

Research Article

A Preliminary Study of Proteomic Analysis on Caseins and Whey Proteins in Donkey Milk from Xinjiang and Shandong of China

Wahafu Luoyizha^{1,2,3}, Bo Zeng⁴, Hui Li^{1,*}, Xiaojun Liao²

¹Institute of Quality Standard and Testing Technology for Agro-Products, Chinese Academy of Agricultural Sciences, Beijing 100081, PR China

²College of Food Science and Nutrition Engineering, China Agricultural University, Beijing 100083, PR China

³College of Life Science and Technology, Xinjiang University, Xinjiang 830046, PR China

⁴Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, PR China

ARTICLE INFO

Article History

Received 27 January 2021

Accepted 22 February 2021

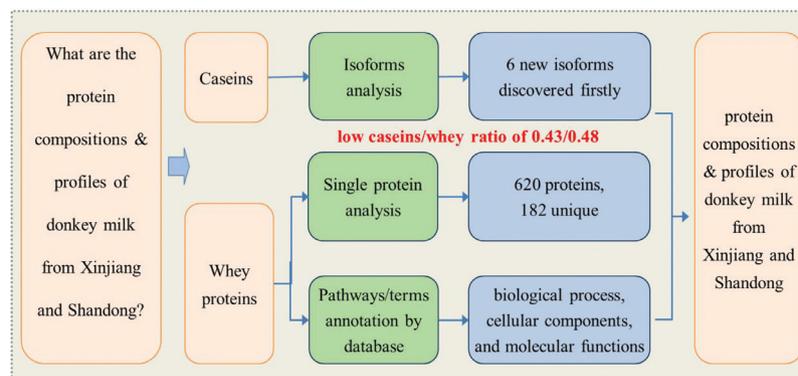
Keywords

Donkey milk
caseins
whey proteins
proteomics
gene ontology (GO)
Kyoto Encyclopedia of Genes and Genomes (KEGG)

ABSTRACT

The proteomics of donkey milk from China's major producing areas have not been well documented. To explore the protein profiles of donkey milk from China, this study preliminarily investigated caseins and whey proteins in products from two major producing areas using quantitative proteomics approach. Total 15 caseins genetic variants and 620 whey proteins were identified and relatively quantified. Functional categorization of whey proteins was clustered into 46 terms based on Gene ontology (GO) and 42 Brite B classes based on Kyoto Encyclopedia of Genes and Genomes (KEGG). Six of the seven identified isoforms of caseins were firstly observed in donkey milk. Whey proteins including α -lactalbumin, β -lactoglobulins and lysozyme were identified, but no immunoglobulins or lactoferrin was determined. An unprecedented low ratio of casein/whey indicated that the products were very good sources with high biological value. This study enhanced knowledge of protein composition and established proteins profiles of differentially expressed proteins in the selected products of donkey milk from China.

GRAPHICAL ABSTRACT



© 2021 The Authors. Publishing services by Atlantis Press International B.V.

This is an open access article distributed under the CC BY-NC 4.0 license (<http://creativecommons.org/licenses/by-nc/4.0/>).

1. INTRODUCTION

Donkey milk is among the most widely used alternatives with favorable nutrition and excellent palatability to reduce the protein allergy for children [1–3]. The relatively low amount of the total protein content in donkey milk inhibits it to produce an excessive renal load of solute [1,4,5]. Moreover, donkey milk has the favorable characteristics of low allergenicity of high whey proteins and

low caseins content, the average ratio of which is higher than the value of human milk but much lower than the value of cow, goat and sheep milk [2–4,6].

Casein fraction in donkey milk is represented mainly by α_{s1} - and β -caseins, which showed a great variability due to the degree of phosphorylation and the presence of genetic variants, with a less amount of α_{s2} - and κ -casein, which is far lower than that of bovine milk [2,3,6,7]. At least four α_{s1} -casein main variants, six β -casein variants, three α_{s2} -casein variants, and 11 κ -casein variants have been reported to present in donkey milk [7,8]. Similar to human milk, the relative minor expression of α_{s1} -Casein (CN) together

*Corresponding author. Email: lihui05@caas.cn

Peer review under responsibility of the International Association of Dietetic Nutrition and Safety

with the absence of α_2 -CN in some donkey milk represents a scientific basis for the use of donkey milk especially for infants with allergies to cow's milk proteins [9].

The higher ratio of whey proteins in donkey milk with 35–50% of the total nitrogen content, compared with the only 20% in bovine milk, makes it more easily absorbed by human body [3,10]. Major components of whey proteins present in donkey milk include β -lactoglobulin, α -lactalbumin, lactoferrin, serum albumin, lysozyme, and immunoglobulins, and other low-abundance proteins such as enzymes and metal-binding proteins [6,11]. As one of the most important milk allergens, β -lactoglobulin accounts for less than 30% of the total whey proteins, which has been suggested to be lower than that of bovine milk (>50%) but higher than that of human milk (negligible) [3,12,13]. The second major whey protein α -lactalbumin is a small molecule protein that binds to calcium ions, with a very close amount in donkey milk to that of human milk, and functions for anti-inflammatory, anti-nociceptive, anti-virus, anti-tumor and anti-emergency properties [1,7,14,15]. Moreover, a high content of lysozyme is present in donkey milk (0.67–4.00 g/kg), which is much higher than that of human (0.04–0.5 g/kg) and bovine milk (negligible) [16,17]. Together with immunoglobulins, lactoferrin and lactoperoxidase, lysozyme functions as one of the most important antimicrobials to reduce the harmful bacteria in donkey milk, and to reduce the incidence of infant's gastrointestinal infections [18].

Despite a large amount of information has been related to the characterization of proteins in donkey milk, proteomics of donkey milk from China's major producing areas have not been well documented. Protein compositions including caseins and whey proteins in donkey milk may differentially expressed among products from different areas or countries, or between samples from different areas. With the development of proteomic technologies, quantitation approach by mass spectrometry based on Tandem Mass Tag (TMT) enables multiplex identification and relative quantification of different fractions of milk proteome, especially low-abundant proteins [19]. In the year of 2019, about 16.5% of the total stock of donkeys (*Equus asinus*) in China were distributed in provinces of Xinjiang and Shandong, with Yopurga and Dongge County as the most famous and major donkey-breeding bases, respectively (data.stats.gov.cn). Therefore, in this study, caseins and whey proteins in donkey milk from Yopurga County in Western China and Dongge County in Eastern China were preliminarily determined using proteomic analysis. The results enhanced knowledge of protein composition in donkey milk from different areas of China, and established profiles of differentially expressed proteins between the different products.

2. MATERIALS AND METHODS

2.1. Preparation of Milk Samples

Samples of donkey milk were collected from Xinjiang Yukunlun Natural Food Engineering Co. Ltd. in Yopurga County of Xinjiang Province (XJ) in Western China, and from Shandong Dongge ejiao Company Limited in Dongge County of Shandong Province (SD) in Eastern China, respectively. The breed of XJ donkey was Jianguye donkey, a hybrid breed by crossing local donkeys with the introduced Guanzhong donkey from Shanxi province, while SD donkey was Dezhou donkey, a breed produced in the counties along the

Bohai Sea in the northern Shandong plain. Samples were collected one batch per day from a collected mixture of donkey milk from 20 donkeys, and within five consecutive days, a total 10 batches of donkey milk were collected from the two plants. All samples were collected from bulk milk containers into sterile tubes, transported to the laboratory immediately after sampling, and then stored at -80°C until further analysis.

