corning Inc, Lowell, MA, USA). Each flow-through was then acidified with formic acid and cleaned of contaminants by solid phase extraction on an Empore 96-well disk plate with C18-SD sorbent (3 M, St. Paul, MN, USA). Eluates were lyophilized and resuspended in 0.1% trifluoroacetic acid for analysis.

Samples were analysed by nano LC-MS/MS with a Waters NanoACQUITY HPLC system (Waters Corporation, Milford, MA, USA) interfaced to a ThermoFisher Q Exactive (ThermoFisher Scientific Inc., Canoga Park, CA, USA). Peptides were loaded on a trapping column and eluted over a 75 μm analytical column with a 1 h gradient at a flow rate of 350 nL min^{-1}, no analytical replicates were used. Both columns were packed with Luna C18 resin (Phenomenex, Torrance, CA, USA). The mass spectrometer was operated in data-dependent mode, with MS and MS/MS performed in the Orbitrap at 70,000 FWHM and 17,500 FWHM resolution respectively. From the MS scan, the fifteen most intense ions were selected for MS/MS.

2.3. Peptide sequence searches

Raw data from the Q Exactive were processed by MaxQuant version 1.5.3.17 (Cox and Mann, 2008) with the Andromeda search engine. HM MS/MS spectra were searched against the Swiss-Prot Human database, while spectra from FM samples were searched against the Uniprot Bovine database. All searches were carried out in unspecific digestion mode with a minimum peptide length of 7 and a fragment mass tolerance of 20 ppm. A false discovery rate (FDR) of 0.01 was selected for both peptides and proteins. Database searches were performed with no fixed modifications and with oxidation (M) as a variable modification.

2.4. Bioinformatic data analysis

For all samples, we considered peptide sequences and their corresponding abundance profiles as returned by MaxQuant (Cox and Mann, 2008). In our data analyses, unique peptides across HM samples were firstly pooled by mother (to yield a total of 3 HM samples), while similarly, FM samples were pooled according to brand (to yield a total of 2 FM samples). Additionally, single representative HM and FM samples were generated by retaining unique peptides across all individual HM and FM samples, where any common peptides were represented only once. Peptide abundance for these representative samples was calculated by summing the peptide intensities in each individual sample, divided by the number of samples containing the observed peptide. N & C terminal abundance for the HM and FM representative samples was calculated as per Holton et al. (2016).

All data analyses were carried out using in-house Python scripts. Briefly, for the HM mother data and FM brand data, Venn diagrams were generated using the MatplotLib Venn package (https://pypi.python.org/pypi/matplotlib-venn), while graphs illustrating peptide physiochemical properties, peptide counts and clustering dendrograms were created using the 'ggplot2' R package (Wickham, 2009). Following the approach of Holton et al. (2016), heatmap representations of the HM and FM representative data were achieved using the heatmap.2 function of the 'gplots' R package, using Pearson correlation as the distance function

and complete linkage as the clustering method. To ensure robust grouping, a random outgroup sample (RS) was created, containing 3000 randomly generated peptides 6–30 amino acids (AA) in length, where AA composition was assigned according to the SwissProt ProtScale tool (https://web.expasy.org/protscale/pscale/A.A.Swiss-Prot.html). A pairwise correlation matrix was generated using the MatplotLib package.

For major milk proteins, multiple sequence alignment between bovine and human orthologs was performed using MUSCLE with default parameters (Edgar, 2004). Resultant alignments were used as the precursor parent proteins for peptide alignment mapping of the HM and FM representative data. Visualisations of such "positional homology peptide alignment maps" were generated using the 'ggplot2' R package (Wickham, 2009), where peptide abundance (indicated in green) was calculated according to Manguy et al. (2017).

To explore potential peptide bioactivity unique to our samples, we compiled a bioactivity database from the public repositories, Milk Bioactive Peptide Database (MBPDB, http://mbpdb.nws.oregonstate.edu/) and BIOPEP (Minkiewicz et al., 2008). A BLAST search was performed using human specific peptides common to all mothers (618) and specific peptides common to all FM samples (1006) against the combined MBPDB (558) and BIOPEP (3315) database. A bioactive peptide was retained if an exact match was obtained or if an e value less than 10 and $\geq$ 80% coverage of the query sequence was detected.

3. Results and discussion

3.1. Peptide distribution in FM and HM milk

We carried out LC-MS/MS on 11 samples of HM from 3 mothers. From the MS data, we determined the unique peptide count for each mother to be 1978 (M1), 1417 (M2) and 1795 (M3). Notably, here our unique peptide yield for each mother was higher than those reported in a recent LC-MS/MS study of HM by Nielsen et al. (2017) (Nielsen et al., 2017b), however, this may be accounted for by experimental variation in mass spectrometry methods. For example; our study used larger samples of HM (100 μ vs. 8–20 μL) for analysis; Nielsen et al. (2017) utilised an extensive sample preparation method which would result in a greater loss of peptides; different gradient lengths and the data base search parameters were also unique to each study (20 ppm vs. 10 ppm mass error). However, a greater sample size of HM and similar analytical methods would be required to confirm the observed differences. There were 637 common peptides found among our 3 mothers (Fig. 1A). The peptide profiles of all HM samples were merged to generate a representative HM peptide profile that contained a total of 3131 unique HM peptides (Fig. 1C).

We performed LC-MS/MS on two brands of FM, using three separate production batches for each. When data for each batch were merged, the total unique peptide count for each brand was 1634 (FM – B1) and 1464 (FM – B2) respectively, while a total of 1031 peptides were found common to both (Fig. 1B). Since FM is quite regulated and therefore standardised, a greater number of shared peptides in FM would be expected compared to in HM, which varies from mother to mother. The peptide profiles of all FM samples were also merged to generate a representative FM peptide profile containing 2067 individual FM peptides (Fig. 1C and D).

When we analysed the peptide overlap between our representative HM (combined peptide yield from each mother; 3131) and FM peptide profiles (combined peptide yield from each FM sample; 2067), we found that only 51 peptides were shared between HM and FM (Fig. 1C). In order to look at the differences between shared HM peptides and FM, we compared the representative FM peptide profile to human-specific peptides common to all three mothers (618) and observed an overlap of 19 peptides between shared HM peptides and FM (Fig. 1D). Whilst it is known that HM and bovine milk differ in terms of protein and AA composition (Khaldi et al., 2014; Lu et al., 2018), we did find this overlap to be unexpectedly low. In total, we found 618 peptides unique to HM (Fig. 1D), these shared HM specific peptides indicate a substantial amount of peptide data that remains unexplored, which may prove important in addressing the functional gap between FM and HM.

