



**HAL**  
open science

## Impact of human milk pasteurization on the kinetics of peptide release during in vitro dynamic digestion at the preterm newborn stage

Amélie Deglaire, Samira Cassia de Oliveira, Julien Jardin, Valérie Briard-Bion, Florian Kroell, Mathieu Emily, Olivia Ménard, Claire Bourlieu-Lacanal, Didier Dupont

### ► To cite this version:

Amélie Deglaire, Samira Cassia de Oliveira, Julien Jardin, Valérie Briard-Bion, Florian Kroell, et al.. Impact of human milk pasteurization on the kinetics of peptide release during in vitro dynamic digestion at the preterm newborn stage. *Food Chemistry*, 2019, 281, pp.294-303. 10.1016/j.foodchem.2018.12.086 . hal-01984298

HAL Id: hal-01984298

<https://hal.science/hal-01984298>

Submitted on 21 Oct 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

1 **Impact of human milk pasteurization on the kinetics of peptide release during *in vitro***  
2 **dynamic digestion at the preterm newborn stage**

3 Amélie Deglaire<sup>1</sup>, Samira De Oliveira<sup>1</sup>, Julien Jardin<sup>1</sup>, Valérie Briard-Bion<sup>1</sup>, Florian Kroell<sup>1,2</sup>,  
4 Mathieu Emily<sup>2</sup>, Olivia Ménard<sup>1</sup>, Claire Bourlieu<sup>1</sup>, Didier Dupont<sup>1</sup>

5

6 <sup>1</sup> STLO, Agrocampus Ouest, INRA, 35042 Rennes, France

7 <sup>2</sup> IRMAR, Agrocampus Ouest, CNRS, 35042 Rennes, France

8 Corresponding author : Amélie Deglaire. Agrocampus Ouest, 65 rue de Saint-Brieuc, CS  
9 84215, 35042 Rennes Cedex, France

10

11 Abbreviations: GRAVY, Grand Average of Hydropathy Value; MS, Mass spectrometry;

12 PHM, pasteurized human milk; RHM, raw human milk.

13

14 Keywords: digestion, human milk, pasteurization, peptidomics, preterm

15 Running title: Peptide release during preterm digestion of pasteurized human milk

16 Highlights :

- 17 • Pasteurization increased the pre-proteolysis of most of the  $\beta$ -casein sequence  
18 • Pasteurization impacted the peptide release during digestion of preterm human milk  
19 • Pasteurization impacted **some** bioactive peptide release before and during digestion

20

21 Number of words: 5510

22

23

24 **Abstract**

25 Holder pasteurization (62.5°C, 30 min) of human milk denatures beneficial proteins. The  
26 present paper aimed to assess whether this can affect the kinetics of peptide release during  
27 digestion at the preterm stage. Raw (RHM) or pasteurized (PHM) human milk were digested  
28 in triplicates using an *in vitro* dynamic system. Mass spectrometry and multivariate statistics  
29 were conducted. Pre-proteolysis occurred mostly on  $\beta$ -casein, for which cumulative peptide  
30 abundance was significantly greater in PHM over 28% of the hydrolysed sequence. Eight  
31 clusters resumed the kinetics of peptide release during digestion, which differed on seven  
32 clusters (69% of the 1134 peptides). Clusters associated to the heat-denaturated proteins,  
33 lactoferrin and bile salt-stimulated lipase, presented different kinetics of release during  
34 digestion, unlike that for  $\beta$ -casein. Some bioactive peptides from  $\beta$ -casein presented  
35 significant different abundances between PHM and RHM before digestion (1-18, 185-211) or  
36 in the intestinal phase (154-160, 161-166). Further physiological consequences should be  
37 investigated.

38

## 39 1. Introduction

40 Human milk is considered as the ideal food for neonates, supporting not only their nutrition  
41 but also their optimal growth and development (Victora, Bahl, Barros, França, Horton,  
42 Krasevec, et al., 2016). When vulnerable hospitalized neonates such as preterm infants  
43 receive donor human milk, it has to be first pasteurized (62.5 °C, 30 min) in the milk bank for  
44 sanitary reasons (Wight, 2001). Together with the associated freeze/thaw cycles,  
45 pasteurization induces some protein denaturation, which reduces the content of some native  
46 proteins and thus affect their bioactivity, such as for immunoglobulins (loss of 20 to 100%),  
47 lysozyme (loss of 20 to 69%), lactoferrin (loss of 35 to 90%), and bile salt-stimulated lipase  
48 (complete loss) (O'connor, Ewaschuk, & Unger, 2015; Peila, Moro, Bertino, Cavallarin,  
49 Giribaldi, Giuliani, et al., 2016). These physical treatments also impact the human milk  
50 microstructure, as previously observed (De Oliveira, Bourlieu, Ménard, Bellanger, Henry,  
51 Rousseau, et al., 2016a; De Oliveira, Deglaire, Ménard, Bellanger, Rousseau, Henry, et al.,  
52 2016b). In the latter studies, we have demonstrated that human milk pasteurization affected its  
53 proteolysis kinetics, at different extent, during *in vitro* gastro-intestinal digestion at the  
54 preterm and term stages (De Oliveira, et al., 2016a; De Oliveira, et al., 2016b), and during *in*  
55 *vivo* gastric digestion in preterm neonates (De Oliveira, Bellanger, Ménard, Pladys, Le Gouar,  
56 Dirson, et al., 2017). More precisely, at the preterm stage, pasteurization did not significantly  
57 impact the kinetics of disappearance of the major proteins in the gastric phase, but during the  
58 intestinal phase, amino acid bioaccessibility was selectively modulated (De Oliveira, et al.,  
59 2016a). Knowledge on the kinetics of peptide release during digestion is crucial for a better  
60 understanding of the potential peptide bioactivity released in the gastrointestinal tract and on  
61 the amino acid absorption rate, the latter of which has been demonstrated to be an  
62 independent factor of modulation of the postprandial regional metabolism, at least in adult  
63 (Deglaire, Fromentin, Fouillet, Airinei, Gaudichon, Boutry, et al., 2009). To date, only our

64 study has investigated and showed that there was an impact of milk pasteurization on the  
65 kinetics of peptide release during *in vitro* dynamic digestion of term milk at the term newborn  
66 stage (Deglaire, De Oliveira, Jardin, Briard-Bion, Emily, Ménard, et al., 2016). We assume  
67 that such an impact should also be found for preterm milk digested in the preterm newborn  
68 condition but certainly at a different extent given the digestive immaturity of the preterm  
69 newborn combined to a higher protein content in the preterm human milk. The objective of  
70 the present paper was to verify this hypothesis. The kinetics of peptide release was followed  
71 during gastro-intestinal digestion of a pool of raw or pasteurized preterm human milk at the  
72 preterm newborn stage. Digestion was simulated using an *in vitro* dynamic model, where  
73 digestive enzymes, pH and digestive compartment emptying were continuously controlled.  
74 Nature and abundance of peptides were determined by mass spectrometry and data analyzed  
75 by multivariate statistics.

## 76 2. Material and methods

77 All chemicals and enzymes were from commercial origin (Sigma-Aldrich, Saint Quentin  
78 Fallavier, France), except for the Rabbit Gastric Extract, obtained from GERME SA  
79 (Marseille, France).

### 80 2.1 Human milk collection and digestion

81 Collection of Human milk has been previously detailed by De Oliveira et al. (2016a). Briefly,  
82 mature human milk samples were obtained frozen from five women (26-34 years old) on  
83 average  $6.6 \pm 2.4$  weeks after preterm delivery *via* a donor milk bank (University Hospital  
84 Center, Rennes, France). The gestation length was in average  $29.8 \pm 3.0$  weeks of  
85 amenorrhea, ranging from 26 up to 34 weeks of gestation. After milk thawing (at 4°C, over 16  
86 h) and pooling, half of the pool was pasteurized (62.5 °C, 30 min). Raw (RHM) and  
87 pasteurized (PHM) human milk were then stored at -20°C. Milk was analyzed by a Human  
88 Milk Analyzer (Miris, AB, Uppsala, Sweden) and consisted of  $35.7 \pm 0.7$  g/L of fat,  $13.2 \pm$   
89  $0.8$  g/L of proteins and  $78.3 \pm 2.3$  g/L of carbohydrates (De Oliveira, et al., 2016a).  
90 Gastrointestinal digestion was performed on 100 mL of whole milk, as RHM or PHM, in a bi-  
91 compartmental *in vitro* dynamic system (DIDGI®, INRA, France) **simulating the preterm**  
92 **newborn digestive conditions, such as reviewed in Bourlieu et al. (2014). The rationale of the**  
93 **parameters used in the present model has been detailed in De Oliveira et al. (2016a). Briefly,**  
94 fluid secretions, pH and emptying were continuously monitored. The transit time of the  
95 chyme followed the Elashoff equation with  $t_{1/2} = 36$  min and  $\beta = 1.15$  for the gastric emptying  
96 and  $t_{1/2} = 200$  and  $\beta = 2.2$  for the intestinal emptying (De Oliveira, et al., 2016a). The gastric  
97 pH followed the equation:  $pH = -0.0155 \times time (min) + milk pH$ . The intestinal pH was  
98 maintained at 6.2. Gastric enzymes were added as rabbit gastric extract (pepsin and lipase:  
99 120 and 8.6 U/mL of gastric content, respectively) and intestinal enzymes as porcine  
100 pancreatin (trypsin, chymotrypsin and lipase: 16, 2 and 59 U/mL of intestinal content) in

101 addition to bovine bile salts. Enzyme units were determined based as a function of the mean  
102 body weight, estimated at 1.9 kg for a preterm newborn at a gestational age of 28 weeks and a  
103 postnatal age of four weeks (De Oliveira, et al., 2016a).

104 Digestions were conducted in triplicate for each matrix over 3 hours with digesta sampling  
105 every 30 or 60 min. Samples were stored at -20°C after addition of the antiprotease  
106 phenylmethylsulfonyl fluoride (PMSF) to avoid further proteolysis.

107 Ethical approval was granted by the Hospital Ethics Committee (n° 13-12) and donors gave  
108 their informed written consent.

