International Journal of Food Science +Technology

Original article

Identification of β-casein phenotypes (A^1/A^2) in the milk of the Indian Jersey crossbreed bovine using the high-resolution accurate mass spectrometer

Manish Kumar Singh,[1]* Arvind Kumar,[1] Dinesh Chandra Rai,[1] Ankur Aggarwal[1] & Mohit Malik[2]

1 Department of Dairy Science and Food Technology, Institute of Agricultural Sciences, Banaras Hindu University, Varanasi, Uttar Pradesh, India

2 Department of Basic and Applied Sciences, National Institute of Food Technology Entrepreneurship and Management, Kundli, Sonipat, Haryana, India

(Received 1 June 2023; Accepted in revised form 10 August 2023)

Summary Foodomics is an emerging probing method of phenotype investigation of the different milk proteins and their subtypes. The polymorphic nature of the β-casein (β-Cn) protein has shown fourteen different protein variants to date in bovines. The analysis of the β-Cn genetic polymorphism from the milk of the crossbred dairy animals is crucial for the quality assurance of the consumers from the various health concerns, especially those linked with the A1 phenotype which yields β-casomorphin-7 on *in vivo* digestion. Jersey-crossed Indian cattle have been widely utilised in dairy because of their better milk production and survival performance trait. In this investigation, an SDS-PAGE coupled with a high-resolution accurate mass spectrometry-based proteomics approach has been applied to identify the presence of specific phenotype of the β-Cn protein in the milk of the 24 Indian crossbred (Jersey crossed) animals. Amino acid sequential analysis has been done using different search modules, as MS Amanda and Sequest HT showed 17 cows are producing A^2 β-Cn (Pro~67) while only seven animals yielded the A^1 variant (His~67). The maximum number of Indian Jersey-crossed animals are lactating milk having A^2 β-Cn. The A^2 milk from the crossbred animals is free from the negative impact on health caused by β-casomorphin-7 (BCM-7) released during digestion of the A^1 phenotype. Among the molecular biology techniques, top-down proteomics has been an intriguing technique for the identification of protein genetic polymorphic products.

Keywords A2 milk, Electrophoresis, Jersey crossbreed, Mass spectrometry, β-casein, β-casomorphin-7.

Introduction

As India is leading in milk production globally, the contribution of cross-bred animals is very significant to avail this liquid food for millions of people. Jersey (*Bos taurus*) is a European bovine breed widely utilised in dairy programs worldwide for milk lactation (Singh *et al*., 2023a, 2023b). In India, Jersey has been crossbred with different zebu cattle of indicine origin (such as Sahiwal, Tharparkar, Kankrej, Vechur, etc.) for better production, fat content, and survival outcomes. Therefore, the protein genetic polymorphism evaluation of the milk from these crossbred animals is crucial for the milk quality assurance of the consumers from the various health risks related to A^1 milk (Daniloski *et al*., 2021).

*Correspondent: E-mail: manish04@bhu.ac.in

Proteins in the milk (~3.5–4.0%; Bovine milk) are a variable constituent, widely influenced by numerous factors. Caseins have made the 80% of milk proteins in the form of phosphoprotein complex micelle and the other 20% are whey proteome of a massive class of proteins (Chopra *et al*., 2020). The major protein of milk, casein constitutes four subtypes termed as αS1-Cn (*CSN1-S1*), αS2-Cn (*CSN1-S2*), β-Cn (*CSN2*), and κ-Cn (*CSN3*) with an approximate proportion of 4:1:3:1 (Daniloski *et al*., 2022). The heterogeneity of milk proteins includes genetic polymorphisms that have received great interest in the last decade among the dairy industries and also the consumer because of their possible association with milk composition, processing attributes, and human health concerns (The BCM-7 release). Nowadate, a total of fourteen different genetic variants of beta-casein (β-Cn) have been identified in several investigations; A (A^1, A^2, A^3, and

doi:10.1111/ijfs.16655

A^4), B, C, D, E, F, G, H (H^1 and H^2), I and J (Patel *et al*., 2020). The peptic digestion of the A^1 β-Cn chain yields β-casomorphin-7 (BCM-7), which is significantly correlated with many health issues like autism at an early age, cardiovascular risks, diabetes, ischaemic heart disease, etc. unlike A^2 milk (Liu *et al*., 2023).