3.2. Distinct sample grouping observed in HM

Recently, there have been advances in computational methods that have allowed for the analysis and comparison of peptidomic profiles of milk samples (Lambers et al., 2015; Holton et al., 2016). In a similar approach to Holton et al. (2016), we adopted the use of Pearson correlation analysis to determine the sample groupings within our HM, FM and a random *in silico* generated peptide sample as an outgroup (RS). In

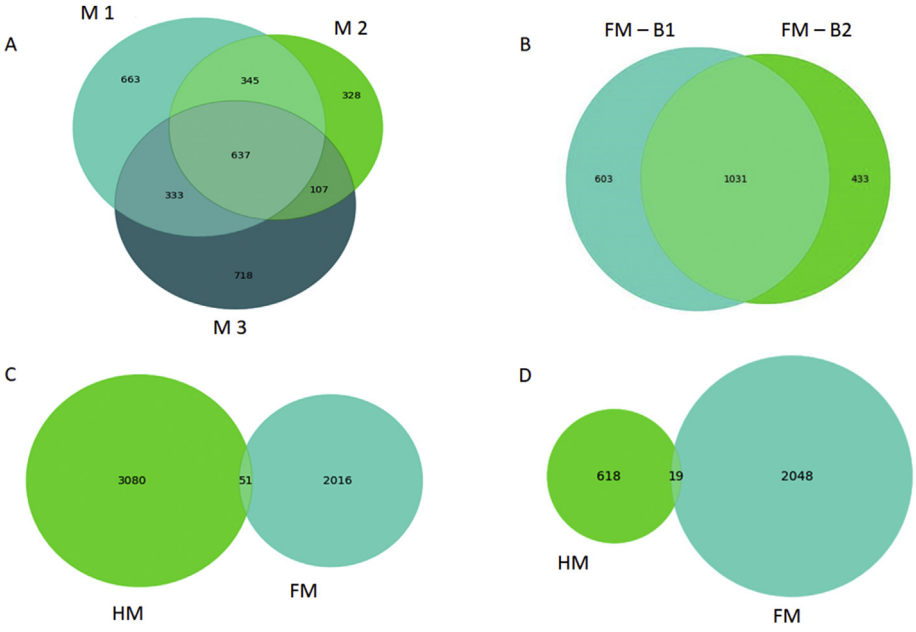

Fig. 1. Venn diagrams illustrating peptide counts in our HM and FM samples. Venn diagram representation of (A) unique and common peptides observed in each mother's samples (M1, M2 and M3), (B) unique and common peptides to both FM brands (FM – B1 and FM – B2) and (C) the number of peptides unique to HM and FM samples and the intersection of peptides common to both and (D) the intersection between FM peptides and peptides common to HM.

our dendrogram plot, HM samples are seen to cluster together, demonstrating a strong correlation between their respective peptide profiles (Fig. 2A). Within this HM cluster, we see that the individual HM samples are grouped according to mother, highlighting the unique peptide profile associated with each mother's breastmilk. Our clustering analysis also demonstrated that samples from each FM brand were extremely similar, as previously demonstrated by the large shared peptide profile of FM shown (Fig. 1B). We noted that different production batches of commercially available formula also exhibit high correlation, in line with the results of Lambers et al. (2015).

Our peptide distribution analysis illustrated little overlap between HM and FM peptides (Fig. 1C), while our dendrogram illustrated clear sample clustering within HM and FM (Fig. 2A). To quantify these differences and for the basis of comparison, we compared our randomly generated sample (RS) to HM and FM and generated a pairwise correlation matrix. For HM, we noted correlation values as high as 0.95 within samples from the same mother, with some variability observed between mothers (0.32–0.84) (Fig. 2B). Similarly, FM samples exhibited strong correlation values within each brand (0.88), with both brands displaying similarities as high as 0.71. However, in all cases HM and FM samples displayed negative correlations between each other (<0) suggesting distinctly different profiles. Moreover, when HM was compared to our randomly generated sample (RS), we similarly observed a negative correlation (<0). Importantly, these results indicate that in terms of peptide composition and abundance, FM is as distinct from HM as it is to a completely random set of amino acids. This serves to illustrate that at present FM fails to account for the peptide profile of HM.

3.3. Physicochemical properties

We analysed the physicochemical properties of the representative peptide profiles in HM and FM which can be informative when investigating peptide bioactivity (Nielsen et al., 2017b). Specifically, we calculated the relative hydrophobicity, sequence length, charge and isoelectric point dispersion in HM and FM (Fig. 3).

It has been reported that the hydrophobicity of amino acids within a peptide can potentially contribute to determining peptide distribution and how a peptide interacts with a cell (Nielsen et al., 2018). To measure hydrophobicity, we calculated the percentage of hydrophobic residues (G, I, V, L, F, C, M, A and W) within each peptide. We found that HM peptides exhibited a median hydrophobicity range of 50–60% (Fig. 3A). While FM peptides exhibited a lower median hydrophobicity range of 40–50% (Fig. 3B). This result suggests the dissimilar hydrophobicity dispersion between HM and FM, which may impact peptide distribution or peptide-cell interaction in a species specific way (Nielsen et al., 2018).

It is known that net peptide charge can be a hallmark of certain bioactivities, for example, positive charge is commonly associated with anti-microbial peptides. (Nielsen et al., 2017a). Here, we found that HM samples had a median negative charge of −1, while FM had a median neutral charge (Fig. 3C and D). This observed difference may be an indication of how the range of peptide bioactivities in HM and FM, differ. The length of a peptide may change the influence of amino acids at certain positions and can be a common characteristic of a type of bioactivity, i.e. the ACE-inhibitory activity of a peptide with an amino acid of a certain size at the C-terminus region was only observed in peptides up to 6 AA in length, the longer the sequence, ACE inhibitory activity decreased (Nielsen et al., 2017a; Pripp et al., 2004). We noted that HM derived peptides contained a larger proportion of longer sequences compared to FM, having a median peptide length of 16 (Fig. 3E). FM consisted mostly of shorter peptide sequences, predominating in the 10–11 residue range; however, this may be an artefact of the FM production process (Fig. 3F).

Together with charge, the isoelectric point (*pI*) is a physicochemical property involved in determining the solubility, localisation and interaction of a protein (Khaldi and Shields, 2011). A previous study showed that *pI* values of certain proteins shift between species, for example, the *pI* value of κ-casein is basic in human compared to acid/neutral in cow (Khaldi and Shields, 2011). As a peptide's function in an acidic environment may be different to its function in a neutral environment, potentially impacting bioactivity (Mohanta et al., 2019), these intrinsic differences in physiochemical properties will likely influence protein function or interaction in an interspecies context. Here, we saw a proportionally higher number of HM peptides with *pI* [4–6 [, compared with a greater proportion of FM peptides with a *pI* range of [4,6 [and [6,8 [(Fig. 3G and H). The observed differences between HM and FM samples may be due to adaptive changes to species-specific needs and digestive systems. Recently, the proteolytic release of peptides from proteins, such as β-casein and α-lactalbumin, has been shown to be pH specific, where only a small amount of HM peptides released from major proteins survived, regardless of pH (Gan et al., 2019). These results illustrate differences in dispersion properties between HM and FM and may offer insights into potential functional variation in their constituent peptides when conferring benefits to infants (Zhu and Dingess, 2019). Combined, these observations may prove important when attempting to render functional benefits such as known protective HM bioactive peptides to FM (Dallas, 2013).