## 109 **2.2 Mass spectrometry**

### 110 *2.2.1 Peptide identification*

111 Mass spectrometry (MS) analysis was conducted as described previously (Deglaire, et al.,  
112 2016). Briefly, a nano-RSLC Dionex U3000 system fitted to a Q-Exactive mass spectrometer  
113 (Thermo Scientific, San Jose, USA) equipped with a nanoelectrospray ion source was used.  
114 Digesta samples were diluted 100 times in the injection buffer and filtered (0.45 µm cut-off)  
115 before concentration on a µ-precolumn pepMap100 (C18 column, 300 µm i.d. × 5 mm length,  
116 5 µm particle size, 100 Å pore size; Dionex, Amsterdam, The Netherlands) and separation on  
117 a PepMap RSLC column (C18 column, 75 µm i.d. x 150 mm length, 3µm particle size, 100 Å  
118 pore size; Dionex).

119 Peptide separation was performed at a flow rate of 0.3 µL.min<sup>-1</sup> using solvents A [2% (v/v)  
120 acetonitrile, 0.08% (v/v) formic acid and 0.01% (v/v) TFA in HPLC gradient grade water] and  
121 B [95% (v/v) acetonitrile, 0.08% (v/v) formic acid and 0.01% (v/v) TFA in HPLC gradient  
122 grade water]. The elution gradient first rose from 5 to 35% solvent B over 87 min, then up to  
123 85% solvent B over 3 min before column reequilibration. The mass spectra were recorded in  
124 positive mode using the m/z range 250-2000. The resolution of the mass analyzer for m/z of  
125 200 amu (atomic mass unit) was set in the acquisition method to 70 000 for MS and 17 500

126 for MS/MS. For each MS scan, the ten most intense ions were selected for MS/MS  
127 fragmentation and excluded from fragmentation for 15 s.

128 Peptides were identified from the MS/MS spectra using the X!TandemPipeline software  
129 (Langella, Valot, Balliau, Blein-Nicolas, Bonhomme, & Zivy, 2017) against a human milk  
130 protein database (Molinari, Casadio, Hartmann, Livk, Bringans, Arthur, et al., 2012) to which  
131 was added the common Repository of Adventitious Protein (<http://thegpm.org/crap>). The  
132 possible post-translational modifications were serine or threonine phosphorylation,  
133 methionine oxidation, lysine or arginine lactosylation, cyclisation of glutamine or glutamic  
134 acid into pyroglutamic acid and acetylation of cysteine, serine, lysine or any amino acid in N-  
135 ter position. Peptides identified with an e-value < 0.01 were automatically validated. The  
136 peptide false discovery rate was less than 0.3%.

#### 137 *2.2.1.1 Quantification of peptides*

138 Each identified peptide was quantified **by label-free MS** using the MassChroQ software  
139 (Valot, Langella, Nano, & Zivy, 2011). A m/z width of 10 ppm was used to perform extracted  
140 ion chromatograms (XIC) of peptides in time-aligned chromatograms and the area under the  
141 curve was then quantified. When a peptide was measured with several charge states, all ion  
142 intensities were summed.

#### 143 *2.2.2 Biochemical characteristics of the peptides*

144 Several biochemical characteristics were determined for each peptide, particularly the amino  
145 acids at cleavage sites in positions P1 (C-ter end) and P1' (N-ter end), the number of essential  
146 amino acids (histidine, isoleucine, leucine, lysine, methionine, phenylalanine, threonine,  
147 valine), the isoelectric point, the molecular weight and the Grand Average of Hydropathy  
148 Value (GRAVY) as defined by the ExPASy "ProtParam tool" (Gasteiger, Hoogland, Gattiker,  
149 Duvaud, Wilkins, Appel, et al., 2005). A positive GRAVY corresponds to a hydrophobic  
150 peptide and a negative one to a hydrophilic peptide (Kyte & Doolittle, 1982).



151 Bioactive peptides were searched within the databases of BIOPEP (Minkiewicz, Dziuba,  
152 Iwaniak, Dziuba, & Darewicz, 2008) and MBPDB (Nielsen, Beverly, Qu, & Dallas, 2017)  
153 both accessed on January 2018. Only exact matching between sequences was considered.  
154 Prediction of antimicrobial activities was examined using CAMPR3 based on the algorithms  
155 of Random Forest (RF) and Support Vector Machines (SVM), previously shown to have a  
156 good prediction performance (Gabere & Noble, 2017). A peptide sequence was determined to  
157 have a greater likelihood of antimicrobial functionality with scores approaching 1. All  
158 peptides with scores > 0.5 on both prediction methods were considered significant for  
159 antimicrobial functionality.

### 160 **2.3 NH<sub>2</sub> quantification**

161 As described earlier (Deglaire, et al., 2016), the method was a fluorescent microplate analysis  
162 based on the reaction of ortho-phthaldialdehyde (OPA) and dithiothreitol (DTT) with primary  
163 amines resulting in 1-alkylthio-2-alkylisondole detected at 340 nm.

### 164 **2.4 Enzymes predicted to be active in milk before gastrointestinal digestion**

165 The peptide sequences as identified in milk before digestion were submitted to the online  
166 software Enzyme Predictor (<http://bioware.ucd.ie>), which examined the cleavage patterns and  
167 predicted the enzymes responsible for this profile based on the cleavage specificities of 35  
168 enzymes (Vijayakumar, Guerrero, Davey, Lebrilla, Shields, & Khaldi, 2012). Enzymes were  
169 then ranked according to their likelihood of having being used to generate the hydrolysate  
170 thanks to the calculation of an odds ratio (OR). The same human enzymes as those previously  
171 considered (Deglaire, et al., 2016) were included. Enzymes presenting only one potential  
172 cleavage were not considered. Cleavage sites of cytosol aminopeptidase and carboxypeptidase  
173 B2 were examined manually over  $\beta$ -casein, such as performed previously (Dallas, Smink,  
174 Robinson, Tian, Guerrero, Parker, et al., 2015).

## 175 2.5 Statistical analyses

176 Data analyses were performed using the R software, version 3.3.1 (R Core Team, 2014). All  
177 graphical representations were based on the log<sub>10</sub>-transformed abundances  
178 [log<sub>10</sub>(abundance+1)]. An in-house program allowed the peptide mapping and their average  
179 cumulative abundances onto the parent protein. For each digestion and each phase (milk prior  
180 to digestion, gastric phase or intestinal phase), abundances of the peptides were averaged over  
181 the different digestion times of the corresponding phase, summed amino acid by amino acid  
182 and log<sub>10</sub>-transformed [log<sub>10</sub>(abundance+1)]. Missing abundances were set at 0. Log<sub>10</sub>-  
183 transformed abundances were averaged by type of milk in order to be plotted along the amino  
184 acid sequence of the protein; they were compared between RHM and PHM for each amino  
185 acid using a t-test.

186 For each peptide identified at least on two digestions of RHM or PHM (n = 1134), a two-way  
187 analysis of variance was fitted (*lm* function; pasteurization, time and their interaction as  
188 factors) on the transformed abundances of the 9 time points (before digestion, 3 time points in  
189 the gastric phase and 5 in the intestinal phase) over the three digestions of RHM and PHM,  
190 such as performed previously (Deglaire, et al., 2016). Abundance transformation consisted in  
191 abundance scaling after log-transformation [log(abundance + 1)] followed by setting the  
192 maximum abundance to 1. From the two-way analysis of variance, the “Pasteurization x  
193 Digestion time” interaction coefficients, for which the sum equalized zero over time within  
194 each milk and at each time for RHM and PHM, were used for subsequent analyses. In  
195 particular, the eight independent interaction coefficients of PHM were used for the ascending  
196 hierarchical clustering (*hclus* function, stats package) of the 1134 peptides based on the  
197 minimum within-cluster variance Ward’s agglomeration. The number of clusters was  
198 determined thanks to the bar heights at one of the most marked jump. The function  
199 *heatmap.2* was used to display the heatmap and its dendrogram. Cluster characterization was

200 undertaken (*catdes* function, FactoMiner package) using the biochemical characteristics. A  
201 Chi-square test was undertaken for qualitative variables (protein origin and cleavage sites),  
202 and a one-way analysis of variance for quantitative variables (isoelectric point, molecular  
203 weight, number of essential amino acids, GRAVY). V-test indicated whether the modality  
204 category was significantly overrepresented ( $v > 2$ ) or underrepresented ( $v < -2$ ).

205 A two-way analysis of variance (with pasteurization, time and their interaction as factors) was  
206 fitted to the abundances of each bioactive peptide. When the interaction was statistically  
207 significant, a *post-hoc* test was performed on the model residuals (*testinteractions* function,  
208 *phia* package) using the p-value adjustment method of Benjamini & Hochberg (1995).  
209 Statistical significance level was considered for p-value  $< 0.05$ .

## 210 3. Results

### 211 3.1 Identified proteins and peptides

212 Twenty-seven human milk proteins (**Supplementary Material 1**) were at the origin of 1531  
213 peptides (**as detailed in Supplementary Material 2**) identified in RHM and PHM prior to and  
214 during gastro-intestinal digestion. Identified peptides contained 6 to 48 amino acids. Some  
215 peptides (n = 162) presented modifications that were serine phosphorylation (n = 84),  
216 methionine oxidation (n = 43), lactosylation of lysine (n = 8) or arginine (n = 2),  
217 glutamine/glutamic acid cyclisation (n = 4), or acetylation (n = 21). A large proportion of the  
218 modified peptides (73%) had a sequence that was also identified either without modification  
219 or with a different modification.

