

## Article

# A MALDI-TOF MS approach for mammalian, human, and formula milks' profiling

Laura Di Francesco<sup>1,§</sup>, Francesco Di Girolamo<sup>1,§</sup>, Maurizio Mennini<sup>2</sup>, Andrea Masotti<sup>3</sup>, Guglielmo Salvatori<sup>4</sup>, Fabrizio Signore<sup>5</sup>, Emanuela Pietrantoni<sup>5</sup>, Margherita Scapaticci<sup>6</sup>, Isabella Lante<sup>6</sup>, Bianca Maria Goffredo<sup>7</sup>, Oscar Mazzina<sup>2</sup>, Ahmed Ibrahim Elbousify<sup>8</sup>, Paola Roncada<sup>9</sup>, Andrea Dotta<sup>4</sup>, Alessandro Fiocchi<sup>2,#</sup> and Lorenza Putignani<sup>1,10,\*,#</sup>

<sup>1</sup>Unit of Human Microbiome, Bambino Gesù Children's Hospital, IRCCS, V.le San Paolo 15, 00146, Rome, Italy. E-mail: [lau.difra@gmail.com](mailto:lau.difra@gmail.com); [di\\_girolamo.f@hotmail.it](mailto:di_girolamo.f@hotmail.it)

<sup>2</sup>Allergy Unit, Bambino Gesù Children's Hospital, IRCCS, Piazza Sant'Onofrio 4, 00165, Rome, Italy. E-mail: [maurizio.mennini@opbg.net](mailto:maurizio.mennini@opbg.net); [oscar.mazzina@opbg.net](mailto:oscar.mazzina@opbg.net); [agiovanni.fiocchi@opbg.net](mailto:agiovanni.fiocchi@opbg.net)

<sup>3</sup>Gene Expression-Microarrays Laboratory, Bambino Gesù Children's Hospital, IRCCS, V.le San Paolo 15, 00146, Rome, Italy. E-mail: [andrea.masotti@opbg.net](mailto:andrea.masotti@opbg.net)

<sup>4</sup>Neonatal Intensive Care Unit, Dept. of Medical and Surgical Neonatology, Bambino Gesù Children's Hospital, IRCCS, Piazza Sant'Onofrio 4, 00165, Rome, Italy. E-mail: [guglielmo.salvatori@opbg.net](mailto:guglielmo.salvatori@opbg.net); [andrea.dotta@opbg.net](mailto:andrea.dotta@opbg.net)

<sup>5</sup>Dept. of Obstetrics and Gynaecology, San Camillo Forlanini Hospital, Circonvallazione Gianicolense 87, 00151, Rome, Italy. E-mail: [fabri64@me.com](mailto:fabri64@me.com); [anguela@libero.it](mailto:anguela@libero.it)

<sup>6</sup>Dept. of Laboratory Medicine, San Camillo Hospital, V.le Vittorio Veneto 18, 31100, Treviso, Italy. E-mail: [scapaticci.m@gmail.com](mailto:scapaticci.m@gmail.com); [Isa.Lante@gmail.com](mailto:Isa.Lante@gmail.com)

<sup>7</sup>Metabolic Unit, Dept. of Pediatric Medicine, Bambino Gesù Children's Hospital, IRCCS, V.le San Paolo 15, 00146, Rome, Italy. E-mail: [biancamaria.goffredo@opbg.net](mailto:biancamaria.goffredo@opbg.net)

<sup>8</sup>University of Tripoli, Tripoli Children Hospital, Tripoli, Libya. E-mail: [elbousify@gmail.com](mailto:elbousify@gmail.com)

<sup>9</sup>Universita' degli Studi "Magna Græcia" di Catanzaro, Department of Health Sciences, Catanzaro, Italy. E mail: [roncada@unicz.it](mailto:roncada@unicz.it)

<sup>10</sup> Unit of Parasitology, Bambino Gesù Children's Hospital, IRCCS, Piazza Sant'Onofrio 4, 00165, Rome, Italy. E-mail: [lorenza.putignani@opbg.net](mailto:lorenza.putignani@opbg.net)

§ These authors are co-first authors on this work.

# These authors share joint senior authorship.

\* Correspondence: [lorenza.putignani@opbg.net](mailto:lorenza.putignani@opbg.net); Tel.: +39-06-68592210-2176-2598

**Abstract:** Human milk composition is dynamic and substitute formulae are intended to mimic its protein content. The purpose of this study was to investigate the potentiality of matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry (MALDI-TOF MS) followed by multivariate data analyses as a tool to analyze peptide profiling of mammalian, human and formula milks. Breast milk samples from women at different lactation stages (2 (n = 5), 30 (n = 6), 60 (n = 5), and 90 (n = 4) days postpartum), and milk from donkeys (n = 4), cows (n = 4), buffaloes (n = 7), goats (n = 4), ewes (n = 5), and camels (n = 2) were collected. Different brands (n = 4) of infant formulae were also analyzed. Protein content (<30 kDa) was analyzed by MS and data were exported for statistical elaborations.

Mass spectra for each milk closely clustered together, whereas different milk samples resulted well separated. Human samples formed a cluster in which colostrum constituted a well-defined subcluster. None of the milk formulae correlated with animal or human milk, although specifically characterized and well correlated each other.

These findings propose MALDI-TOF MS milk profiling as an analytical tool to discriminate, in a blinded way, different milk types. As each formula has a distinct specificity, shifting a baby from

one to another formula implies a specific proteomic exposure. These profiles may assist in milk proteomics for easiness of use and low cost consuming, suggesting that the MALDI-TOF MS pipelines may result useful for milk adulteration assessment but also for the characterization of banked milk specimens in paediatric clinical settings.

**Keywords:** Infant nutrition, breast milk, mammalian milk, formula milk, protein similarity profiling, MALDI-TOF mass spectrometry.

---

## 1. Introduction

Breast milk (BM) is the primary food source for newborn mammals and World Health Organization recommends that infants should be exclusively breastfed for the first six months of life [1,2]. BM synthesis is subtly regulated at a local level [3] and its composition is influenced by several factors such as animal species and genetics, environmental conditions, and animal nutritional status [4]. Human milk (HM) composition varies with gestational age, lactation stage (transition from colostrum to late lactation), within feeds, diurnally, and amongst mothers [5,6]. HM provides proteins, tricalcium phosphate, lipids, vitamins, salt and lactose; it also contains many hundreds to thousands of distinct bioactive molecules, which protect against infection and inflammation and contribute to immune maturation, organ development, and healthy microbial colonization [7].