3.4. Distinct cleavage patterns of HM and FM

Cleavage patterns of proteases for HM-derived peptides in infants have been shown to be dictated by HM which persists in the infants stomach (Gan et al., 2019; Holton et al., 2014), studying proteolytic cleavage patterns and peptide termini characteristics, can potentially enhance our understanding of the peptide profile (Klein et al., 2018) and bioactivities such as anti-microbial and cell proliferation conferred by HM (Dallas et al., 2014; Nielsen et al., 2017b; Beverly et al., 2019). Holton et al. (2016) previously examined the potential grouping of milk hydrolysates via cleavage patterns and peptide abundance at the N and C termini. Following this approach, we generated a heatmap to explore N and C terminal abundance for our representative HM and FM samples (Fig. 4). Our results demonstrate that HM exhibits a high frequency of C terminal cleavage at K, P, and V residues, while D was the only residue to show substantial N terminal cleavage. In FM, we detected a high proportion of C terminal cleavage at K, V and L residues, with an elevated abundance of N terminal cleavage seen at S residues. Overall, we observed a difference in C terminal cleavage at P residues in HM and L residues in FM. It is known that certain amino acids are responsible for important functional roles in proteins and peptides. In line with this, proline (P) is known to confer structure to peptides, and has been shown to enhance stability in gastric digestion (Nielsen et al., 2017a). Further investigation may demonstrate that the HM specific pattern of C-terminal cleavage observed at P residues may similarly stabilise peptides in HM. Taken collectively, our results highlight that HM and FM peptides diverge substantially in terms of cleavage signatures indicating the considerable scope for FM adaption.

3.5. Positional homology peptide alignment maps

Visualisation of peptide distribution across samples allows us to look at regions of interest within protein sequences to examine variations in peptide coverage and abundance (Manguy et al., 2017). Using a peptide alignment map, we plotted the coverage of peptides in our representative HM and FM samples to their regions of origin in their parent proteins (Manguy et al., 2017). In a novel approach, we additionally incorporated the positional homology between proteins (α-CN, β-CN, κ-CN and α-LA) of human and bovine origin to facilitate appropriate between species comparison of peptide coverage and abundance.

In agreement with recent investigations of HM, and bovine milk (FM and commercial dairy samples) peptides (Suet al., 2017; Bhattacharya et al., 2019), we found the highest number of peptides in HM and FM were derived from 3 major caseins (α_{s1}-CN, β-CN, κ-CN). In both HM and FM, the highest peptide coverage was reported for β-CN (Fig. 5A and B).

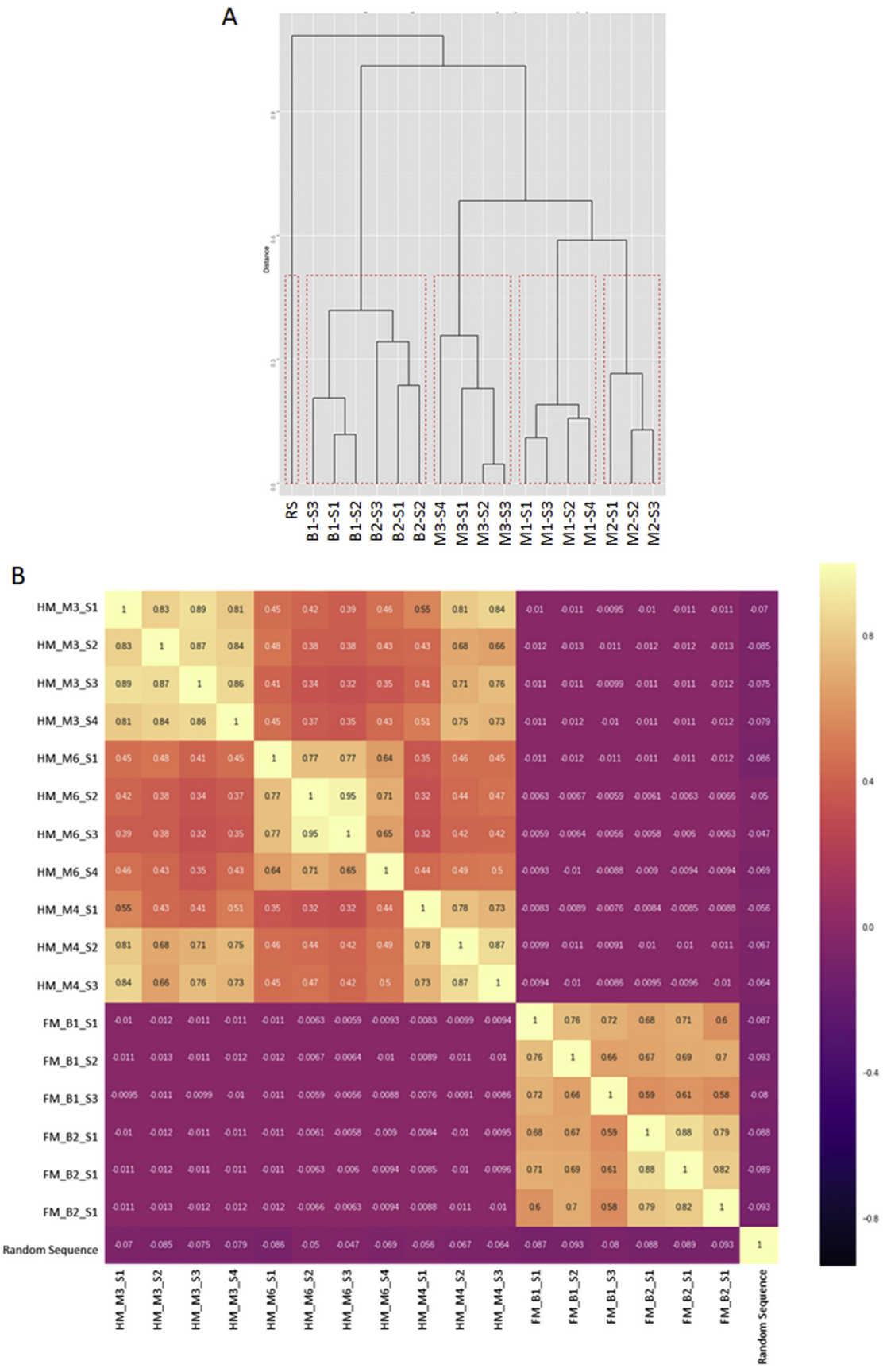

Fig. 2. Distinct HM and FM profiles. (A) Dendrogram illustrating clustering of peptide profiles of all HM (M1, M2, M3; S1-S4) and FM (B1, B2; S1-S3) samples based on Pearson correlation coefficients. A random *in silico* generated sample (RS) was used as an outgroup for clustering. (B) Pairwise correlation matrix of our HM, FM samples and a random *in silico* generated sample (RS). Colour indicates correlation values as per scale on the right of the table.

