

Antibacterial potential of donkey's milk disclosed by untargeted proteomics

Valentina Spada^a, Pasquale Ferranti^b, Lina Chianese^b, Elisabetta Salimei^c, Francesco Addeo^b, Gianluca Picariello^{a,*}

^a Istituto di Scienze dell'Alimentazione, Consiglio Nazionale delle Ricerche (CNR), Via Roma 64, 83100 Avellino, Italy

^b Dipartimento di Agraria, Università di Napoli "Federico II", Parco Gussone 80055 Portici, Napoli, Italy

^c Dipartimento Agricoltura, Ambiente, Alimenti, Università del Molise, Via F. De Sanctis, 86100 Campobasso, Italy



ARTICLE INFO

Keywords:

Donkey milk

Proteomics

Antibacterial properties

Passive immunity

L-amino-acid oxidase

ABSTRACT

Donkey's milk (DM) has been extensively investigated as a valuable substitute of breast milk, often suitable to manage cow's milk protein allergy in infants. DM exhibits potent inhibitory properties against numerous microbial species. Although oligosaccharides and lipids might contribute to the antimicrobial potential, the current inventory of proteins is not able to justify the low count of microorganisms generally observed in DM.

The shotgun proteomic analysis of fractionated DM disclosed a set of 94 gene products, 41% of which have documented antimicrobial activity or are involved in transferring the passive immunity to the donkey offspring. The concerted action of lysozyme, lactoferrin, immunoglobulins provides the molecular basis for part of the DM antibacterial potential. The pH 4.6 insoluble fraction contained significant levels of L-amino acid oxidase, identified with 11 unique peptides matching the horse homologue gene product. This enzyme catalyses the oxidative deamination of amino acids into ketoacids, producing ammonia and H₂O₂. κ-casein, likely occurring as a fully O-glycosylated protein, may concur to inhibit the adhesion of pathogenic microorganisms, along with other glycoproteins.

Proteomics supports the alimentary use of DM not only as a substitute of human milk in early infancy, but also for growing children, convalescent, elderly people and general population.

Significance: Donkey's milk (DM) is acquiring increasing popularity because it is a suitable substitute of the human milk, when breastfeeding is not possible and infants suffer from cow's milk allergy.

DM is characterized by a much lower microbial load compared to ruminants' milk. This feature has been traditionally attributed to the high content of lysozyme. DM exhibits potent activity against a broad range of bacteria, viruses and fungi, suggesting that other protein components can be responsible of the antimicrobial potential. The gel-free proteomic analysis of pH 4.6-insoluble and soluble (whey) fractions demonstrated that DM contains a large number of gene products involved in antimicrobial mechanisms and in transferring passive immunity to the donkey offspring.

DM contains relatively high levels of L-amino acid oxidase that catalyses the oxidative deamination of amino acid substrates into ketoacids, with production of ammonia and H₂O₂. In combination with lysozyme, lactoferrin and immunoglobulins, the presence of L-amino acid oxidase provides the molecular basis of the antibacterial potential observed for DM. Considered the low microbial load, DM can be sanitized at mild conditions, thereby preserving many of the native nutritional traits.

Thus, DM can be considered a safe and nutritionally valid alimentary resource for growing children, convalescent, elderly people and general population.

Data of this study represent the largest inventory of proteins identified in *Equidae* milk, so far.

1. Introduction

Donkey's milk (DM) has been used since ancient times for alimentary and even cosmetic purposes, although at a far lower extent than cow's and other ruminants' milk, which can be produced more easily at

large amount. At present, the estimated contribution of *Equidae* milk to the global milk production does not exceed 0.1% [1]. A renewed interest in DM for human consumption has been mainly fuelled by its relative similarity to human milk in terms of biomolecular composition. With a total protein content of 1.5–1.8% (w/w) and a conventionally

* Corresponding author.

E-mail address: picariello@isa.cnr.it (G. Picariello).

<https://doi.org/10.1016/j.jprot.2020.104007>

Received 4 June 2020; Received in revised form 10 August 2020; Accepted 18 September 2020

Available online 08 October 2020

1874-3919/ © 2020 Elsevier B.V. All rights reserved.

estimated whey protein-to-casein ratio of 1.3, the gross composition of DM approaches better that of human milk, compared to the milk of the main dairy species. On the other hand, human and donkey's milk significantly differ for the lipid content (3.8–4.5% vs. 0.3–1.8%, respectively) and for the composition of the protein fractions, since DM contains α_{s2} -casein and relatively high levels of β -lactoglobulin, both missing in the human counterpart [2]. DM is well tolerated at oral challenge by more than 80% of the children suffering from IgE-mediated cow's milk allergy and it does not elicit allergic response even in individuals intolerant to extensively hydrolysed milk-based formulas [3]. For this reason, DM is commonly considered as a valuable alternative to hypoallergenic infant formulas for managing cow's milk allergy in infants and children [4]. The favourable whey protein-to-casein ratio with respect to cow's milk might justify on a compositional basis the lower risk of DM to trigger allergic reactions [5]. However, DM proteins exhibit very limited cross-reactivity with monoclonal antibodies raised against cow's milk [6]. Therefore, more likely the low allergenic potential depends on the divergence between the protein sequences of DM and those of the main dairy species, rather than on the compositional similarity with human milk [7].

Donkey's caseins and major whey proteins have been characterized in detail. However, research on donkey's and *Equidae* milk is still at its early stage compared to bovine milk [8–12]. Beyond the milk composition and the specific nature of nutrients, several aspects might appear peculiar to DM and could support its superior nutritional value. Apart from the relatively high content of polyunsaturated fatty acids and the release of presumed bioactive peptides [6], DM enzymes chiefly rule the interplay among composition and nutritional/health-promoting properties. Compared to the milk of other species, DM has extremely low somatic cell and total microbial counts [12,13]. As recently reviewed, many studies have supported a broad-spectrum activity of DM against viruses and bacteria [14]. Interestingly, DM exerts antimicrobial action against fungi and foodborne bacteria [15]. The antimicrobial activity of DM has been primarily attributed to the lysozyme, which occurs in the 0.3–1.1 g/L range [16], and to lactoferrin, which however is much less abundant in donkey's than in human milk (0.08 vs 0.3–4.2 g/L). Protein hydrolysates of DM maintain in part their antimicrobial activity, probably owing to the release of bioactive peptides or to the proteolytic resistance of lysozyme [17,18]. The synergic activity of whey lysozyme, lactoferrin and immunoglobulins in inhibiting microbial growth could have relevant health implications, since it would contribute to reduce the incidence of gastrointestinal infections during infancy. However, since DM inhibited *Listeria monocytogenes* more effectively than *Staphylococcus aureus* [19] and was active against Gram-negative bacteria as well [20], additional components should contribute to its overall antimicrobial potential.

Most of the information now available on the DM proteome has entailed the previous purification of protein components by gel electrophoresis or HPLC. Two-dimensional electrophoresis (2DE)-mass spectrometry (MS) proteomics is still virtually irreplaceable to investigate the qualitative and quantitative polymorphisms, post-translational modifications (i.e. phosphorylation and glycosylation) of milk proteins and alternatively spliced casein variants [21]. On the other hand, the limited dynamic range of the “classical” proteomic approaches enables the detection of a handful of high-abundance milk proteins, while many functionally active low-abundance gene products may remain undisclosed. The gel-free proteomic approaches overcome in part these drawbacks and expand the coverage of the “deep proteome”.

Almost ten years ago, DM has been analyzed with a discovery driven proteome-wide approach making use of the ProteoMiner technology to narrow down the protein dynamic range [22]. Very recently, proteins [23] or more specifically milk fat globule membrane proteins [24] of donkey's milk and colostrum have been compared by proteomic analysis. In this work, we performed a shotgun proteomic analysis of skimmed DM, after protein fractionation by isoelectric precipitation of

casein. Herein, we specifically aimed at identifying minor proteins that could justify the previously demonstrated antimicrobial activity, in the light of several recently published research articles and updates in the genomic annotation of *Equus* spp.

2. Materials and methods

Dithiothreitol (DTT), iodoacetamide (IAA), guanidine, trifluoroacetic acid (TFA), ammonium bicarbonate, Tris-base, EDTA and other chemicals were purchased from Sigma-Aldrich (St. Louis, MI, USA). HPLC-MS grade solvents were from Carlo Erba (Milan, Italy).

2.1. Sampling

Bulk DM was a pooled sample obtained from 8 lactating jennies reared in a dairy donkey farm located in Oliveto Citra (Salerno, Italy). All jennies were pluriparous, in early lactation (30–90 days lactation) and in healthy clinical conditions. Milk was manually expressed and collected in sterile plastic tubes. Soon after collection, milk was added with phenylmethanesulfonyl fluoride (PMSF) serine-protease inhibitor (Sigma-Aldrich) at final 1 mM concentration, immediately refrigerated and then stored at -26°C until used. Total protein content was assessed by Kjeldahl method ($N \times 6.38$). Total bacteria count and somatic cell count were monitored performing two analysis of bulk raw milk weekly for two months, using a Bactoscan 8000 and a Fossomatic 360 (both from Foss, Hillerød, Denmark), respectively.

2.2. Protein fractionation

Milk (10 mL aliquots) was thawed in an ice-cold bath and skimmed by centrifugation twice at 3500 $\times g$, 20 min, 4°C . The floating fat layer was hardened 10 min at -20°C and removed with a spatula. The casein fraction was precipitated from skimmed milk at pH 4.6 with 0.2 M sodium acetate/acid acetic buffer 4 h in an ice bath and separated from the supernatant by centrifugation at 4500 $\times g$ for 30 min, 4°C . To remove residual whey proteins, the pellet was resuspended into 50 mM Tris pH 8.0 and re-precipitated with 0.2 M sodium acetate/acid acetic buffer, pH 4.6. Finally, the protein isoelectric precipitate was washed twice with 50 mL of -20°C cold acetone and lyophilized. Whey was used for hydrolysis of whey proteins. Hereinafter, the two protein pools will be also referred to as insoluble and soluble fractions.