### 220 3.2 Proteolysis of human milk prior to gastrointestinal digestion

221 Before digestion, 204 peptides were identified and derived at 82.3% from  $\beta$ -casein; the  
222 remaining peptides derived from  $\alpha_{s1}$ -casein (7.4%), osteopontin (4.9%) and polymeric  
223 immunoglobulin receptor (4.4%). PHM and RHM presented 43 and 33 specific peptides,  
224 respectively (at 84 and 70% from  $\beta$ -casein) and shared 128 common peptides. The average  
225 abundances of all the peptides identified prior to digestion were plotted onto their parent  
226 protein sequences, as shown in **figure 1A**. These peptides covered 95% of the  $\beta$ -casein  
227 sequence in both RHM and PHM, with **28% of this hydrolysed sequence having a non-**  
228 **transformed abundance significantly higher for PHM than for RHM (4 times; P < 0.05)**. The  
229 greatest difference was observed **between the residues 136-138 residues with a PHM/RHM**  
230 **ratio of 12.**

231 For  $\alpha_{s1}$ -casein (figure 1A), peptides were identified over 53% and 71% of the sequence for  
232 PHM and RHM, respectively. Non-transformed abundances were **significantly** lower for  
233 PHM than for RHM **only between residues 84 and 109** (2 times; P < 0.05). **For osteopontin**  
234 **and polymeric immunoglobulin receptor**, peptides were identified on a limited portion of the

235 sequence (figure 1A) and abundances were not significantly different between RHM and  
236 PHM.

237 Based on their high odds ratio and total number of cleavages, plasmin (which shares the same  
238 cleavage sites as trypsin) was predicted to be the most active enzyme in human milk  
239 proteolysis prior to gastrointestinal digestion (Table 1), with a similar number of cleavages  
240 between RHM and PHM (87 and 85, respectively). Cathepsin D was also assumed to be  
241 responsible of a high number of cleavages (61 and 71 for RHM and PHM), but due to its large  
242 specificity, the odds ratio was below 1. Cytosol aminopeptidase was assumed to have mainly  
243 cleaved sites located on the C-ter side of  $\beta$ -casein for both RHM and PHM (Supplemental  
244 material 3), while carboxypeptidase B2 mainly cleaved at the residue 96.

### 245 3.3 NH<sub>2</sub> release after digestion of human milk

246 The concentration of NH<sub>2</sub> released after gastric and intestinal digestion was not significantly  
247 different ( $P > 0.05$ ) between RHM and PHM, with average values of  $47.5 \pm 15.3$  and  $611.2 \pm$   
248  $14.7$  mg of NH<sub>2</sub>/L of milk at 90 min of gastric digestion and 180 min of intestinal digestion,  
249 respectively.

### 250 3.4 Peptide release during gastro-intestinal digestion of RHM and PHM

#### 251 *Peptidomic profiles of proteins*

252 The first eight proteins being at the origin of the identified peptides were included in Figure  
253 1, showing the average cumulative peptide abundances onto the protein sequence during  
254 digestion. The proportion of the sequence covered by peptides was at 100% during all the  
255 digestion only for  $\beta$ -casein (figure 1A), while it decreased from the gastric to the intestinal  
256 phase for polymeric immunoglobulin receptor and bile-salt stimulated lipase (-12 and -21%,  
257 respectively) and increased for the other proteins (+13% for  $\alpha_{s1}$ -casein up to +117% for  $\alpha$ -  
258 lactalbumin). During the gastric phase, only abundances from bile-salt stimulated lipase  
259 significantly differed between RHM and PHM with greater abundances for RHM on 16% of

260 the hydrolysed sequence (**Figure 1B**). During the intestinal phase, this relationship was  
261 opposite for bile-salt stimulated lipase with abundances being significantly higher for PHM (6  
262 times higher for non-transformed abundances) on 59% of the hydrolysed sequence. A similar  
263 relationship was observed for lactoferrin during the intestinal digestion (10 times significantly  
264 higher for non-transformed abundances of PHM; 32% of the hydrolysed sequence), while for  
265  $\alpha_{s1}$ -casein cumulative abundance was higher for RHM (6 times) but on a limited proportion of  
266 the hydrolysed sequence (15%).

### 267 *Kinetics of peptide release during gastrointestinal digestion*

268 Among the 1531 peptides, 26 % of them were not included in the clustering analyses, because  
269 they were detected exclusively in RHM (n = 172) or in PHM (n= 73), or sporadically in both  
270 RHM and PHM (i.e. only detected on one out of the three digestions, n = 152). This did not  
271 greatly affect the parent protein distribution (**Figure 2A**).

272 Eight clusters resumed the kinetics of release of the 1134 peptides (**Figure 2B**) and the  
273 qualitative variable being the most significantly associated to the clustering was the parent  
274 protein (p-value = 6.1 E-21). One cluster (n° 4; 31% of the peptides) included peptides  
275 exerting similar kinetics along the entire gastro-intestinal digestion. These peptides were  
276 originating mostly from  $\beta$ -casein and to a lesser extent from  $\kappa$ -casein, while there was a very  
277 limited number of peptides from bile-salt stimulated lipase and lactoferrin (Table 2).

278 Clusters 6 and 8 (36% of the peptides) included peptides mainly produced during the gastric  
279 phase and showed a lower abundance for PHM than for RHM (**Figure 2C**). In cluster 6,  
280 leucine (cleavage specificity for pepsin) in position P1 was significantly more represented  
281 unlike lysine and arginine (cleavage specificity for plasmin/trypsin) significantly less  
282 represented than in the overall dataset; peptides had in average a higher molecular weight than  
283 in overall (**Table 2**). Cluster 8 was significantly associated to bile salt-stimulated lipase.

284 Clusters 1 to 3 (20% of the peptides) included peptides mostly produced in the intestinal  
285 phase, where the abundance was lower for PHM than for RHM **at some digestion time** (Figure  
286 2C). In cluster 1, the proportion of  $\alpha_{s1}$ -casein peptides and of arginine in position P1 was  
287 greater than in overall (Table 2). In cluster 3, there was significantly less peptides from  $\beta$ -  
288 casein than in overall and no peptides from bile salt-stimulated lipase (Table 2). Cluster 5 (5%  
289 of the peptides) showed peptides produced only during the intestinal phase **at a constant level**  
290 **for RHM while decreasing for PHM** (Figure 2C). These peptides were more originating from  
291 lactoferrin (46% of the cluster peptides) and less from  $\beta$ -casein and  $\alpha_{s1}$ -casein (24% of the  
292 cluster peptides); **they contained** more lysine and arginine in position P1 than in overall  
293 (Table 2). Cluster 7 (9% of the peptides) was the only cluster where PHM abundance was  
294 higher during intestinal digestion (at 90 and 120 min). It included more peptides from bile  
295 salt-stimulated lipase but less from  $\beta$ -casein and  $\alpha$ -lactalbumin; their molecular weight was  
296 lower than in overall (Table 2).

### 297 **3.5 Bioactive peptides**

298 **The exact matching sequence search over the 1531 peptides resulted in 36 peptides having a**  
299 **predicted antimicrobial activity; they were originating mostly from bile-salt stimulated lipase**  
300 **(39% of them) and lactoferrin (19%) (data not shown). Eight peptides were found to have a**  
301 **proved bioactivity according to the literature, such as indexed by BIOPEP and MBPDB**  
302 **databases. They were all from  $\beta$ -casein, except for one peptide from lactoferrin (Table 3). The**  
303 **antimicrobial peptide  $\beta$ -casein(185-211) had an abundance six times lower ( $P < 0.05$ ) for**  
304 **PHM prior to digestion but its abundance was then similar ( $P > 0.05$ ) between RHM and**  
305 **PHM during the first hour of gastric digestion before being null for both types of milk. Two**  
306 **angiotensin-converting enzyme inhibitor peptides [ $\beta$ -casein(154-160) and  $\beta$ -casein(161-166)]**  
307 **were present only in the intestinal phase and had abundances significantly higher ( $P < 0.05$ )**  
308 **for PHM than for RHM. Non-transformed abundances were two times higher for  $\beta$ -**

309 casein(154-160) in PHM during the intestinal digestion (except at 60 min) and four times  
310 higher for  $\beta$ -casein(161-166) in PHM only at 180 min of digestion. An additional bioactive  
311 peptide, not reported within the bioactive peptide databases, was the  $\beta$ -casein(1-18), a  
312 caseinophosphopeptide (CPP) homologue (Wada & Lonnerdal, 2015), which was present in  
313 the human milk prior to digestion in a lower proportion in PHM than in RHM (3 times lower).  
314 This difference was maintained along the gastric digestion, while the peptide was hydrolysed  
315 rapidly in the intestinal phase.



## 316 **4. Discussion**

317 The present paper demonstrates that pasteurization **impacted selectively** the peptide release  
318 from preterm human milk prior to digestion and during gastro-intestinal digestion at the  
319 preterm newborn stage, **mainly depending on the protein origin**. This is the first time that a  
320 complete view of the peptides identified and released during digestion is provided, from semi-  
321 quantitative peptidomic maps up to peptide clustering based on their kinetics of release.

### 322 **4.1 Human milk pre-proteolysis**

323 Before digestion, endogenous peptides were detected in human milk and originated  
324 essentially from  $\beta$ -casein, and to a lesser extent from  $\alpha_{s1}$ -casein, osteopontin and polymeric  
325 immunoglobulin receptor, such as observed previously (Dallas, Guerrero, Khaldi, Castillo,  
326 Martin, Smilowitz, et al., 2013; Deglaire, et al., 2016; Dingess, De Waard, Boeren, Vervoort,  
327 Lambers, Van Goudoever, et al., 2017; Ferranti, Traisci, Picariello, Nasi, Boschi, Siervo, et  
328 al., 2004; Khaldi, Vijayakumar, Dallas, Guerrero, Wickramasinghe, Smilowitz, et al., 2014;  
329 Nielsen, Beverly, & Dallas, 2017). This confirms that indigenous milk proteases (e.g.  
330 plasmin, carboxypeptidase, cytosol aminopeptidase, kallikrein) can hydrolyse human milk  
331 proteins, particularly  $\beta$ -casein, within the mammary gland (Demers-Mathieu, Nielsen,  
332 Underwood, Borghese, & Dallas, 2017; Nielsen, Beverly, & Dallas, 2017) but also during  
333 milk processing (freeze-thaw cycle, pasteurization), as qualitative and quantitative differences  
334 occurred on the same pool of milk differing only by their processing (raw or pasteurized).