Despite many campaigns for the promotion of breastfeeding, only 38% of infants in the world are exclusively breastfed [8,9]. When HM becomes unsuitable or inadequate, its ideal substitutions are the infant formulae, defined as “a breast milk substitute specially manufactured to satisfy, by itself, the nutritional requirements of infants during the first months of life up to the introduction of appropriate complementary feeding” [10]. Unlike dynamic composition of HM, infant formulae are standard products with a composition highly regulated by the authorities. To date, the most commonly recommended infant formulae are based on cow milk [11]. The FDA in the US and the European Society for Pediatric Gastroenterology Hepatology and Nutrition (ESPGHAN) in Europe recommended the formulae to be enriched in whey protein fractions and lowered in caseins [9,12]. The worldwide recommendations are primarily based on the chemical analysis of human milk and manufacturers are continually modifying their products to make them more similar to human breast milk, the gold standard to estimate the needs of an infant [13], and to obtain health benefits, including iron, nucleotides, prebiotics and compositions of fat blends [14]. In this context, during the last half century the number of studies on HM and its protein composition has dramatically increased [15–19]. For children with cow’s milk allergy (CMA) whose mother cannot breastfeed, milk from different mammals has been evaluated, but no milk from animals different from cow has been formulated, and cross-reactivity is possible between cow’s and other mammalian milk’s proteins [20]. Thus, other animal-milk-based formulas are currently not recommended [21].

The recommendations on substitute formulae in case of lactation failure are based on many aspects, but nutritional considerations are prominent. Among the nutritional factors, proteins are the most important. This study aims to investigate the potentiality of linear matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) as a tool to assess the diversity and oddness of different artificial and mammalian kind of milk compared to the reference human milk. MALDI-TOF MS is a platform adopted from many healthcare clinical laboratories worldwide owing to its simplicity of use, high reproducibility of the mass spectra and low-cost of the analysis [22]. Recently, this technique has been proposed as a powerful tool to obtain informative fingerprints of milk proteins [23]. Identification of diversity and similarity of several artificial and animal milk compared to the references human milk at different stage of lactation could assist the design of infant formulae. Furthermore, MALDI-TOF MS-based approach coupled to multivariate statistical assessment of MS data could represent a versatile workflow to evaluate quality and safety of sample milk in blind for non-specialized mass-spectrometric laboratories.

## 2. Materials and Methods

### 2.1. Milk Sampling and Pre-treatment

The human milk (HM) samples were collected from twenty healthy breastfeeding mothers at the Department of Obstetrics and Gynecology, San Camillo Forlanini Hospital of Rome, Italy, at four different periods of lactation: 2 days (colostrum, HC), 30 days (HM30), 60 days (HM60), and 90 days (HM90). The study protocol was approved by the Ethics Committee of the San Camillo-Forlanini Hospital (protocol 460/CE; 27/03/2012) and by the Institutional Review Board of the Bambino Gesù Children's Hospital (protocol 295 LB; 16/05/2012). Informed written consent was obtained from all mothers. Raw donkey milk (DM) from four she-donkeys belonging to Amiata, Viterbese and Martina Franca breeds, cow milk (CM) from four cows belonging to Frisona breed, buffalo milk (BM) from seven buffaloes belonging to Mediterranean Italian breed, goat milk (GM) from four goats belonging to Maltese breed, and ewe's milk (EM) from five ewes belonging to Tuscolania breed, were collected from Italian farms (Lazio and Puglia). Camel milk (CAM, from two *Camelus Dromedarius*) was collected from Libyan Desert farms. Commercially available infant formula milk samples from four different brands: Aptamil 1 (A) [Mellin SpA, Milan, Italy], Humana 1 (H) [Humana Italia SpA, Milan, Italy], Formulac 1 (F) [Dicofarm, S.A., Rome, Italy] and Nidina 1 (N) [Nestlé, S.A., Milan, Italy] were also studied. For each brand, we got four samples from different batches produced over a period of two years. All the animal milk samples were mechanically milked at middle lactation stage, into sterile polystyrene containers, immediately frozen, and stored at  $-80^{\circ}\text{C}$  until use to prevent undesired proteolysis. After thawing, raw milk samples were defatted by two-step centrifugation using the Eppendorf Centrifuge 5417 R. A first centrifugation was performed at  $3,000\times g$  for 10 min at  $4^{\circ}\text{C}$ . The skimmed milk was then centrifuged at  $20,000\times g$  for 20 min at  $4^{\circ}\text{C}$  to remove bacteria and cell debris. The skimmed milk's fractions were subsequently diluted 1:100 with ultrapure water (Milli-Q Millipore) and subjected to mass spectrometry analysis.

### 2.2. MALDI-TOF Spectra Acquisition

An aliquot (1  $\mu\text{L}$ ) of each skimmed milk's fraction was directly spotted onto a MSP 96 polished steel target (Bruker Daltonics, Bremen, Germany), overlaid with 1 L of matrix, represented by a solution of 10 mg/mL of sinapinic acid (Sigma-Aldrich, St. Louis, MO, USA) in 50% acetonitrile, containing 0.1% trifluoroacetic acid (v/v) and allowed to dry at room temperature. MALDI-TOF analysis was performed with a Microflex LT linear mass spectrometer (Bruker Daltonics) equipped with the FlexControl software package, version 3.0 (Bruker Daltonics), for spectra recording in the positive linear mode (laser frequency 20 Hz; ion source 1 voltage, 20 kV; ion source 2 voltage, 18.4 kV; lens voltage, 9.1 kV; mass range, 2,000 to 30,000 Da). Four independent spectra (500 shots one-step from different positions of the target spot, for spectrum) for each skimmed milk's fraction were manually collected, externally calibrated by using Bacterial Test Standard (Bruker Daltonics) and subsequently analyzed.

### 2.3. Statistical Analysis

200 mass spectra, as reported in Table 1, were manually acquired and visually inspected before statistical analysis. Subsequently, each spectra was loaded into FlexAnalysis software, version 3.0 (Bruker Daltonics), to perform mass adjustment (spectra were compressed by a factor of 10 in the total mass range), smoothing (mass data were adjusted by the Savitsky-Golay algorithm with a frame size of 25 Da), baseline subtraction (was applied the minimum value for finding the baseline), normalization (was applied the maximum norm to normalize the baseline subtracted data), and peak picking (was applied spectra differentiation algorithm for finding the peaks, maximum peaks 100, threshold 0.1, method Peak Fitting). The total pre-processed raw datasets of the 200 milk spectra were imported into R Bioconductor (<http://www.bioconductor.org/>) [24] for Pearson's correlation analysis and hierarchical clustering. The package pvclust was applied for bootstrapping. For each cluster generated by hierarchical computation, p-values (between 0 and 1) were calculated via

multiscale bootstrap resampling. Two different p-values were provided by the package pvclust: Approximately Unbiased (AU) and Bootstrap Probability (BP). AU was computed by multiscale bootstrap resampling and represents a better approximation to unbiased p-value than BP value computed by normal bootstrap resampling. The same pre-processed raw datasets were imported into ClinProTools™ bioinformatics software, version 2.2 (Bruker Daltonics) [25], and converted into a virtual gel-like format. The mass values (m/z) were reported on the X axis, while the gray scale bar, reported on the Y axis, showed the relationship between the colour intensity and the peak intensity. Finally, Principal Component Analysis (PCA) was performed via ClinProTools™ software, version 2.2 (Bruker Daltonics), which employs for the calculation only those peaks resulted statistically significant after group classification. Based on a Welch's t-test, a P value for each peak was calculated. This value indicates the probability that the observed intensity differences among

Milk source	Samples (n)	Spectra replicates (n)	Total spectra analyzed (n)
Human milk			
Breastfeeding women at 2 days	5	4	20
Breastfeeding women at 30 days	6	4	24
Breastfeeding women at 60 days	5	4	20
Breastfeeding women at 90 days	4	4	16
Commercial brands of infant formula	4	4	16
Animal milk <sup>1</sup>			
cows	4	4	16
buffaloes	7	4	28
goats	4	4	16
ewes	5	4	20
she-donkeys	4	4	16
camels	2	4	8

the various peaks are due to chance. These calculations have been done independently for peak heights and peak areas.