2.3. RP-HPLC

Aliquots (40 μg) of the protein fractions were analyzed by RP-HPLC to profile the most abundant gene products. To this aim, the insoluble fraction (1 mg) was dissolved in 1 mL of 10% aqueous acetonitrile, while whey was injected without any previous treatment. The HPLC separation was carried out using an HP 1100 modular chromatographer (Agilent technology, Palo Alto, CA, USA) equipped with a C4 reversed-phase Jupiter 300 column (Phenomenex, Torrance, CA, USA), 5 μm particle size, 250×2.1 mm internal diameter. Solvent A was 0.1% TFA in LC-MS grade water (v/v) and solvent B 0.1% (v/v) TFA in acetonitrile. Proteins were separated using a 30–50% or a 35–55% gradient of solvent B in 60 min for the insoluble fraction and whey proteins, respectively, at a flow rate of 0.2 mL/min. Effluents were monitored using a multi-waves UV-Vis detector at 214 and 280 nm. The most abundant peaks of whey proteins were manually collected and some of them individually analyzed as intact proteins by MS.

2.4. Reduction/alkylation and trypsin digestion

The isoelectric precipitate (10 mg) and 1 mL of the isoelectric supernatant were dissolved or diluted, respectively, in 10 mL of denaturing and reducing buffer (6 M guanidine HCl, 50 mM Tris, 1 mM EDTA, 10 mM DTT, pH 8.0) and incubated at 55°C for 1 h under N_2 .

After reduction, cysteine residues were alkylated with 55 mM IAA, for 40 min at room temperature in the dark.

Afterwards, proteins were quantified with the micro-Lowry modified method (kit from Sigma-Aldrich), diluted 10 fold with 50 mM ammonium bicarbonate pH 7.8 and 1 mg of the two fractions was sequentially digested with 1/100 sequencing grade Lys-C, overnight at 37 °C, and with 1/50 sequencing grade modified trypsin (enzymes from Promega, Madison, WI, USA) overnight at 37 °C. Resulting peptides were solid-phase extracted and purified from 3 mL of the solution with C18 Sep-Pak prepacked cartridges (Waters, Wilford, MA, USA), washing extensively with 0.1% (v/v) TFA and eluting with 3 mL of 70% acetonitrile/0.1% (v/v) TFA. Peptides were vacuum dried, lyophilized and re-constituted in 0.1% formic acid for analysis.

2.5. Nanoflow-high performance liquid chromatography-electrospray-tandem mass spectrometry (nHPLC-MS/MS)

nHPLC-MS/MS analyses were performed using an Ultimate 3000 nanoflow ultrahigh performance liquid chromatography (Dionex/Thermo Scientific, San Jose, CA) coupled to a Q-Exactive Orbitrap mass spectrometer (Thermo Scientific). Nearly 2 µg of the peptide pools were loaded through Acclaim PepMap 100 trap columns (75-µm i.d. × 2 cm; Thermo Scientific) using a FAMOS autosampler (Thermo Scientific). Eluent A was 0.1% formic acid (v/v) in LC-MS grade water; eluent B was 0.1% formic acid (v/v) in 80% aqueous acetonitrile. Peptides were separated using an EASY-Spray™ PepMap C18 column (25 cm × 75 µm) with 2 µm particles and 100-Å pore size (Thermo Scientific). The separation was done with a 2–50% gradient of B over 120 min after 20 min of isocratic elution at 2% B, at a constant flow rate of 300 nL/min. MS1 precursor spectra were acquired in the positive ionization mode scanning the 1800–300 *m/z* range with resolving power of 70,000 full width at half maximum (FWHM), an automatic gain control (AGC) target of 1×10^6 ions, and maximum ion injection time of 100 ms. The spectrometer operated in full scan MS1 and data-dependent acquisition mode, selecting up to the 10 most intense ions for MS/MS fragmentation and applying a 12-s dynamic exclusion. Fragmentation spectra were obtained at a resolving power of 17,500 FWHM. Ions with charge +1 or greater than +6 were excluded from the MS/MS fragmentation. Spectra were elaborated using the software Xcalibur version 3.1 (Thermo Scientific). Both the protein fractions were analyzed in triplicate.

2.6. Database search and protein identification

Raw or .mgf files generated from the nHPLC-MS/MS runs were searched against the *Equus* NCBI database (downloaded on November 2019) using open source Protein Prospector Batch-Tag Web tool (<http://prospector2.ucsf.edu>) or licensed Mascot MS/MS ion search (version 2.6, <http://matrixscience.com>), respectively. Outputs are reported according to the scoring system of Protein Prospector Batch-Tag Web. The protein variants were assigned with the Proteome Discoverer 2.1 software (Thermo Scientific) making use of SEQUEST search engine. In this case raw spectra were searched against a home compiled database containing the sequence of already characterized isoforms. For all the searches, parameters were the following: Cys carbamidomethylation as a static modification; Met oxidation, pyroglutamic acid at N-terminus Gln, Asn deamidation and Ser/Thr phosphorylation as variable modifications; mass tolerance value of 10 ppm for precursor ion and 0.02 Da for MS/MS fragments; trypsin as the proteolytic enzyme, including the possibility of semi-trypsin cleavages; missed tryptic cleavage up to 2. The dataset was also reprocessed selecting the “no cleavage specificity” option, to identify proteins from possible non-trypsin peptides. Protein identification scores were calculated by Target Decoy Peptide Spectrum Matches (PSMs) filtering working at 1% peptide-level false discovery rate (FDR) that were matched to proteins identified at 1% protein-level FDR. Searches arising from three technical replicates

were merged and only the entries identified with at least two PSMs and protein score > 22 were accepted. Exponentially modified Protein Abundance Index (emPAI) values were obtained from the Mascot search engine. The emPAI values < 0.01 were not reported.

2.7. Functional classification

Protein entries were categorized according to different functional classes using the bioinformatics resource KEGG (Kyoto Encyclopedia of Genes and Genomes). Multifunctional proteins were assigned to a class based on KO (KEGG Orthology) “context-dependent” functional orthologs. Proteins were grouped according to the classification by Bevan et al. (1998) [25]. Functional classification of genes coding for the identified protein entries was carried out also with the PANTHER classification system (release 2020) in Gene Ontology (<http://geneontology.org/>).

3. Results and discussion

3.1. Protein amount, total microbial and somatic cell count

The protein content of DM accounted for 1.52 g/L, which is consistent with values reported in literature [26]. DM pH was 7.11 ± 0.04 and it did not significantly vary within three days storage at 4 °C. Interestingly, DM did not coagulate within five days after milking at room temperature, at variance of a control raw bovine milk sample, which visibly coagulated already three days after milking. Total bacterial count (colony forming units, CFU) and somatic cell count (SCC) averaged 3.94 ± 0.07 log CFU/mL and 3.68 ± 0.09 log SCC/mL, respectively, confirming good both milk hygiene and mammary health status. Equine milk total bacterial count and SCC can vary within an ample range depending on management procedures and hygiene at farm level, for animals and milk [26]; however, values observed in the present study are low and abundantly below the threshold allowed for commercial raw cow's milk, as established by EU Regulation No. 853/2004 (total bacterial count and SCC should not exceed 100,000/mL CFU and 400,000/mL, respectively).

3.2. RP-HPLC

The RP-HPLC analysis of the protein fractions obtained by isoelectric precipitation enabled the detection of only the most abundant components, which were assigned by comparison with previous literature data [7,12,27] or by direct injection MS (Fig. 1). Casein fractions are highly polymorphic due to multiple phosphorylations and alternative splicing events. A detailed characterization of HPLC-isolated casein isoforms can be found in Chianese et al. [12]. Like other biological fluids, DM contains a few abundant proteins and a high number of gene products covering an almost continuous range of lower concentrations. The “peptide-centric” techniques, combined with the recently developed bioinformatic tools for spectral matching, expand the dynamic range and achieve a much higher proteome coverage compared to “protein-centric” electrophoresis- and HPLC-based strategies [28]. Both insoluble and whey protein fractions were individually digested by trypsin and the resulting peptides analyzed by nHPLC-MS/MS, using a high resolution/high sensitivity mass spectrometer.