335 In the present study, plasmin was predicted, to be the most active protease in human milk,  
336 such as reported previously (Deglaire, et al., 2016; Nielsen, Beverly, & Dallas, 2017). The  
337 number of cleavages from plasmin between PHM and RHM was not different, unlike in our  
338 previous study (Deglaire, et al., 2016), however PHM abundance was greater at 6 out of the  
339 11 Lys residues (plasmin cleavage specificity) of the  $\beta$ -casein sequence. As discussed  
340 previously (Deglaire, et al., 2016), pasteurization may have activated the milk plasmin system

341 through inactivation of plasminogen activator inhibitors and plasmin inhibitors (Ismail &  
342 Nielsen, 2010). Besides, denaturation of lactoferrin by pasteurization, as confirmed by reverse  
343 HPLC (signal decreased by 3 after pasteurization; unpublished data), may have contributed to  
344 this activation, as it acts as a natural inhibitor of plasminogen activation (Zwirzitz, Reiter,  
345 Skrabana, Ohradanova-Repic, Majdic, Gutekova, et al., 2018). It should be noted that the  
346 arginine and lysine residues are also cleavage specificities of kallikrein and thrombin, which  
347 have been shown to be active in preterm human milk (Demers-Mathieu, Nielsen, Underwood,  
348 Borghese, & Dallas, 2017). In addition, the arginine and lysine residues are the cleavage sites  
349 for trypsin that has been shown to be potentially present in human milk but not active  
350 (Demers-Mathieu, Nielsen, Underwood, Borghese, & Dallas, 2017).

351 A greater abundance for PHM was observed between the  $\beta$ -casein residues 130-145, although  
352 this was statistically significant only between residues 136-138. This area includes the  
353 hydrophobic amino acids that are leucine, valine, isoleucine, which are cleavage sites for both  
354 cathepsin D and elastase (Vijayakumar, Guerrero, Davey, Lebrilla, Shields, & Khaldi, 2012).  
355 However, cathepsin D, which is an acid protease, is mostly present in human milk as  
356 procathepsin D, i.e. the inactive zymogen (Vijayakumar, Guerrero, Davey, Lebrilla, Shields,  
357 & Khaldi, 2012). On the contrary, elastase has been shown to be active in human milk  
358 recently (Demers-Mathieu, Nielsen, Underwood, Borghese, & Dallas, 2017) and may have  
359 been more active in PHM in this area due to the heat-induced structural changes of  $\beta$ -casein  
360 (Guyomarc'h, Law, & Dalgleish, 2003) unmasking resistant area and possibly compensating  
361 for the loss of activity likely induced by heat treatment. For the same reason, cystosol  
362 aminopeptidase may also have been more active in PHM particularly at the residues 136 and  
363 137 (Supplemental Material 3), where the abundance was 12 times greater for PHM. Because  
364 cystosol aminopeptidase is a minor protease in milk, particularly in bovine milk, few  
365 information are available on the impact of heat treatment on its activity (O'Mahony, Fox, &

366 Kelly, 2016). Some pepsin and chymotrypsin activities have been predicted in the present  
367 paper, likely due to their overlapping specificity with elastase (Vijayakumar, Guerrero,  
368 Davey, Lebrilla, Shields, & Khaldi, 2012) as they have not been shown to be present in  
369 human milk (Demers-Mathieu, Nielsen, Underwood, Borghese, & Dallas, 2018).

370 Overall, as reported earlier (Demers-Mathieu, Nielsen, Underwood, Borghese, & Dallas,  
371 2017), the present work suggests that indigenous enzymes, such as plasmin, elastase and  
372 probably cytosol aminopeptidase, were more active in pasteurized milk. As a result, some  
373 peptides were more actively produced (i.e. more abundant) in PHM or, on the contrary, more  
374 extensively degraded, such as observed for the antimicrobial peptide,  $\beta$ -casein(185-211). The  
375 latter was less abundant in PHM prior to digestion, maybe due to a greater activity of cytosol  
376 amino peptidase in this area (Supplemental Material 3). Nevertheless, this does not indicate  
377 whether the proteolysis degree prior to digestion was different between RHM and PHM.

378 Pre-proteolysis of  $\beta$ -casein may have contributed to its further digestion, in particular on the  
379 first 20 amino acids where the cumulative abundance remained similar before and after 30  
380 min of gastric digestion. Indigenous enzymes, in particular the acid enzyme cathepsin D,  
381 could have contributed to gastric digestion although a minimal contribution has been  
382 previously assumed (Demers-Mathieu, Nielsen, Underwood, Borghese, & Dallas, 2018).

## 383 4.2 Human milk digestion

384 The distribution of the protein origin of the common peptides was similar to that observed in  
385 our previous study at the term stage (Deglaire, et al., 2016), with the major protein of human  
386 milk,  $\beta$ -casein, generating 42% of the peptides, whilst  $\alpha$ -lactalbumin, the other major protein  
387 in human milk, generating only 3% of the peptides. This can reflect its proteolysis resistance  
388 as observed by SDS-PAGE (De Oliveira, et al., 2016a) but also probably the non-  
389 identification of the peptides engaged in a disulphide bond (four out of 123 amino acids of the  
390 protein sequence) with our non-reducing analysis. Such as observed previously (Deglaire, et

391 al., 2016), despite being minor proteins in terms of concentrations, bile-salt stimulated lipase  
392 and osteopontin generated almost the same number of peptides as lactoferrin and  $\alpha_{s1}$ -casein.  
393 Nevertheless, this is not indicative of the proteolysis degree of these proteins.

394 In the present study, the proportion of peptides predominantly released in the gastric phase  
395 (36% of the peptides in 2 clusters) was lower than in the intestinal phase, while the opposite  
396 was true in the term study (Deglaire, et al., 2016). This is certainly resulting from the  
397 decreased level of pepsin (2 times lower) combined to a greater protein content (1.2 times  
398 higher) and a faster gastric emptying ( $t_{1/2}$  of 36 min) for the preterm model compared to the  
399 term one ( $t_{1/2}$  of 47 min) (Bourlieu, Menard, Bouzerzour, Mandalari, Macierzanka, Mackie, et  
400 al., 2014).

401 In the gastric phase, peptide abundances were generally lower for PHM than for RHM,  
402 particularly for peptides from bile salt-stimulated lipase, such as observed on both peptidomic  
403 mapping and clustering. On the contrary, intestinal peptide abundances from bile-salt  
404 stimulated lipase became in average higher for PHM than for RHM, as observed on the  
405 protein sequence and in clusters 7 and 8 (gathering 50% of its peptides). These gastro-  
406 intestinal kinetics of peptide release are in line with that previously at the term stage  
407 (Deglaire, et al., 2016). This indicates a delay in the proteolysis of bile salt-stimulated lipase  
408 from PHM unlike that for RHM, maybe due to some increased resistance to gastric  
409 proteolysis due to some heat-induced structural changes, probably masking pepsin cleavage  
410 sites. This heat-denaturation is already known to drastically inactivate this indigenous lipase  
411 (Henderson, Fay, & Hamosh, 1998). Besides, it can be noted that no peptide was found from  
412 the residue 526 until the C-ter end of bile-salt stimulated lipase. This region includes a high  
413 number of proline and has been described, from residue 538, as a mucin-like flexible C-  
414 terminal tail with a high level of glycosylation (Downs, Xu, Tang, Wang, 1994), probably  
415 preventing it from proteolysis but also from MS/MS peptide identification.

416

417 Regarding lactoferrin, data suggested a greater and faster release of its peptides in PHM,  
418 resulting in the intestinal phase in peptides being less abundant in PHM than in RHM (cluster  
419 5 and protein sequence in supplemental Material 4). This can indicate a lower resistance to  
420 intestinal proteolysis for the heat-denaturated lactoferrin, unlike that for the native lactoferrin,  
421 a globular protein, known to be resistant to proteolysis such as discussed earlier (Deglaire et  
422 al., 2016). The concentration of native lactoferrin reaching the intestinal epithelium is thus  
423 likely to differ between PHM and RHM, which can modulate its bioactivity. As reported by  
424 Buccigrossi et al. (2007), high lactoferrin concentrations induced epithelial cell proliferation  
425 while low lactoferrin concentrations stimulated intestinal differentiation. In addition, the  
426 greater persistence of intact peptides in the gastro-intestinal tract may have physiological  
427 consequences, as it is known to be important for gut-associated-lymphoid tissue (GALT)  
428 education and microbiota shaping (Chatterton, Nguyen, Bering, & Sangild, 2013; Newburg &  
429 Walker, 2007). It should be noted that the end of our intestinal digestion may not be fully  
430 representative of the *in vivo* condition, as brush border peptidases, not included in our model,  
431 can further hydrolyse the present peptides. Thus, pasteurization may impact the regulation of  
432 the early postnatal intestinal development, through modulation of lactoferrin proteolysis, but  
433 further investigations are needed.

434 Regarding  $\beta$ -casein, data suggested that its peptide abundances were in average not different  
435 between RHM and PHM and its kinetics of peptide release were opposite to those for  
436 lactoferrin or bile salt-stimulated lipase. This is well in agreement with our previous  
437 observations in the term study (Deglaire, et al., 2016). Nevertheless, there was no emerging  
438 cluster significantly associated to  $\beta$ -casein with an increased gastric abundance for PHM in  
439 the present study, while this was true in the term study for a limited proportion of the peptides

440 (11% of the total peptides, 17% of the  $\beta$ -casein peptides) (Deglaire, et al., 2016). The reasons  
441 for this remain unknown.

442 Overall, the different kinetics of peptide release between PHM and RHM may have been  
443 somewhat compensating, together with smaller and larger peptides not taken into account  
444 here, as the overall degree of gastro-intestinal proteolysis (based on the primary amine  
445 content) was not significantly different but the intestinal bioaccessibility of some amino acids  
446 (lysine, phenylalanine, tyrosine) was increased by pasteurization (De Oliveira, et al., 2016a).

