Emirates Journal of Food and Agriculture. 2022. 34(8): 675-687
doi: 10.9755/ejfa.2022.v34.i8.2882
http://www.ejfa.me/

RESEARCH ARTICLE

Genotyping, physicochemical characterization, and protein isoform quantification of β-casein A2 milk in chinese simmental and Angus cattle

Adhita Sri Prabakusuma[1], Hamzah Aleryani[1], Xiaoyan Kong[2], Xianwei Shi[2], and Aixiang Huang[1]*

[1]College of Food Science and Technology, Yunnan Agricultural University, Kunming, China – 650201, [2]College of Animal Science and Technology, Yunnan Agricultural University, Kunming, China – 650201

ABSTRACT

The recent findings on β-casein (β-CN) A2 cattle's milk variant have triggered global research interest, as it is claimed to be a safer choice for human health. Only a few studies have been performed to investigate the presence of β-CN A1 and A2 alleles of the CSN2 gene in dual-purpose Simmental and Angus cattle in Yunnan province, China. This present study aimed to genotype cattle producing A2 milk, predict the physicochemical properties, and quantify the isoform concentration of β-CN A2 protein. Blood samples were collected from 286 cattle (201 Simmental and 85 Angus breeds). Genotyping was performed using polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) coupled with the sequencing method. Physicochemical properties were predicted using several web-based bioinformatics tools. A sandwich enzyme-linked immunosorbent assay (ELISA) was used to measure β-CN A2 protein isoform concentration in fresh A2 milk. Allele frequencies of β-CN A2 were more dominant in Simmental (0.642) and Angus (0.835) than A1. The β-CN A1A2 genotype was most frequent in Simmental (44.8%), while A2A2 genotype was primarily found in Angus (70.6%). The predicted monoisotopic mass, molecular weight, isoelectric point (pI), net charge at pH 7, total number of atoms, instability index, and grand average of hydropathicity (GRAVY) of β-CN A2 were 23,567.23 Da, 23,582.30 Da, 5.24, −6.5, 3,357, 96.53, and −0.355, respectively. The β-CN A2 protein isoform concentration in Angus (17.8 ng/ml) was higher than in Simmental (16.7 ng/ml). Identifying β-CN A2 milk provides new insights for screening favorable alleles and establishing breeding programs based on marker-assisted selection.

Keywords: Dairy cattle; Genetic variation; Genotyping; Milk casein; β-CN A2 milk

INTRODUCTION

Cattle's milk and its by-products are nutrient-dense foods, making them an essential source of protein, vitamins, and minerals required for human health (Clegg et al., 2021; Miluchová et al., 2018; Sebastiani et al., 2020). Technical efforts to increase milk safety, quality, and chemical composition by implementing standard animal breeding practices, controlling environmental factors, and selecting favorable haplotypes are significant in the modern dairy industry (Brito et al., 2021; Caroli et al., 2009; Marchitelli et al., 2013). In ruminant milk, casein accounts for nearly 80% of its total protein and is physiochemically encoded by four casein haplotype genes, which are α-S1-CN (CSN1S1), α-S2-CN (CSN1S2), β-CN (CSN2), and κ-CN (CSN3), located in a 250-kb spanning region on bovine chromosome 6 (Caroli et al., 2009; Kolenda et al., 2021; Vanvanhossou et al., 2021). Among those haplotypes, the CSN2 gene plays an essential role in the determination of the surface properties of the casein micelles, influences milk yield, physicochemical properties, quality, composition, processability, and nutrition balance in dairy products (Cieślińska et al., 2019; Dai et al., 2016; Gai et al., 2021; St-gelais and Haché, 2005). In cattle's milk, β-CN is observed as the most polymorphic gene region and comprises 36–45% of total casein with 209 amino acid residues and a molecular weight of 23,983 kDa (Cieślińska et al., 2019; Dai et al., 2016; Fuerer et al., 2019; Yasmin et al., 2020).