3.3. Protein identification and gross quantification

Overall, 94 gene products were confidently identified in the skimmed DM using shotgun proteomics (Table 1), which grossly corresponded to the 106 entries previously identified after ProteoMiner enrichment of minor proteins [22]. Many of the proteins were identified by homology with *Equus caballus*, as they are not annotated yet in the sequence databases of *Equus asinus*. Recently, almost 300 gene products have been identified in donkey's milk and colostrum [23] and,

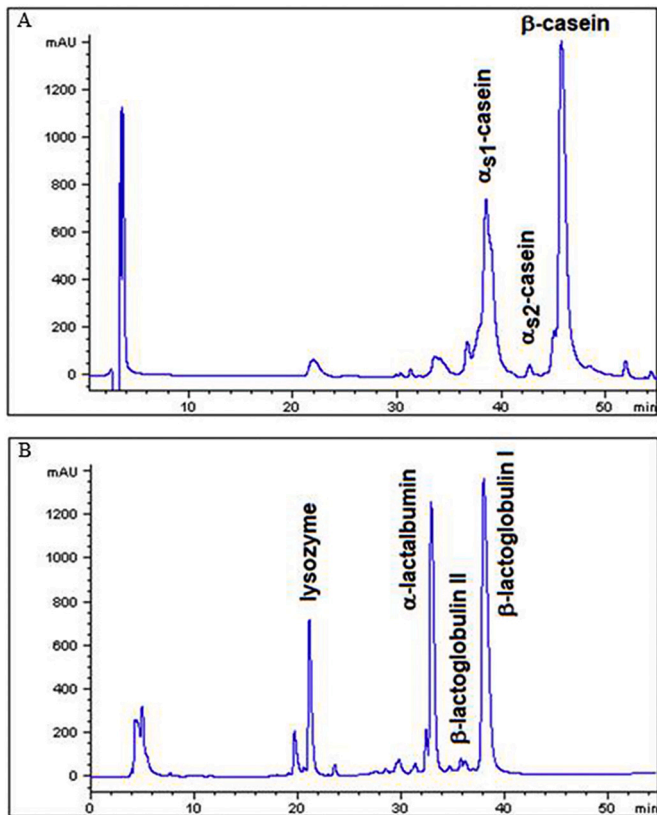


Fig. 1. RP-HPLC separation of the DM pH 4.6-insoluble (panel A) and soluble (panel B) fractions.

strikingly, more than 900 protein entries in donkey's milk fat globule membrane [24]. Different inventories interesting the minor protein components can be the result of slightly divergent analytical and bioinformatic workflows, rather than breed-dependent heterogeneity. Interestingly, 40 gene products of Table 1 – after exclusion of protein isoforms – are described for the first time in DM, to the best of our knowledge. As expected, caseins, especially high-abundance β- and αs1-casein, were the most represented proteins in the insoluble fraction, although the casein content of DM is almost three-fold lower than bovine milk. β-lactoglobulin (β-Lg) I and II, α-lactalbumin (α-La), serum albumin and lysozyme dominated the whey protein fraction. Both protein fractions contained a relatively high number of non-tryptic peptides, most of which derived from either caseins or whey proteins, as already reported for human milk proteins [29]. For this reason, the search-assisted peptide identification was carried out with “no enzymatic” specificity, thus strikingly expanding the coverage of the protein sequences. For instance, β- and αs1-casein were identified in the insoluble fraction through 150 and 122 unique PSMs, respectively, as well as whey β-Lg I and serum albumin were identified through 153 and 151 PSMs, respectively. Thus, likewise the human counterpart, DM proteins are rather extensively hydrolysed by endogenous proteases (e.g. plasmin, cathepsins). In contrast, the formation of amino acid ladders by effects of secondary proteolysis mediated by bacterial exopeptidases was a marginal event in DM, because of the low microbial load.

The PSMs mapped almost entirely the sequence of the dominant proteins, except for the phosphorylated casein regions. Due to their low intrinsic efficiency to ionization, the identification of casein phosphopeptides would require a dedicated analytical workflow involving a previous selective purification of phosphorylated regions or an enzymatic removal of phosphate groups.

As introduced above, 2DE-MS proteomics is the optimal approach to

identify genetic polymorphism and non-allelic variants. However, shotgun proteomics can be exploited in a targeted fashion to monitor the presence of previously characterized proteoforms.

For instance, Cunsolo et al. [11] described two constitutive variants of the previously described donkey's αs1-casein: one variant resulted from the deletion of Gln₈₃ (A¹ variant), which arose from the alternative splicing of exon 11 in the mRNA transcript; the other showed the insertion of the pentapeptide HTPRE between the amino acid residues Leu₃₃ and Glu₃₄ (B variant), in analogy with mare αs1-casein, as a consequence of exon 5 skipping. Accordingly, we detected the semi-tryptic peptide αs1-casein f(80–89) with inserted/deleted Gln₈₃ ($M + H^+ = 1156.93$ and 1028.57 , respectively) (Supplementary material, Fig. S1) and αs1-casein f(28–41) both in the canonical and longer forms carrying the insertion of HTPRE ($M + H^+ = 1712.87$ and 2333.18 , respectively) (Supplementary material, Fig. S2). Similarly, the presence of the non-allelic αs2-casein variant resulting from the internal deletion of Asn₁₇₆-Gln₁₈₀ encoded by exon 17, which is also constitutively spliced in mare αs2-casein [30], was confirmed by the detection of the junction tryptic peptide, namely αs2-CN f(174–184) lacking the internal N₁₇₆KINQ₁₈₀ sequence (Supplementary material, Fig. S3).

Clearly, with the “bottom-up” approach the information about the entire protein chains is lost. Thus, it was not possible to confirm the presence of the αs1-casein proteoform carrying both the sequence modifications (B¹ variant). The identification of other splicing variants from αs2- and β-casein would be possible in principle with the gel-free approach, but it would require an additional step of hydrolysis of phosphate groups or, in alternative, the selective enrichment of relevant phosphopeptides, because the expected junction peptides are both phosphorylated.

As concerns whey proteins, the occurrence of β-Lg II D and E (differing from each other by three punctual amino acid replacements, namely Asp₂ → Asn, Arg₁₈ → Lys and Val₂₅ → Ala) was supported by the detection of diagnostic peptides derived from both the genetic variants (Supplementary material, Fig. S4) [27]. Determination of the intact molecular weight (MW) of HPLC-isolated β-Lg II confirmed the presence of the D (measured 18,311.16 Da, calculated 18,311.12 Da) and E (measured 18,256.13 Da, calculated 18,256.08 Da) variants. Sequence analysis of peptides demonstrated that donkey β-Lg I corresponds to the horse homologue carrying the single conservative amino acid substitution V₁₅₀→I, typical of the donkey β-Lg I variant B. The measured molecular weight of unhydrolyzed β-Lg I (18,510.18 Da) closely matched the calculated one (18,510.28 Da).

The nHPLC-ESI-MS/MS runs were processed with three independent search engines, using the same searching parameters. Protein entries in Table 1 are the common outputs of Protein Prospector and Mascot search engines. The Mascot search engine provided the emPAI values (Table 1), which constitute a gross quantitative ranking of the protein abundance. The emPAI value is defined by the formula:

$$\text{emPAI} = \frac{N_{\text{observed}}}{10^{N_{\text{observable}} - 1}}$$

where N_{observed} is the number of experimental PSMs and $N_{\text{observable}}$ is the calculated number of observable PSMs [31]. Therefore, the emPAI values exponentially increases with the number of PSMs.

The emPAI values observed for the most abundant DM proteins spanning over three orders of magnitude (10^7 – 10^{10}) arose from the high number of additional PSMs identified with respect to the number of theoretical tryptic peptides, due to the occurrence of unexpected non-tryptic peptides. The emPAI values obtained in this study spanned over 10 orders of magnitude, thereby reflecting the ample dynamic range of protein abundance in DM, just like milk from other species and biological fluids in general.

Table 1

List of the identified proteins in donkey's milk ($n = 94$). Protein entries are categorized according to the KEGG functional classification. Protein entries identified for the first time in DM are highlighted in bold. Immunoglobulin (Ig) isoforms have not been considered as newly identified entries.

Protein Name (Organism)	Gene Symbol	Accession Number NCBI	Accession Number UniProt	Coverage (%)	Peptide Spectrum matches (max. number)	Protein Score	MW [kDa]	Casein protein- MASCOT emPAI	Whey protein- MASCOT emPAI
Nutrition and Endocrine system									
α_{S1} -casein (<i>E. asinus</i>)	CSN1S1	238055131 NP_001075352	Q8SPR1	82.7	122	3474.5	24.4	2.39E + 08	58.62
α_{S2} -casein (<i>E. asinus</i>)	CSN1S2	387942467	B7VGF9	69.1	37	778.2	27.7	88.9	3.83
Apolipoprotein A-II (<i>E. caballus</i>)	APOA2	149759787 XP_001503846	F6RM73	26	2	94.8	11.3		0.02
Apolipoprotein A-I-like 1 (<i>E. caballus</i>)	APOA1	149716548 XP_005611649	F6Z2L5	50	10	279.4	30.3	0.11	0.02
β -casein (<i>E. asinus</i>)	CSN2	270118591 NP_001075321	P86273 Q9GKK3	73.4	150	4148.5	27.1	5.21E + 08	83.7E + 02
κ -casein (<i>E. asinus</i>)	CSN3	323714054 NP_001075353	P82187 Q28400	53.8	37	970.4	21.0	13.1	0.34
Perilipin-2 (<i>E. caballus</i>)	PLIN2	338719732 XP_001917546		10.9	3	136.9	55.1		0.02
Genetic information processing									
45 kDa calcium-binding protein (<i>E. caballus</i>)	SDF4	149758374 XP_001496575		9.3	2	54.0	41.5	0.08	0.01
Granulin isoform X1 (<i>E. caballus</i>)	GRN	149723703 XP_001489791	F7AB03	18.6	7	247.7	62.6		0.01
Nuclear receptor coactivator 6 isoform X1 (<i>E. caballus</i>)	NCOA6	149733205 XP_023482163		2.9	3	57.5	219.2		X
Nucleobindin-1 (<i>E. caballus</i>)	NUCB1	194215746 XP_023505840		15	5	196.3	42.0		0.5
Peptidyl-prolyl cis-trans isomerase B (<i>E. caballus</i>)	PPIB	153792484 NP_001093231	A5YBL8	28.7	4	163.0	23.8		0.05
Immune system									
α_1 -antitrypsin (<i>E. caballus</i>)	SPI2	100065158 NP_001108005		6.2	4	24.9	46.9		0.05
Antithrombin-III (<i>E. caballus</i>)	SERPIN- C1	149708147 XP_014594947		5.6	2	61.0	52.4	0.5	0.06
β_2 -microglobulin (<i>E. asinus</i>)	B2M	75072805 NP_001075971	P30441 K9K497	15.3	3	127.0	13.5	0.56	2.06
C-C motif chemokine 28-like (<i>E. caballus</i>)	CCL28	338718753 XP_005604346	F6VWX6	31.3	7	243.3	16.6	0.23	0.10
CD5 antigen-like (<i>E. caballus</i>)	CD5L	338725177 XP_005610044	F7BXD8	7	2	66.0	39.3		0.01
Clusterin precursor (<i>E. caballus</i>)	CLU	126352584 NP_001075413	Q29482	10.9	5	121.8	52.2		X
Complement factor B (<i>E. caballus</i>)	CFB	149732066 XP_001492602	A0A3Q2IDD2	4.8	2	49.1	86.0		X
Fibrinogen α chain isoform 2 (<i>E. caballus</i>)	FGA	338722639 XP_003364581	P02672	12.7	4	161.4	67.7		X
Fibrinogen γ chain (<i>E. caballus</i>)	FGG	338722637 XP_001914833		6.6	3	75.2	49.8		0.01
Gelsolin (<i>E. caballus</i>)	GSN	126352530 NP_001075422	Q28372	11.4	4	195.5	80.8		0.17
IgGfC-binding protein (<i>E. caballus</i>)	FCGBP	1333554384 XP_023505501		3	5	28.56	273.2		0.0
Ig μ heavy chain constant chain secreted form (<i>E. caballus</i>)	IGHM	51831151 1199748903	AAU09792	35.3	10	355.8	49.4		0.05
Ig heavy chain Mem5-like (<i>E. asinus</i>)	No gene	958789758 XP_014723377		21	8	100.58	28.4		X
Ig heavy chain variable region (<i>E. caballus</i>)	No gene	347825890	AEP27026	63.2	4	150.8	12.3		0.05
Ig heavy chain V-D-J region, partial (<i>E. caballus</i>)	No gene	300386798	ADK09424	16.9	2	71.3	19.6		0.01
Ig J chain (<i>E. caballus</i>)	JCHAIN	338723570 664743142 NP_001271464	A0A0B4J1C4	48.7	17	699.0	20.3		0.95