### 447 **4.3 Bioactive peptides**

448 The present bioactive peptides were identified only after exact matching sequence, although  
449 other bioactive peptides can be encrypted within larger peptides and further released *in vivo*  
450 by brush border enzymes not included in the present model. Most potential bioactive peptides  
451 were predicted, but not experimentally demonstrated, as having an antimicrobial activity.  
452 Only a limited number of bioactive peptides, mostly from  $\beta$ -casein, was based on a  
453 demonstrated activity in the literature. Prior to digestion, two bioactive peptides, the  
454 caseinophosphopeptide (CPP) homologue  $\beta$ -casein(1-18) (Wada & Lonnerdal, 2015) and the  
455 antimicrobial peptide  $\beta$ -casein(185-211), were present in human milk as observed previously  
456 (Deglaire, et al., 2016; Dingess, et al., 2017; Wada & Lonnerdal, 2015). Their abundance  
457 were lower in PHM than in RHM prior digestion, unlike that observed in the term study  
458 (Deglaire, et al., 2016). Whether this is due to the term vs. preterm condition remains  
459 unknown. Both peptides were hydrolysed rapidly in the intestinal phase in both types of milk,  
460 unlike that observed by Wada et al. (2015). This can result from milk variability but also from  
461 the simpler and shorter model of digestion (static conditions) used by Wada et al. (2015).  
462 Regarding lactoferrin, only one opioid antagonist peptide was released during the gastric  
463 phase in a similar manner between RHM and PHM. No peptide was detected between the  
464 residues 1-49 that includes the antimicrobial lactoferricin peptide (Zwirzitz, et al., 2018),

465 probably due to the presence of disulphide bonds between Cys<sup>19</sup> and Cys<sup>36</sup> and between Cys<sup>9</sup>  
466 and Cys<sup>45</sup> (Farnaud & Evans, 2003; Wada & Lonnerdal, 2015). However, looking at  
467 individual bioactive peptides and at their abundances should be considered with care; further  
468 confirmation would be required.

## 469 5. Conclusion

470 To conclude, the present study demonstrated that pasteurization impacted the human milk  
471 peptidome prior to digestion mainly originating from  $\beta$ -casein and induced different kinetics  
472 of peptide release during gastro-intestinal digestion mainly for the heat-denaturated proteins  
473 (bile-salt stimulated lipase and lactoferrin), such as observed in the term stage. When  
474 considering the overall proteolysis, there was no impact of pasteurization; however, it may  
475 induce different concentrations of intact peptides within the gastro-intestinal, which may have  
476 further physiological consequences.

477 **Acknowledgments**

478 The authors thank Frederic Carrière and GERME SA for providing rabbit gastric extract. The  
479 authors also thank the volunteer mothers for their  
480 donations of breast milk samples as well as Emelyne Dirson and her colleagues from the bank  
481 milk of the Rennes University Hospital Center for their involvement in the project. This work  
482 was supported by INRA (French National Institute for Agricultural Research) as well as by  
483 CNPq (Brazilian National Council for Scientific and Technological Development) by funding  
484 a Ph.D. scholarship for the author S. D. O. Most authors were participants of the Food and  
485 Agriculture COST (European Cooperation in Science and Technology) Action FA1005  
486 “Improving health properties of food by sharing our knowledge on the digestive process  
487 (INFOGEST).” The authors have declared no conflict of interest.



488 **List of tables and figures**

489 **Table 1:** Predicted enzymes and cleavage counts for proteolysis of human milk as raw or  
490 pasteurized prior to digestion.

491 **Table 2.** Peptide biochemical characteristics significantly associated ( $p < 0.05$ ) with clusters.

492 **Table 3.** Bioactive peptides identified within the entire peptide dataset ( $n=1531$ ) using an  
493 exact matching search the databases of BIOPEP and MBPDB.

494 **Figure 1.** Mapping of the average cumulative peptide abundances of RHM and PHM onto  
495 their parent proteins **A**) in the undigested milk and over the gastric (G) and intestinal (I)  
496 phases or **B**) only in the G and I phases. Peptide abundances, summed at each amino acid, were  $\log_{10}$ -  
497 transformed and averaged over the three digestion replicates of RHM or PHM and over the digestion times of G  
498 or I phase (digestion time in min). T-tests were conducted at each amino acid between RHM and PHM  $\log_{10}$ -  
499 transformed abundances. (\*,  $P < 0.05$ , \*\*,  $P < 0.01$ , \*\*\*,  $P < 0.001$ )

500 **Figure 2.** **A**) Protein origin of the peptides by cluster obtained during gastrointestinal  
501 digestion of raw (RHM) and pasteurized (PHM) human milk, **B**) Heatmap of the “  
502 Pasteurization x Digestion time ” interaction coefficients as classified by ascending  
503 hierarchical classification , **C**) average abundances of the peptides by cluster.  $n = 1134$   
504 peptides; G: gastric phase, I: intestinal phase, digestion time in min.

505 **Supporting information**

506 **Supplemental material 1.** Proteins ( $n = 27$ ) from which derived the identified peptides  
507 before and during digestion of raw and pasteurized milk.

508 **Supplemental material 2.** Dataset of all the unique peptides ( $n = 1531$ ) identified before and  
509 during digestion of raw and pasteurized milk.

510 **Supplemental material 3.** Mapping of  $\beta$ -casein hydrolysis prior to digestion with cleavage  
511 sites of plasmin, cytosol amino peptidase and carboxypeptidase B2.

512 **Supplemental material 4.** Mapping of the average peptide abundances of RHM and PHM  
513 onto their parent proteins at 180 min of intestinal digestion. Peptide abundances, summed at each  
514 amino acid, were averaged over the three digestion replicates of RHM or PHM at 180 min of intestinal digestion.

**Table 1:** Predicted enzymes and counts for proteolysis of human milk as raw or pasteurized before gastrointestinal digestion<sup>1</sup>

Enzymes	Raw human milk (n = 161 peptides)				Pasteurised human milk (n = 171 peptides)					
	Total number of cleavages	Number of expected cleavages within peptides	Number of proteins cleaved	Odds Ratio <sup>2</sup>	Standard Error <sup>3</sup>	Total number of cleavages	Number of expected cleavages within peptides	Number of proteins cleaved	Odds Ratio <sup>2</sup>	Standard Error <sup>3</sup>
Plasmin	87	137	5	4.7	1.2	85	137	5	4.8	1.2
Trypsin 1	87	137	5	4.7	1.2	85	137	5	4.8	1.2
Cathepsin D	61	769	4	0.4	1.2	71	788	3	0.5	1.1
Pepsin (pH1.3) 1	56	238	4	1.5	1.2	58	254	4	1.6	1.2
Chymotrypsin low 1	48	300	4	1.0	1.2	44	304	3	0.9	1.2
Elastase	45	413	2	0.6	1.2	39	428	3	0.6	1.2
Pepsin (pH1.3) 2	34	232	3	0.9	1.2	27	241	2	0.7	1.2
Proline endopeptidase	18	492	2	0.2	1.3	21	505	2	0.2	1.3

<sup>1</sup> Results obtained from the online software Enzyme Predictor. Cleavages from cytosol aminopeptidase and carboxypeptidase B2 were examined manually on  $\beta$ -casein and appear only in supplementatl figure 3.

<sup>2</sup> Odds ratio (OR) were calculated on the present enzymes as  $OR = (a * d) / (b * c)$ , where a is the total number of sites cleaved by enzyme “A” at the peptide termini (“total number of cleavages”), b is the total number of all sites cleaved by all enzymes except “A”; c is the total number of sites enzyme “A” could have cleaved within the current peptides (“number of expected cleavages within peptides”), and d is the number of sites all enzymes, aside from A, could have cleaved within the current peptides.

<sup>3</sup> Standard errors (SE) were calculated as :  $SE = \exp(1/a + 1/b + 1/c + 1/d)^{1/2}$  (Vijayakumar, Guerrero, Davey, Lebrilla, Shields, & Khaldi, 2012).