2.2. Protein Extraction, and Digestion TMT Labeling

Skimmed milk was prepared and then centrifuged at 10,000g for 30 min to obtain a supernatant of whey proteins, which was consequently determined by the bicinchoninic acid method and sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) electrophoresis [20,21]. Protein lysate (8M urea, 1% SDS, containing protease inhibitor) was used for sample lysis [22].

Protein of 200 μg was supplemented to the volume of 180 μl with lysate. The samples were mixed with Tris(2-carboxyethyl)phosphine (final concentration was 10 mM) and were then incubated at 37°C for 60 min, followed by a reaction with appropriate iodoacetamide (final concentration was 40 mM) at room temperature for 40 min in the dark. Each tube of sample was added by precooled acetone, and incubated at -20°C for 4 h before centrifugation at 10,000g for 20 min. Trypsin digestion was performed at a substrate ratio of 1:50 (w/w), followed by an overnight incubation of enzymatic hydrolysis at 37°C [22,23].

The digested peptides were labeled with a TMT Reagent Kit (Thermo Fisher Scientific, Wilmington, USA, No. 90111) according to the manufacturer's protocol. Samples after tagging were prepared for further analysis.

2.3. One-dimensional Separation of UPLC with High pH

Sample were fractionated using high pH reverse phase separation. The TMT-tagged peptides were re-suspended with ultra-performance liquid chromatography (UPLC) loading buffer, and separated by an ACQUITY UPLC BEH C18 Column (1.7 μm , $2.1 \times 150 \text{ mm}^2$, Waters Corporation, Milford, USA) in high PH liquid phase. Separation was performed at 200 $\mu\text{L}/\text{min}$ with the gradient increased for 76 min (Phase A: 2% acetonitrile, pH 10; Phase B: 80% acetonitrile, pH 10), and the ultraviolet detection wavelength was 214 nm.

2.4. LC-MS/MS Analysis

Experiments were performed on a Q-Exactive mass spectrometer (Thermo Fisher Scientific) that was coupled with Easy-nLC 1200 (Waters Corporation), and Thermo Xcalibur 4.0 (Thermo Fisher Scientific) was used as the data acquisition software. A 4- μL of each fraction was injected for nano liquid chromatography with tandem mass spectrometry (LC-MS/MS) analysis. The peptide mixture (2 μg) was loaded onto a C18-reversed phase column (75 $\mu\text{m} \times 25 \text{ cm}$, Thermo Fisher Scientific) in buffer A (2% acetonitrile and 0.1% Formic acid) and separated with a linear gradient

of buffer B (80% acetonitrile and 0.1% Formic acid) at a flow rate of 300 nL/min for 66 min. Q-Exactive mass spectrometer was operated in the data-dependent mode to switch automatically between MS and MS/MS acquisition. Survey full-scan MS spectra (350–1300 m/z) were acquired with a mass resolution of 70 K, followed by 20 sequential high energy collisional dissociation MS/MS scans with a resolution of 35 K. In all cases, one microscan was recorded using dynamic exclusion of 18 s.

2.5. Sequence Database Searching for Protein Annotation

The software version for library search is Proteome Discoverer Software 2.1. Acquired MS/MS spectra were analyzed for protein identification and relative quantification using Proteome Discoverer (Thermo Fisher Scientific, version 2.1) against the protein database from NCBI (version 29/08/2018; donkey), with a precursor mass tolerance of 20 ppm and fragment mass tolerance of 0.02 Da. The highest score for a given peptide mass (best match to that predicted in the database) was used to identify parent proteins. Trypsin digestion was specified with up to two missed cleavages. Peptide spectral matches were validated based on q -values at a 1% false discovery rate.

2.6. Statistics

The t -test function in R language was used to calculate the p -value of the significant difference between samples. In this project, the screening standard of the significant differentially expressed proteins was

$p < 0.05$ & (fold change < 0.83 or $FC > 1.20$). Principal Component Analysis (PCA) was carried out against proteomic compositions to observe differences between XJ and SD. Gene Ontology (GO) analysis of identified proteins was carried out by BLAST2GO software (version 2.5.0). Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis was conducted by KOBAS (version 2.1.1) according to the KEGG database (<http://www.genome.jp/kegg/pathway.html>).

3. RESULTS

Overall, the amount of total protein, casein and whey protein in the XJ samples were 2.84%, 0.94%, and 1.18%, respectively, and the amount of them in the SD samples were 2.51%, 0.71%, and 1.00%, respectively.

3.1. Casein Proteins Profiles in Donkey Milk

Total 15 casein proteins were identified from the two sample groups (Table 1). Among them, three proteins including κ -casein, α_2 -casein and casein kinase were overexpressed in the SD samples (3.2, 1.4, and 1.3 times of that in the XJ samples, respectively), while only one protein predicted κ -casein was found to be overexpressing in the XJ samples. Proteins including α_1/α_2 -casein and precursors, casein kinases and substrates, β -casein, and caseinolytic peptidase were expressed in both sample groups (Table 1). Our study revealed a characteristic caseins pattern in the two donkey samples consisting of: (i) three α_1 -CN components; (ii) three α_2 -CN components; (iii) one β -CN components; (iv) two κ -CN components; and (v) four CN kinase and one caseinolytic peptidase (Table 1).

Table 1 Expression of casein proteins in donkey milk from XJ and SD

Description	No. of AA	FC (XJ/SD)	p -value	Amino acid sequence
α_1 -CN ^b	202	1.01	0.832	RPKLPHRHPEIIQNEQDSREKVLKERKFPSPFALHTPREEYINELNRQRELLKEKQK DEHKEYLIEDPEQQESSSTSSEEVVPIINTEQKRIPREDMLYQHTLEQLRRLSKYN QLQLQAIYAQEQLIRMKENSQRKPMRVVNQEAYFYLEPFQPSYQLDVYPYAAWFH PAQIMQHVAYSPPFHDTAKLIASENSEKTDIPEW
α_1 -CN precursor ^a	212	1.14	0.065	MKLLILTCLVAVALARPKLPHRHPEIIQNEQDSREKVLKERKFPSPFALEYINELN RQRELLKEKQKDEHKEYLIEDPEQQESSSTSSEEVVPIINTEQKRIPREDMLYQHTLE ALRRLSKYNQLQLQAIYAQEQLIRMKENSQRKPMRVVNQEAYFYLEPFQPSYQLD VYPYAAWFHPAQIMQHVAYSPPFHDTAKLIASENSEKTDIPEW
Predicted: α_1 -CN isoform X2 ^b	208	1.09	0.154	MKLLILTCLVAVALARPKLPHRHPEIIQNEQDSREKVLKERKFPSPFALEYINELN RQRELLKEKQKDEHKDTSNESTEEYLIEDPEQQESSSTSSEEVVPIINTEKRI QHTLEQLRRLSKYNQLQLQAIYAQRKPMRVVNQEAYFYLEPFQPSYQLDVYPYAAW FHPAQIMQHVAYSPPFHDTAKLIASENSEKTDIPEW
α_2 -CN ^b	236	0.96	0.37	MKFFIFTCLLAVALAKHNMEHRSSSEDSVNISQEKQEKYVVIPTSKESICSTSCEE ATRNINEMESAKFPTEVYSSSSSEESAKFPTEREEKEVEEKHHLKQLNKNQFYEKL NFLQYLQALRQPRIVLTPWDQTKTGASPFPIVNTQQLFTSEEIPKKTVDMESTE VTEKTELTEEEKNYLKLKLYEKFVTLQYFKIVH QHQTMDPQSHSKTNSYQIIPVLR RYF
Predicted: α_2 -CN isoform X2 ^b	231	0.69	0.001**	MKFFIFTCLLAVALAKHNMEHRSSSEDSVNISQEKQEKYVVIPTSKESICSTSCEE ATRNINEMESAKFPTEVYSSSSSEESAKFPTEREEKEVEEKHHLKQLNKNQFYEKL NFLQYLQALRQPRIVLTPWDQTKTGASPFPIVNTQQLFTSEEIPKKTVDMESTE VTEKTELTEEEKNYLKLKLYEKFVTLQYFKIVH QHQTMDPQSHSKTNSYQIIPVLR RYF
α_2 -CN B precursor ^b	160	1.21	0.186	MKFFIITCLLAVALAKHEIKHVSSEESTNISQEKYKQDNNVAFQTSQESSGSSSEETD SLTDEKEHHSSEFTSISQEKTSKKTVDMGSTEIFPEEIIELSDDEEKNYLKQLKQLV KINPKFPSPQYFQAVHPQIPMPSPWNRSKENTYPIITLRFN
Predicted: β -CN ^b	233	1.08	0.048*	MKILILACLVALALAREKEELNVSSSETVESLSSNEPDSSEKQVQKFKHEGQQQREVE HQDKISRFVQPQPVVYPYAEVPPYAVVPQNILPLAQPPVIFLQPEIMEVVSQAKETLL PKRKVMPFLKSPVIFSERQILNPTNGENLRPLVHLLIQPFMHQVPSQLLQTLMLP SQPVLSPPQSKVAPFPQPVVYPQRDTPVQAFLLYQDPQLGLTGEFDPATQPIVPHNPVIV