**Table 1.** Milk samples analyzed and relative mass spectra acquired.

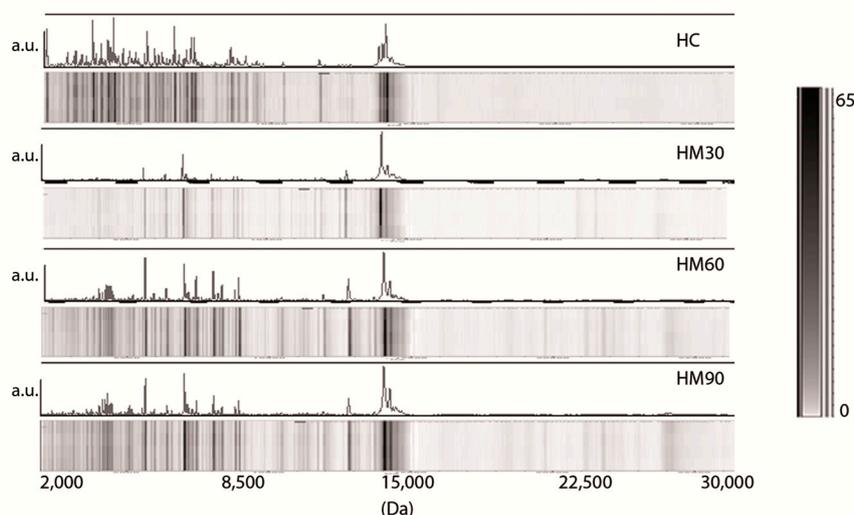
<sup>1</sup>All animal milk samples were collected at middle lactation stage.

### 3. Results

#### 3.1. Low-molecular weight protein profiles from crude milk by MALDI-TOF MS

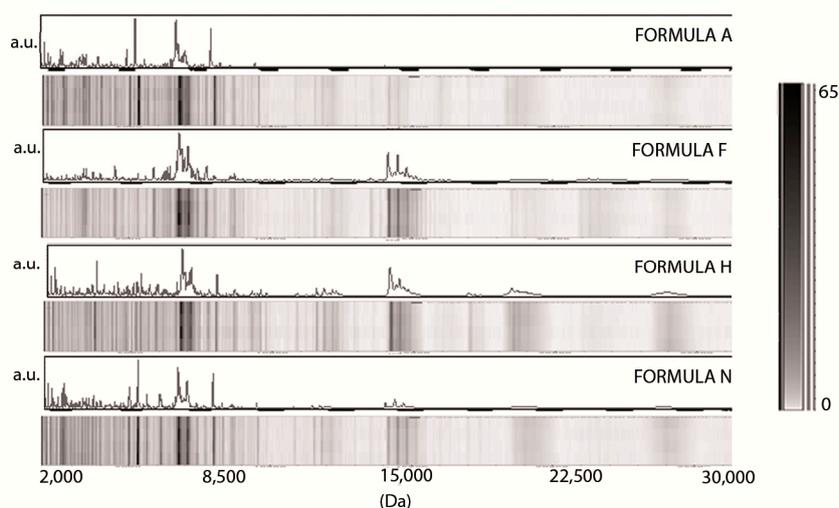
Analysis of peptides and low-molecular weight proteins (2,000–30,000 Da) present in the defatted crude milk was performed by a bench-top linear MALDI-TOF mass spectrometer. We obtained complex mass spectra not affected by signal background problems. Four independent MALDI-TOF MS protein profiles from each milk sample were recorded in order to ascertain a high level of analytical reproducibility of the analysis. Mass spectra obtained from different milk samples from the same source prepared and run on the same day were virtually indistinguishable and relative intensities of protein species detected in each replicate were constant (data not shown). Each spectrum was visually inspected and the resulting flattened profiles were compared by gel-like representations with spectra from different samples. **Figure 1** reports the MALDI-TOF MS profile and the pseudo-gel view of human milk samples analyzed at 2 (colostrum, HC), 30 (HM30), 60 (HM60), and 90 (HM90) days postpartum. In all human samples, many peaks are visible in the left part of the mass spectrum. After conversion of the obtained mass spectra to gel-like format, it appeared clear that many species below a molecular weight of 5,000 Da were present especially in the HC samples. In the middle of the spectra (medium-mass range), many peaks between 8,500 and 16,000 Da were present in all samples, whereas in the rightmost part of the spectra (>16,000 Da) no

peaks were detected, although the sinapinic acid matrix, which is beneficial for ionization of higher molecular weight proteins, was used during the sample deposition on the MALDI target. Only in mature milk (HM60 and HM90) a mass value around 24,000 Da became detectable.



**Figure 1.** Representative MALDI-TOF MS profiling and pseudo-gel view of crude human milk at 2, 30, 60 and 90 days of lactation, indicated respectively as HC, HM30, HM60 and HM90. The mass-to-charge ratios ( $m/z$ ) are reported on the X axis (Da), while the peak intensities are indicated as arbitrary units (a.u.) in the gray scale on the Y axis.

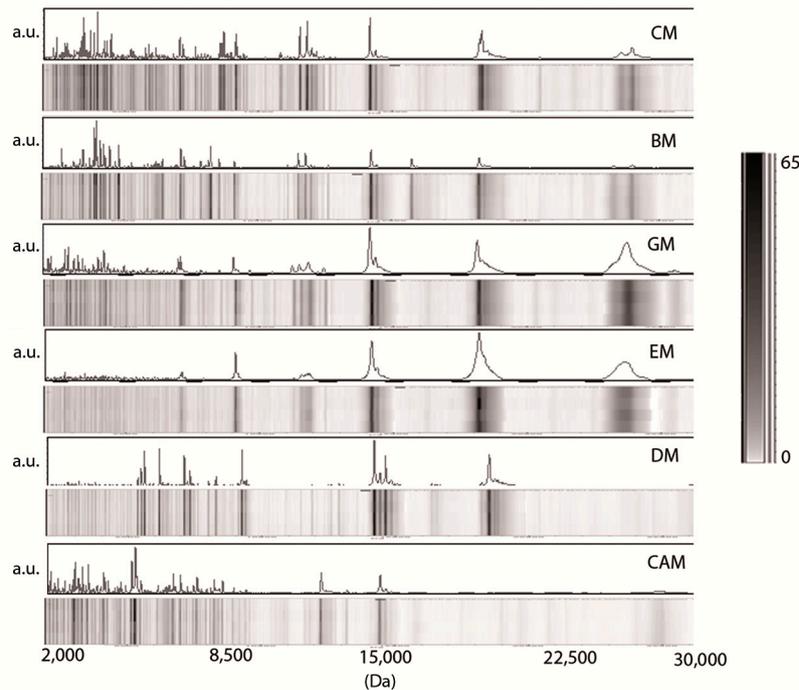
**Figure 2** shows the MALDI-TOF MS profile and the pseudo-gel view relative to four commercial infant formulas (A, H, F and N). The distribution of molecular weights observed in the mass range between 2,000 and 30,000 Da was similar in formula F, H and N while formula A displayed peaks only in the first part of the spectra, and no detectable signals were observed in the high-mass range, although is not a hydrolysate formula. A faint signal above background around  $m/z$  18,000 can be detectable for both formula H and N.