**Table 2.** Peptide biochemical characteristics significantly associated ( $p < 0.05$ ) with clusters<sup>1,2</sup>.

	Qualitative modalities					Quantitative modalities					
	Modality	Cl/Mod <sup>3</sup>	Mod/Cl <sup>4</sup>	Overall <sup>5</sup> (n= 1134)	v.test	p.value	Modality	Mean $\pm$ SD in category	Mean $\pm$ SD in overall (n= 1134)	v.test	p.value
<b>Cluster 1</b>	protein = CASA1	14.43	19.44	8.55	3.01	0.0027					
	P1' = M	19.23	6.94	2.29	2.24	0.0251					
	P1' = S	15.22	9.72	4.06	2.18	0.0290					
	P1 = R	11.02	19.44	11.20	2.13	0.0336					
	protein = FAS	40.00	2.78	0.44	2.08	0.0373					
	protein = TRFL	11.11	16.67	9.52	1.97	0.0483					
	P1 = T	0.00	0.00	4.14	-2.02	0.0429					
	P1' = K	0.00	0.00	4.59	-2.16	0.0304					
P1' = D	0.00	0.00	4.59	-2.16	0.0304						
<b>Cluster 2</b>	protein = AMY2B	42.86	2.86	0.62	2.28	0.0224	aliphatic index	94.41 $\pm$ 42.47	87.58 $\pm$ 37.07	1.98	0.0476
	P1' = L	13.82	20.00	13.40	1.98	0.0472	MW <sup>6</sup>	1.75 $\pm$ 0.67	1.92 $\pm$ 0.8	-2.26	0.0235
	P1' = P	1.67	0.95	5.29	-2.31	0.0212	isoelectric point	5.4 $\pm$ 1.55	5.85 $\pm$ 1.99	-2.45	0.0143
<b>Cluster 3</b>	protein = BT1A1	22.22	12.50	2.38	3.39	0.0007					
	P1' = G	14.29	12.50	3.70	2.64	0.0082					
	protein = FAS	40.00	4.17	0.44	2.39	0.0168					
	P1' = K	11.54	12.50	4.59	2.25	0.0241					
	P1' = D	11.54	12.50	4.59	2.25	0.0241					
	protein = AMY2B	28.57	4.17	0.62	2.12	0.0344					
	P1 = F	9.21	14.58	6.70	1.98	0.0482					
	P1 = L	1.79	8.33	19.66	-2.12	0.0338	aliphatic index	76.84 $\pm$ 42.38	87.58 $\pm$ 37.07	-2.05	0.0403
	protein = CASB	2.74	27.08	41.89	-2.14	0.0324	GRAVY	-0.7 $\pm$ 0.87	-0.42 $\pm$ 0.74	-2.62	0.0088
	protein = CEL	0.00	0.00	7.50	-2.29	0.0219	isoelectric point	4.9 $\pm$ 1.43	5.85 $\pm$ 1.99	-3.37	0.0007
	P1' = L	0.66	2.08	13.40	-2.64	0.0084					
<b>Cluster 4</b>	protein = CASB	38.53	51.55	41.89	4.43	<0.0001					
	protein = CASK	55.81	6.76	3.79	3.36	0.0008					
	P1' = N	53.85	5.92	3.44	2.94	0.0033					
	P1' = H	54.84	4.79	2.73	2.72	0.0066					
	P1 = M	18.87	2.82	4.67	-2.04	0.0412	GRAVY	-0.52 $\pm$ 0.76	-0.42 $\pm$ 0.74	-3.04	0.0024
	P1 = F	17.11	3.66	6.70	-2.86	0.0042					
	protein = CEL	8.24	1.97	7.50	-5.23	<0.0001					
	protein = TRFL	6.48	1.97	9.52	-6.51	<0.0001					
<b>Cluster 5</b>	protein = TRFL	25.00	45.76	9.52	7.53	<0.0001					
	P1 = K	13.37	38.98	15.17	4.58	<0.0001					
	P1' = T	14.29	10.17	3.70	2.27	0.0232					
	P1 = R	9.45	20.34	11.20	2.10	0.0355					
	protein = CASA1	1.03	1.69	8.55	-2.11	0.0347	essential AA	5.07 $\pm$ 1.57	7.02 $\pm$ 3.65	-4.23	<0.0001
protein = CASB	2.95	23.73	41.89	-2.96	0.0031	MW	1.42 $\pm$ 0.44	1.92 $\pm$ 0.8	-4.97	<0.0001	
<b>Cluster 6</b>	P1 = D	65.00	5.70	1.76	4.34	<0.0001	essential AA	8.62 $\pm$ 3.98	7.02 $\pm$ 3.65	7.39	<0.0001
	P1 = L	29.15	28.51	19.66	3.63	0.0003	MW	2.24 $\pm$ 0.85	1.92 $\pm$ 0.8	6.63	<0.0001
	P1' = P	40.00	10.53	5.29	3.62	0.0003	isoelectric point	6.46 $\pm$ 2.21	5.85 $\pm$ 1.99	5.20	<0.0001
	P1 = M	33.96	7.89	4.67	2.42	0.0157	GRAVY	-0.29 $\pm$ 0.63	-0.42 $\pm$ 0.74	2.98	0.0030
	protein = ALBU	50.00	2.63	1.06	2.27	0.0232					
	P1' = K	9.62	2.19	4.59	-2.01	0.0440					
	P1' = N	7.69	1.32	3.44	-2.07	0.0383					
	protein = OSTP	10.87	4.39	8.11	-2.41	0.0160					
	P1 = Y	9.68	3.95	8.20	-2.77	0.0056					
	protein = CASK	4.65	0.88	3.79	-2.84	0.0046					
	P1 = R	11.02	6.14	11.20	-2.84	0.0045					
P1 = K	11.05	8.33	15.17	-3.37	0.0007						
<b>Cluster 7</b>	protein = CEL	20.00	18.68	7.50	3.66	0.0002					
	P1' = K	19.23	10.99	4.59	2.65	0.0082	isoelectric point	5.4 $\pm$ 1.61	5.85 $\pm$ 1.99	-2.24	0.0254
	P1' = G	19.05	8.79	3.70	2.33	0.0198	essential AA	5.51 $\pm$ 2.58	7.02 $\pm$ 3.65	-4.14	<0.0001
	protein = LALBA	19.35	6.59	2.73	2.04	0.0414	MW	1.57 $\pm$ 0.65	1.92 $\pm$ 0.8	-4.39	<0.0001
	protein = CASB	6.11	31.87	41.89	-2.03	0.0425					
<b>Cluster 8</b>	protein = CEL	32.94	15.91	7.50	4.17	<0.0001					
	P1' = T	28.57	6.82	3.70	2.19	0.0286					
	P1' = P	25.00	8.52	5.29	1.96	0.0498					
	P1' = G	2.38	0.57	3.70	-2.68	0.0073					

<sup>1</sup> Abbreviations: AA: amino acid, ALBU : serum albumin, AM2YB :  $\alpha$ -amylase 2B, BT1A1 : butyrophilin subfamily 1 member A1, CASA1:  $\alpha$ s1-casein, CASB :  $\beta$ -casein, CASK : K-casein, CEL : carboxyl ester lipase or bile salt-stimulated lipase, FAS: fatty acid synthase, LALBA :  $\alpha$ -lactalbumin, OSTP: osteopontin, TRFL: lactoferrin.

<sup>2</sup> Grey lines correspond to modalities that are negatively associated to a cluster

<sup>3</sup> Cla/Mod: occurrence of the modality category in the cluster divided by its occurrence in the entire dataset

<sup>4</sup> Mod/Cla: proportion of the modality category within the cluster

<sup>5</sup> Overall proportion of this modality category within the entire dataset

<sup>6</sup> MW: Molecular Weight, expressed in kDa.

**Table 3.** Bioactive peptides identified within the entire dataset (n = 1531) using an exact matching search within the databases of BIOPEP and MBPDB

Parent protein	Peptide position <sup>1</sup>	Peptide sequence	Activity	Pasteurization effect <sup>2</sup>	Time effect <sup>2</sup>	Pasteurization x Time effect <sup>2</sup>
	51-58	YPFVEPIP	opioid agonist	NS	***	NS
	54-59	VEPIPY	immunomodulating	NS	***	NS
β-casein	105-117	SPTIPFFDPQIPK	stimulating proliferation	NS	***	NS
	125-130	HLPLPL	anti-amnestic	NS	***	NS
	154-160	WSVPQPK	ACE inhibitor/ antioxidant	***	***	***
	161-166	VLPIPQ	ACE inhibitor	NS	***	*
	185-211	QELLLNPTHQIYPVTQPLAPVHNPISV	antimicrobial	NS	***	**
Lactoferrin	319-324	YLGSGY	opioid antagonist	NS	***	NS

<sup>1</sup> Peptide position on the parent protein without taking into account the signal peptide.

<sup>2</sup> Effects were determined using a two-way analysis of variance with Pasteurization, Time and their interaction as factors.

## List of references

- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B*, 57, 289–300.
- Bourlieu, C., Menard, O., Bouzerzour, K., Mandalari, G., Macierzanka, A., Mackie, A. R., & Dupont, D. (2014). Specificity of infant digestive conditions: Some clues for developing relevant in vitro models. *Critical Reviews in Food Science and Nutrition*, 54(11), 1427-1457.
- Buccigrossi, V., de Marco, G., Bruzzese, E., Ombrato, L., Bracale, I., Polito, G., & Guarino, A. (2007). Lactoferrin induces concentration-dependent functional modulation of intestinal proliferation and differentiation. *Pediatric Research*, 61, 410.
- Chatterton, D. E., Nguyen, D. N., Bering, S. B., & Sangild, P. T. (2013). Anti-inflammatory mechanisms of bioactive milk proteins in the intestine of newborns. *International Journal of Biochemistry and Cell Biology*, 45(8), 1730-1747.
- Dallas, D. C., Guerrero, A., Khaldi, N., Castillo, P. A., Martin, W. F., Smilowitz, J. T., Bevins, C. L., Barile, D., German, J. B., & Lebrilla, C. B. (2013). Extensive in vivo human milk peptidomics reveals specific proteolysis yielding protective antimicrobial peptides. *Journal of Proteome Research*, 12(5), 2295-2304.
- Dallas, D. C., Smink, C. J., Robinson, R. C., Tian, T., Guerrero, A., Parker, E. A., Smilowitz, J. T., Hettinga, K. A., Underwood, M. A., Lebrilla, C. B., German, J. B., & Barile, D. (2015). Endogenous human milk peptide release is greater after preterm birth than term birth. *Journal of Nutrition*, 145(3), 425-433.
- de Oliveira, S. C., Bellanger, A., Ménard, O., Pladys, P., Le Gouar, Y., Dirson, E., Kroell, F., Dupont, D., Deglaire, A., & Bourlieu, C. (2017). Impact of human milk pasteurization on gastric digestion in preterm infants: A randomized controlled trial. *American Journal of Clinical Nutrition*, 105(2), 379-390.

- de Oliveira, S. C., Bourlieu, C., Ménard, O., Bellanger, A., Henry, G., Rousseau, F., Dirson, E., Carrière, F., Dupont, D., & Deglaire, A. (2016a). Impact of pasteurization of human milk on preterm newborn in vitro digestion: Gastrointestinal disintegration, lipolysis and proteolysis. *Food Chemistry*, *211*, 171-179.
- de Oliveira, S. C., Deglaire, A., Ménard, O., Bellanger, A., Rousseau, F., Henry, G., Dirson, E., Carrière, F., Dupont, D., & Bourlieu, C. (2016b). Holder pasteurization impacts the proteolysis, lipolysis and disintegration of human milk under in vitro dynamic term newborn digestion. *Food Research International*, *88*, 263-275.
- Deglaire, A., De Oliveira, S., Jardin, J., Briard-Bion, V., Emily, M., Ménard, O., Bourlieu, C., & Dupont, D. (2016). Impact of human milk pasteurization on the kinetics of peptide release during in vitro dynamic term newborn digestion. *Electrophoresis*.
- Deglaire, A., Fromentin, C., Fouillet, H., Airinei, G., Gaudichon, C., Boutry, C., Benamouzig, R., Moughan, P. J., Tome, D., & Bos, C. (2009). Hydrolyzed dietary casein as compared with the intact protein reduces postprandial peripheral, but not whole-body, uptake of nitrogen in humans. *American Journal of Clinical Nutrition*.
- Demers-Mathieu, V., Nielsen, S. D., Underwood, M. A., Borghese, R., & Dallas, D. C. (2017). Analysis of milk from mothers who delivered prematurely reveals few changes in proteases and protease inhibitors across gestational age at birth and infant postnatal age. *Journal of Nutrition*, *147*(6), 1152-1159.
- Demers-Mathieu, V., Nielsen, S. D., Underwood, M. A., Borghese, R., & Dallas, D. C. (2018). Changes in proteases, antiproteases, and bioactive proteins from mother's breast milk to the premature infant stomach. *Journal of Pediatric Gastroenterology and Nutrition*, *66*(2), 318-324.
- Dingess, K. A., de Waard, M., Boeren, S., Vervoort, J., Lambers, T. T., van Goudoever, J. B., & Hettinga, K. (2017). Human milk peptides differentiate between the preterm and term infant and across varying lactational stages. *Food & Function*, *8*(10), 3769-3782.