(Continued)

Table 1 | Expression of casein proteins in donkey milk from XJ and SD—Continued

Description	No. of AA	FC (XJ/SD)	p-value	Amino acid sequence
κ -CN ^a	182	0.31	0.001**	MKSFFLVVNILALTLPLFLGAEVQNQEPTCRKNDEFFDLKTVKYIPIYYVLNSSPR NEPIYYQHRLAVLINNQHMPYQYYARPAAVRPHVQIPQWQVLPNIYPTVVRH PRPHPSFIAIPPKKLQEKTVIPKINTIATVEPTPIPTPEPTVNNAVIPDASSEFIIAST PETTTVPVTPSPV
Predicted: κ -CN ^b	182	1.34	0.033*	MKSFFLVVNILALTLPLFLGAEVQNQEPTCRKNDEFFDLKTVKYIPIYYVLNSSPR NEPIYYQHRLAVLINNQHMPYQYYARPAAVRPHVQIPQWQVLPNIYPTVVRH PRPHPSFIAIPPKKLQEKTVIPKINTIATVEPTPIPTPEPTVNNAVIPDASSEFIIAST PETTTVPVTPSPV
Predicted: protein kinase C and CN kinase substrate in neurons protein 2 isoform X1 ^b	486	0.6	0.102	MSVTYDDSVGVEVSSDSFWEVGNKRTVKRIDDGHRCTDLMNCLHERARIEKAYA QQLTEWARRWRQLVEKGPQYGTVEKAWIWMSEAEVSELHLEVRASLMNKDFEK IKNWQKEAFHKQMMGSFKETKEAEDGFRKAQKPWAKKLEVEAAKKAYHAACK EEKLAISRETNKGDPSLNPEQLKQLQDKVEKCKQDVLKTKKEYEKSLKELDQGTPOY MENMEQMFEQCQFEEKRLRFREVLLEVQKHLDSLNVSSYKTIYQDLEQSIRAA DAVEDLRWFRVNHGPGMSMNWPQFEESADLNRTLRSREKKKNADGVTTLTGINQT GDQALQNKPSDSLSPNSPARSAQLLPSYNPFDEDDTGSTVSEKEDIKAKNVSSYEK TQSYPADWSDDESNNPFSSTDPSPGDSNPFDEDATSGTEVVRALYDYEGQEHEDELSFK AGDELTKIEDEDEQGWCKGRDLNGTVGLYPANYVEAIQ
Predicted: CN kinase I isoform α isoform X1 ^b	380	0.76	0.001**	MASSSGSKAEFIVGGKYKLVKRKIGSGSFGDIYLAINITNGEEVAVKLESQKARHPQLLY ESKLYKILQGGVGIPHIRWYQGEKDYNVLMDDLGPSLEDFNFCRRFTMKTVLM LADQMISRRIEYVHTKNFIHRDIKPDNFLMGIGRHCNKLESFVGGKRRKRSMTVSTSQD PSFSGNLQFLIDFLAKKYRDNRRTRQHIPPYREDKNLTGTARYASINAHLGIEQSRDD MESLGYVLMYFNRTSLPWQGLKAATKKQKYEKISEKMMSTPVEVLCKGFPAEFAM YLYCRGLRFEEAPDYMYLRQLFRILFRTLNHQYDYTFDWTMLKQKAAQQAASS GGGQQAQTPGTGKQTDKTKSNMKGTVLVTQQRNPNHRGRP
Predicted: CN kinase II subunit α ^b	213	0.89	0.074	MYELLKALDYCHSKGIMHRDVKPHNV MIDHQKQLRLIDWGLAEFYHPAQEYN VRVASRYFKGPELLVDYQMYDYSLDMWSLGCMLASMIFRKEPFFHGQDNYD QLVRIAKVLGTDELYGYLKKYHIDLDPHFNDILGQHSRKRWFENFIHSENRLHV SPEALDLLDKLLRYDHDQRLTAKEAMEHPYFYPVVKESQSTCTDNAVLSGLTAAR
Predicted: protein kinase C and CN kinase substrate in neurons protein 3 ^b	424	1.08	0.853	MAPEEDAGGEALEGSFWEAGNYRRTVQRVEDGHRCLGDLVSCFQERARIEKAYA QQLADWARKWRGAVEKGPQYGTLEKAWHAFFTAERLSVLHLEVREKLQGD SERVRAWQRGAFHRPVLGGFRESRAEDGFRKAQKPWLKRLKEVEASKKSYHAAR KEEKTAQTRESHAKADS AVSQEQLRKLQERVERCTKEAEKVKTQYEQTLAELHRYT PRYMEDMEQAFESCQAERQRLFFKDMLLTFHQHLDLSSEKQELHRDLH QSIEAASDEEDLRWWRSTHGPGMAMNWPQFEESLDTQRTISRKEKGGRS DEVTLTSIVATRDGAVPPSQSLGSPGGGQEEEWSDEESSRKAATGVRVRYALDYAGQ EAELEFRAGEELLKMSSEDEQWCGQLQSGRIGLYPANYVECVGA
Predicted: CN kinase I isoform γ -1 isoform X1 ^b	459	1.54	0.056	MDHPSREKDERQRTTKPMAQRSAHCSRPSGSSTASGVLVMPNFRVGGKIGCGNF GELRLGKNLYTNEYVAIKLEAIKSRAPQLHLEYRFYKQLGSTGEGLPQVYFPGC GKYNAMVLELLGPSLEDFDLCDRTFTLKTVMIAIQLLSRMEYVHSKNLIYRDVK PENFLIGRQGNKKEHVIHIDFGLAKEYIDPETKKHIPYREHKSLGTARYMSINTHL GKEQSRDDLEALGHMFMYFLRGLSPWQGLKADTLKERYKIGDTRKNTPEAL CENFPEEMATYLRVRRLDFFKPDYELRALTDLFERKGYTFDYAYDWWGRPIPT PVGSVHVDSGASAITRESHTHDRPSQQQLRNQTASSERRGEWEIQPSRQTNT SYLTSHLAADRHGGSVQVVSSTNGELNVDDPTGAHSNAPITAHAEVEVVEEAKC CCFFKRKRKKTAKRHK
Predicted: caseinolytic peptidase B protein homolog isoform X1 ^b	706	0.65	0.136	MLGSTVLRRTALAPRLQLLRSSSLRSHGGASGLSVTTGGRGEPQWLRAAVGGRPGTS PALLTRGGAATGGRQGGRTETQCLAAATWGRLPIPEETLPGQDSWNGVPSRAGL GLWALATALVVHCYSKSPSNKDAALMEARANNVQEVRRLLSEGADVNRHRL GWTALMVA AISRNDSSVQVLLAAGADPNLGDFFSSVYKTAKEQGIHSLDGD GQDDASWHITNQWTSALFERRWLGVPVTLITREDDFNRLNRRASFQKGTAL HYAVLADDYRTVKELLDGGANPLQRNEMGH TPLDYAREGEMKLLRTSEAKY QEKQRKREAEERRRPLEQRLKEHIIGQESAIATVGA AIRKENGWYDEEHPLVFL FLGSSGIGKTELAKQTAKYMHKDAKKG FIRLDMSEFQERHEVAKFIGSPPGYIGHEE GGQLTKLQKCPNAVVLDFEVDKAHPDVLTIMLQLFDEGRITDGGKGTIDCK DAIFIMTSNVASDEIAQHALQLRQEALMSHNRIANLGDVQINDKITISKNFEN VIRPILKAHFRRDEFLGRINEIVYFLPFCHSELIQLVKNELNFWAKRAKQRHNITLLW DREVADVLDGYNVHYGARSIKHEVERRVNNQLAAAYEQDLLPGGCTLRITVEDSD KQLLKSPELSPKAEKRPPKLRLEIIDKSKTRKLDIRAPLHPEKVCHTL