In recent years, the genetic polymorphisms of the CSN2 gene have received considerable attention in the animal breeding and dairy industry due to their correlation with milk production traits, quality, and β-casomorphins (BCMs) (Čítek et al., 2019; Nuhriawangsa et al., 2021; Pedrosa et al., 2021). CSN2 gene polymorphisms in the

***Corresponding author:**
Aixiang Huang, Professor Dr. Aixiang Huang, College of Food Science and Technology, Yunnan Agricultural University, Kunming, China – 650201. **Tel.:** (+86)-13769132115. **Fax:** (+86)-0871-63649220. **E-mail:** aixianghuang@126.com

Received: 24 December 2021; **Accepted:** 13 July 2022

Bos genus have 12 known genetic variants; these are A1, A2, A3, B, C, D, E, F, G, H1, H2, and I, of which variants A1 and A2 are predominant and considered an ancestor type (Antonopoulos et al., 2021; Cieślińska et al., 2019; Gai et al., 2021; Sebastiani et al., 2020). Substitution of C→A in the single nucleotide polymorphism (SNP) site of the CSN2 gene at g.8101C>A in exon 7 in the amino acid codon 67 causes the alteration of proline to histidine (Kumar et al., 2019). In the CSN2 gene variant A1, the A-nucleotide releases His67, producing an β-CN A1 milk that has the potential to release β-casomorphin-7 (BCM-7), a bioactive peptide with opioid properties. Meanwhile, the C-nucleotide cleaves Pro67 in CSN2 gene variant A2, resulting in β-CN A2 milk with higher protein and fat concentrations as well as a higher milk yield (Brooke-Taylor et al., 2017; Demirel and Çak, 2018; Rangel et al., 2016).

Previous studies reported that the presence of BCM-7 in cattle's milk is claimed to have an association with a risk factor for several human diseases, including cattle's milk protein allergy (CMPA), ischemic heart disease, cardiovascular diseases, type 1 diabetes, atherosclerosis, sudden infant death syndrome, autistic disorders, inflammatory response, gastrointestinal discomfort reaction, and delayed gastrointestinal transit time (Park and Haenlein, 2021; Rangel et al., 2016; Ul Haq, 2020). Due to the potential adverse implications of consumption of milk containing BCM-7 on human health, increasing food safety awareness, and the significant growth of market demand for β-CN A2 milk in China, it is necessary to screen for the presence of β-CN A1 and A2 alleles of the CSN2 gene in the dairy cattle population. The objectives of this present study were to determine the frequency of β-CN A2 allele compared to A1 allele of the CSN2 gene at the SNP site of g.8101C>A in dual-purpose Simmental and Angus cattle breeds using PCR-RFLP coupled with the sequencing method, to predict the physicochemical protein properties of β-CN A2, and to measure β-CN A2 protein isoform concentration in A2 milk samples using the sandwich ELISA approach. This present study is the first investigation of the allele frequencies and detailed *in silico* physicochemical properties of β-CN A2 protein compared with β-CN A1 in Simmental and Angus cattle breeds in Yunnan province, China.

MATERIALS AND METHODS

Animal ethics approval

Experimental procedures and animal welfare in this study were carried out in compliance with the Care and Use Guide of Laboratory Animals approved by the Institutional Animal Ethical Committee (IAEC) of Yunnan Agricultural University, China (July 14, 2020).

Genomic DNA extraction

All observed animals, including Simmental (n = 201) and Angus (n = 85), were randomly selected for the genomic DNA-based analysis from Shilin dairy farms in Kunming city, which is one of the biggest centers of dairy farming in Yunnan province, southwestern China. Approximately 3–5 ml of fresh blood samples from each cattle's jugular vein were collected, stored in sterile ethylenediaminetetraacetic acid (EDTA) vacutainer tubes, and maintained at –18°C in the freezer for further analysis (Fig. 1a). The DNA was extracted from 200 µl of whole blood sample using the TIANamp blood DNA kit DP318 (Beijing Tiangen Biotechnology Co., Ltd., China), following the manufacturer's guidelines. Genomic DNA quality results from Simmental and Angus cattle breeds were qualitatively documented on 1% gel electrophoresis in 1X TAE buffer after a 30 min reaction at 100 V (Fig. 1b). The purity and concentration of the genomic DNA were quantitatively measured using the NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific Co., Ltd., USA), which indicated that the optical density ratio of all samples ranged from 1.6 to 1.9 at OD_{260}/OD_{280} nm, and the concentration varied from 100 to 150 ng/µl. DNA samples were kept at –40°C until further analysis.