**Figure 2.** Representative MALDI-TOF MS profiling and pseudo-gel view of commercial starting formula from four different companies (Aptamil 1, A; Formulac 1 from Dicofarm, F; Humana 1, H; Nidina 1 from Nestlé, N). The mass-to-charge ratios ( $m/z$ ) are reported on the X axis (Da), while the peak intensities are indicated as arbitrary units (a.u.) in the gray scale bar on the Y axis.

**Figure 3** reports the mass spectra and the relative pseudo-gel view of other mammalian milk analyzed in the present study: cow milk (CM), buffalo milk (BM), goat milk (GM), ewe milk (EM), donkey milk (DM) and camel milk (CAM). All animal milk samples contain several peaks in the low-medium- and high-mass range.

In the region of the mass spectra <10,000 Da many spectrometric signals were detectable except from EM. As in human and formula milk, a mass value around 14,000 Da predominated in all animal samples. Conversely, mass spectra of animal milk samples showed two intense peaks around 18,000 (missing in human samples as well as CAM) and 24,000 Da (missing in HC, HM30 as well as in DM and CAM). In our conditions, CAM milk showed a peculiar profile, displaying a reduced number of peaks in the high-mass range compared to other milk.



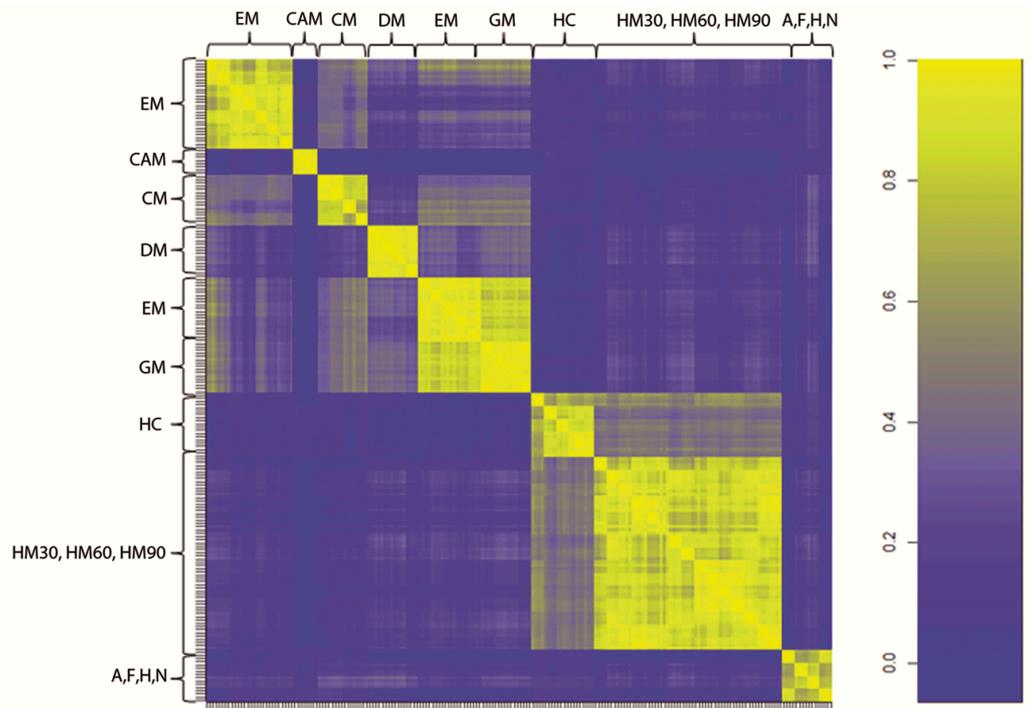
**Figure 3.** Representative MALDI TOF MS profiling and pseudo-gel view of crude milk from cow milk (CM), buffalo milk (BM), goat milk (GM), ewe milk (EM), donkey milk (DM) and camel milk (CAM). The mass-to-charge ratios ( $m/z$ ) are reported on the X axis (Da), while the peak intensities are indicated as arbitrary units (a.u.) in the gray scale bar on the Y axis.

### 3.2. Milk spectra properties and similarities

To evaluate similarities and differences among peptide and protein compositions of different milk samples, we use an external statistical software (R Bioconductor) to perform a correlation analysis on spectral values ( $m/z$  and intensities) extracted after ClinProTools™ bioinformatics software pre-processing.

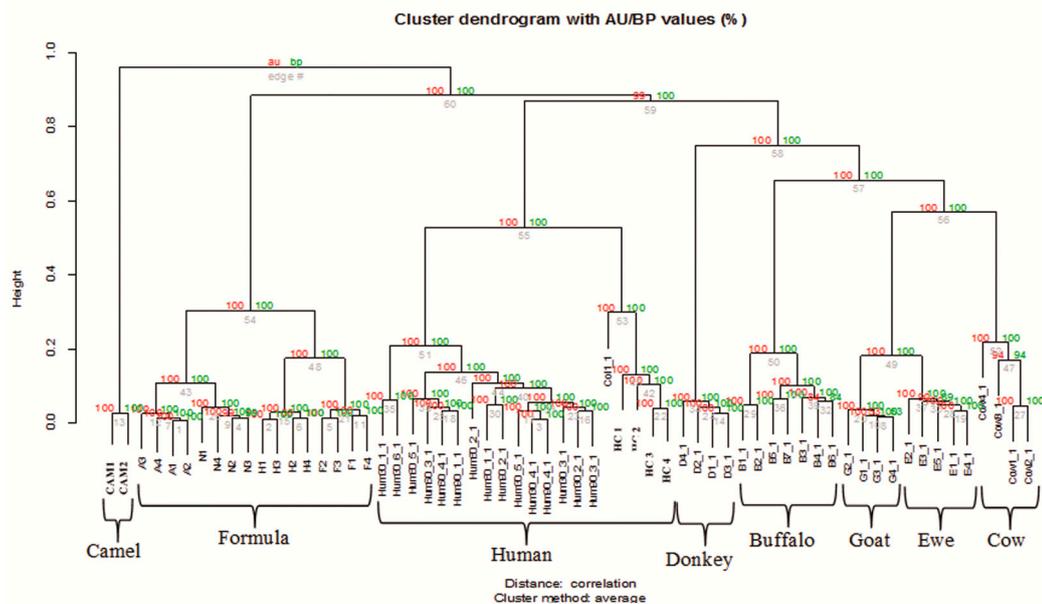
**Figure 4** displays the correlation matrix obtained from all of the spectra. The figure represents three wide sub-groups: the first group of animal milk (BM, CAM, CM, DM, EM and GM), the second of human milk (HC, HM30, HM60 and HM90), and the third of formula milk (A, F, H and N). From a visual analysis, animal milk do not have an appreciable correlation with human or formula milk. CAM displayed a very poor correlation with all of the considered milk. This suggest that this milk might have a different protein profile compared to other animal milk. BM has a quite good correlation with CM and to a less extent with EM and GM, in agreement with our previous study [26]. Human colostrum has a good correlation with mature human milk although HM30, HM60 and HM90 are more correlated between them. Of note, although well correlated each other, none of the

milk formulae analyzed in this study displayed an appreciable correlation with animal or human milk.