^aProteins from *Equus asinus africanus*. ^bProteins from *Equus asinus*. * and ** represent significant difference at $p < 0.05$ and $p < 0.01$ respectively. AA, amino acids; FC, fold change.

3.2. Overview of Whey Proteins and Functions in Donkey Milk

The overall whey proteins and functions were analyzed by GO and KEGG analysis, respectively (Figure 1). GO analysis was performed to identify the Biological Process (BP), Cellular Components (CC), and Molecular Functions (MF) (Figure 1A), while analysis of KEGG pathways was performed to identify the functions including six of Brite A class and 42 of Brite B class (Figure 1B). Among all of the identified whey proteins, total 46 terms of the three GO classes were clustered, with 10 different molecular functions, 15 kinds of cellular components, and 21 diverse biological processes. To be noticed, the summation of the numbers of proteins presented within the 46 terms was 1982, which was much higher than the actual identified number of proteins 620, indicating that some whey proteins performed several functions and thus belonged to multiple GO classes. For the KEGG analysis, 184, 141, 117, 87, 84, and 35 proteins were related to the six Brite A functions of organismal systems, human diseases, cellular processes, metabolism, environmental information processing and genetic information processing, respectively.

The top 10 KEGG Brite B functions included immune system, transport and catabolism, global and overview metabolic pathway maps, signal transduction, cancers, endocrine system, infectious viral diseases, infectious bacterial diseases, digestive system, and cell growth and death. Further phylogenetic classification for orthologous proteins were analyzed by annotation based on Cluster of Orthologous Groups of proteins. Among the total 25 functions in the database, 23 functions were annotated (Supplementary Figure S1), with the top five functions including posttranslational modification, protein turnover, chaperones (O), General function prediction only (R), Signal transduction mechanisms (T), Carbohydrate transport and metabolism (G), and Amino acid transport and metabolism (E).

3.3. Common and Different Whey Proteins in Donkey Milk from the Two Groups

The comparison of proteome data of whey proteins between the XJ and SD samples were initially analyzed by PCA and correlation heatmap of all proteins detected from the two groups (Figure 2).

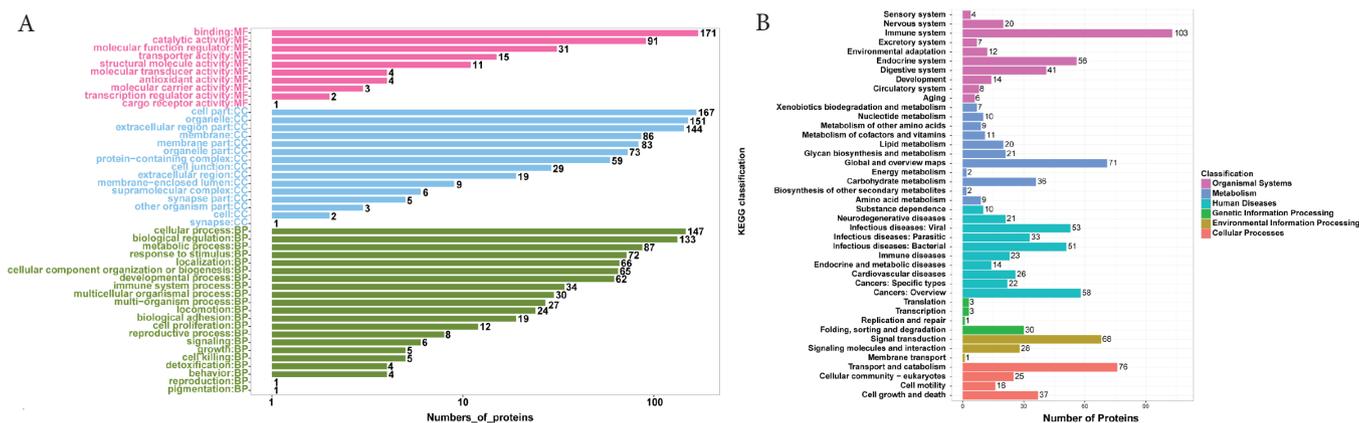


Figure 1 | Functional analysis based on gene ontology level 2: Biological Process (BP), Cellular Components (CC), and Molecular Functions (MF) (A), and Kyoto Encyclopedia of Genes and Genomes (B) of whey proteins in donkey milk from China.

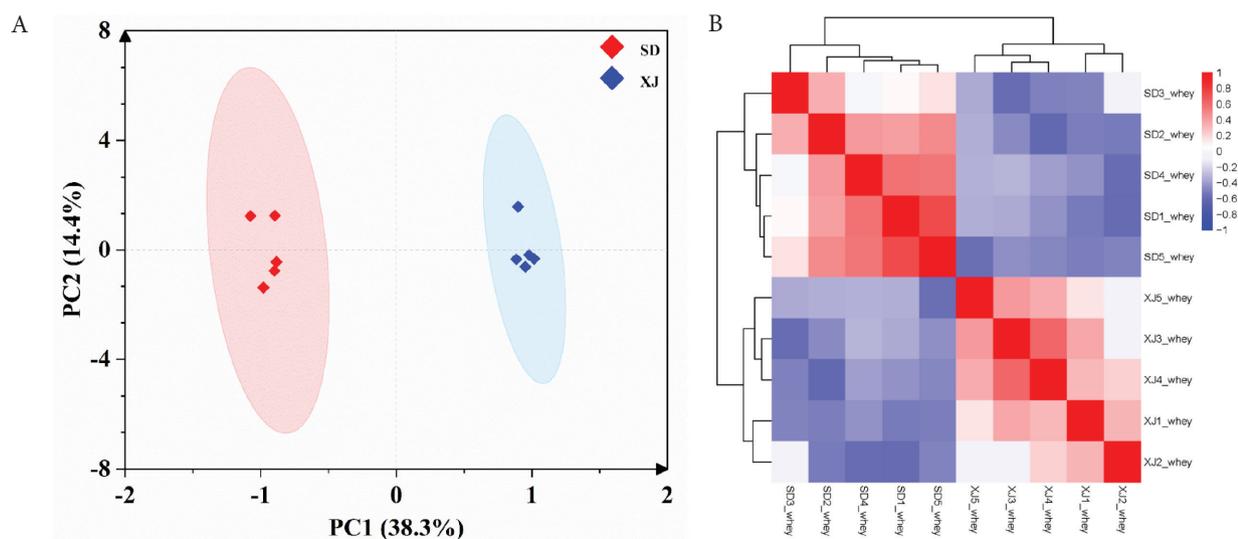


Figure 2 | Correlations of expression of proteins between the Xinjiang Province (XJ) and Shandong Province (SD) samples of donkey milk. (A) Principle component analysis of proteomic data. (B) Correlation heatmap of whey proteins between the two sample groups.