(continued on next page)

Table 1 (continued)

Protein Name (Organism)	Gene Symbol	Accession Number NCBI	Accession Number UniProt	Coverage (%)	Peptide Spectrum matches (max. number)	Protein Score	MW [kDa]	Casein protein- MASCOT emPAI	Whey protein- MASCOT emPAI
Ig α constant heavy chain, partial (<i>E. caballus</i>)	IGHA	32331167	AAP80145	37.6	26	1084.1	37.0		0.63
Ig γ 1 heavy chain constant region, partial (<i>E. caballus</i>)	IGHC1	15020816	CAC44624	19	5	69.97	37.4		0.10
Ig γ 2 heavy chain constant region, partial (<i>E. caballus</i>)	IGHC2	15026999	CAC44761	15	4	31.10	37.5		0.07
Ig γ 3 heavy chain constant region, partial (<i>E. caballus</i>)	IGHG3	18996193	CAC86339	11	3	39.45	38.6		0.05
Ig γ 5 heavy chain constant region (<i>E. caballus</i>)	IGHG5	18996195	CAC86340	63.8	23	940.5	35.9		0.27
Ig γ 6 heavy chain constant region, partial (<i>E. caballus</i>)	IGHG6	18996197	CAC86341	39	7	91.26	35.9		0.11
Ig γ 7 heavy chain (<i>E. caballus</i>)	IGHG7	42528291	AAS18414	45.1	18	732.9	35.7		0.13
Ig λ chain V–I region BL2-like isoform 2 (<i>E. caballus</i>)	No gene	338727471 XP_003365493		19.2	14	530.5	24.4		X
Ig λ constant 7 (<i>E. caballus</i>)	No gene	1333703395 XP_023503816		21	4	58.25	17.0		X
Ig λ light chain V–J region (<i>E. caballus</i>)	No gene	300387570 ADK09799		25.9	9	45.7	23.2	X	0.15
Ig λ -1 light chain (<i>E. caballus</i>)	No gene	1333703565 XP_023503887		20	4	40.12	24.2		1.35
Lactotransferrin (<i>E. asinus</i>)	LTF	958736226 NP_001157446	O77811F6XLB1	31	13	107.22	77.3		1.27
L-amino-acid oxidase (<i>E. caballus</i>)	No gene	1333593388 XP_023487776		29.8	11	445.7	59.0	1.01	0.04
λ -immunoglobulin (<i>E. caballus</i>)	IgL	291462	AAA50971	27.1	8	260.7	20.4		0.07
Lysozyme C, milk isozyme (<i>E. asinus</i>)	LYZ	958727973 XP_001491556	P11375 , P11376 F6PPX7	96	83	3373.0	16.8	1.75	6.53E+07
Membrane cofactor protein (<i>E. caballus</i>)	No gene	194210457 XP_001915894		32	5	182.0	20.2		0.04
Monocyte differentiation antigen CD14 precursor (<i>E. caballus</i>)	CD14	126352663 NP_001075396	Q9TTT3	10	2	69.7	39.1		X
Neutrophil gelatinase-associated lipocalin-isoform X1 (<i>E. caballus</i>)	LCN2	338720560 XP_023484750		27.6	4	117.3	23.1		0.03
Osteopontin (<i>E. asinus</i>)	SPP1	149701529 958701516 XP_001496202	F7AYC1	40.6	26	929.0	35.0	0.57	0.60
Platelet glycoprotein 4 isoform X1 (<i>E. caballus</i>)	CD36	149704540 XP_001487957		5.5	2	39.0	52.8	X	
Polymeric immunoglobulin receptor (<i>E. asinus</i>)	PIGR	958675515 NP_001271468		43	28	1014.1	83.2		0.10
Prothrombin (<i>E. caballus</i>)	F2	338712085 XP_001490892		8.7	2	57.6	70.3		X
Semaphorin-7A (<i>E. caballus</i>)	SEMA7A	194206477 XP_001917710	F7C7V8	7.2	3	80.1	74.5		0.01
Proteases									
Cathepsin B (<i>E. caballus</i>)	CTSB	149698064 XP_001498242	P07688	9.7	2	80.3	37.7		0.03
Cathepsin L2 (<i>E. caballus</i>)	CTSV	149755226 XP_001494409	Q5E998	21.3	5	144.6	37.4		0.02
Cathepsin S (<i>E. caballus</i>)	CTSS	149751225 XP_001490531	F6VUW2	14.2	3	111.6	37.3		0.02
Plasminogen (<i>E. caballus</i>)	PLG	130315 XP_001500552	F6USP9	13.3	2	61.4	37.1		0.02

(continued on next page)

Table 1 (continued)

Protein Name (Organism)	Gene Symbol	Accession Number NCBI	Accession Number UniProt	Coverage (%)	Peptide Spectrum matches (max. number)	Protein Score	MW [kDa]	Casein protein- MASCOT emPAI	Whey protein- MASCOT emPAI
Metabolism									
α -1B-glycoprotein (<i>E. asinus</i>)	A1BG	958781274 338710440 XP_023507318		29	5	218.7	60.2		0.02
α -2-HS-glycoprotein-like (<i>E. caballus</i>)	AHSG	194222675 XP_005601929	F7C450	13.2	6	201.4	38.9		0.04
α -2-macroglobulin-like (<i>E. caballus</i>)	A2ML	194211675 XP_014596074		2	2	50.3	164.0		0.00
α -lactalbumin (<i>E. caballus</i>)	LALBA	126000 194211931 125991 XP_001915824	P08896 P08334 A0A5K1V2X0	100	126	5571.9	16.2	5.50	1.25E+9
β -1.4- galactosyltransferase 1 (<i>E. caballus</i>)	B4GALT1	149737071 1333610968 XP_023483236		8	2	84.2	31.7		0.02
Cystatin-C (<i>E. caballus</i>)	CST3	194224089 XP_023481760		12.5	2	74.9	17.8		X
Haptoglobin (<i>E. asinus</i>)	No gene	149699777 958695472 XP_001497860	F6XWM5	55.3	12	424.6	38.5		0.05
Histidine-rich glycoprotein-like (<i>E. caballus</i>)	HRG	338715996 XP_023479252		17	4	153.2	49.2		X
Xanthine dehydrogenase/ oxidase isoform X3 (<i>E. caballus</i>)	No gene	149727656 XP_023475059		3.8	4	142.5	146.7		0.01
Nonspecific or unknown function									
Collagen α -1(I) chain (<i>E. sp.</i>)	COL1A1	XP_023472734	C0HJN9	30	9	43.17	79.2	X	X
Core protein VP3 (African horse sickness virus)	Segment- 3		P32509	5	2	22.1	103.3		X
Glycodelin (<i>E. asinus</i>)	PAEP	958774826 XP_014719222		18	2	26.09	21.7	X	
Glycoprotein gp2 (Equine herpesvirus)	GP2	XP_005598938	Q6S6W0 F6WVV0	6	2	23.8	86.5		X
Golgin subfamily A member 3 (<i>E. caballus</i>)	GOLGA3	149756806 XP_001493460		2.1	2	39.6	169.3	X	
Hypothetical protein LOC100057656 (<i>E. caballus</i>)	LO- C100057- 656	338724846 XP_001490989		8.3	4	94.2	71.9	X	X
Neutral α -glucosidase AB isoform x2 (<i>E. caballus</i>)	GANAB	194218317 XP_001502746		20.3	15	423.9	106.6	0.48	
Protein broad-minded (<i>E. caballus</i>)	TBC1D32	149723152 XP_023506762		1.8	2	31.5	145.2		X
Sushi repeat-containing protein SRPX (<i>E. caballus</i>)	SRPX	338728990 XP_005614121		2.6	2	48.1	50.4	0.06	
Transforming growth factor- β -induced protein ig-h3- like. Partial (<i>E. caballus</i>)	TGFBI	338713612 XP_005599767		4.8	3	86.4	76.5	0.09	X
Uncharacterized protein LOC103566764 (<i>E. przewalskii</i>)	No gene	664767961		23	3	22.13	17.1		0.07
Signaling and cellular processes									
Butyrophilin subfamily 1, member A1 (<i>E. caballus</i>)	BTN1A1	149754992 1333600148 XP_023480339		24.9	15	536.7	67.7	0.31	0.04
Chondroitin sulfate proteoglycan 4 (<i>E. caballus</i>)	CSPG4	338717372 XP_005602958	F6T201	2.7	2	38.6	250.6		X
Chordin-like protein 2 (<i>E. caballus</i>)	CHRD12	194213448 XP_014597408		26.8	8	255.3	49.9	0.14	0.02
Fibroblast growth factor- binding protein 1-like (<i>E. caballus</i>)	No gene	338723498 XP_005608931		30.4	5	164.5	27.0		0.02