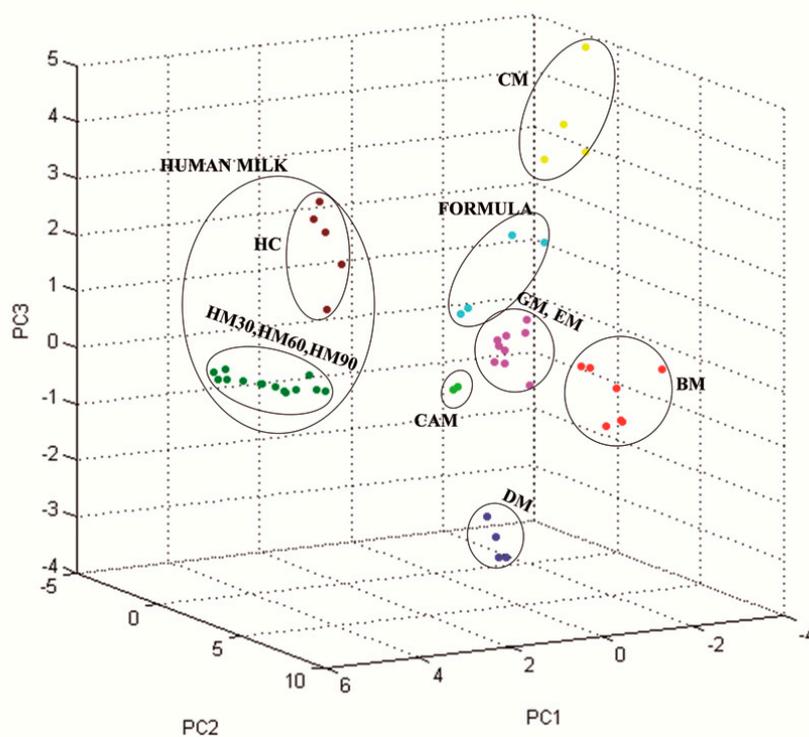


**Figure 4.** Pearson's correlation matrix of all spectral replica datasets for animal milk (BM, CAM, CM, DM, EM and GM), human milk at 2, 30, 60 and 90 days (HC, HM30, HM60 and HM90, respectively) and infant formula (A, F, H, N). Correlation coefficients are represented with decreasing blue and yellow colors according to a scale ranging from 0 to 1, respectively.

These considerations were also supported by the hierarchical clustering tree (**Figure 5**) determined by the statistical software R Bioconductor. Camel milk cluster separates from the other three clades that represent the group of formula milk, animal milk and human milk. Formula milk from four different companies displays a quite homogeneous clustering, suggesting quite common spectral characteristics. However, differences among the different brands allow to identify a common sub-cluster for each brand, even across its different batches. The group of human mature milk clustered near the colostrum clade, whereas the animal milk group is quite well separated. The spectra were then analysed by principal component analysis (PCA) using the integrated software ClinProTools™. As shown in **Figure 6**, the milk samples of the same species closely clustered together, whereas the different milk species resulted well separated each other. The 3D scatter plot image obtained from the PCA analysis indicates that seven MALDI-TOF MS profiles can be grouped, again corresponding to breast milk (with the two subgroups of colostrum and the other stages of breast milk), starting formulae, CM, BM, GM/EM, DM and CAM.



**Figure 5.** Hierarchical clustering tree (bootstrap  $n=1000$ ) of all MALDI-TOF spectral replica from animal milk (CM, BM, GM, EM, DM and CAM), human milk at 2, 30, 60 and 90 days (HC, HM30, HM60 and HM90, respectively) and infant formula (A, F, H, N). Red values (left) are Approximately Unbiased (AU) p-values, green values (right) are Bootstrap Probability (BP) values and grey values are cluster labels (bottom).



**Figure 6.** 3D scatter plot image from the PCA analysis for human milk at 2, 30, 60 and 90 days (HC, HM30, HM60 and HM90, respectively), infant formula milk (A, F, H, N) and animal milk (CM, BM, GM, EM, DM and CAM). Each spot represents one milk sample.

#### 4. Discussion

In the current study, we focused our interest on protein content that in milk serves diverse biological activities, such as providing essential amino acids to growing infants, supplying newborns of enzymatic activity and making available vitamins and hormones. Proteins are present in milk with a very large dynamic range in their concentrations [27]. After a defatting operation consisting in two-step centrifugation, the milk fat globule membrane proteins were lost in great part together with their lipid counterpart and the protein content of all samples resulted principally in caseins and soluble whey proteins. We then conducted a qualitative MALDI-TOF MS-based analysis of defatted milk to evaluate similarities and differences in low molecular weight profiles composition of milk from different source.

Multiple components were detected as clear signals in the mass range of 2,000-30,000 Da. This region is known to represent milk proteins and peptides with pivotal roles in infants health and development, such as antimicrobial activities (e.g., lysozyme, 16 kDa and - lactalbumin, 14 kDa) and mineral absorption functions (e.g., caseins, between 20 and 25 kDa). Furthermore, most of the bioactive factors are peptides originally present in milk, which may exert their biologic activity in the upper gastrointestinal tract regardless of digestive processes [19].

Our data indicate that the human colostrum profile (<30,000 Da) is more complex than every other kind of milk (Figures 1-3). Colostrum, secreted in the few days after birth, is reported to contain higher amount of peptides, proteins and vitamins compared to mature milk (28). The unique characteristics of HC, with additional nutrients, immune and growth factors, make it interesting as a therapy to promote neonatal health [28,29]. We found that HC spectra are particularly rich in low-molecular weight proteins and are dominated by spectrometric signals with a mass <10,000 Da and mainly <7,000 Da, a significant fraction of which may be involved in its physiologic characteristics. Based on the linear MALDI-TOF MS, our analysis does not allow to identify any milk proteins, but it clearly indicates that milk protein profile changes gradually over the following 30 days after birth. The amount of peptides and proteins decreases rapidly during the first month of lactation and then stabilizes in "mature" milk after 60 days (HM60 and HM90; Figure 1). By contrast, polypeptides around 15,000 Da remain stable across colostrum and mature milk. Noteworthy, the appearance of spectrometric signals with a m/z value around 24,000 Da in mature milk (HM60 and HM90) concomitant to a decrease of components with a molecular weight <12,000 Da. In agreement with these data, proteomic studies have shed light on the dynamic composition of human milk throughout lactation stages [18,30–32]. In particular, whey proteins implicated in the modulation of immune system and in the maturation of the gastrointestinal tract of neonates are overrepresented in the human milk during the first days of lactation.