The samples from different groups clustered separately, while the samples from the same group clustered together, suggesting that there were great differences in expression of whey proteins between the XJ and SD samples, and the samples within each group had good repeatability.

Profiles of whey proteins in donkey milk were compared between the XJ and SD samples. A total of 620 proteins were identified and relatively quantified by proteomic analysis, including 438 of the common proteins in both sample groups (Figure 3A and 3B). There are 26 whey proteins with a statistically significant greater than twofold change in the samples of donkey milk from XJ and SD. Compared with those in the SD samples, 20 proteins such as Whey-acidic-protein (WAP) four-disulfide core domain protein and elafin were up-regulated, and six proteins including GDP-mannose 4,6-dehydratase and protein S100-G were down-regulated in the XJ samples (Figure 3A and Supplementary Table S1). The common proteins with the similar abundance ($p > 0.05$) in the two sample groups occupied approximately 70.6% of the total identified whey proteins, which indicates that whey proteins in the two groups of donkey milk shared a big part of fractions of whey proteins. However, 105 and 77 proteins were identified to overexpress in XJ and SD samples, respectively ($p < 0.05$), suggesting that each group had unique fraction of proteins profile (Figure 3B). The hierarchical clustering analysis of differentially expressed whey proteins further revealed two different clusters corresponding to the two sample groups, with total 182 proteins significantly more abundant in either of the XJ or in the SD samples (Supplementary Figure S2).

Functions of biological process, cellular components, and molecular functions associated with the differentially expressed proteins were identified based on the analysis of GO level 2 (Figure 4). Functional categorization showed that the numbers of differently expressed proteins in each of the GO term were all bigger than that of the SD samples, indicating that more proteins involved in the three targeted functions in the XJ samples rather than the

SD samples. For example, 31 proteins of cellular process related to biological process, 35 extracellular region part proteins of the cellular components, and 33 binding proteins related to molecular functions were overexpressing in the XJ samples, respectively, while only another nine, nine and 18 proteins of the same GO functions were overexpressing in the SD samples, respectively (Figure 4).

In addition to a single protein and its function, identification of biological network is very important for understanding how proteins interact with each other and function within the cell. The interactions of differentially expressed whey proteins in donkey milk were investigated using the Search Tool for the Retrieval of Interacting Genes/Proteins database (Supplementary Figure S3). The combined score of the interaction networks ranged from 150 to 999, and the higher score reflects the higher probability that the interaction is supported by several types of evidences and biologically meaningful. Total 112 proteins were enriched in networks. Important proteins overexpressed in the XJ samples with a high fold change included tissue-type plasminogen activator (big and red node, accession no.: XP_014693260.1), which involved in protein-protein interactions with 74 proteins. Neuroserpin and apolipoprotein B-100 (big and green nodes, accession no.: XP_014701927.1 and XP_014721916.1, respectively) were highly expressed in the SD sample rather than in the XJ samples, which had interactions with 58 and 68 proteins, respectively.

4. DISCUSSION

4.1. Favorable Properties Indicated by a Low Casein/Whey Protein Ratio

Caseins were characterized with allergenic properties, and a low casein/whey protein ratio made it more favorable for human nutrition

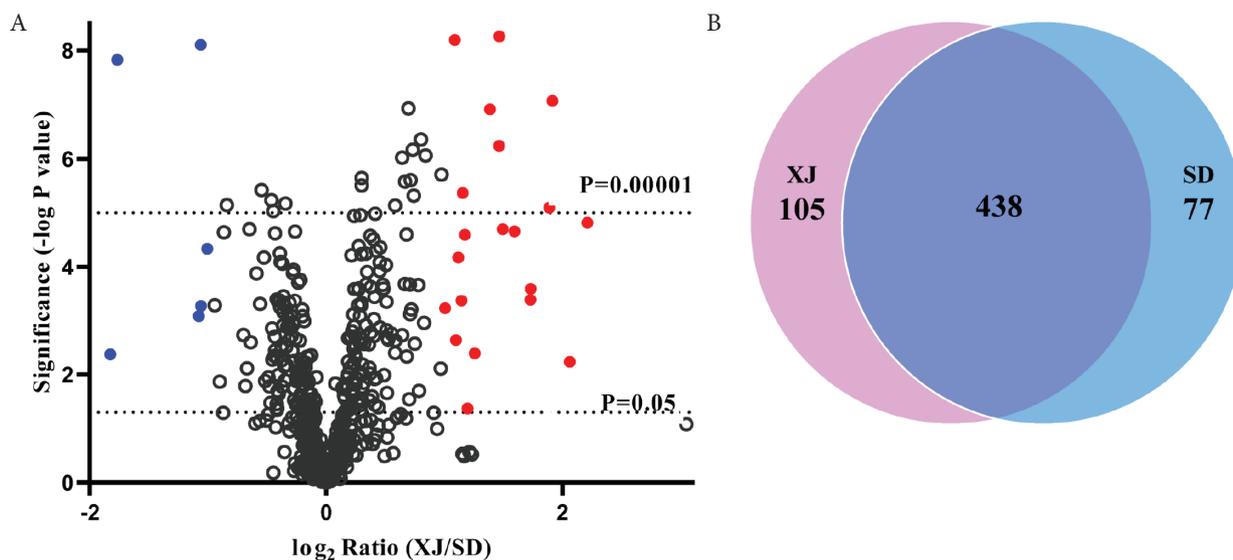


Figure 3 | Common and different expression of whey proteins in donkey milk from Xinjiang Province (XJ) and Shandong Province (SD). (A) Expression of whey proteins. Shown is the \log_2 transformed ratio of the protein abundance in the XJ samples relative to the abundance in the SD samples. Grey dots: background proteome. The level of significance was evaluated based on $p < 0.05$ and the fold change two times cut-off. Red dots: proteins expressed uniquely high in the XJ samples; blue dots: proteins expressed uniquely high in the SD samples. (B) Venn diagram of number of whey proteins expressed in the two sample groups. The fold change 1.2 times cut-off.