In the world market A^2 milk and milk products are being sold with the label "A^2 products" and have their targeted consumer choice, hence the identification of A^2 and A^1 variants is of great importance for the food industries, dairy farmers, and consumers too (Liu *et al*., 2023). In previous decades, proteomic-based analysis coupled with mass spectroscopy has been exceedingly applied in the analytical studies of the milk proteome of various dairy animals (Roy *et al*., 2020). Although studies on the detection of β-Cn genetic variation have been carried out to identify the phenotypes with the application of ultra-performance liquid chromatography (UPLC) and high-resolution mass spectrometry (HRMS) (Fuerer *et al*., 2020). To our knowledge, this is the first-time probe of β-Cn genetic variant profiling of a crossbred cattle (Indian Jersey) milk using protein-based sequencing by the application of the robust technique of electrophoresis prior to the high-resolution accurate mass spectrometry (HRAMS) and proteome database search.

Materials and methods

Acetonitrile, bovine serum albumin (BSA), prestained protein marker, iodoacetamide (IAA), trichloroacetic acid (TCA), trypsin, formic acid, tris-base, tris (2-carboxyethyl)phosphine (TCEP), PBS, glycerol, methanol, sodium dodecyl sulfate (SDS), ammonium persulphate, tetra-methyl-ethylenediamine (TEMED), bis-acrylamide, 2-mercaptoethanol, glycine, bromophenol blue, coomassie G-250, and Folin's-phenol reagent were acquired from Sigma and Himedia of molecular biology grade. Analytical grade glacial acetic acid, sodium potassium tartrate, copper sulfate, hydrochloric acid, sodium carbonate, and caustic soda were procured from SRL chemicals.

Sample collection and casein isolation

Twenty-four Jersey-crossed bovines ($N = 24$) of an almost homologous period of lactation were selected from the animal shed of the Banaras Hindu University, Varanasi, India. The animal husbandry practices of these animals are similar. Each sample of 200 mL raw milk was drawn separately and was stored at a cooling temperature of −20°C until analysis. As Hollar *et al*. (1991) described, a slightly altered iso-electric precipitation was done to obtain pure casein. Briefly, the samples were heated to 60°C before centrifugation at 2600 × ***g*** for 25 min at 30°C for the skimming of milk. Then, the pH of obtained fatless milk was reduced to 4.6 by adding 10% (v/v) glacial acetic acid. Samples were centrifuged subsequently to iso-electric coagulation, and the resulting supernatant undergoes decantation. The casein pellet was washed two times in acidic water (pH:4.0). The whole pellet was resuspended in deionised water of neutral pH. The Lowry estimation procedure was applied to quantitatively analyse the protein concentration with the BSA standard.

SDS-PAGE protein separation

Andrews (1983) described the procedure which has been used to perform the gel electrophoretic run of obtained casein with few alterations. In short, casein samples were added 1:1(v/v) with a 125 mM Tris HCl, 4% SDS, 0.4% bromophenol blue, 10% 2-mercaptoethanol, and 10% glycerol at pH 6.8. At 90°C, the protein working solutions were undergoes heat denaturation for 10 min. On the run, 10 μg of whole caseins from the crossbred Jersey cows were loaded into the gel wells. The 5 μL of prestained protein marker was loaded in the first well. The stacking gel solution with a pH value of 6.8 differs from the resolving gel (pH 8.3) in terms of the ratio of acrylamide and SDS 1:2 and 3:2, respectively. A Mini PROTEAN system (Bio-Rad) was fully poured with 5X running buffer with 10% concentration of SDS, pH 8.3. The gel electrophoresis was completed using the voltage-stepped method, and staining was done with Coomassie (G-250) dye followed by documentation on the gel doc Go image system.