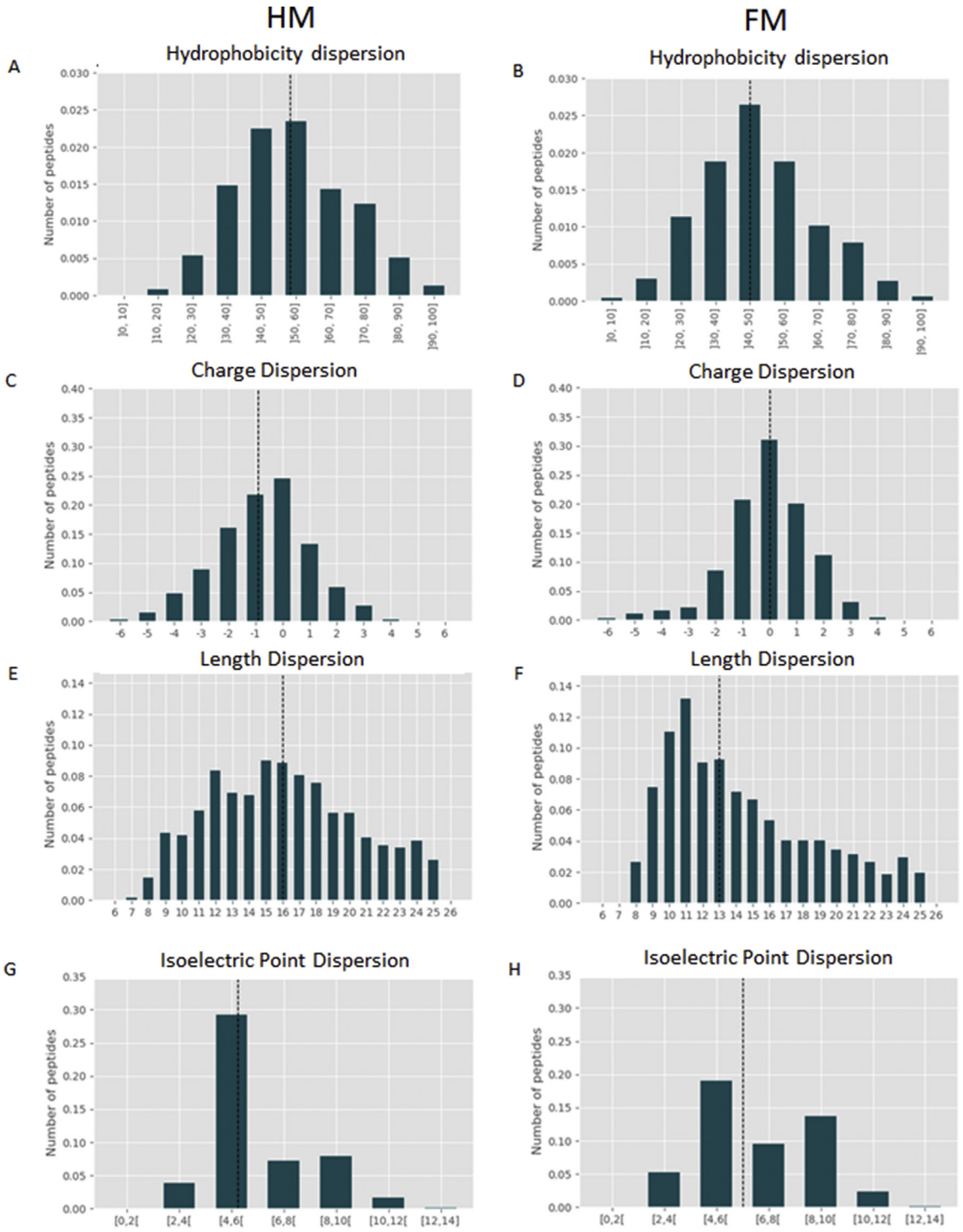

Fig. 3. Physicochemical properties of peptide profiles of HM and FM. Hydrophobicity dispersion based on the percentage of hydrophobic residues in each representative peptide for (A) HM and (B) FM. Amino acid charge dispersion for (C) HM and (D) FM. Peptide sequence length dispersion for (E) HM and (F) FM. Bar charts representing *pI* values for (G) HM and (H) FM peptides. All data is normalised; dashed lines indicate the median.

Notably, HM contained over twice the number of β-CN peptides than FM, with 974 and 447 peptides detected respectively (Fig. 5C). This is in line with a recent study that reported higher abundance of β-CN proteins in HM compared to various ruminant milks (Lu et al., 2018). Overall, we found that peptide coverage and general regions of abundance in HM and FM followed a similar pattern in β-CN. However, HM exhibited considerably elevated peptide abundance at the N terminus and mid sequence regions compared to FM (Fig. 5A and B). This observed increase in peptide coverage unique to HM derived β-CN could be a potential source for important bioactive peptides lacking in FM and is an example of the opportunities that exist for FM humanisation.

In HM, we found 192 α_{s1}-CN peptides with considerable coverage towards the N terminus at α_{s1}-CN$_{(16\text{-}77)}$ (Fig. S2A). We detected 383 α_{s1}-CN peptides in FM displaying similar peptide coverage to analogous regions in HM where present; however, there was a sizeable increase in peptide abundance in FM compared to HM (Figs. S2B and C). Elsewhere