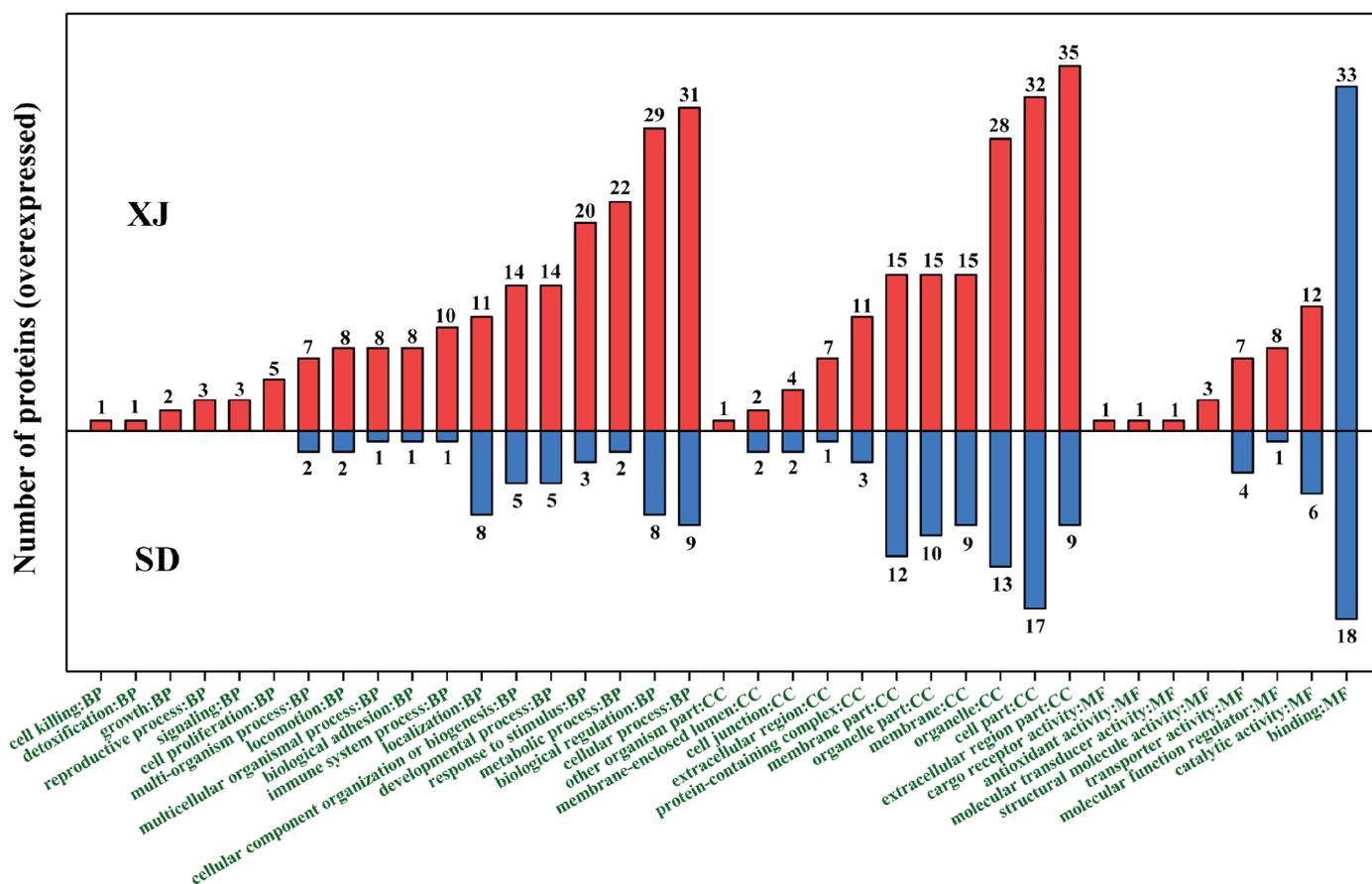


Figure 4 | Functional categorization based on Gene Ontology (GO) level 2 analysis of differentially abundant whey proteins in donkey milk from Xinjiang Province (XJ) and Shandong Province (SD).

[2–4,6,24]. The casein/whey ratio in our study was similar between the samples from XJ and SD ($p > 0.05$), with a low level of 0.48 and 0.43 (Supplementary Figure S4), respectively, both much lower than the previously reported ratio of 1.3 in donkey milk collected from the same area of XJ, and also lower than the mean ratio of 0.67 in donkey milk from Italy [4,9]. The low level of casein/whey ratio is much more lower than other milk (such as cow, sheep, goat and mare), very close to the value of human milk (0.4), and thus plays an important role in reducing its allergenicity [2,3,6]. The lower ratio of casein/whey relating to the lower ratio of soluble calcium also provides the special property of soft curd formation for donkey milk, and thus makes donkey milk a good substitute of human milk for bottle-fed infants [2,3,6]. Therefore, the products studied in our study are suggested to be very good sources of human nutrition with a favorable casein/whey ratio.

4.2. Specificity of Important Caseins and New Isoforms

In most milk types, four gene products including α_{s1} -, α_{s2} -, β -, and κ -caseins are synthesized in mammary gland in response to lactogenesis hormones and other stimuli, with α_{s1} -casein as the most important protein causing milk allergy [9]. Within the previously identified isoforms of α_{s1} -casein, only one reported type with 202 amino acids was observed in this study. Actually, numbers of amino acids of α_{s1} -casein were found to be diverse in different samples,

for example, four α_{s1} -casein isoforms with 202, 201, 197 and 196 amino acids were detected in a dephosphorylated milk sample collected from an individual donkey at middle lactation stage, while a higher molecular mass with 210 amino acids long was observed in donkey milk in another research [9,25]. To be noticed, there is one new isoform of α_{s1} -casein consisting of 208 amino acids in our study, which has not been reported in donkey milk before. The existence in an individual milk sample of a full-length α_{s2} -casein with 221 amino acids and its four different isoforms of 216, 185, 214, and 183 amino acids have been reported in previous studies [3,8,26]. However, the two α_{s2} -casein isoforms identified in our study had 236 and 231 amino acids, which were firstly observed in donkey milk. Actually, α_{s2} -casein is a very minor component in donkey milk, thus the absence of these two α_{s2} -casein components in previous studies were very likely due to the very low amount of them in donkey milk [2,3,24]. The primary structure of two β -casein isoforms as the most abundant casein component, with 226 and 218 amino acids, respectively, have been characterized in donkey milk in literatures [25,27]. Unlikely in our study, one new isoform of β -casein with 233 amino acids was observed in both of the samples with the similar abundance. As the amount of κ -casein in the fraction of *Equidae* is also very low, two components of κ -casein both with 182 amino acids were identified in our study, indicating that our methods had a very high accuracy [3]. Moreover, the discovery of the two κ -casein components in our study was the first time on account of that only one isoform of κ -casein with 162 amino acids long has been reported in donkey milk previously [3].

Thus among the total seven isoforms of caseins, only one has been reported in previous studies, and the other six of them were firstly identified here in donkey milk.