(continued on next page)

Table 1 (continued)

Protein Name (Organism)	Gene Symbol	Accession Number NCBI	Accession Number UniProt	Coverage (%)	Peptide Spectrum matches (max. number)	Protein Score	MW [kDa]	Casein protein- MASCOT emPAI	Whey protein- MASCOT emPAI
Guanine nucleotide-binding protein G(t) subunit α -2 (<i>E. caballus</i>)	GNAT2	149708712 XP_001495066	egfVZ1	4.2	2	44.7	40.1	X	
Hyaluronan and proteoglycan link protein 3-like (<i>E. caballus</i>)	HAPLN3	338717539 XP_023506021		8.5	2	47.2	39.6	0.08	
Lactadherin (<i>E. caballus</i>)	MFGE8	194206303		35.8	15	509.7	45.7	1.38	0.15
Milk fat globule-EGF factor 8 splice variant. Partial (<i>E. caballus</i>)	MFGE-8	330688384	AEC32944	47.9	20	678.0	43.3	2.91	
Serum amyloid A protein isoform X2 (<i>E. caballus</i>)	No gene	149719501 XP_001505010	F7C0Y4	81.5	26	894.1	14.3	0.53	2.74
Serum amyloid A protein-like (<i>E. asinus</i>)	SAA1	958720353 958720357 XP_001505006	P19857	56	9	157.19	14.5		0.60
Tetranectin (<i>E. caballus</i>)	CLEC3B	338714954 XP_003363176		10.4	2	42.5	22.3		X
Transporters									
β -lactoglobulin-1 (<i>E. asinus</i>)	LGB1	126723662 125913 958685684 NP_001075962	P02758 P13613	83.3	153	6716.5	24.9	22.39	8.97E+10
β -lactoglobulin-2 (<i>E. asinus</i>)	LGB2	125904 NP_001075963	P19647 P07380	94	14	361.29	20.1	1.71	3.37E+02
Minor allergen Can f 2-like (<i>E. przewalskii</i>)	No gene	664772923 XP_023485146		24	38	2605.73	34.3		X
Prosaposin isoform X1 (<i>E. caballus</i>)	PSAP	149689950 XP_001503814		5.3	2	74.6	58.5		X
Serotransferrin precursor (<i>E. caballus</i>)	TF	126352628 823630974 NP_001075415	P27425	48.2	32	1072.9	78.1		1.93
Serum albumin (<i>E. asinus</i>)	ALB	76363596 664734900 NP_001075972	P35747 Q5XLE4	95.9	151	6565.1	68.5		2.32E+03
Transcobalamin-1 (<i>E. asinus</i>)	TNC1	958768443 149725169 XP_005598248	F6T7K7	8	3	29.1	48.1		X
Vitamin D-binding protein (<i>E. asinus</i>)	GC	958718239 149701606 XP_001489400	F6T0P6	9.6	4	163.2	54.3		0.02

3.4. Functional classification

The deep human milk proteome consists of thousands minor components that have been recently catalogued [32]. The DM proteome is expected to be extended comparably. Nevertheless, only a limited number of proteins were identified in skimmed DM mainly depending on the incompleteness of *Equus* databases. Moreover, our study did not consider the gene products associated with the milk fat globule membrane or milk extracellular vesicles.

The DM protein entries identified were categorized in eight classes according to the functional classification of KEGG, including a broad group of gene products with nonspecific or unknown function (Table 1). Interestingly, a glycoprotein and a core protein of the inner capsid from equine herpes virus have been included in this latter group.

Caseins were classified primarily as nutrition-relevant proteins, while β -Lg has been included within the transporters, as it can serve as a nonspecific carrier of lipids. Endogenous proteases of plasmatic (*i.e.* plasminogen) or lysosomal (*i.e.* cathepsins) origin with possible multi-functional activities were categorized in a specific class. In terms of protein number, the gene products involved in the immune protection of the donkey's offspring were the richest class making up 41% total

(Fig. 2). Alternatively, DM proteins were classified in three categories according to GO classification, namely biological processes, molecular function, and cellular components (Fig. 3). Notably, several protein entries are catalogued with multiple biological processes or molecular functions, thereby justifying why the total number of GO catalogued proteins apparently exceeded the total number of proteins identified. Based on the cellular localization, the most numerous group of protein entries was represented by extracellular components, as expected. In agreement with recent results, a significant fraction of DM proteins is involved in binding and transport of metabolites [23]. In apparent contrast with our hypothesis (see below), only a relatively low number of components was categorized within the immune system processes, since many of the components also involved in the host defence system are catalogued as proteins exerting different primary functions (*e.g.* osteopontin, κ -casein, milk fat globule membrane-associated glycoproteins).

Lysozyme is strikingly high in DM, ranging between 0.3 and 1.1 g/L [16], as also confirmed from its relatively high emPAI values. According to other authors, concentration of lysozyme can even exceed 3.0 g/L in local donkey breeds [14,33]. Lysozyme has enzymatic activity catalysing the hydrolysis of the 1,4- β -linkages between

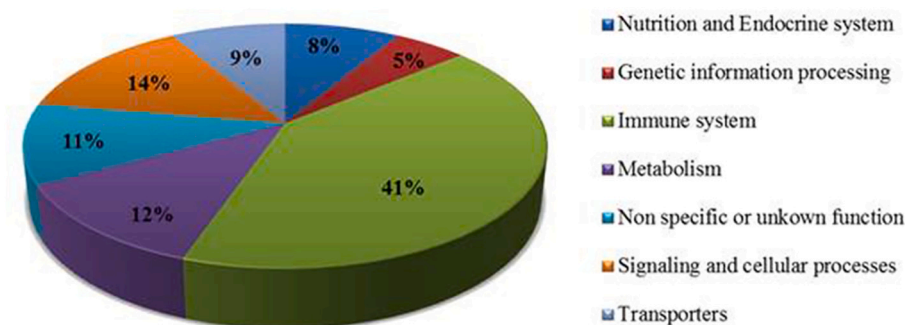


Fig. 2. KEGG (Kyoto Encyclopedia of Genes and Genomes) functional classification of all proteins listed in Table 1.

acetylmuramic acid and *N*-acetylglucosamine in peptidoglycan, which is the main cell wall constituent of Gram-positive bacteria. Besides causing bacterial lysis, lysozyme is also able to inhibit viruses, parasites and fungi [34]. Noteworthy, the highest activity of lysozyme has been recorded at the early stages of the jenny lactation, to meet the augmented requirements of passive immunity for the donkey foal [15]. The RP-HPLC chromatogram of the whey protein extract confirmed that lysozyme occurred at levels comparable with those of α -La and β -Lg (Fig. 1B).

Lysozyme is accessioned as C-type in UniprotKB, indicating the structural homology with α -La and chicken (C) egg lysozyme. Two genetic variants have been characterized in DM so far, namely lysozyme A and B [35]. Lysozyme A and B differ by three amino acid substitutions ($N_{49} \rightarrow D$, $Y_{52} \rightarrow S$, and $S_{61} \rightarrow N$), all localized within the tryptic peptide 47–62 [36]. The UniprotKB sequence of lysozyme was entirely mapped except for the protein fragment 47–62, which was manually sequenced as the tryptic peptide of lysozyme B variant, carrying the three characteristic amino acid substitutions (Supplementary material, Fig. S5).