High-resolution accurate mass spectrometry

At the end of the SDS-PAGE electrophoretic run, the in-solution digestion process was carried out, *i.e*., different fractions were excised from the gel and cut into small parts before destaining. The reducing reaction was done using 0.01 M tris (2-carboxyethyl) phosphine, joined by alkylation through 0.05 M IAA, followed by trypsin cleavage was performed for 16 h at 37 °C (1:50, enzyme: lysate). Vacuum drying of the peptides was done at the end of the reaction with a 1/10-part solution of TCA. Thermo Fischer Scientific orbitrap eclipse tribrid mass spectrometer with nano LC and UHPLC system coupled with QE plus was employed for the mass spectrum detection of the peptides mixture. In the C^{18} column, the gradient of chromatography was performed for 120 min. The resolution of the orbitrap was set up to the level of 70 K to obtain the mass spectrum. Every charged state of the precursor had been removed after a 10-s dynamic exclusion. After samples were analysed, a file was compiled, and a proteome discoverer was utilised to compare it to the UniProt proteome reference

database. The average protein false discovery rate (FDR) was adjusted to 0.04 after matching the peptide spectrum (Singh *et al.*, 2023a, 2023b). All samples were studied in triplicates for statistical analysis.

Results and discussion

The present investigation aims to analyse the β-Cn genetic variant in the 24 indicine crossbred bovine (Jersey crossed) milk samples by electrophoretic separation and HRAMS is the first study to identify various β-Cn phenotypes (A^1 and A^2) of Indian crossed cattle using a proteomic approach. The SDS-PAGE run of the isolated whole casein samples gives the casein fraction a good separation on a 15% resolving gel with an average retention factor (Rf) value of 0.48. The proteomics investigation generates a mass spectrum in relation to the relative abundance and mass: charge (m/z) (Fig. 1). The UniProt proteome database search reveals the types of the β-Cn variants in the individual samples by sequential analysis. The maximum number of sequences (Nilsson *et al.*, 2020) of β-Cn from the Indian Jersey crossbred cow milk has shown the presence of the A^2 (Pro~67 AA) type of genetic variant, and only seven samples showed the A^1 (His~67 AA) type. In the NCBI search database, β-Cn A^2 has UniProt KB ID number P02666 and WAS60671 for A^1 (Table 1). Validation of this casein subtype phenotyping results was done using pure β-Cn protein standard (Sigma-Aldrich) along with the previously genotyped milk from pure A^1 and A^2 lactating cows.

Polymorphism in the bovine β-Cn has been detected with high-resolution mass spectroscopy with ultra-performance liquid chromatography (Fuerer *et al.*, 2020). Duarte-Vázquez *et al.* (2018) fractionated Holstein Friesian milk β-Cn variants by urea gel electrophoresis and liquid chromatography-mass spectroscopy sequencing. In different studies, various detection probes such as polymerase chain reaction (PCR), isoelectric focussing, mass spectroscopy, mid-infrared spectroscopy, nuclear magnetic resonance, and Fourier transform infrared spectroscopy have been used in genetic variant identification (Mayer *et al.*, 2021; Daniloski *et al.*, 2022). Although molecular methods such as different types of PCR are the most commonly applied technique for the analysis of β-Cn subvariants, but proteomics-based phenotyping is becoming an option for DNA-based analysis, which is fast, reliable, and economical too but have a long procedure that limits their wide spectrum applicability (Vigolo *et al.*, 2022). The proteomics method has more environmental sustainability for the dairy industry in comparison to other molecular techniques because no blood samples were used.