4.3. Variations of Major Whey Proteins and Functions

As the major constituents of whey proteins, β -lactoglobulins and α -lactalbumin were found to have similar expression profile with fold change of no more than 1.2 (Table 2). Lysozyme was recognized as the third major whey protein in donkey milk, and much more abundant than that in other milk such as cow, goat, sheep and even human milk [2–4]. In our study, donkey milk from SD overexpressed lysozyme in comparison with that of XJ while expressed less lysozyme precursor, which would be turned into more lysozyme after post-translational modification. Even though most gram-positive bacteria can be initiated lysis by lysozyme, some coccus-shaped species are resistant to lysozyme due to the O-acetylation of its peptidoglycan by O-acetyltransferase [1,6,28]. Interestingly, our previous work found that gram-positive bacteria *Macrococcus* and *Streptococcus* were identified as the top predominant genera in the same samples of donkey milk (data not published), which is consistent with the observation of lysozyme in this study because of their resistance to lysozyme. Moreover, absence of *Listeria*, *Salmonella*, or *Shigella* in the same samples of donkey milk (data not published), is very likely related to antimicrobial effects of at the presence of lysozyme and lactoperoxidase in it (Table 2) [18,29]. Immunoglobulins (mainly IgG, IgA, IgM) and lactoferrin was reported to occupy about 11.5% and 4.5% of the total whey proteins in donkey milk, however, our proteomic analysis has no observation of these two enzymes but only five immunoglobulins-related components as shown in Table 2 [4,5]. Among them, IgGFc-binding protein has been suggested to be important in the immunological defense for cell protection and anti-inflammation in tissues based on its immunoglobulin G-fragment crystallizable (IgG-Fc)

Table 2 | Expression of important whey proteins in donkey milk from XJ and SD

Description	FC (XJ/SD)	p-value
Predicted: β -lactoglobulin-1, partial ^a	0.87	0.00083***
β -Lactoglobulin-1	0.86	0.00017***
β -Lactoglobulin-2	0.93	0.05821
β -Lactoglobulin II variant B ^a	0.98	0.82671
α -Lactalbumin	0.83	0.00002***
Lysozyme C precursor ^a	1.63	0.02871*
Predicted: lysozyme C, milk isozyme ^a	0.88	0.01090*
IgGFc-binding protein ^a	1.91	8.46×10^{-8} ***
Predicted: V-set and immunoglobulin domain-containing protein 10-like ^a	1.78	0.00112**
Predicted: immunoglobulin J chain ^a	0.88	0.00104**
Predicted: polymeric immunoglobulin receptor ^a	0.79	0.00001***
Predicted: high affinity immunoglobulin α and immunoglobulin mu Fc receptor isoform X1 ^a	0.77	0.00008***
Serum albumin precursor	0.99	0.79207
Predicted: lactoperoxidase ^a	0.50	0.00067***

^a*Equus asinus*. *, **, and *** represent significant difference at $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively. FC, Fold change.

binding property and tissue distribution [30,31]. The J chain is an important polypeptide which regulates polymer formation of IgA and IgM, and its incorporation into polymeric IgA and pentameric IgM endows them with several special features including agglutination of bacteria, non-inflammatory reaction with limited complement-activating activity, and a high affinity for the polymeric immunoglobulin receptor of transporting them into milk [32]. In addition, serum albumin was reported to account for 6.2% of total whey proteins in donkey milk previously, however, only its precursor was determined in this study, the reason of which is unclear [2,4,5]. Noticeably, serum albumin is not synthesized in the mammary gland but may enter the milk by uptake with immunoglobulins, which may partially explain the absence of serum albumin and immunoglobulins, but further studies are necessary to do for more precise explanations [3].

4.4. Potential Factors for Proteins Differentiation between the Two Provinces

Differentiation on expression of whey proteins between the XJ and SD maybe ascribed to several factors including donkey breed, feed, climate, and lactation period [33,34]. The detailed information of the two sample groups have been summarized and shown in Supplementary Table S2. The most dominant protein WAP Four-Disulfide Core domain protein 2 (WFDC2) in XJ belongs to a large family of WFDC proteins with multiple functions including anti-protease, anti-bacterial, anti-viral and anti-inflammatory properties [35–37]. It has been shown that the expression of WFDC2 was related to lactation period, for example, the expression is elevated in early lactation, down regulated in mid lactation, subsequently increase slowly and markedly during involution [38]. Our samples were all collected from jennies with mature milk, however, specific lactation period of them might be various, which reasonably induced the significant difference on the expression of it between the two groups. Moreover, different expression of some proteins may be also related to the health status of donkey breast, for example, the higher expression proteins in XJ including clusterin, C-reactive protein and serum amyloid A protein-like have been proposed to up-regulate in milk from donkey bodies with mastitis [39]. Component of feed is another important factor which may affect the expression of some proteins, for example, protein S100-G, the expression of which was much higher in SD rather than in XJ, has been indicated to have multiple binding partners and functions and utilize calcium to propagate a response through a signaling pathway as a calcium-binding protein [40]. The up regulation of this protein in SD may be related to the addition of vitamin D3 and calcium ions in feed (Supplementary Table S2) [41]. It should be noticed that the differentiation of these proteins may not only result from single factor but simultaneous effects of multiple factors, thus further studies of concrete analysis on combined factors are necessary to be performed in the future.

5. CONCLUSION

This study preliminarily determined caseins and whey proteins in donkey milk from two major producing areas of China using proteomic analysis. Total 620 whey proteins were clustered according to their functional categorization based on GO and KEGG.

Six new isoforms of caseins were identified, and a very low ratio of casein/whey was discovered in donkey milk. This study established the profiles of protein composition in donkey milk from China, enhancing knowledge of protein expression in the potential high value products of donkey milk, and providing significant guidance for the future production and consumption of donkey milk.

CONFLICTS OF INTEREST

All authors declare that this research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflicts of interest.

AUTHORS' CONTRIBUTION

WL contributed to experimental design, samples and data collection, and the drafting of the manuscript. BZ contributed to the data analysis, and critical revision of the manuscript. HL contributed to data analysis, figures plots, drafted the manuscript and revised it critically for important intellectual content. XL designed the study and contributed to the editing of the manuscript.

FUNDING

The work was supported by National Natural Science Foundation of China (Grant No. 31901798), and the Special item of China Xinjiang Uygur Autonomous regional synergy innovation - Science and Technology assistance project in Xinjiang 2019 (Project No. 2019E0272).

ACKNOWLEDGMENTS

We are very grateful for the supply of raw donkey milk by the Xinjiang Yukunlun Natural Food Engineering Company Limited and Shandong Dong'e jiao Company Limited.

SUPPLEMENTARY MATERIALS

Supplementary data related to this article can be found at <https://doi.org/10.2991/efood.k.210222.001>.