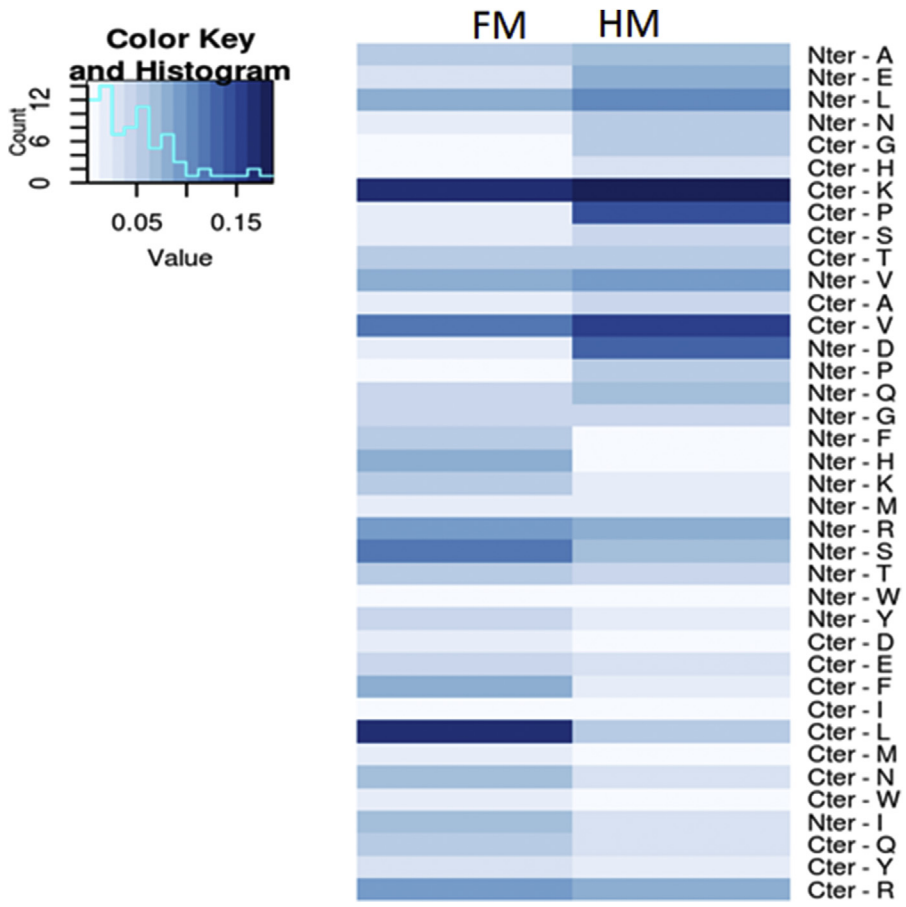

Fig. 4. Heatmap of cleavage patterns in HM and FM. Heatmap representing the abundance weighted N and C terminal amino acid frequencies of our representative HM and FM samples. The colour scale indicates the summed peptide abundance of a given residue.

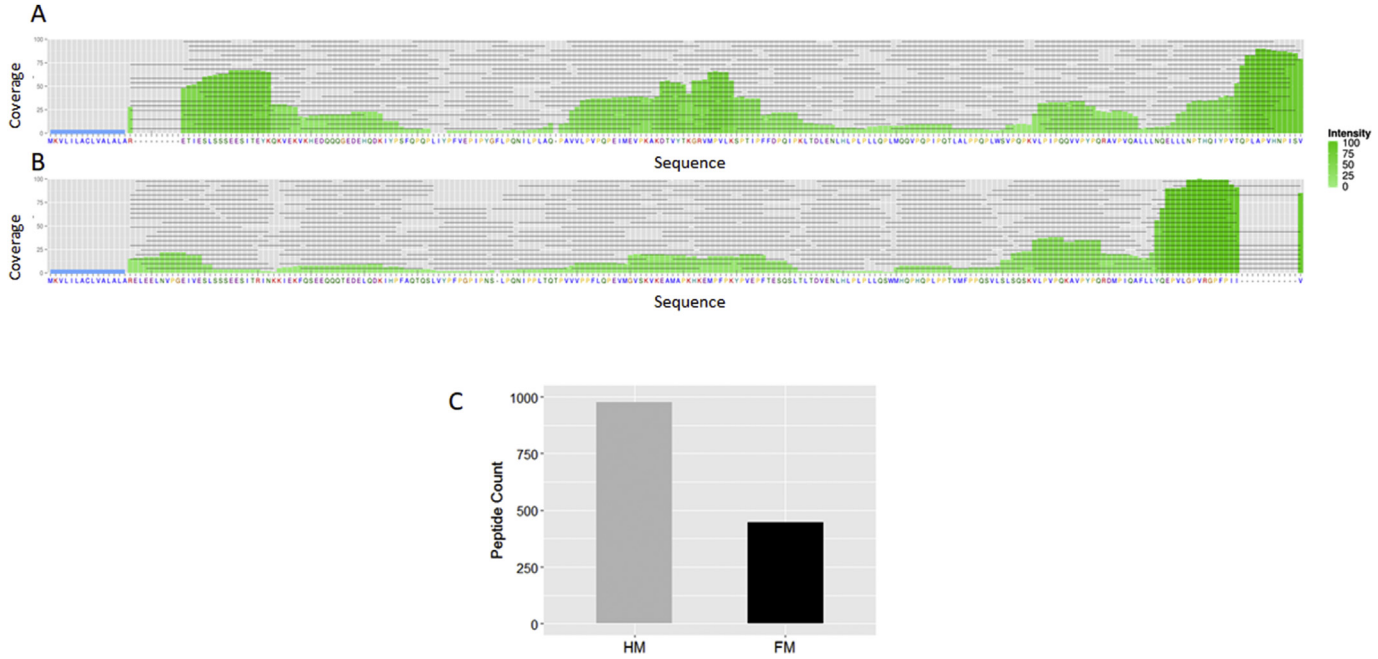

Fig. 5. Positional homology peptide alignment map of β-CN. Peptide coverage and abundance of β-CN from representative (A) HM and (B) FM samples. Black lines represent β-CN detected peptides. Blue bars indicate the signal peptide region, while the green colour scale indicates peptide abundance. Dashes along the protein sequence indicate alignment gap regions inserted to preserve positional homology between bovine and human proteins. (C) Bar chart illustrating β-CN peptide count for HM (grey bar) and FM (black bar).

in FM, there was moderate peptide coverage throughout the α_{s1}-CN sequence, with areas of rich abundance observed in the N terminal and mid sequence regions. In this study, we found that κ-CN in HM contained 76 peptides, while 117 peptides were detected for this protein in FM (Fig. S3C). In HM, peptide coverage was localised to the mid region of κ-CN, with an area of rich abundance seen at κ-$CN_{(80\text{-}120)}$ (Fig. S3A). FM exhibited considerable peptide coverage throughout the sequence of κ-CN, with regions of intense abundance seen towards the C-terminus at κ-$CN_{(127\text{-}158)}$ and κ-$CN_{(170\text{-}191)}$ (Fig. S3B). Our results indicate distinct differences in the mid region of HM and FM κ-CN, further study of HM peptides in this region of κ-CN warrant further study.

Whey is the largest protein complex in HM with α-LA being the principal component (Chatterton et al., 2013). Despite this, we observed 3 peptides in HM towards the centre of α-$LA_{(58\text{-}72)}$ that displayed a high overall intensity (Figs. S4A and C). Dallas (2013) suggested that the low peptide counts found in HM may be due to the tightly packed structure of α-LA giving it increased resistance to proteolysis. As bovine derived α-LA is compositionally very similar to HM α-LA (Lönnerdal, 2014), predictably we recorded a single peptide displaying low intensity at position α-$LA_{(54\text{-}73)}$ in FM (Figs. S4B and C). Overall our peptide alignment results further substantiate the considerable variance in peptide content between FM and HM. Not only are overall peptide properties different between species but there are extensive differences in their relative positions and abundances.