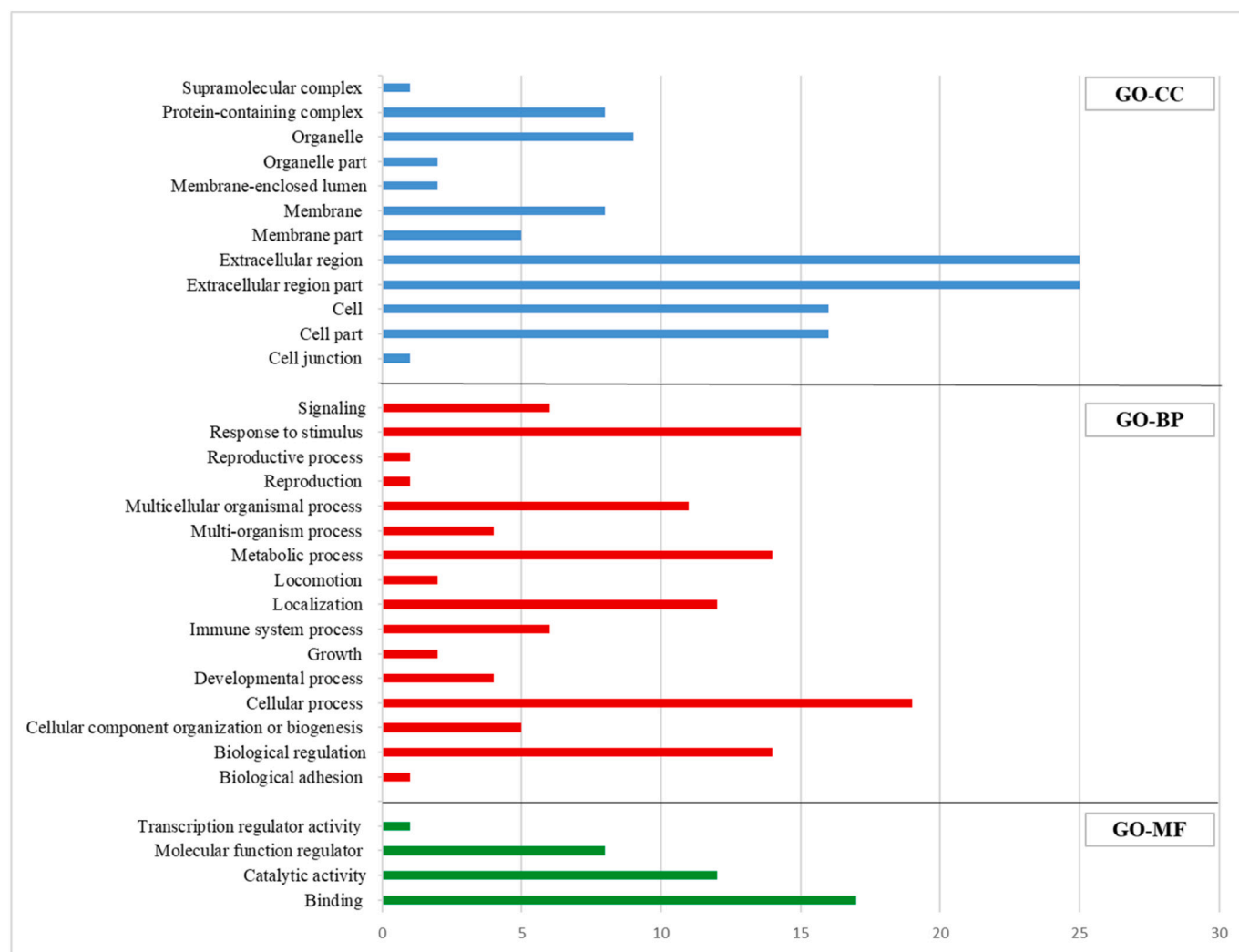


Fig. 3. Gene Ontology (GO) annotation of proteins in DM. GO-BP: GO-biological processes; GO-MF, GO-molecular function; GO-CC, GO-cellular components. Notably, proteins entries can be catalogued in GO with multiple biological or molecular functions.

The donkey's peptide 47–62 matched the peptide of horse's lysozyme and occurred in two almost co-eluting forms, containing alternatively N₄₉ or D₄₉, suggesting that the modification N₄₉→D could be the result of a spontaneous deamidation, rather than the effect of an amino acid substitution under genetic control. Horse's and donkey's lysozyme differ by the D₈₆→N additional amino acid substitution. In this case, only the tryptic peptide 83–96 of donkey was identified. The MS analysis of intact HPLC-isolated lysozyme confirmed that it corresponded to the variant B (measured 14,630.91 Da, calculated 14,630.94 Da). The presence of an unexpected HPLC peak (Fig. 1B) eluting slightly earlier than lysozyme indicates the possible presence of a yet uncharacterized variant.

Interestingly, DM lysozyme is relatively thermostable [32] and significantly resistant to gastrointestinal digestion [37], so that it may reach the gut partly intact.

Lactotransferrin is an iron (Fe³⁺)-binding multifunctional enzyme that in milk and other mucosal secretions exhibits potent bacteriostatic and bactericidal properties. The average concentration of lactoferrin in DM is not as high as in human milk, but it is much more abundant than in bovine milk [35]. Serotransferrin sourcing from plasma, distinct from lactotransferrin (62.5% sequence homology) and occurring at comparable level based on emPAI values of 1.93 and 1.27, respectively (Table 1), can contribute to the bacteriostatic activity, because of the iron-sequestering capability. Jointly, serotransferrin and lactoferrin make up a relevant antimicrobial system in the whey. The emPAI-based concentration of these enzymes might be underestimated, considered that they are heavily N-glycosylated and the resulting glycopeptides escaped the MS detection.

Just like milk from other species, the soluble fraction of DM contains several immunoglobulin chains, which play a central role in conferring the passive immunity to offspring, as they bind and neutralize antigens. Secretory IgA (sIgA), which is crucial within the “first line” defensive system of biological fluids, was identified with 26 PSMs belonging to the constant segment of the heavy chain and had a relatively high emPAI value (0.65, Table 1). In effects, sIgA is predominant over the IgG immunoglobulins in mature equine milk, similarly to what found in human milk and colostrum [37]. The joining (J) immunoglobulin chain, expressed by mucosal and glandular plasma cells, regulates polymer formation of IgA and IgM. J-chain was identified with 17 PSMs and exhibited an emPAI value higher than sIgA (0.95).

A primary antimicrobial role in DM has been attributed to lactoperoxidase, a 78 kDa glycoprotein secreted in many biological fluids, milk included [2]. Lactoperoxidase catalyses the formation of hypothiocyanite (SCNO⁻) from hydrogen peroxide (H₂O₂) and thiocyanate (SCN⁻), both naturally found in milk. The simultaneous presence of H₂O₂ and SCNO⁻ is essential for the antimicrobial activity. Combined with lactoperoxidase these two anions constitute the ‘lactoperoxidase system’ that is involved in the natural host defence against pathogenic microorganisms [38]. In DM lactoperoxidase is reported to occur at low concentration comparable to that of human milk [39]. Nevertheless, we did not detect any lactoperoxidase-derived peptide in our sample, although several protein entries are catalogued in the UniprotKB and NCBI databases for *Equus* peroxidase isoforms. In contrast, DM milk contained L-amino acid oxidase, which was identified with 11 unique peptides matching the horse homologue gene product retrieved in the insoluble milk fraction, ranking among the relatively abundant milk enzymes according to its emPAI value (1.01 and 0.04 in insoluble and soluble fractions, respectively, Table 1). L-amino acid oxidase is a flavoprotein that catalyses the oxidative deamination of amino acids, converting them into ketoacids, ammonia and H₂O₂. This 59 kDa enzyme has been already identified in mouse milk [40] as a mammary tissue specific gene product and detected in DM as well [22]. Nevertheless, the ProteoMiner equalization strategy failed to appreciate its relatively significant abundance.

L-amino acid oxidase, widely distributed across diverse *phyla* from bacteria, fungi to mammals and also described in venom of several

snake species, plays a primary role in host defence in the mammary glands, as it is active against a large spectrum of pathogenic microorganisms and viruses through the production of H₂O₂ [41].

Thus, L-amino acid oxidase could be one of the most potent antimicrobial enzymes of DM, operating concertedly with other protein components, including lysozyme, which in the presence of Ca²⁺/Mg²⁺ might exert activity even against Gram-negative bacteria [34,42]. Combined with the low activity of lactoperoxidase [39], the concentration of H₂O₂ in DM might be significantly higher than in milk of other species. Further studies would be required to quantify accurately L-amino acid oxidase and to confirm that DM contains levels of H₂O₂ compatible with a relevant antimicrobial activity.

3.5. Glycoconjugates: the case of κ -casein

Milk contains a large variety of oligosaccharides and glycoconjugates (glycolipids, glycoproteins, mucins, and glycosaminoglycans) that protect the newborn against microbes, viruses, and toxins [43]. In fact, milk glycans antagonize competitively the binding of enteropathogens to their host receptors. Alternatively, glycans can block the mannose-binding lectins in bacteria. The specificity of carbohydrate structures appears correlated with the need of protecting the offspring against specific pathogens [44]. Glycosylated proteins constitute more than 70% of the minor DM proteome. Aside from immunoglobulins and lactoferrin that are involved in direct antimicrobial functions, glycoproteins such as lactadherin, butyrophilin, α -1B-glycoprotein, α -2-HS-glycoprotein-like, polymeric immunoglobulin receptor and semaphorin 7A potentiate the milk-mediated passive immunity. Some specific glycoproteins relatively abundant in human milk [45], including tenascin C, bile salt stimulated lipase and macrophage mannose receptor 1, were not detected in DM. The failure in the search of those missing proteins might be ascribed to a defective annotation in the *Equus* databases, rather than indicating their unique presence in human milk.

Conversely, osteopontin (also referred to as secreted phosphoprotein 1), found in human milk as well, was identified in DM with 26 PSMs. Osteopontin is a multifunctional phosphorylated glycoprotein that can act as a regulator of the immune response at multiple levels, modulating cell adhesion and migration, promoting the action of immunocompetent cells and tissue repair [46].

Nearly 10–15% of α -La in the milk of ruminants is N-glycosylated at the level of canonical consensus triplets N-X-S/T (where X is any amino acid except P). The primary structure of donkey's α -La lacks consensus sequences, but, like the human homologue [47], it could be 1% N-glycosylated at the N₇₁ belonging to the N-X-C triplet [48]. However, in absolute terms glycosylated α -La could contribute to the overall compute of milk glycoproteins to a non-negligible extent.

In general, κ -casein is the only glycosylated among the four casein families. A fraction of the bovine and human milk κ -casein carries a variable number of O-linked glycans localized in the casein-macropeptide (CMP) domain. In the past, several authors reported the absence of κ -casein in equine milk (for a review on the topic see reference no. [2]). Iametti et al. (2001) [49] demonstrated the presence of κ -casein in mare's milk, albeit at trace amount.

Our analysis unequivocally demonstrated the occurrence of κ -casein in DM, as it was identified through 37 and 10 PSMs, in pH 4.6-insoluble and whey protein fractions, respectively. Considering the PSMs from the insoluble and soluble protein fractions, the 1–104 region κ -casein, corresponding to the protein entry F0V6V5 of the UniprotKB, was entirely mapped (Supplementary information, Table S1). The mapped region covered the para- κ -casein (sequence 1–96) and it extended beyond the F₉₆-I₉₇, that is the presumed cleavage site of chymosin. No PSMs matched the 105–162 CMP region. However, the 113–162 protein domain does not contain any tryptic cleavage site. In a previous work, DM κ -casein escaped both HPLC and SDS-electrophoretic identification, while it showed a complex pattern of at least nine immunoreactive isoforms spread over the 4.2–6.0 pI range, as evidenced by isoelectric

focusing and immunostaining [12].