REFERENCES

- [1] Vincenzetti S, Polidori P, Mariani P, Cammertoni N, Fantuz F, Vita A. Donkey's milk protein fractions characterization. *Food Chem* 2008;106:640–9.
- [2] Aspri M, Economou N, Papademas P. Donkey milk: an overview on functionality, technology, and future prospects. *Food Rev Int* 2017;33:316–33.
- [3] Cunsolo V, Saletti R, Muccilli V, Gallina S, Di Francesco A, Foti S. Proteins and bioactive peptides from donkey milk: the molecular basis for its reduced allergenic properties. *Food Res Int* 2017;99:41–57.
- [4] Guo HY, Pang K, Zhang XY, Zhao L, Chen SW, Dong ML, et al. Composition, physiochemical properties, nitrogen fraction distribution, and amino acid profile of donkey milk. *J Dairy Sci* 2007;90:1635–43.
- [5] Salimei E, Fantuz F, Coppola R, Chiofalo B, Polidori P, Varisco G. Composition and characteristics of ass's milk. *Anim Res* 2004;53:67–78.
- [6] Uniacke-Lowe T, Huppertz T, Fox PF. Equine milk proteins: chemistry, structure and nutritional significance. *Int Dairy J* 2010;20:609–29.
- [7] Vincenzetti S, Amici A, Pucciarelli S, Vita A, Micozzi D, Carpi FM, et al. A proteomic study on donkey milk. *Biochem Anal Biochem* 2012;1:1–8.
- [8] Chianese L, Calabrese MG, Ferranti P, Mauriello R, Garro G, De Simone C, et al. Proteomic characterization of donkey milk "caseome". *J Chromatogr A* 2010;1217:4834–40.
- [9] Costanzo A. Characterization of donkey milk proteins by a proteomic approach. Ph.D. Thesis, Università degli Studi di Napoli Federico II, Napoli, Campania, 2013. Retrieved from: http://www.fedoa.unina.it/9122/1/costanzo_angela_25.PDF.
- [10] Ng-Kwai-Hang KF. Milk proteins: heterogeneity, fractionation, and isolation. In: Fuquay JW, editor. *Encyclopedia of dairy sciences*. San Diego, USA: Academic Press; 2011, pp. 751–64.
- [11] Massouras T, Triantaphyllopoulos KA, Theodossiou I. Chemical composition, protein fraction and fatty acid profile of donkey milk during lactation. *Int Dairy J* 2017;75:83–90.
- [12] Martin P, Grosclaude F. Improvement of milk protein quality by gene technology. *Livest Prod Sci* 1993;35:95–115.
- [13] Restani P, Ballabio C, Di Lorenzo C, Tripodi S, Fiocchi A. Molecular aspects of milk allergens and their role in clinical events. *Anal Bioanal Chem* 2009;395:47–56.
- [14] Permyakov EA, Berliner LJ. α -Lactalbumin: structure and function. *FEBS Lett* 2000;473:269–74.
- [15] Yamaguchi M, Yoshida K, Uchida M. Novel functions of bovine milk-derived α -lactalbumin: anti-nociceptive and anti-inflammatory activity caused by inhibiting cyclooxygenase-2 and phospholipase A₂. *Biol Pharm Bull* 2009;32:366–71.
- [16] Miranda G, Mahé MF, Leroux C, Martin P. Proteomic tools to characterize the protein fraction of *Equidae* milk. *Proteomics* 2004;4:2496–509.
- [17] Brumini D, Criscione A, Bordonaro S, Vegarud GE, Marletta D. Whey proteins and their antimicrobial properties in donkey milk: a brief review. *Dairy Sci Technol* 2016;96:1–14.
- [18] Šarić LČ, Šarić BM, Kravić SŽ, Plavšić DV, Milovanović ILJ, Gubić JM, et al. Antibacterial activity of domestic Balkan donkey milk toward *Listeria monocytogenes* and *Staphylococcus aureus*. *Food Feed Res* 2014;41:47–54.
- [19] Bilic P, Guillemin N, Kovačević A, Ljubić BB, Jović I, Galan A, et al. Serum proteome profiling in canine idiopathic dilated cardiomyopathy using TMT-based quantitative proteomics approach. *J Proteomics* 2018;179:110–21.
- [20] Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 1976;72:248–54.
- [21] Lu J, Zhang S, Liu L, Pang X, Ma C, Jiang S, et al. Comparative proteomics analysis of human and ruminant milk serum reveals variation in protection and nutrition. *Food Chem* 2018;261:274–82.
- [22] Abdelmegid S, Murugaiyan J, Abo-Ismael M, Caswell JL, Kelton D, Kirby GM. Identification of host defense-related proteins using label-free quantitative proteomic analysis of milk whey from

- cows with *Staphylococcus aureus* subclinical mastitis. *Int J Mol Sci* 2018;19:78.
- [23] Chen L, Wang J, Jiang P, Ren F, Lei X, Guo H. Alteration of the colostrum whey proteome in mothers with gestational hypothyroidism. *PLoS One* 2018;13:e0205987.
- [24] Nazzaro F, Orlando P, Fratianni F, Coppola R. Isolation of components with antimicrobial property from the donkey milk: a preliminary study. *Open Food Sci J* 2010;4:43–7.
- [25] Cunsolo V, Cairone E, Fontanini D, Criscione A, Muccilli V, Saletti R, et al. Sequence determination of α_{s1} -casein isoforms from donkey by mass spectrometric methods. *J Mass Spectrom* 2009;44:1742–53.
- [26] Saletti R, Muccilli V, Cunsolo V, Fontanini D, Capocchi A, Foti S. MS-based characterization of α_{s2} -casein isoforms in donkey's milk. *J Mass Spectrom* 2012;47:1150–9.
- [27] Criscione A, Cunsolo V, Bordonaro S, Guastella AM, Saletti R, Zuccaro A, et al. Donkeys' milk protein fraction investigated by electrophoretic methods and mass spectrometric analysis. *Int Dairy J* 2009;19:190–7.
- [28] Moynihan PJ, Clarke AJ. O-Acetylated peptidoglycan: controlling the activity of bacterial autolysins and lytic enzymes of innate immune systems. *Int J Biochem Cell Biol* 2011;43:1655–9.
- [29] Šarić L, Pezo L, Šarić B, Plavšić D, Jovanov P, Karabasil N, et al. Calcium-dependent antibacterial activity of donkey's milk against *Salmonella*. *Ann Microbiol* 2017;67:185–94.
- [30] Kobayashi K, Ogata H, Morikawa M, Iijima S, Harada N, Yoshida T, et al. Distribution and partial characterisation of IgG Fc binding protein in various mucin producing cells and body fluids. *Gut* 2002;51:169–76.
- [31] Gazi MH, He ML, Cheville JC, Young CYF. Downregulation of IgG Fc binding protein (FcγBP) in prostate cancer. *Cancer Biol Ther* 2008;7:70–5.
- [32] Johansen FE, Braathen R, Brandtzaeg P. Role of J chain in secretory immunoglobulin formation. *Scand J Immunol* 2000;52:240–8.
- [33] Gubić J, Tomić J, Torbica A, Iličić M, Tasić T, Šarić L, et al. Characterization of several milk proteins in domestic Balkan donkey breed during lactation, using lab-on-a-chip capillary electrophoresis. *Chem Ind Chem Eng Q* 2016;22:9–15.
- [34] Li SS, Shen JS, Ren DX, Liu JX. Effects of the processing methods of corn grain and soybean meal on milk protein expression profiles in dairy cows. *Animal* 2015;9:267–74.
- [35] Small DM, Doherty DE, Dougan CM, Weldon S, Taggart CC. The role of whey acidic protein four-disulfide-core proteins in respiratory health and disease. *Biol Chem* 2017;398:425–40.
- [36] Henrick BM, Yao XD, Nasser L, Roozrogousheh A, Rosenthal KL. Breastfeeding behaviors and the innate immune system of human milk: working together to protect infants against inflammation, HIV-1, and other infections. *Front Immunol* 2017; 8:1631.
- [37] Reading JL, Meyers AFA, Vyakarnam A. Whey acidic proteins (WAPs): novel modulators of innate immunity to HIV infection. *Curr Opin HIV AIDS* 2012;7:172–9.
- [38] Watt AP, Sharp JA, Lefevre C, Nicholas KR. WFDC2 is differentially expressed in the mammary gland of the tamarin wallaby and provides immune protection to the mammary gland and the developing pouch young. *Dev Comp Immunol* 2012;36: 584–90.
- [39] Swanson KM, Stelwagen K, Dobson J, Henderson HV, Davis SR, Farr VC, et al. Transcriptome profiling of *Streptococcus uberis*-induced mastitis reveals fundamental differences between immune gene expression in the mammary gland and in a primary cell culture model. *J Dairy Sci* 2009;92:117–29.
- [40] Dempsey BR, Rintala-Dempsey AC, Shaw GS. S100 proteins. In: Choi S, editor. *Encyclopedia of signaling molecules*. New York, NY: Springer; 2016, pp. 1–10.
- [41] Dong YP, Yang LP, Luo WT, Zhao Q. Mechanisms of vitamin D in regulation of calcium binding protein. *Int J Pediatr* 2017;44: 59–63.