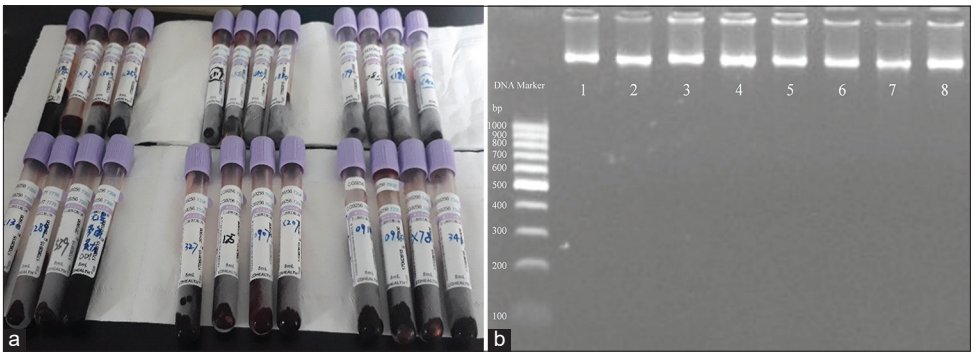

Fig 1. (a) Whole blood samples collected from Simmental and Angus cattle breeds; (b) Genomic DNA quality results from Simmental and Angus cattle breeds. Lane 0: DNA marker (1000 bp). Lane 1–4: genomic DNA of Simmental cattle and lane 5–8: genomic DNA of Angus cattle (4 µl).

Amplification protocols

A pair of forward and reverse primers were designed using Primer3 version 4.1.0 software (https://bioinfo.ut.ee/primer3/) to amplify the specific fragment of CSN2 gene at the SNP site of g.8101C>A in exon 7 (GeneBank accession number M55158); forward primer 5'-CAC CTG TGA AGA AAG TGG GTT A-3' and reverse primer 5'-ACT GTG CTG TTT AAC TTC TGA TG-3' (Table S1) (Cieślińska et al., 2019; Untergasser et al., 2012). Besides, the reference sequences of polymorphic CSN2 gene, including β-CN A1 allele (codon CAT) were accessed from GeneBank accession number JX273430.1, and β-CN A2 allele (codon CCT) were accessed from JX273429.1. The PCR amplification protocol was performed in a gradient thermal cycler A2000 (Hangzhou LongGene Scientific Instrument Co., Ltd., China) according to the following program described elsewhere by Cieślińska et al. (2019), with slight modifications: 95°C for 3 min (pre-denaturation) as the first phase, followed by 95°C for 30 sec (denaturation), 60°C for 45 sec (annealing) with 1°C stepwise temperature intervals per cycle, 72°C for 1 min (elongation), 72°C for 5 min (extension), where the number of cycles was 25 and cooling at 4°C. The contents of the 25-µL reaction mixture volume consisted of 2X Taq PCR mix (Beijing Tsingke Biological Technology Co., Ltd., China), 150 ng of DNA, 1 µl of primer forward and reverse, and free-nuclease water. The amplified products were evaluated by electrophoresis on a 2% agarose gel (w/v) at 120 V for 20 min with ethidium bromide (EtBr) staining.

Genotyping alleles polymorphism and sequencing

PCR-RFLP was used in genotyping the presence of β-CN A1 and A2 allele polymorphisms according to Antonopoulos et al. (2021), with slight modifications. A total of 6 µl amplified touchdown PCR products were digested using 1 µl NEB 2 U *DdeI* R0175S restriction enzyme (New England Biolabs Co., Ltd., UK). Furthermore, PCR products were then purified using a BigDye Xterminator purification kit (Applied Biosystems, Carlsbad, CA, USA) and further sequenced on an automatic Genetic Analyzer ABI 3730xl (Applied Biosystems, Carlsbad, CA, USA), following standard protocols described by the manufacturer.

Sequence analysis

Chromas software version 2.6.6 (http://technelysium.com.au/wp/chromas/) was used to interpret the obtained sequence data (Fig. S1) (Goodstadt and Ponting, 2001). Multiple sequence alignments were analyzed using the ClustalW tool in the BioEdit v7.2.5 software (Fig. S1) (http://www.mbio.ncsu.edu/BioEdit/bioedit.html) (Alzohairy, 2011). The amino acid sequence translated from the DNA coding sequence was analyzed using the EDIT Seq tool in the LASERGENE genomic software (DNASTAR, Inc., USA) (Burland, 2000). Polymorphic sites were determined by visual observation of the chromatogram (Fig. 2). Allele and genotype frequencies were measured using the formula described by Nei and Kumar (2000). Meanwhile, the allele and genotype frequencies and the conformity with Hardy-Weinberg Equilibrium (HWE) were calculated using GenAlEx software version 6.51b2 (https://biology-assets.anu.edu.au/GenAlEx/Download.html) running on Excel 2016 for Windows (Microsoft Corp., Redmond, USA) (Smouse et al., 2017). In this framework, the allele frequency of each cattle breed was analyzed using a χ^2 standard Pearson's goodness-of-fit chi-squared test to determine departure from HWE based on the frequency of observed and expected genotypes ($\alpha = 0.05$) (Nei and Kumar, 2000; Rousset, 2008). The null hypothesis of the right-tailed probability of the chi-squared distribution of allele and genotype frequencies is rejected when the *p*-value is lower than 0.05.