Hierarchical cluster analysis of the mass spectra was used to group milk samples according to the similarity of their spectral profiles. In this unsupervised analysis, group assignment of the protein/peptides expression patterns is generated based on similarities of spectral patterns in the automatic selected peaks. This analysis demonstrated that all analysed milk samples formed four main clusters (Figure 5). All human samples (n = 20) formed a cluster in which milk at 30, 60 and 90 days constitute a well-defined subcluster. These results indicate that colostrum could be clearly differentiated by signal patterns of their MALDI-TOF mass spectra as an out-group. Even more clearly, the PCA analysis of the qualitative characteristics generated seven principal-components (PCs) (Figure 6). Human milk was segregated into a single PC, in which HC (brown dots) was identified as different from human milk samples at different stage of lactation (green dots).

Spectra relative to commercial starting formulae reported in Figure 2 are from four different brands in which bovine milk is the only source of protein (e.g., casein alone, whey proteins together with caseins, etc.) and fat content derives from a mixture of vegetable oil. Each group includes four different samples of the same formula, coming from different batches. Linear MALDI-TOF MS technique is able to identify a consistent similarity among the profiles of different milk samples. All formulae displayed a profile consistently different from cow's milk, which is their parent protein source. This is consistent with the ESPGHAN recommendations for the protein composition of infant formulae (12), which prescribe a modification: the formulae studied are all enriched in whey

protein fractions and lowered in caseins. However, each brand is characterized by a specific protein profile. Although all 16 samples of infant formula formed a tight cluster (**Figure 5**), each brand can be differentiated, indicating that the proteomic asset of each formula is stable across batches and can be identified at a blinded, unsupervised analysis.

Animal milk analysed in the present study showed similar protein peak profiles ranging from 14,000 to 18,000 Da (**Figure 3**). In particular, these spectrometric signals are conserved across CM, BM, GM, EM, and DM, while they are less evident in CAM. The spectrometric signals at 18,000 Da and 24,000 Da correspond, respectively, to the theoretical molecular weights of lactoglobulin and casein, two components of cow milk responsible for cow's milk allergy (CMA), the most common food allergy affecting children [33]. The absence of these spectrometric signals in the CAM sample and in all analysed human milk would be in agreement with previous studies that report the lack of these proteins in camel milk [34,35]. For this reason, CAM, although not extensively used, was recommended by pediatricians for cow's milk-allergic children and hence evaluated in the present work [34,36–39].

GM and EM clearly display two additional peaks at 20,000 and 25,000 Da, less evident in CM and BM. They are lacking in donkey milk, characterized by a protein content notoriously low (1.3 - 2.8 g/100 mL) and by a high whey protein/casein ratio [40]. The hierarchical cluster analysis of peak profiles indicated that 24 samples of five animals milk could be subcategorized by source: a subcluster identified DM vs. bovidae milk (buffalo, goat, ewe, cow), another included all seven BM samples, and a third was composed by the 13 samples of CM, EM and GM. The results indicate that cow, ewe and goat have a homogeneous milk proteome, while milk proteins from donkey and cow share a low-sequence similarity due to the genetic distance between Equidae and Bovidae families. Indeed, there is evidence of cross-reactivity between cow milk and proteins from goat, sheep and buffalo milk (20), while substantial differences in the IgE-binding epitope of cow milk protein and the corresponding domains of donkey milk protein may contribute, besides the low content in caseins, to explain the demonstrated reduced allergenicity of DM [41]. Its distinctive proteins composition is also evident in the results of our PCA, able to differentiate between BM and CM samples and also between these ones and DM. The only two kinds of milk that clustered together were from ewe and goat samples, not a surprising finding if we consider the taxonomic proximity between goat and sheep (i.e., Caprinae subfamily) and the well-known clinical cross-reactivity among their milk [42]. By converse, camel milk is completely different from any other mammalian milk from a proteomic point of view. The same observation stems from the PCA reported in Figure 6, where CAM is segregated with a different colour. These results confirm the quite different protein composition of camel milk compared to other animals milk [43]. Among mammalian species proposed to be suitable as a valid substitute of cow's milk-based formulas, the CAM has a unique spectra profile that could have interesting properties in children nutrition.

These findings seem to suggest that the choice of an alternative to breast milk cannot be made exclusively on the basis of macronutrient composition, but that the proteomic profile can be a useful evaluation tool [44,45]. Their results may be even more relevant if they will be replicated in extensively Hydrolyzed milk Formulae (eHFs). These formulae are the first choice milk substitutes in CMA, but may carry residual allergens able to cause reactions in sensitive infants [21]. A recent proteomic analyses revealed that also the peptide profiles of commercially available eHFs provide a descriptive and distinct signature [46].

## 5. Conclusions

MALDI-TOF MS profiling of milk proteins in combination with statistic tools proved to be a high throughput and low-cost approach with promising applications as analytical tool to quickly assess similarities and differences of low-molecular weight proteins present in milk from different sources with a high level of accuracy and sensitivity. Our data point to differences between the potential alternative sources of infant formula milk and create the basis for further proteomic investigations to achieve more conclusive results on protein content of the milk types herein evaluated.

Moreover, a MALDI-TOF mass spectral database compilation can assist non-specialized mass spectrometric laboratories for a rapid screening and characterization of milk samples. The screening procedures could become a powerful method for analyzing milk in a blinded way, in order to evaluate animal milk adulterations, milk samples present in human donor breast milk banking and matching between human-milk formula composition. Rapid MALDI- TOF MS assay could also become an instrument for interpreting the individuality in phenotypic expression of allergy to cow's milk proteins and beyond.

**Acknowledgments:** This work was supported by San Camillo Hospital, Treviso, Italy Fellowship Program to Lorenza Putignani.