3.7. Bioactive peptides in HM

In this present study we have established that there is a definite and substantial distinction between HM and FM in terms of peptide coverage and properties. A recent study observed striking differences in the functional activities of N-glycoproteins in human and bovine milk, where an overlap of just 12% of functional N-glycoproteins was found common to both (Cao et al., 2019). Considering this, we wanted to further explore the functional properties associated with peptides that were expressly specific to HM. By additionally focusing only on human-specific peptides unanimously detected in each mother's milk, we aimed to capture peptide activities that are likely to be advantageously conserved in humans. Bioactivities for these shared human-specific peptides were determined by searching against public databases MBPDB (http://mbpdb.nws.oregonstate.edu/) and BIOPEP (Minkiewicz et al., 2008).

Firstly, we investigated HM-specific peptides with exact matches to MBPDB and BIOPEP peptides, where a limited number of anti-microbial peptides (2) and anti-cancer peptides (2) were returned (Fig. 6). In light of the low number of exact matches, we additionally performed a BLAST homology (≥80%) search against MBPDB and BIOPEP. From this analysis, the range of activities found within the HM-specific peptide set included anti-microbial, angiotensin converting enzyme-inhibition (ACE_i), immunomodulatory, and anti-cancer (Fig. 6). In particular, we observed 17 anti-microbial peptides (AMPs), which are widely reported to protect the neonate from infection (Mohanty et al., 2016). Moreover, it is known that AMPs can also be released pre-digestion in the mammary gland, conveying protection from infection for mother and baby (Dallas, 2013).

Our HM-specific set contained 14 peptides with ACE-inhibition activity, which is believed to be beneficial for early human growth and may confer long-term benefits into adulthood. For instance, Wada and Lönnerdal (2014) described a possible ameliorating effect of milk-derived ACEi peptides on cardiovascular disease in adulthood as a result of ACE serum levels in infancy (Wada and Lönnerdal, 2014). A number of other benefits additionally reported for ACE-inhibitory peptides include treatment of hypertension and reduced intensity of necrotizing enterocolitis *in vitro* and in murine models; however, the role of these peptides in HM is not yet fully understood (Chatterton et al., 2013; Nielsen et al., 2017a). We observed 3 peptides with immunomodulatory activity in our HM-specific set. Immunomodulatory peptides are known to convey protection and support growth

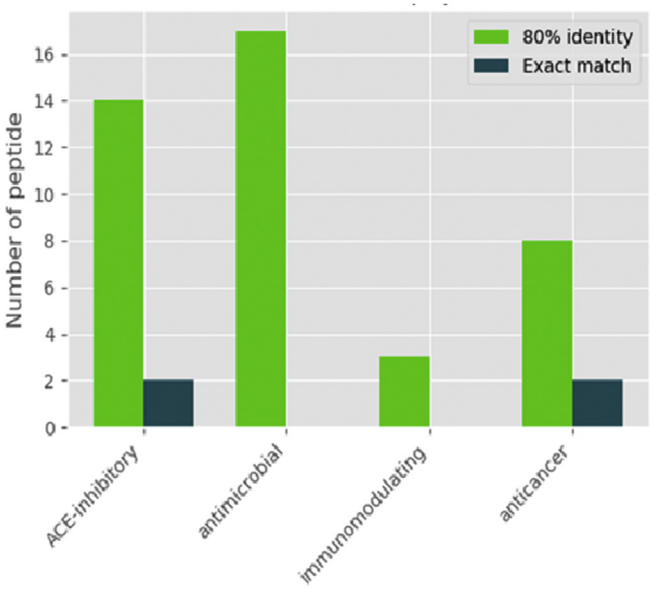

Fig. 6. Bar chart illustrating the number of human specific peptides with reported bioactivities in MBPDB and BIOPEP. Number of peptides found in HM (Exact match and ≥80% match) conveying anti-microbial, ACE-inhibitory, immunomodulating or anticancer activities.

(Nielsen et al., 2017b). Finally, we found 8 peptides with anti-cancer activity that are unique to HM. It has been shown previously that milk derived peptides exert a preventative role in cancer with anti-carcinogenic, anti-proliferative, antioxidant, anti-inflammatory and apoptosis inducing actions (Hsieh et al., 2015).

We similarly examined common peptides specific to FM for bioactivities using the same public databases. The total number of identified bioactive peptides was greater for FM (401) compared to HM (46). There was also a greater number of ACE-inhibitory, antimicrobial and immunomodulatory associated peptides with FM, however, we recorded a higher amount of anticancer HM-specific peptides. Although we attributed a greater number of bioactive peptides and bioactivities to FM peptides, which include antioxidative, antithrombotic, and anxiolytic activity (Fig. S5), these differences may be due, in part, to the amount of peptidomic research that has been carried out on bovine derived milk to date compared to HM (Nielsen et al., 2017a). As per MBPDB, 129 HM derived bioactive peptides compared to 545 FM derived bioactive peptides have been reported (http://mbpdb.nws.oregonstate.edu/). [24 September 2020]. Importantly, this highlights the need for further research to elucidate the marked differences between HM and FM peptide profiles and the activities HM peptides may confer to the infant.

Functional bioactivities such as those reported in HM specific peptides here are of great importance when addressing enhancement of FM given the important role that HM peptides play in the healthy development of an infant (Chatterton et al., 2013; Dallas, 2013; Mohanty et al., 2016; Wada and Lönnerdal, 2014). Notably, however, the majority of our HM-specific peptide set did not correspond to any known bioactive peptide featured in BIOPEP or MBPDB. While broadening our database selection is likely to ascribe bioactivities to some of this remaining HM-specific set, a larger set is likely to remain unclassified in terms of bioactivity due to limitations associated with traditional bioactive peptide discovery. Artificial Intelligence (AI) has recently been shown to successfully discover bioactive peptides from naturally derived sources (Rein et al., 2019). As such, AI represents a promising approach for elucidating the unknown bioactivities unique to HM in order to address the functional gap in FM.

Using peptidomics and bioinformatic techniques, we performed comprehensive comparisons between the peptide profiles of HM and

FM for the first time. In this unique analysis, we detected marked differences in HM compared to commercially available FM. When analysing peptide distribution in our samples, we found little overlap between HM and FM, indicating a substantial amount of peptide data to be explored in HM. By examining physicochemical properties, we aimed to explore differences in HM and FM that may potentially impact on bioactivity (Mohanty et al., 2016). Accordingly, we observed species-specific differences in terms of hydrophobicity, amino acid charge, sequence length and in the distribution of *pIs* for HM compared to FM. In performing clustering analyses between HM and FM, we found a strong correlation between samples of each species milk, and little correlation at the inter species level. Further analysis revealed that FM peptide profiles failed to show any correlation to those of HM. Furthermore, FM displayed as little correlation to HM as it did to a randomly generated peptide sample, starkly illustrating how distinctly different FM is to HM and the considerable need for greater alignment of the FM peptide profile to that of HM.