In silico analysis of β-CN protein variants A1 and A2

In this study, the physicochemical properties, amino acid compositions, and allergenicity probability of β-CN protein variants A1 and A2 protein were predicted *in silico* using publicly available web-based bioinformatics tools. Obtained sequences from previous sequencing reactions were translated into amino acid sequence proteins using the BIOPEP tool (http://www.uwm.edu.pl/biochemia/index.php/en/biopep/) and compared with the amino acid sequence database in BIOPEP (ID 1097 for β-CN A1 and 1098 for β-CN A2) and NCBI GeneBank (Minkiewicz et al., 2019). In BIOPEP, theoretical chemical mass and monoisotopic mass can also be predicted. According to Pooja et al. (2017), the protein-peptide was then inputted to ProtParam tool (https://web.expasy.org/protparam/) to obtain estimated data about the theoretical isoelectric point, amino acid composition, and molecular weight. The prediction of allergenicity was performed using the Allergen F.P. tool (https://ddg-pharmfac.net/AllergenFP/). The peptide extinction coefficient, net charge at pH 7, aliphatic index, grand average of hydropathicity, estimated solubility in water, and chemical formula, were calculated using the PepCalc tool (https://pepcalc.com/).

β-CN A2 protein isoform concentration measurement

Fresh β-CN A2 milk samples were collected in triplicate from each cattle breed carrying the A2A2 genotype. A total of 500 ml of milk samples were collected in sterilized plastic bottles, labelled, stored under 4°C, and transferred to the College of Food Science and Technology laboratory for further analysis. Milk samples were routinely analyzed for chemical analyses, including protein, fat, total solids, solids-non-fat/SNF, ash content, and pH using an automatic MilkoScan analyzer FT2 (FOSS Electric A/S Analytics, Hillerød, Denmark), according to the manufacturer's guidelines. Total protein in A2 milk samples was quantified