In recent years, several studies have reported on the health implications of A^1 phenotype containing milk on human physiology, especially the biosynthesis of a μ-opioid termed BCM-7 during proteolytic digestion of A^1 proteoform of β-Cn which is linked with multiple non-communicable health discomforts like autism, diabetic mellitus, ischemic heart disease, sudden infant death syndrome, and atherosclerosis (Daniloski *et al.*, 2021). The A^2 milk from the crossbred animals is free from the negative impact on health because BCM-7 cannot be released during digestion (Lambers *et al.*, 2021). Čítek *et al.* (2021) have shown that a greater β-Cn A^1 phenotype is observed in crossed dairy animals in comparison to native animals, implying an outcome of genetic selection in cross-breeding, on the other hand, the A^1 variant is related to a higher yield of protein (Rangel *et al.*, 2017). Many probes have shown the effect of casein polymorphic products in relation to coagulation attributes, *i.e.*, the β-Cn A^2 variant is showing an unsuitable milk coagulation ability, while the A^1 variant leads to good coagulum output and cheese yield (Nilsson *et al.*, 2020). In recent years, similar LC/MS-based proteomics approaches have been widely utilised in the molecular characterisation of casein proteoforms in different dairy animals (De Poi *et al.*, 2020; Broadbent *et al.*, 2021; Auzino *et al.*, 2022; Guo *et al.*, 2022; Singh *et al.*, 2023a, 2023b). The results of this study are showing homology with aforesaid works in the characterisation of β-Cn phenotype with the successful application of gel

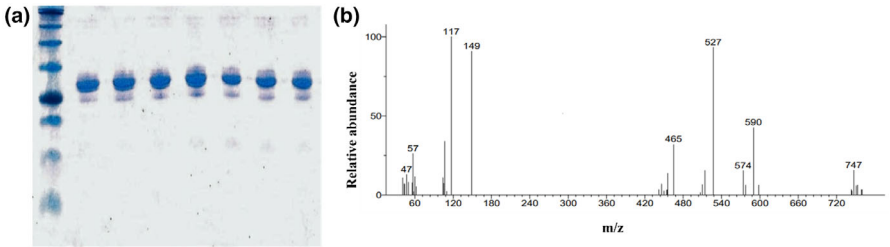

Figure 1 Illustrative representation of SDS PAGE gel (a) and a mass spectrum (b) obtained after proteomics-based analysis of the individual sample for the β-Cn genetic variant detection using electrophoresis and HRAMS.

Table 1 The results of AAs sequential analysis of Indian Jersey crossbred bovine milk β-Cn fraction by HRAMS and proteome database search from UniProt KB. Proline and histidine residues are highlighted in the A^2 and A^1 β-Cn chains

Accession	Description	Sum PEP score	Coverage [%]	Gene	PSM	Calc. pI	Sequence	Pfam ID	Score MS Amanda 2.0	Score sequest HT
P02666	Beta-casein OS = *Bos taurus* (A2)	131.605	33	CSN2	302	5.35	MKVLILACLVALALAREL ELNVPGEIVESLSSSEESITRI NKKIEKFQSEEQQQTED ELQDKIHPFAQTQSLV YPFPGPI**P**NSLPQNIPPLTQT PVVVPPFLQPEVMGVSKVK EAMAPKHKEMPFPKYPVE PFTESQSLTDVENLHLPLP LLQSWMHQPPLPPTVMFP PQSVLSLSQSKVLPVPQKA VPYPQRDMPIQAFLLYQE PVLGPVRGPFPIIV	Pf00363	34586.7	298.86
WAS60671	Beta-casein OS = *Bos taurus* (A1)	104.773	34	CSN2	224	5.35	DELQDKIHPFAQTQSLVYPF PGPI**H**NSLPQNIPPLTQTPVVV PPFLQPEVMGVSKVKE AMAPKHKEMPFPKYPVEP FTESQSLTLTDVENLHLPLP LLQSWMHQPHQPLPPTVMFPP QSVLSLSQSKVLPVPQKAVP YPQRDMPIQAFLLYQEPVLGPFPII	Pf00363	28409.78	280.48

electrophoresis and high-resolution accurate mass spectroscopy by obtaining the amino acid sequences in the Jersey crossbred animals, as these animals are the source of milk for a large population in India. Same the case of Australia, A2 milk products is covering a good share of the dairy market in developing countries like Brazil, China, India, and South Korea and the production system of A2 milk needs fast and accurate phenotypes screening methods (Fernández-Rico *et al*., 2022). Protein-based identification of A^2 milk is a key procedure regarding this attempt.