**Author Contributions:** F.D.G. conceived and designed the study, acquired mass spectra, interpreted the data and drafted the manuscript; L.D.F. and M.M. wrote the paper with input from all authors; A.M. performed statistical analysis; G.S., F.S., E.P., M.S., B.M.G., O.M. and A.I.E. recruited samples and interpreted the data; I.L. contributed to performing the pre-analytical procedures and to data generation; P.R., A.D. and I.L. contributed to interpretation of data and to draft the manuscript; A.F. and L.P. contributed to conceiving and designing the study and to drafting the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. PAHO/WHO. Guiding principles for complementary feeding of the breastfed child. Division of Health Promotion and Protection. Food and Nutrition Program. **2003**.
2. Lessen, R.; Kavanagh, K. Position of the Academy of Nutrition and Dietetics: Promoting and Supporting Breastfeeding. *J. Acad. Nutr. Diet.* **2015**, *115*, 444–449, doi:10.1016/j.jand.2014.12.014.
3. Weaver, S. R.; Hernandez, L. L. Autocrine-paracrine regulation of the mammary gland. *J. Dairy Sci.* **2016**, *99*, 842–853, doi:10.3168/jds.2015-9828.
4. Prentice, P.; Ong, K. K.; Schoemaker, M. H.; van Tol, E. A. F.; Vervoort, J.; Hughes, I. A.; Acerini, C. L.; Dunger, D. B. Breast milk nutrient content and infancy growth. *Acta Paediatr.* **2016**, *105*, 641–647, doi:10.1111/apa.13362.
5. Ballard, O.; Morrow, A. L. Human Milk Composition. *Pediatr. Clin. North Am.* **2013**, *60*, 49–74, doi:10.1016/j.pcl.2012.10.002.
6. Roncada, P.; Stipetic, L. H.; Bonizzi, L.; Burchmore, R. J. S.; Kennedy, M. W. Proteomics as a tool to explore human milk in health and disease. *J. Proteomics* **2013**, *88*, 47–57, doi:10.1016/j.jprot.2013.04.008.
7. Del Chierico, F.; Vernocchi, P.; Petrucca, A.; Paci, P.; Fuentes, S.; Praticò, G.; Capuani, G.; Masotti, A.; Reddel, S.; Russo, A.; Vallone, C.; Salvatori, G.; Buffone, E.; Signore, F.; Rigon, G.; Dotta, A.; Miccheli, A.; de Vos, W. M.; Dallapiccola, B.; Putignani, L. Phylogenetic and Metabolic Tracking of Gut Microbiota during Perinatal Development. *PLOS ONE* **2015**, *10*, e0137347, doi:10.1371/journal.pone.0137347.
8. Cai, X.; Wardlaw, T.; Brown, D. W. Global trends in exclusive breastfeeding. *Int. Breastfeed. J.* **2012**, *7*, 12, doi:10.1186/1746-4358-7-12.
9. Demonstration of the Quality Factor Requirements Under 21 CFR 106.96(i) for “Eligible” Infant Formulas. 2014.
10. Food and Agriculture Organization of the United Nations World Health Organization. CODEX alimentarius: standard for infant formula and formulas for special medical purposes intended for infants. Revised standard for infant formula. 1981.
11. Rossen, L. M.; Simon, A. E.; Herrick, K. A. Types of Infant Formulas Consumed in the United States. *Clin. Pediatr. (Phila.)* **2016**, *55*, 278–285, doi:10.1177/0009922815591881.
12. Koletzko, B.; Baker, S.; Cleghorn, G.; Neto, U. F.; Gopalan, S.; Hernell, O.; Hock, Q. S.; Jirapinyo, P.; Lonnerdal, B.; Pencharz, P.; Pzyrembel, H.; Ramirez-Mayans, J.; Shamir, R.; Turck, D.; Yamashiro, Y.; Zong-Yi, D. Global standard for the composition of infant formula: recommendations of an ESPGHAN coordinated international expert group. *J. Pediatr. Gastroenterol. Nutr.* **2005**, *41*, 584–599.
13. Stam, J.; Sauer, P. J.; Boehm, G. Can we define an infant’s need from the composition of human milk? *Am. J. Clin. Nutr.* **2013**, *98*, 521S–528S, doi:10.3945/ajcn.112.044370.
14. Ryan, A. S.; Hay, W. W. Challenges of infant nutrition research: a commentary. *Nutr. J.* **2015**, *15*, doi:10.1186/s12937-016-0162-0.
15. Gura, T. Nature’s first functional food. *Science* **2014**, *345*, 747–749, doi:10.1126/science.345.6198.747.
16. Weber, M.; Grote, V.; Closa-Monasterolo, R.; Escribano, J.; Langhendries, J.-P.; Dain, E.; Giovannini, M.; Verduci, E.; Gruszfeld, D.; Socha, P.; Koletzko, B.; for The European Childhood Obesity Trial Study Group Lower protein content in infant formula reduces BMI and obesity risk at school age: follow-up of a randomized trial. *Am. J. Clin. Nutr.* **2014**, *99*, 1041–1051, doi:10.3945/ajcn.113.064071.
17. Hidayat, K.; Du, H.-Z.; Yang, J.; Chen, G.-C.; Zhang, Z.; Li, Z.-N.; Qin, L.-Q. Effects of milk proteins on blood pressure: a meta-analysis of randomized control trials. *Hypertens. Res.* **2017**, *40*, 264–270, doi:10.1038/hr.2016.135.