Overall, these data suggest that donkey's κ -casein could be fully glycosylated, at difference of the bovine and human homologues. In line with this hypothesis, it should be underlined that donkey's κ -casein contains a higher number of potential O-glycosylation than the bovine counterpart (12 vs 9). Because of the intense glycosylation of κ -casein, DM used for the infant nutrition could contribute to inhibit the adhesion of *Helicobacter pilori* to gastric mucosa, as supposed for human milk [50].

4. Conclusions

Shotgun proteomics enables the large-scale identification of proteins in complex biological systems. Modern MS instruments and bioinformatic tools suited for protein identification have greatly enlarged the inventory of gene products identifiable in biological fluids, milk included. However, the limited dynamic range of the MS technique in relation to the window of protein concentrations rarely allows delineating a global proteomic perspective of complex biological fluids. In general, the “depth” of the qualitative analysis and the accuracy of the quantitative determinations represent a compromise between complexity of the workflows (e.g. multiple pre-fractionation steps, time of analysis, use of internal standards for quantification etc.) and comprehensiveness of the analytical response.

In this work, we obtained one of the most inclusive lists of the DM proteins. It can be envisaged that the description of *Equidae* milk proteins will increase in the next future because of upcoming advances in fractionation methods, proteomic techniques and genomic annotation.

Our analysis disclosed further details of a robust enzymatic apparatus in DM, which is deputed to transfer passive immunity and prevent infections in donkey's offspring. Realistically, like in the other species, the evolutionary pressure has naturally selected genes and, hence, gene products to fulfil specific physiological requirements of the offspring. In this case, the need of donkey foal to receiving milk endowed with such a powerful antibacterial potential still awaits to be elucidated.

DM shares some of the compositional and functional traits with human milk, although many others appear peculiar. Human milk unleashes great part of its antibacterial potential through lactoferrin, while DM could do it, probably even more effectively, through lysozyme and L-amino acid oxidases, along with a series of other protein components. Remarkably, sanitation of DM might be carried out at milder conditions compared to ruminants' milk, so better preserving its nutritional attributes.

Because of its properties, DM could be proposed not only as a hypoallergenic substitute of human milk in the dietary treatment of allergy to cow's milk during early infancy, but also as a safe and nutritionally valid alimentary resource for growing children, convalescent, elderly people and general population.

Acknowledgements

The authors gratefully acknowledge the farm LatteA located in Oliveto Citra (Salerno), Italy, for providing donkey's milk samples analyzed in this study. This research was not supported by any specific fund by public or private institutions.

Data availability statement

Raw data are available upon request to the corresponding author.

Credit Author Statement

G.P., P.F. and L.C. conceptualized the research. V.S., G.P. and E.S. designed and supervised the experiments. V.S. and E.S. elaborated data. G.P. and F.A. wrote the original draft. All authors reviewed and approved the paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jprot.2020.104007>.

References

- [1] I.K. Karabagias, E.Z. Halatsi, Donkey Milk exploitation: health benefits, potential applications and prospective invigoration of the Hellenic economy, *Agri. Res. & Tech.: Open Access J.* 12 (2017) 139–141, <https://doi.org/10.19080/ARTOAJ.2017.12.555863>.
- [2] T. Uniacke-Lowe, T. Huppertz, P.F. Fox, Equine milk proteins: chemistry, structure and nutritional significance, *Int. Dairy J.* 20 (2010) 609–629, <https://doi.org/10.1016/j.idairyj.2010.02.007>.
- [3] G. Monti, E. Bertino, M.C. Muratore, A. Coscia, F. Cresi, L. Silvestro, C. Fabris, D. Fortunato, M.G. Giuffrida, A. Conti, Efficacy of donkey's milk in treating highly problematic cow's milk allergic children: an in vivo and in vitro study, *Pediatr. Allergy Immunol.* 18 (2007) 258–264, <https://doi.org/10.1111/j.1399-3038.2007.00521.x>.
- [4] P. Polidori, S. Vincenzetti, Use of donkey Milk in children with Cow's Milk protein, allergy, *Foods* 2 (2013) 151–159, <https://doi.org/10.3390/foods2020151>.
- [5] F. Lara-Villoslada, M. Olivares, J. Xaus, The balance between caseins and whey proteins in cow's milk determines its allergenicity, *J. Dairy Sci.* 88 (2005) 1654–1660, [https://doi.org/10.3168/jds.S0022-0302\(05\)72837-X](https://doi.org/10.3168/jds.S0022-0302(05)72837-X).
- [6] P. Restani, C. Ballabio, C. Di Lorenzo, Molecular aspects of milk allergens and their role in clinical events, *Anal. Bioanal. Chem.* 395 (2009) 47–56, <https://doi.org/10.1007/s00216-009-2909-3>.
- [7] V. Cunsolo, R. Saletti, V. Muccilli, S. Gallina, A. Di Francesco, S. Foti, Proteins and bioactive peptides from donkey milk: the molecular basis for its reduced allergenic properties, *Food Res. Int.* 99 (2017) 41–57, <https://doi.org/10.1016/j.foodres.2017.07.002>.
- [8] H.Y. Guo, K. Pang, X.Y. Zhang, L. Zhao, S.W. Chen, M.L. Dong, F.Z. Ren, Composition, physicochemical properties, nitrogen fraction distribution, and amino acid profile of donkey milk, *J. Dairy Sci.* 90 (2007) 1635–1643, <https://doi.org/10.3168/jds.2006-600>.
- [9] S. Vincenzetti, P. Polidori, P. Mariani, N. Cammertoni, F. Fantuz, A. Vita, Donkey's milk protein fractions characterization, *Food Chem.* 106 (2008) 640–649, <https://doi.org/10.1016/j.foodchem.2007.06.026>.
- [10] A. Criscione, V. Cunsolo, S. Bordonaro, A.M. Guastella, R. Saletti, A. Zuccaro, D. Marletta, Donkeys' milk protein fraction investigated by electrophoretic methods and mass spectrometric analysis, *Int. Dairy J.* 19 (2009) 190–197, <https://doi.org/10.1016/j.idairyj.2008.10.015>.
- [11] V. Cunsolo, E. Cairone, D. Fontanini, A. Criscione, V. Muccilli, R. Saletti, S. Foti, Sequence determination of alphas1-casein isoforms from donkey by mass spectrometric methods, *J. Mass Spectrom.* 44 (2009) 1742–1753, <https://doi.org/10.1002/jms.1683>.
- [12] L. Chianese, M.G. Calabrese, P. Ferranti, R. Mauriello, G. Garro, C. De Simone, M. Quarto, F. Addeo, G. Cosenza, L. Ramunno, Proteomic characterization of donkey milk caseome, *J. Chromatogr. A* 1217 (2010) 4834–4840, <https://doi.org/10.1016/j.chroma.2010.05.017>.
- [13] E. Malissiova, G. Arsenos, P. Papademas, D. Fletouris, A. Manouras, M.A. Nikolopoulou, A. Giannopoulou, I.S. Arvanitoyannis, Assessment of donkey milk chemical, microbiological and sensory attributes in Greece and Cyprus, *Int. J. Dairy Technol.* 69 (2016) 1–8, <https://doi.org/10.1111/1471-0307.12245>.
- [14] D. Brumini, A. Criscione, S. Bordonaro, G.E. Vegarud, D. Marletta, Whey proteins and their antimicrobial properties in donkey milk: a brief review, *Dairy Sci. Technol.* 96 (2016) 1–14, <https://doi.org/10.1007/s13594-015-0246-1>.
- [15] M. Koutb, M. Khider, E.H. Ali, N.A. Hussein, Antimicrobial activity of donkey milk against dermatomycotic fungi and foodborne bacteria, *Int. J. Biomed. Mat. Res.* 4 (2016) 11–17, <https://doi.org/10.11648/j.ijbmr.20160403.11>.
- [16] M. Martini, F. Salari, R. Licita, C. La Motta, I. Altomonte, Lysozyme activity in donkey milk, *Int. Dairy J.* 96 (2019) 10–98, <https://doi.org/10.1016/j.idairyj.2019.04.009>.
- [17] F. Nazzaro, P. Orlando, F. Fratianni, R. Coppola, Isolation of components with antimicrobial property from the donkey milk: a preliminary study, *Open Food Sci. J.* 4 (2010) 43–47, <https://doi.org/10.2174/1874256401004010043>.
- [18] F. Tidona, C. Sekse, A. Criscione, M. Jacobsen, S. Bordonaro, D. Marletta, G.E. Vegarud, Antimicrobial effect of donkeys' milk digested in vitro with human gastrointestinal enzymes, *Int. Dairy J.* 21 (2011) 158–165, <https://doi.org/10.1016/j.idairyj.2010.10.008>.
- [19] L.C. Šarić, B.M. Šarić, S.T. Kravić, D.V. Plavšić, I.L. Milovanović, J.M. Gubić, N.M. Nedeljković, Antibacterial activity of domestic Balkan donkey milk toward *Listeria monocytogenes* and *Staphylococcus aureus*, *Food Feed Res.* 4 (2014) 47–54, <https://doi.org/10.5937/FFR1401047S>.
- [20] X. Zhang, L. Zhao, L. Jiang, M. Dong, F. Ren, The antimicrobial activity of donkey milk and its microflora changes during storage, *Food Control* 19 (2008) 1191–1195, <https://doi.org/10.1016/j.foodcont.2008.01.005>.
- [21] A. Rogowska-Wrzesinska, M.C. Le Bihan, M. Thaysen-Andersen, P. Roepstorff, 2D gels still have a niche in proteomics, *J. Proteomic.* 88 (2013) 4–13, <https://doi.org/10.1016/j.jprot.2013.01.010>.
- [22] V. Cunsolo, V. Muccilli, E. Fasoli, R. Saletti, P.G. Righetti, S. Foti, Poppea's bath liquor: the secret proteome of she-donkey's milk, *J. Proteome* 74 (2011) 2083–2099, <https://doi.org/10.1016/j.jprot.2011.05.036>.
- [23] W. Li, M. Li, X. Cao, H. Han, F. Kong, X. Yue, Comparative analysis of whey proteins in donkey colostrum and mature milk using quantitative proteomics, *Food Res. Int.* 127 (2020) 108741, <https://doi.org/10.1016/j.foodres.2019.108741>.
- [24] W. Li, M. Li, X. Cao, M. Yang, H. Han, F. Kong, X. Yue, Quantitative proteomic analysis of milk fat globule membrane (MFGM) proteins from donkey colostrum