Downs, D., Xu, Y. Y., Tang J., Wang C. S. (1994). Proline-rich domain and glycosylation are not essential for the enzymatic activity of bile salt-activated lipase. Kinetic studies of T-BAL, a truncated form of the enzyme, expressed in *Escherichia coli*. *Biochemistry*, 33(26), 7979–7985

Farnaud, S., & Evans, R. W. (2003). Lactoferrin—a multifunctional protein with antimicrobial properties. *Molecular Immunology*, 40(7), 395-405.

Ferranti, P., Traisci, M. V., Picariello, G., Nasi, A., Boschi, V., Siervo, M., Falconi, C., Chianese, L., & Addeo, F. (2004). Casein proteolysis in human milk: Tracing the pattern of casein breakdown and the formation of potential bioactive peptides. *Journal of Dairy Research*, 71(1), 74-87.

Gabere, M. N., & Noble, W. S. (2017). Empirical comparison of web-based antimicrobial peptide prediction tools. *Bioinformatics*, 33(13), 1921-1929.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein identification and analysis tools on the expasy server. In J. M. Walker (Ed.), *The proteomics protocols handbook*, (pp. 571-607 ): Humana Press.

Guyomarc'h, F., Law, A. J. R., & Dalgleish, D. G. (2003). Formation of soluble and micelle-bound protein aggregates in heated milk. *Journal of Agricultural and Food Chemistry*, 51(16), 4652-4660.

Henderson, T. R., Fay, T. N., & Hamosh, M. ( 1998). Effect of pasteurization on long chain polyunsaturated fatty acid levels and enzyme activities of human milk. *Journal of Pediatrics*(132), 876-878.

Ismail, B., & Nielsen, S. S. (2010). Invited review: Plasmin protease in milk: Current knowledge and relevance to dairy industry. *Journal of Dairy Science*, 93(11), 4999-5009.

Khaldi, N., Vijayakumar, V., Dallas, D. C., Guerrero, A., Wickramasinghe, S., Smilowitz, J. T., Medrano, J. F., Lebrilla, C. B., Shields, D. C., & German, J. B. (2014). Predicting the

- important enzymes in human breast milk digestion. *Journal of Agricultural and Food Chemistry*, 62(29), 7225-7232.
- Kyte, J., & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*, 157(1), 105-132.
- Langella, O., Valot, B., Balliau, T., Blein-Nicolas, M., Bonhomme, L., & Zivy, M. (2017). X!Tandempipeline: A tool to manage sequence redundancy for protein inference and phosphosite identification. *Journal of Proteome Research*, 16(2), 494-503.
- Minkiewicz, P., Dziuba, J., Iwaniak, A., Dziuba, M., & Darewicz, M. (2008). Biopep database and other programs for processing bioactive peptide sequences. *Journal of AOAC International*, 91, 965-980.
- Molinari, C. E., Casadio, Y. S., Hartmann, B. T., Livk, A., Bringans, S., Arthur, P. G., & Hartmann, P. E. (2012). Proteome mapping of human skim milk proteins in term and preterm milk. *Journal of Proteome Research*, 11(3), 1696-1714.
- Newburg, D. S., & Walker, W. A. (2007). Protection of the neonate by the innate immune system of developing gut and of human milk. *Pediatric Research*, 61(1), 2-8.
- Nielsen, S. D., Beverly, R. L., & Dallas, D. C. (2017). Milk proteins are predigested within the human mammary gland. *Journal of Mammary Gland Biology and Neoplasia*, 22(4), 251-261.
- Nielsen, S. D., Beverly, R. L., Qu, Y. Y., & Dallas, D. C. (2017). Milk bioactive peptide database: A comprehensive database of milk protein-derived bioactive peptides and novel visualization. *Food Chemistry*, 232, 673-682.
- O'Connor, D. L., Ewaschuk, J. B., & Unger, S. (2015). Human milk pasteurization: Benefits and risks. *Current Opinion in Clinical Nutrition and Metabolic Care*, 18(3), 269-275.
- O'Mahony, J. A., Fox, P. F., & Kelly, A. L. (2016). Indigenous enzymes of milk. In P. L. H. McSweeney & P. F. Fox (Eds.), *Advanced dairy chemistry. Proteins : Basic aspects.*, vol. 1A (pp. 337-386): Springer.

- Peila, C., Moro, G. E., Bertino, E., Cavallarín, L., Giribaldi, M., Giuliani, F., Cresi, F., & Coscia, A. (2016). The effect of holder pasteurization on nutrients and biologically-active components in donor human milk: A review. *Nutrients*, 8(8).
- R Core Team. (2014). R: A language and environment for statistical computing. In *R Foundation for Statistical Computing*. Vienna, Austria.
- Valot, B., Langella, O., Nano, E., & Zivy, M. (2011). Masschroq: A versatile tool for mass spectrometry quantification. *Proteomics*, 11(17), 3572-3577.
- Victora, C. G., Bahl, R., Barros, A. J. D., França, G. V. A., Horton, S., Krasevec, J., Murch, S., Sankar, M. J., Walker, N., & Rollins, N. C. (2016). Breastfeeding in the 21st century: Epidemiology, mechanisms, and lifelong effect. *The Lancet*, 387(10017), 475-490.
- Vijayakumar, V., Guerrero, A. N., Davey, N., Lebrilla, C. B., Shields, D. C., & Khaldi, N. (2012). Enzymepredictor: A tool for predicting and visualizing enzymatic cleavages of digested proteins. *Journal of Proteome Research*, 11(12), 6056-6065.
- Wada, Y., & Lonnerdal, B. (2015). Bioactive peptides released from in vitro digestion of human milk with or without pasteurization. *Pediatric Research*, 77(4), 546-553.
- Wight, N. E. (2001). Donor human milk for preterm infants. *Journal of Perinatology*, 21(4), 249-254.
- Zwirzitz, A., Reiter, M., Skrabana, R., Ohradanova-Repic, A., Majdic, O., Gutekova, M., Cehlar, O., Petrovcikova, E., Kutejova, E., Stanek, G., Stockinger, H., & Leksa, V. (2018). Lactoferrin is a natural inhibitor of plasminogen activation. *The Journal of Biological Chemistry*.