Our N and C termini analysis further demonstrated a clear separation in the cleavage profiles of HM and FM, indicating distinct hydrolysis histories and peptide compositions which may have implications for important factors such as stability, durability, bioavailability and function (Klein et al., 2018). By examining peptide alignment maps it is possible to compare peptide coverage and abundance in our representative HM and FM samples. It is known that HM and FM peptide profiles differ (Raikos and Dassios, 2014), but until now there has been no side by side comparison of peptide coverage and abundance across all major whey and casein proteins. Here, we were able to visualise the endogenous peptide profiles of HM mapped relative to their parent protein, aligned to their corresponding protein and its associated peptide profile in FM. Again, this analysis served to highlight important incongruences between FM and HM, demonstrating that variance in peptide profiles extend to both protein position and abundance.

Finally, by comparing all our HM samples we were able to identify 618 peptides unique to HM that were shared by all mothers. When compared to public bioactive peptide databases, we detected 46 peptides with known bioactivities including anti-microbial, ACE-Inhibitory, immunomodulatory and anti-cancer effects. Although similar bioactivities were attributed to common FM peptides, importantly, the HM-specific peptides exhibiting these bioactivities cannot be accounted for in bovine derived FM, indicating the potential in exploring the retention of function from HM specific peptides in an infant compared to FM derived peptides. Despite determining a small set of peptides with known function, there is a larger set of HM-specific peptides not identified. Since these peptides are conserved across all mothers, they warrant further investigation and may be key to determining noteworthy functionalities which may be lacking in FM.

4. Conclusion

Even with efforts to humanise FM, there is still a sizeable gap between FM and HM, indicating a need for more HM based research (Lönnerdal, 2014). Recent technological advances in mass spectrometry have allowed for the application of peptidomics to milk research in an attempt to address this issue (Khaldi et al., 2014; Lambers et al., 2015; Holton et al., 2016; Dallas, 2013). To that end, there have been efforts to develop a full database of peptides in HM through LC-MS (Nielsen et al., 2017a, b; Dallas, 2013) and we also now know more about endogenous peptide profiles and proteolytic enzymes involved in milk digestion (Khaldi et al., 2014). Despite limitations of this study, such as sample size of HM and FM, our results, combined with other recent findings, could assist with the classification of HM peptides and their unknown bioactivities. However, further research, with a larger cohort of lactating mothers, is required to explore the marked differences in the HM and FM peptide profile and to help elucidate possible novel human milk-specific bioactivities.

Credit author statement

Cyril Lopez: Software Programming, Formal Analysis, Visualisation, Data Curation Alessandro Adelfio: Software Programming, Formal Analysis, Visualisation, Investigation, Data Curation Audrey M. Wall: Writing - Original Draft, Writing - Review & Editing Brendan Molloy: Methodology, Investigation, Validation Thérèse A. Holton: Formal Analysis, Visualisation, Investigation, Writing - Review & Editing And Nora Khaldi: Conceptualization, Supervision, Writing - Review & Editing

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:
All authors are employees of Nuritas Limited.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.crfs.2020.07.003.

References

Beverly, R.L., Underwood, M.A., Dallas, D.C., 2019. Peptidomics analysis of milk protein-derived peptides released over time in the preterm infant stomach. J. Proteome Res. 18 (3), 912–922.
Bhattacharya, M., Salcedo, J., Robinson, R.C., Henrick, B.M., Barile, D., 2019. Peptidomic and glycomic profiling of commercial dairy products: identification, quantification and potential bioactivities. NPJ Sci. Food 3 (1).
Boyle, R.J., et al., 2016. Hydrolysed Formula and Risk of Allergic or Autoimmune Disease: Systematic Review and Meta-Analysis. *BMJ*.
Cao, X., Yang, N., Liang, X., Tao, D., Liu, B., Wu, J., Yue, X., et al., 2019. Characterization and comparison of whey N-glycoproteomes from human and bovine colostrum and mature milk. Food Chem. 276 (September 2018), 266–273.
Chatterton, D.E.W., Nguyen, D.N., Bering, S.B., Sangild, P.T., 2013. Anti-inflammatory mechanisms of bioactive milk proteins in the intestine of newborns. Int. J. Biochem. Cell Biol. 45 (8), 1730–1747.
Chichlowski, M., German, J.B., Lebrilla, C.B., Mills, D.A., 2011. The influence of milk oligosaccharides on microbiota of infants: opportunities for formulas. Annu. Rev. Food Sci. Technol. 2, 331–351.
Cox, J., Mann, M., Nov. 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. 26, 1367.
Craik, D.J., Fairlie, D.P., Liras, S., Price, D., 2013. The future of peptide-based drugs. Chem. Biol. Drug Des. 81 (1), 136–147.
Dallas, D.C., et al., 2013. Extensive in vivo human milk peptidomics reveals specific proteolysis yielding protective antimicrobial peptides. J. Proteome Res. 12 (5), 2295–2304.
Dallas, D.C., Guerrero, A., Khaldi, N., Borghese, R., Bhandari, A., Underwood, M.A., Lebrilla, C.B., German, B.J., Barile, D., et al., 2014. A peptidomic analysis of human milk digestion in the infant stomach reveals protein-specific degradation patterns. J. Nutr. 144 (6), 815–820.
Dingess, K.A., van den Toorn, H.W.P., Mank, M., Stahl, B., Heck, A.J.R., 2019. Toward an efficient workflow for the analysis of the human milk peptidome. Anal. Bioanal. Chem. 411 (7), 1351–1363.
Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32 (5), 1792–1797.
Fomon, S.J., 2001. Infant feeding in the 20th century: formula and beikost. J. Nutr. 131, 421S–425S.
Gan, J., Robinson, R.C., Wang, J., Krishnakumar, N., Manning, C.J., Lor, Y., Breck, M., Barile, D., German, J.B., 2019. Peptidomic profiling of human milk with LC-MS/MS reveals pH- specific proteolysis of milk proteins. Food Chem. 15 (274), 766–774.
Hays, T., Ra, W., Sep. 2005. A systematic review of the role of hydrolyzed infant formulas in allergy prevention. Arch. Pediatr. Adolesc. Med. 159 (9), 810–816.
Herrmann, K., Carroll, K., May 2014. An exclusively human milk diet reduces necrotizing enterocolitis. Breastfeed. Med. 9 (4), 184–190.
Holton, T.A., Dillon, E.T., Robinson, A., Wynne, K., Cagney, G., Shields, D.C., 2016. Optimal computational comparison of mass spectrometric peptide profiles of alternative hydrolysates from the same starting material. LWT - Food Sci. Technol. (Lebensmittel-Wissenschaft -Technol.) 73, 296–302.
Holton, T.A., Vijayakumar, V., Dallas, D.C., Guerrero, A., Borghese, R.A., Lebrilla, C.B, German, J.B, Barile, D., Underwood, M.A., Shields, D.C., Khaldi, N., et al., 2014.