Conclusion

Over the last four decades, studies on the characterisation of the genetic variants of casein subtypes have been reported substantially. This work delineates the genetic variant profiling of Indian Jersey crossbreed dairy animals using a proteomic (LC/MS) approach. The HRAMS and electrophoresis-based assay on the twenty-four individual cows of this crossbreed animal has observed the A^2 variant in seventeen samples; only seven bovines have lactated the A^1 milk. Among the molecular biology methodology, top-down proteomics has been an intriguing technique for the identification of protein genetic polymorphic products. As crossbreed animals occupied a large part of cattle resources across the world, there is still much that can be contributed to apply casein phenotypes linked understanding to the animal breeding program.

Acknowledgment

The authors would like to gratefully acknowledge the financial support of the Institute of Eminence; IoE Research Grant, Banaras Hindu University No. R/Dev/IoE/Seed & Incentive Grant-III/2022-23/49243 and Sophisticated Analytical and Technical Help Institute (SATHI), Banaras Hindu University, Varanasi, India, for the HRAMS facility. We also presents our special thanks to the Bio-NEST, InnoRes Tech foundation, Banaras Hindu University for providing laboratory facility.

Author contributions

Manish Kumar Singh: Conceptualization (equal); data curation (equal); formal analysis (equal); investigation (equal); writing – original draft (equal). **Arvind Kumar:** Supervision (equal); writing – review and editing (equal). **Dinesh Chandra Rai:** Resources (equal); supervision (equal). **Ankur Aggarwal:** Visualization (equal); writing – review and editing (equal). **Mohit Malik:** Writing – review and editing (equal).

Conflicts of interest

The authors declare no conflicts of interest.

Data availability statement

Research data are not shared.

References

Andrews, A.T. (1983). Proteinases in normal bovine milk and their action on caseins. *Journal of Dairy Research*, **50**, 45–55.

Auzino, B., Miranda, G., Henry, C. *et al*. (2022). Top-down proteomics based on LC-MS combined with cDNA sequencing to characterize multiple proteoforms of Amiata donkey milk proteins. *Food Research International*, **160**, 111611.

Broadbent, J.A., Condina, M.R. & Colgrave, M.L. (2021). Quantitative mass spectrometry-based analysis of proteins related to cattle and their products–focus on cows' milk beta-casein proteoforms. *Methods*, **186**, 112–118.

Chopra, A., Ali, S.A., Bathla, S. *et al*. (2020). High-resolution mass spectrometer–based ultra-deep profile of milk whey proteome in Indian zebu (Sahiwal) cattle. *Frontiers in Nutrition*, **7**, 150.

This study is included because it has a detailed explanation of milk protein, especially the whey protein class.

Čítek, J., Brzáková, M., Hanusová, L. *et al*. (2021). Gene polymorphisms influencing yield, composition and technological properties of milk from Czech Simmental and Holstein cows. *Animal bioscience*, **34**, 2–11.

Daniloski, D., McCarthy, N.A., Markoska, T., Auldist, M.J. & Vasiljevic, T. (2022). Conformational and physicochemical characteristics of bovine skim milk obtained from cows with different genetic variants of β-casein. *Food Hydrocolloids*, **124**, 107186.

Daniloski, D., McCarthy, N.A. & Vasiljevic, T. (2021). Bovine β-Casomorphins: friends or foes? A comprehensive assessment of evidence from in vitro and ex vivo studies. *Trends in Food Science & Technology*, **116**, 681–700.

This reference is included because it provides the health implication of β-Casomorphin (BCM-7) in length and says about the negative aspects of A^1 milk (or variant).

De Poi, R., De Dominicis, E., Gritti, E., Fiorese, F., Saner, S. & Polverino de Laureto, P. (2020). Development of an LC-MS method for the identification of β-casein genetic variants in bovine milk. *Food Analytical Methods*, **13**, 2177–2187.

Duarte-Vázquez, M.A., García-Ugalde, C.R., Álvarez, B.E. *et al*. (2018). Use of urea-polyacrylamide electrophoresis for discrimination of A1 and A2 beta casein variants in raw cow's milk. *Journal of Food Science and Technology*, **55**, 1942–1947.

This work includes the application of electrophoresis followed by mass spectrometry for the identification of beta-casein variants using the amino acids sequencing approach.