Figure 1  
Deglaire

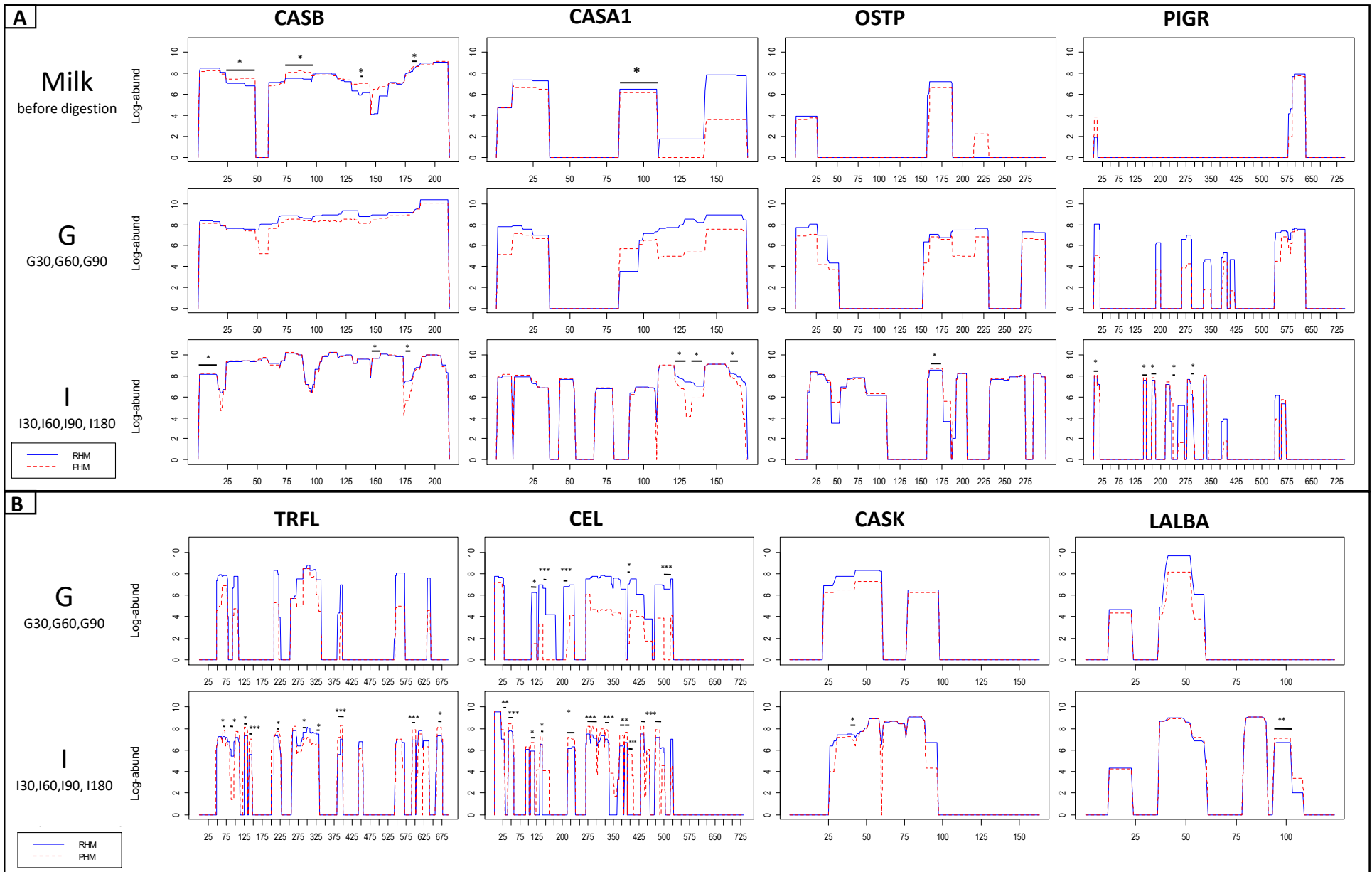
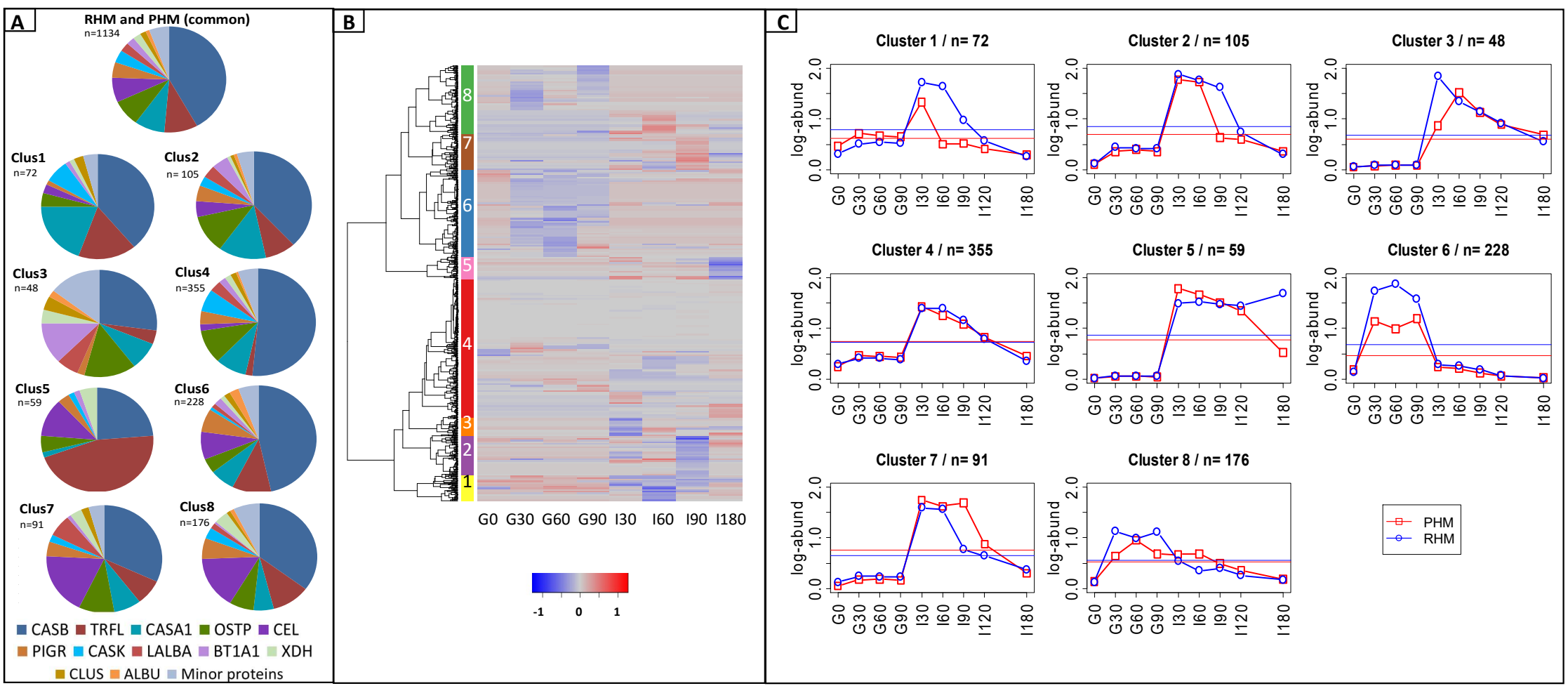


Figure 2  
Deglaire



	Qualitative modalities					Quantitative modalities					
	Modality	Cla/Mod <sup>3</sup>	Mod/Cla <sup>4</sup>	Overall <sup>5</sup> (n=1134)	v.test	p.value	Modality	Mean ± SD in category	Mean ± SD in overall (n=1134)	v.test	p.value
<b>Cluster 1</b>	protein = CASA1	14.43	19.44	8.55	3.01	0.0027					
	P1' = M	19.23	6.94	2.29	2.24	0.0251					
	P1' = S	15.22	9.72	4.06	2.18	0.0290					
	P1 = R	11.02	19.44	11.20	2.13	0.0336					
	protein = FAS	40.00	2.78	0.44	2.08	0.0373					
	protein = TRFL	11.11	16.67	9.52	1.97	0.0483					
	P1 = T	0.00	0.00	4.14	-2.02	0.0429					
	P1' = K	0.00	0.00	4.59	-2.16	0.0304					
	P1' = D	0.00	0.00	4.59	-2.16	0.0304					
<b>Cluster 2</b>	protein = AMY2B	42.86	2.86	0.62	2.28	0.0224		94.41 ±			
	P1' = L	13.82	20.00	13.40	1.98	0.0472	aliphatic index	42.47	87.58 ± 37.07	1.98	0.0476
	P1' = P	1.67	0.95	5.29	-2.31	0.0212	MW <sup>6</sup>	1.75 ± 0.67	1.92 ± 0.8	-2.26	0.0235
<b>Cluster 3</b>	protein = BT1A1	22.22	12.50	2.38	3.39	0.0007					
	P1' = G	14.29	12.50	3.70	2.64	0.0082					
	protein = FAS	40.00	4.17	0.44	2.39	0.0168					
	P1' = K	11.54	12.50	4.59	2.25	0.0241					
	P1' = D	11.54	12.50	4.59	2.25	0.0241					
	protein = AMY2B	28.57	4.17	0.62	2.12	0.0344					
	P1 = F	9.21	14.58	6.70	1.98	0.0482					
	P1 = L	1.79	8.33	19.66	-2.12	0.0338		76.84 ±			
	protein = CASB	2.74	27.08	41.89	-2.14	0.0324	aliphatic index	42.38	87.58 ± 37.07	-2.05	0.0403
	protein = CEL	0.00	0.00	7.50	-2.29	0.0219	GRAVY	-0.7 ± 0.87	-0.42 ± 0.74	-2.62	0.0088
P1' = L	0.66	2.08	13.40	-2.64	0.0084	isoelectric point	4.9 ± 1.43	5.85 ± 1.99	-3.37	0.0007	

<b>Cluster 4</b>	protein = CASB	38.53	51.55	41.89	4.43	<0.0001
	protein = CASK	55.81	6.76	3.79	3.36	0.0008
	P1' = N	53.85	5.92	3.44	2.94	0.0033
	P1' = H	54.84	4.79	2.73	2.72	0.0066
	P1 = M	18.87	2.82	4.67	-2.04	0.0412
	P1 = F	17.11	3.66	6.70	-2.86	0.0042
	protein = CEL	8.24	1.97	7.50	-5.23	<0.0001
protein = TRFL	6.48	1.97	9.52	-6.51	<0.0001	

GRAVY	-0.52 ± 0.76	-0.42 ± 0.74	-3.04	0.0024
-------	--------------	--------------	-------	--------

<b>Cluster 5</b>	protein = TRFL	25.00	45.76	9.52	7.53	<0.0001
	P1 = K	13.37	38.98	15.17	4.58	<0.0001
	P1' = T	14.29	10.17	3.70	2.27	0.0232
	P1 = R	9.45	20.34	11.20	2.10	0.0355
	protein = CASA1	1.03	1.69	8.55	-2.11	0.0347
	protein = CASB	2.95	23.73	41.89	-2.96	0.0031

essential AA	5.07 ± 1.57	7.02 ± 3.65	-4.23	<0.0001
MW	1.42 ± 0.44	1.92 ± 0.8	-4.97	<0.0001

<b>Cluster 6</b>	P1 = D	65.00	5.70	1.76	4.34	<0.0001
	P1 = L	29.15	28.51	19.66	3.63	0.0003
	P1' = P	40.00	10.53	5.29	3.62	0.0003
	P1 = M	33.96	7.89	4.67	2.42	0.0157
	protein = ALBU	50.00	2.63	1.06	2.27	0.0232
	P1' = K	9.62	2.19	4.59	-2.01	0.0440
	P1' = N	7.69	1.32	3.44	-2.07	0.0383
	protein = OSTP	10.87	4.39	8.11	-2.41	0.0160
	P1 = Y	9.68	3.95	8.20	-2.77	0.0056
	protein = CASK	4.65	0.88	3.79	-2.84	0.0046
	P1 = R	11.02	6.14	11.20	-2.84	0.0045
	P1 = K	11.05	8.33	15.17	-3.37	0.0007

essential AA	8.62 ± 3.98	7.02 ± 3.65	7.39	<0.0001
MW	2.24 ± 0.85	1.92 ± 0.8	6.63	<0.0001
isoelectric point	6.46 ± 2.21	5.85 ± 1.99	5.20	<0.0001
GRAVY	-0.29 ± 0.63	-0.42 ± 0.74	2.98	0.0030

<b>Cluster 7</b>	protein = CEL	20.00	18.68	7.50	3.66	0.0002					
	P1' = K	19.23	10.99	4.59	2.65	0.0082					
	P1' = G	19.05	8.79	3.70	2.33	0.0198	isoelectric point	5.4 ± 1.61	5.85 ± 1.99	-2.24	0.0254
	protein = LALBA	19.35	6.59	2.73	2.04	0.0414	essential AA	5.51 ± 2.58	7.02 ± 3.65	-4.14	<0.0001
	protein = CASB	6.11	31.87	41.89	-2.03	0.0425	MW	1.57 ± 0.65	1.92 ± 0.8	-4.39	<0.0001
<b>Cluster 8</b>	protein = CEL	32.94	15.91	7.50	4.17	<0.0001					
	P1' = T	28.57	6.82	3.70	2.19	0.0286					
	P1' = P	25.00	8.52	5.29	1.96	0.0498					
	P1' = G	2.38	0.57	3.70	-2.68	0.0073					

---