18. Gao, X.; McMahon, R. J.; Woo, J. G.; Davidson, B. S.; Morrow, A. L.; Zhang, Q. Temporal Changes in Milk Proteomes Reveal Developing Milk Functions. *J. Proteome Res.* **2012**, *11*, 3897–3907, doi:10.1021/pr3004002.
19. Savino, F.; Sorrenti, M.; Benetti, S.; Lupica, M. M.; Liguori, S. A.; Oggero, R. Resistin and leptin in breast milk and infants in early life. *Early Hum. Dev.* **2012**, *88*, 779–782, doi:10.1016/j.earlhumdev.2012.05.004.
20. Restani, P.; Gaiaschi, A.; Plebani, A.; Beretta, B.; Cavagni, G.; Fiocchi, A.; Poiesi, C.; Velonà, T.; Ugazio, A. G.; Galli, C. L. Cross-reactivity between milk proteins from different animal species. *Clin. Exp. Allergy J. Br. Soc. Allergy Clin. Immunol.* **1999**, *29*, 997–1004.
21. Fiocchi, A.; Brozek, J.; Schünemann, H.; Bahna, S. L.; von Berg, A.; Beyer, K.; Bozzola, M.; Bradsher, J.; Compalati, E.; Ebisawa, M.; Guzmán, M. A.; Li, H.; Heine, R. G.; Keith, P.; Lack, G.; Landi, M.; Martelli, A.; Rancé, F.; Sampson, H.; Stein, A.; Terracciano, L.; Vieths, S.; World Allergy Organization (WAO) Special Committee on Food Allergy World Allergy Organization (WAO) Diagnosis and Rationale for Action against Cow's Milk Allergy (DRACMA) Guidelines. *Pediatr. Allergy Immunol. Off. Publ. Eur. Soc. Pediatr. Allergy Immunol.* **2010**, *21 Suppl 21*, 1–125, doi:10.1111/j.1399-3038.2010.01068.x.
22. Putignani, L.; Del Chierico, F.; Onori, M.; Mancinelli, L.; Argentieri, M.; Bernaschi, P.; Coltella, L.; Lucignano, B.; Pansani, L.; Ranno, S.; Russo, C.; Urbani, A.; Federici, G.; Menichella, D. MALDI-TOF mass spectrometry proteomic phenotyping of clinically relevant fungi. *Mol BioSyst* **2011**, *7*, 620–629, doi:10.1039/C0MB00138D.
23. Di Girolamo, F.; Masotti, A.; Salvatori, G.; Scapatucci, M.; Muraca, M.; Putignani, L. A Sensitive and Effective Proteomic Approach to Identify She-Donkey's and Goat's Milk Adulterations by MALDI-TOF MS Fingerprinting. *Int. J. Mol. Sci.* **2014**, *15*, 13697–13719, doi:10.3390/ijms150813697.
24. Gentleman, R. C.; Carey, V. J.; Bates, D. M.; Bolstad, B.; Dettling, M.; Dudoit, S.; Ellis, B.; Gautier, L.; Ge, Y.; Gentry, J.; Hornik, K.; Hothorn, T.; Huber, W.; Iacus, S.; Irizarry, R.; Leisch, F.; Li, C.; Maechler, M.; Rossini, A. J.; Sawitzki, G.; Smith, C.; Smyth, G.; Tierney, L.; Yang, J. Y. H.; Zhang, J. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **2004**, *5*, R80, doi:10.1186/gb-2004-5-10-r80.
25. Ketterlinus, R.; Hsieh, S.-Y.; Teng, S.-H.; Lee, H.; Pusch, W. Fishing for biomarkers: analyzing mass spectrometry data with the new ClinProTools software. *BioTechniques* **2005**, *Suppl*, 37–40.
26. Di Girolamo, F.; D'Amato, A.; Lante, I.; Signore, F.; Muraca, M.; Putignani, L. Farm Animal Serum Proteomics and Impact on Human Health. *Int. J. Mol. Sci.* **2014**, *15*, 15396–15411, doi:10.3390/ijms150915396.
27. Bislev, S. L.; Deutsch, E. W.; Sun, Z.; Farrah, T.; Aegersold, R.; Moritz, R. L.; Bendixen, E.; Codrea, M. C. A Bovine PeptideAtlas of milk and mammary gland proteomes. *PROTEOMICS* **2012**, *12*, 2895–2899, doi:10.1002/pmic.201200057.
28. Gephart, S. M.; Weller, M. Colostrum as Oral Immune Therapy to Promote Neonatal Health: *Adv. Neonatal Care* **2014**, *14*, 44–51, doi:10.1097/ANC.0000000000000052.
29. Bagwe, S.; Tharappel, L. J. P.; Kaur, G.; Buttar, H. S. Bovine colostrum: an emerging nutraceutical. *J. Complement. Integr. Med.* **2015**, *12*, doi:10.1515/jcim-2014-0039.
30. Liao, Y.; Alvarado, R.; Phinney, B.; Lönnnerdal, B. Proteomic Characterization of Human Milk Whey Proteins during a Twelve-Month Lactation Period. *J. Proteome Res.* **2011**, *10*, 1746–1754, doi:10.1021/pr101028k.
31. Liao, Y.; Alvarado, R.; Phinney, B.; Lönnnerdal, B. Proteomic Characterization of Human Milk Fat Globule Membrane Proteins during a 12 Month Lactation Period. *J. Proteome Res.* **2011**, *10*, 3530–3541, doi:10.1021/pr200149t.
32. Reinhardt, T. A.; Lippolis, J. D. Developmental Changes in the Milk Fat Globule Membrane Proteome During the Transition from Colostrum to Milk. *J. Dairy Sci.* **2008**, *91*, 2307–2318, doi:10.3168/jds.2007-0952.
33. Sackesen, C.; Assa'ad, A.; Baena-Cagnani, C.; Ebisawa, M.; Fiocchi, A.; Heine, R. G.; Von Berg, A.; Kalayci, O. Cow's milk allergy as a global challenge: *Curr. Opin. Allergy Clin. Immunol.* **2011**, *11*, 243–248, doi:10.1097/ACI.0b013e328346566f.

34. Shabo, Y.; Barzel, R.; Margoulis, M.; Yagil, R. Camel milk for food allergies in children. *Isr. Med. Assoc. J. IMAJ* **2005**, *7*, 796–798.
35. Merin, U.; Bernstein, S.; Bloch-Damti, A.; Yagil, R.; van Creveld, C.; Lindner, P.; Gollop, N. A comparative study of milk serum proteins in camel (*Camelus dromedarius*) and bovine colostrum. *Livest. Prod. Sci.* **2001**, *67*, 297–301, doi:10.1016/S0301-6226(00)00198-6.
36. El-Agamy, E. I.; Nawar, M.; Shamsia, S. M.; Awad, S.; Haenlein, G. F. W. Are camel milk proteins convenient to the nutrition of cow milk allergic children? *Small Rumin. Res.* **2009**, *82*, 1–6, doi:10.1016/j.smallrumres.2008.12.016.
37. El-Agamy, E. I. The challenge of cow milk protein allergy. *Small Rumin. Res.* **2007**, *68*, 64–72, doi:10.1016/j.smallrumres.2006.09.016.
38. Järvinen, K. M.; Chatchatee, P. Mammalian milk allergy: clinical suspicion, cross-reactivities and diagnosis: *Curr. Opin. Allergy Clin. Immunol.* **2009**, *9*, 251–258, doi:10.1097/ACI.0b013e32832b3f33.
39. Ehlayel, M. S.; Hazeima, K. A.; Al-Mesaifri, F.; Bener, A. Camel milk: An alternative for cow's milk allergy in children. *Allergy Asthma Proc.* **2011**, *32*, 255–258, doi:10.2500/aap.2011.32.3429.
40. Piovesana, S.; Capriotti, A. L.; Cavaliere, C.; La Barbera, G.; Samperi, R.; Zenezini Chiozzi, R.; Laganà, A. Peptidome characterization and bioactivity analysis of donkey milk. *J. Proteomics* **2015**, *119*, 21–29, doi:10.1016/j.jprot.2015.01.020.
41. Gallina, S.; Cunsolo, V.; Saletti, R.; Muccilli, V.; Di Francesco, A.; Foti, S.; Lorenzten, A. M.; Roepstorff, P. Sequence characterization and glycosylation sites identification of donkey milk lactoferrin by multiple enzyme digestions and mass spectrometry. *Amino Acids* **2016**, *48*, 1569–1580, doi:10.1007/s00726-016-2209-0.
42. Restani, P.; Fiocchi, A.; Beretta, B.; Velonà, T.; Giovannini, M.; Galli, C. L. Meat allergy: III--Proteins involved and cross-reactivity between different animal species. *J. Am. Coll. Nutr.* **1997**, *16*, 383–389.
43. Sabahelkhier, M.; Faten, M.; Omer, F. Comparative Determination of biochemical constituents between animals (goat, sheep, cow and camel) milk with human milk. **2012**, 69–71.
44. Roncada, P.; Piras, C.; Soggiu, A.; Turk, R.; Urbani, A.; Bonizzi, L. Farm animal milk proteomics. *J. Proteomics* **2012**, *75*, 4259–4274, doi:10.1016/j.jprot.2012.05.028.
45. D'auria, E.; Agostoni, C.; Giovannini, M.; Riva, E.; Zetterström, R.; Fortin, R.; Greppi, G.; Bonizzi, L.; Roncada, P. Proteomic evaluation of milk from different mammalian species as a substitute for breast milk. *Acta Paediatr.* **2005**, *94*, 1708–1713, doi:10.1080/08035250500434793.
46. Lambers, T. T.; Gloerich, J.; van Hoffen, E.; Alkema, W.; Hondmann, D. H.; van Tol, E. A. F. Clustering analyses in peptidomics revealed that peptide profiles of infant formulae are descriptive. *Food Sci. Nutr.* **2015**, *3*, 81–90, doi:10.1002/fsn3.196.