Fernández-Rico, S., Mondragón, A.D.C., López-Santamarina, A. *et al*. (2022). A2 milk: new perspectives for food technology and human health. *Food*, **11**, 2387.

Fuerer, C., Jenni, R., Cardinaux, L. *et al*. (2020). Protein fingerprinting and quantification of β-casein variants by ultra-performance liquid chromatography–high-resolution mass spectrometry. *Journal of Dairy Science*, **103**, 1193–1207.

Guo, D., Deng, X., Gu, S., Chen, N., Zhang, X. & Wang, S. (2022). Online trypsin digestion coupled with LC-MS/MS for detecting of A1 and A2 types of β-casein proteins in pasteurized milk using biomarker peptides. *Journal of Food Science and Technology*, **59**, 2983–2991.

Hollar, C.M., Law, A.J.R., Dalgleish, D.G., Medrano, J.F. & Brown, R.J. (1991). Separation of β-casein A1, A2, and B using cation-exchange fast protein liquid chromatography. *Journal of Dairy Science*, **74**, 3308–3313.

Lambers, T.T., Broeren, S., Heck, J., Bragt, M. & Huppertz, T. (2021). Processing affects beta-casomorphin peptide formation during simulated gastrointestinal digestion in both A1 and A2 milk. *International Dairy Journal*, **121**, 105099.

Liu, Z., Pan, S., Wu, P., Li, M. & Liang, D. (2023). Determination of A1 and A2 β-casein in Milk using characteristic thermolytic peptides via liquid chromatography-mass spectrometry. *Molecules*, **28**, 5200.

This recent work on β-casein variant analysis has facts on digestion outcomes of A^1 milk and A^2 labelled products in the market.

Mayer, H.K., Lenz, K. & Halbauer, E.M. (2021). "A2 milk" authentication using isoelectric focusing and different PCR techniques. *Food Research International*, **147**, 110523.

This study has included different methods of assay of A^2 milk and shown their comparison which is relevant to our work.

Nilsson, K., Johansen, L.B., de Koning, D.J. *et al*. (2020). Effects of milk proteins and posttranslational modifications on noncoagulating milk from Swedish red dairy cattle. *Journal of Dairy Science*, **103**, 6858–6868.

Patel, S., Shah, T., Sabara, P. *et al*. (2020). Understanding functional implication of β-casein gene variants in four cattle breeds characterized using AmpliSeq approach. *Understanding functional implication of β-casein gene variants in four cattle breeds characterized using AmpliSeq approach*., **3 Biotech, 10**, 1–8.

Rangel, A.H.N., Zaros, L.G., Lima, T.C. *et al*. (2017). Polymorphism in the Beta casein gene and analysis of milk characteristics in Gir and Guzerá dairy cattle. *Genetics and Molecular Research*, **16**, 1–9.

Roy, D., Ye, A., Moughan, P.J. & Singh, H. (2020). Composition, structure, and digestive dynamics of milk from different species—a review. *Frontiers in Nutrition*, **7**, 577759.

Singh, M.K., Kumar, A., Nimmanapalli, R. & Hooda, A. (2023a). Probing of alpha, beta and kappa-caseins polymorphic variants in Gangatiri cow milk with the use of polyacrylamide gel electrophoresis and HRAMS. *Journal of Dairy Research*, **90**, 54–57.

This study includes a description of the high-resolution accurate mass spectroscopy procedure which is similar to our work to analyse the phenotypes.

Singh, M.K., Kumar, A., Nimmanapalli, R. & Pandey, A.K. (2023b). Proteomics-based milk whey proteome profiling of Indian Jersey cross-breed cows followed by chromosomal mapping. *Journal of the Science of Food and Agriculture*., **103**, 5634–5640.

Vigolo, V., Franzoi, M., Cendron, F. *et al*. (2022). Characterization of the genetic polymorphism linked to the β-casein A1/A2 alleles using different molecular and biochemical methods. *Journal of Dairy Science*, **105**, 8946–8955.

This work has compared the molecular and biochemical methods of β-casein variant analysis, hence it is useful to draw the comparison of the results and proteomics method with other available techniques.