Following the digestion of milk proteins from mother to baby. J. Proteome Res. 13 (12), 5777–5783.
Hsieh, C.C., Hernández-Ledesma, B., Fernández-Tomé, S., Weinborn, V., Barile, D., De Moura Bell, J.M.L.N., 2015. Milk proteins, peptides, and oligosaccharides: effects against the 21st century disorders. BioMed Res. Int. 2015.
Khaldi, N., Shields, D.C., 2011. Shift in the isoelectric-point of milk proteins as a consequence of adaptive divergence between the milks of mammalian species. Biol. Direct 6 (1), 40.
Khaldi, N., Holton, T.A., Shields, D.C., 2014. Amino acid enrichment and compositional changes among mammalian milk proteins and the resulting nutritional consequences. J. Dairy Sci. 97 (3), 1248–1258.
Khaldi, N., Vijayakumar, V., Dallas, D.C., Guerrero, A., Wickramasinghe, S., Smilowitz, J.T., Medrano, J.F., Lebrilla, C.B., Shields, D.C., German, B.J., et al., 2014. Predicting the important enzymes in human breast milk digestion. J. Agric. Food Chem. 62 (29), 7225–7232.
Klein, T., Eckhard, U., Dufour, A., Solis, N., Overall, C.M., 2018. "Proteolytic cleavage - mechanisms, function, and 'omic' approaches for a near-ubiquitous posttranslational modification. Chem. Rev. 118 (3), 1137–1168.
Lambers, T.T., Gloerich, J., van Hoffen, E., Alkema, W., Hondmann, D.H., van Tol, E.A., 2015. Clustering analyses in peptidomics revealed that peptide profiles of infant formulae are descriptive. Food Sci. Nutr. 3 (1), 81–90.
Lönnerdal, B., 2014. Infant formula and infant nutrition: bioactive proteins of human milk and implications for composition of infant formulas. Am. J. Clin. Nutr. 99 (3).
Lu, J., Zhang, S., Liu, L., Pang, X., Ma, C., Jiang, S., Lv, J., et al., 2018. Comparative proteomics analysis of human and ruminant milk serum reveals variation in protection and nutrition. Food Chem. 261, 274–282.
Lucas, A., Morley, R., Cole, T.J., 1998. Randomised trial of early diet in preterm babies and later intelligence quotient. Br. Med. J. 317 (7171), 1481–1487.
Manguy, J., Jehl, P., Dillon, E.T., Davey, N.E., Shields, D.C., Holton, T.A., 2017. Peptigram: a web-based application for peptidomics data visualization. J. Proteome Res. 16 (2), 712–719.
McCann, J., Ames, B., 2005. Is docosahexaenoic acid, an n-3 long-chain polyunsaturated fatty acid, required for development of normal brain function? An overview of evidence from cognitive and behavioral tests in humans and animals. Am. J. Clin. Nutr. 82 (2), 281–295.
Minkiewicz, P., Dziuba, J., Iwaniak, A., Dziuba, M., Darewicz, M., 2008. BIOPEP database and other programs for processing bioactive peptide sequences. J. AOAC Int. 91 (4), 965–980.
Mohanta, T.K., Khan, A., Hashem, A., Fathi, E., Allah, A., Al-harrasi, A., 2019. The molecular mass and isoelectric point of plant proteomes. BMC Genom. 20 (631), 1–14.
Mohanty, D.P., Mohapatra, S., Misra, S., Sahu, P.S., 2016. Milk derived bioactive peptides and their impact on human health – a review. Saudi J. Biol. Sci. 23 (5), 577–583.
Moossavi, S., Miliku, K., Sepehri, S., Khafipour, E., Azad, M.B., 2018. The prebiotic and probiotic properties of human milk: implications for infant immune development and pediatric asthma. Front. Pediatr. 6 (July), 1–7.
Nielsen, S.D., Beverly, R.L., Qu, Y., Dallas, D.C., 2017. Milk bioactive peptide database: a comprehensive database of milk protein-derived bioactive peptides and novel visualization. Food Chem. 232 (November), 673–682.
Nielsen, S.D., Beverly, R.L., Dallas, D.C., 2017. Peptides released from foremilk and hindmilk proteins by breast milk proteases are highly similar. Front. Nutr. 4 (November), 54.
Nielsen, S.D., Beverly, R.L., Underwood, M.A., Dallas, D.C., 2018. "Release of functional peptides from mother's milk and fortifier proteins in the premature infant stomach. PloS One 13 (11), e0208204.
Pripp, A.H., Isaksson, T., Stepaniak, L., Sørhaug, T., 2004. Quantitative structure-activity relationship modelling of ACE-inhibitory peptides derived from milk proteins. Eur. Food Res. Technol. 219 (6), 579–583.
Qawasmi, A., Landeros-Weisenberger, A., Bloch, M.H., 2013. Meta-analysis of LCPUFA supplementation of infant formula and visual acuity. Pediatrics 131 (1), e262–e272.
Raikos, V., Dassios, T., 2014. Health-promoting properties of bioactive peptides derived from milk proteins in infant food: a review. Dairy Sci. Technol. 94, 91–101.
Rein, D., Ternes, P., Demin, R., Gierke, J., Helgason, T., Schön, C., 2019. Artificial intelligence identified peptides modulate inflammation in healthy adults. Food Funct. 6030–6041.
Stevens, E.E., Patrick, T.E., Pickler, R., 2009. A history of infant feeding. J. Perinat. Educ. 18 (2), 32–39.
Su, M.-Y., et al., 2017. Comparative analysis of human milk and infant formula derived peptides following in vitro digestion. Food Chem. 221, 1895–1903.
Vandenplas, Y., 2017. "Prevention and management of cow's milk allergy in non-exclusively breastfed infants. Nutrients 9 (731), 1–15.
Vandenplas, Y., De Greef, E., Veereman, G., 2014. Prebiotics in infant formula. Gut Microb. 5 (6), 681–687.
Wada, Y., Lönnerdal, B., 2014. Bioactive peptides derived from human milk proteins - mechanisms of action. J. Nutr. Biochem. 25 (5), 503–514.
WHO, 2003. Global Strategy on Infant and Young Child Feeding.
Wiciński, M., Sawicka, E., Gębalski, J., Kubiak, K., Malinowski, B., Jan. 2020. Human milk oligosaccharides: health benefits, potential applications in infant formulas, and pharmacology. Nutrients 12 (1), 266.
Wickham, H., 2009. Use R! ggplot2 - Elegant Graphics for Data Analysis. Springer, New York. New York.
Zhu, J., Dingess, K.A., 2019. The functional power of the human milk proteome. Nutrients 11 (8), 1834.