- and mature milk, *Food Funct.* 10 (2019) 4256–4268, <https://doi.org/10.1039/c9fo00386j>.
- [25] M. Bevan, I. Bancroft, E. Bent, K. Love, H. Goodman, C. Dean, et al., Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*, *Nature* 391 (1998) 485–488, <https://doi.org/10.1038/35140>.
- [26] E. Salimei, F. Fantuz, Horse and donkey milk, in: Y.W. Park, G.F.W. Haenlein (Eds.), *Milk and Dairy Products in Human Nutrition: Production, Composition and Health*, John Wiley & Sons, Ltd. Chapter 27, 2013, pp. 594–613, <https://doi.org/10.1002/9781118534168.ch27>.
- [27] L. Chianese, C. De Simone, P. Ferranti, R. Mauriello, A. Costanzo, M. Quarto, G. Garro, G. Picariello, G. Mamone, L. Ramunno, Occurrence of qualitative and quantitative polymorphism at donkey beta-lactoglobulin II locus, *Food Res. Int.* 54 (2013) 1273–1279, <https://doi.org/10.1016/j.foodres.2012.11.005>.
- [28] G. Picariello, P. Ferranti, G. Mamone, I. Klouckova, Y. Mechref, M.V. Novotny, F. Addeo, Gel-free shotgun proteomic analysis of human milk, *J. Chromatogr. A* 1227 (2012) 219–233, <https://doi.org/10.1016/j.chroma.2012.11.005>.
- [29] P. Ferranti, M.V. Traisci, G. Picariello, A. Nasi, V. Boschi, M. Siervo, C. Falconi, L. Chianese, F. Addeo, Casein proteolysis in human milk: tracing the pattern of casein breakdown and the formation of potential bioactive peptides, *J. Dairy Res.* 71 (2004) 74–87, <https://doi.org/10.1017/s0022029903006599>.
- [30] R. Saletti, V. Muccilli, V. Cunsolo, D. Fontanini, A. Capocchi, S. Foti, MS-based characterization of α_2 -casein isoforms in donkey's milk, *J. Mass Spectrom.* 47 (2012) 1150–1159, <https://doi.org/10.1002/jms.3031>.
- [31] Y. Ishihama, Y. Oda, T. Tabata, T. Sato, T. Nagasu, J. Rappsilber, M. Mann, Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein, *Mol. Cell. Proteomics* 4 (2005) 1265–1272, <https://doi.org/10.1074/mcp.M500061-MCP200>.
- [32] M.J. van Herwijnen, M.I. Zonneveld, S. Goerdayal, E.N.M. Nolte-t Hoen, J. Garssen, B. Stahl, A.F. Maarten Altelaar, F.A. Redegeld, M.H.M. Wauben, Comprehensive proteomic analysis of human Milk-derived extracellular vesicles unveils a novel functional proteome distinct from other Milk components, *Mol. Cell. Proteomics* 15 (2016) 3412–3423, <https://doi.org/10.1074/mcp.m116.060426>.
- [33] R. Coppola, E. Salimei, M. Succi, E. Sorrentino, M. Nanni, P. Ranieri, Behaviour of *Lactobacillus rhamnosus* strains in ass's milk, *Ann. Microbiol.* 52 (2002) 55–60.
- [34] N. Benkerroum, Antimicrobial activity of lysozyme with special relevance to milk, *Afr. J. Biotechnol.* 7 (2008) 4856–4867, <https://doi.org/10.5897/AJB08.072>.
- [35] M. Aspri, N. Economou, P. Papademas, Donkey milk: an overview on functionality, technology, and future prospects, *Food Rev. Int.* 33 (2017) 316–333, <https://doi.org/10.1080/87559129.2016.1175014>.
- [36] M. Herrouin, D. Mollé, J. Fauquant, F. Ballestra, J.L. Maubois, J. Léonil, New genetic variants identified in Donkey's milk whey proteins, *J. Protein Chem.* 19 (2000) 105–115, <https://doi.org/10.1023/a:1007078415595>.
- [37] F. Tidona, A. Criscione, T.G. Devold, S. Bordonaro, D. Marletta, G.E. Vegarud, Protein composition and micelle size of donkey milk with different protein patterns: effects on digestibility, *Int. Dairy J.* 35 (2014) 57–62, <https://doi.org/10.1016/j.idairyj.2013.10.018>.
- [38] Z. Koksai, I. Gulcin, H. Ozdemir, An important Milk enzyme: Lactoperoxidase. In: *Milk proteins - from structure to biological properties and health aspects*. Chapter 7, (2016), InTech. doi:<https://doi.org/10.5772/64416>.
- [39] S. Vincenzetti, A. Amici, S. Pucciarelli, A. Vita, D. Micozzi, F.M. Carpi, V. Polzonetti, P. Natalini, P. Polidori, A proteomic study on donkey milk, *Biochem. Anal. Biochem.* 1 (2012) 109, <https://doi.org/10.4172/2161-1009.1000109>.
- [40] Y. Sun, E. Nonobe, Y. Kobayashi, T. Kuraishi, F. Aoki, K. Yamamoto, S. Sakai, Characterization and expression of L-amino acid oxidase of mouse milk, *J. Biol. Chem.* 277 (2002) 19080–19086, <https://doi.org/10.1074/jbc.M200936200>.
- [41] K. Nagaoka, F. Aoki, M. Hayashi, Y. Muroi, T. Sakurai, K. Itoh, M. Ikawa, M. Okabe, K. Imakawa, S. Sakai, L-amino acid oxidase plays a crucial role in host defense in the mammary glands, *FASEB J.* 23 (2009) 2514–2520, <https://doi.org/10.1096/fj.08-126466>.
- [42] S.E. Permyakov, T.I. Khokhlova, V.N. Uversky, E.A. Permyakov, Analysis of Ca^{2+} / Mg^{2+} selectivity in α -lactalbumin and Ca^{2+} -binding lysozyme reveals a distinct Mg^{2+} -specific site in lysozyme, *Proteins* 78 (2010) 2609–2624, <https://doi.org/10.1002/prot.22776>.
- [43] R. Peterson, W.Y. Cheah, J. Grinyer, N. Packer, Glycoconjugates in human milk: protecting infants from disease, *Glycobiology* 23 (2013) 1425–1438, <https://doi.org/10.1093/glycob/cwt072>.
- [44] G. Georgi, N. Bartke, F. Wiens, B. Stahl, Functional glycans and glycoconjugates in human milk, *Am. J. Clin. Nutr.* 98 (2013) 578S–585S, <https://doi.org/10.3945/ajcn.112.039065>.
- [45] G. Picariello, P. Ferranti, G. Mamone, P. Roepstorff, F. Addeo, Identification of N-linked glycoproteins in human milk by hydrophilic interaction liquid chromatography and mass spectrometry, *Proteomics* 8 (2008) 3833–3847, <https://doi.org/10.1002/pmic.200701057>.
- [46] N. Clemente, D. Raineri, G. Cappellano, E. Boggio, F. Favero, M.F. Soluri, C. Dianzani, C. Comi, U. Dianzani, A. Chiocchetti, Osteopontin bridging innate and adaptive immunity in autoimmune diseases, *J. Immunol. Res.* 13 (2016) 1–15, <https://doi.org/10.1155/2016/7675437>.
- [47] M.G. Giuffrida, M. Cavaletto, C. Giunta, B. Neuteboom, A. Cantisani, L. Napolitano, V. Calderone, J. Godovac-Zimmermann, A. Conti, The unusual amino acid triplet Asn-Ile-Cys is a glycosylation consensus site in human α -lactalbumin, *J. Protein Chem.* 16 (1997) 747–753, <https://doi.org/10.1023/a:1026359715821>.
- [48] J.M. Girardet, M.A. N'negue, A.S. Egitto, S. Campagna, A. Lagrange, J.L. Gaillard, Multiple forms of equine α -lactalbumin: evidence for N-glycosylated and deamidated forms, *Int. Dairy J.* 14 (2004) 207–217, <https://doi.org/10.1016/j.idairyj.2003.08.001>.
- [49] S. Iametti, G. Tedeschi, E. Oungre, F. Bonomi, Primary structure of κ -casein isolated from mares' milk, *J. Dairy Res.* 68 (2001) 53–61, <https://doi.org/10.1017/s0022029900004544>.
- [50] M. Stromqvist, P. Falk, S. Bergstrom, L. Hansson, B. Lonnnerdal, S. Normark, O. Hernell, Human-milk κ -casein and inhibition of helicobacter pylori adhesion to human gastric mucosa, *J. Pediatr. Gastroenterol. Nutr.* 21 (1995) 288–296, <https://doi.org/10.1097/00005176-199510000-00006>.