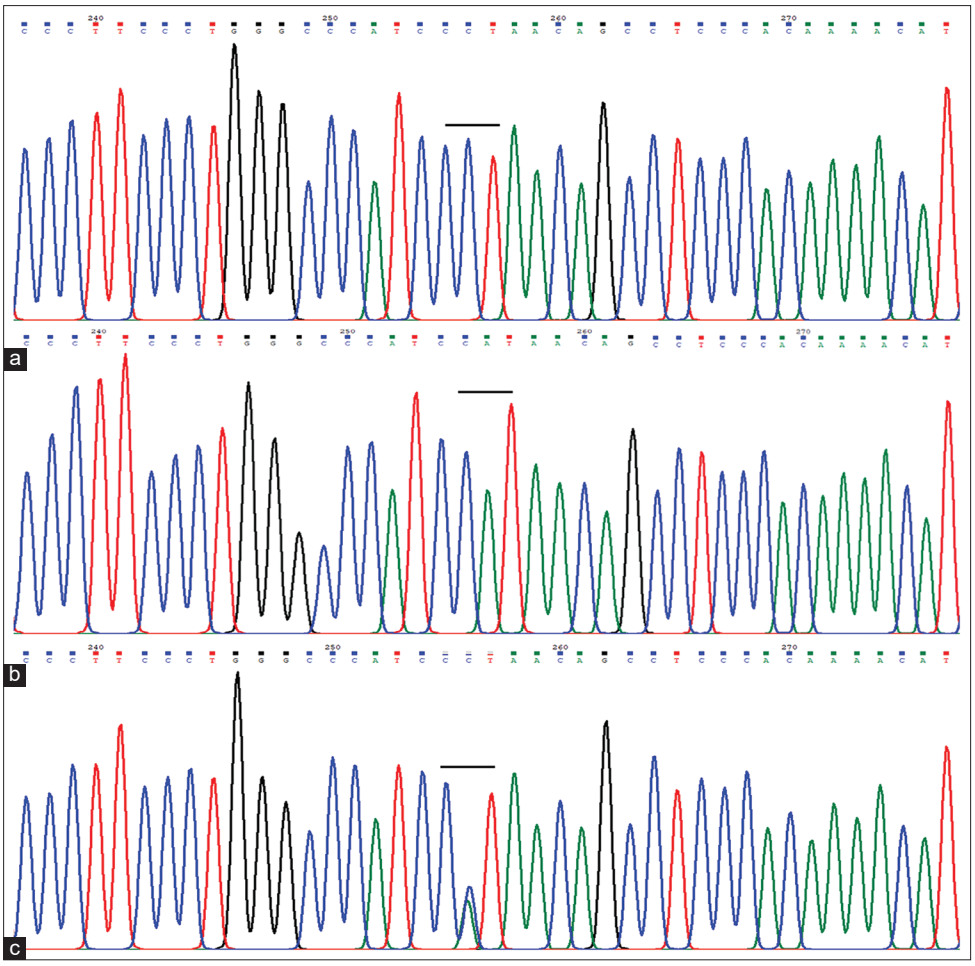

Fig 2. The sequencing results of the polymorphism of CSN2 gene at the SNP site of g.8101C>A in exon 7. Homozygous A1A1 genotype (codon CAT, responsible in producing A1 milk variant) (a), homozygous A2A2 genotype (codon CCT, responsible in producing A2 milk variant) (b), and heterozygous A1A2 genotype (codon CC/AT, responsible in producing A1 milk variant) (c).

using a bicinchoninic acid (BCA) assay kit (Beijing Solarbio Science and Technology Co. Ltd., China), following the manufacturer's guidance. The measurement of β-CN A2 protein isoform concentration in fresh A2 milk samples was directly performed using the bovine β-CN A2 sandwich ELISA kit BEK-2244-1P (Biosensis Pty. Ltd., Australia), following the manufacturer's instructions. The sandwich assay was performed using affinity-purified rabbit anti-bovine β-CN polyclonal antibodies (pAbs) pre-coated on a 96-well plate. The A2 specific chicken pAbs was used for β-CN A2 protein standards were used to determine the β-CN A2 protein concentration based on the standard curve with an assay range of 3.1 ng/ml–200 ng/ml. The intensity of color development that relies on the concentration of β-CN A2 protein isoform in milk samples was measured with a Multiskan FC microplate reader (Thermo Fisher Scientific Co., Ltd., USA) at an absorbance of 450 nm, and the optical density values of standards and samples were interpreted using SkanIt software (Thermo Fisher Scientific Co., Ltd., USA).

RESULTS AND DISCUSSION

Gene and genotypic frequencies

In this present work, we detected the CSN2 gene variants A1 and A2 caused by SNP at rs43703011A>C, g.8101C>A site in exon 7, affecting the nucleotide transversion of codon CCT (proline) by CAT (histidine) at position 67 of the amino acid sequence at the same *Bos taurus* autosome (BTA) position 6:87181619 (Ensembl transcript ARS-UCD1.2: CSN2-201) (Table 1, Fig. 3, and Fig. S3). In the Chinese Simmental and Angus cattle population, three types of genotypes, including A1A1 (a single band in 345 bp and a 121-bp outer product), A1A2 (a double band in 345 bp and 284 bp, with a 121-bp outer product), and A2A2 (a single band in 284 bp and a 121-bp outer product), were observed using PCR-RFLP (Fig. 3). To validate our PCR-RFLP assay with *DdeI*-restriction enzyme, before genotyping for all individuals within the population, we first genotyped 30 cattle with previously identified CSN2 genotypes using genomic sequencing. The PCR-RFLP genotyping results

were completely consistent with the sequencing results, indicating that the A1A1, A1A2, and A2A2 genotypes could be accurately screened. As PCR-RFLP facilitated us to determine the specific targeted allele in real time, we could genotype the CSN2 gene locus more quickly and easily compared to high-throughput methods, such as the sequencing approach. The genotyping method with blood sampling and the PCR-RFLP showed high performance for routine screening of β-CN genotypes. Thus, we concluded that the performance of the PCR-RFLP assay with *DdeI* could be used for CSN2 genotyping. Blood sampling from observed animals, commonly considered the "gold standard" method for genetic testing quality, has been a main key to genomic sequencing and studies that expand the understanding of the biomarker for A2 milk variant (de Vries et al., 2022; Goud et al., 2018). However, it requires appropriate medical equipment and skillful veterinarians.

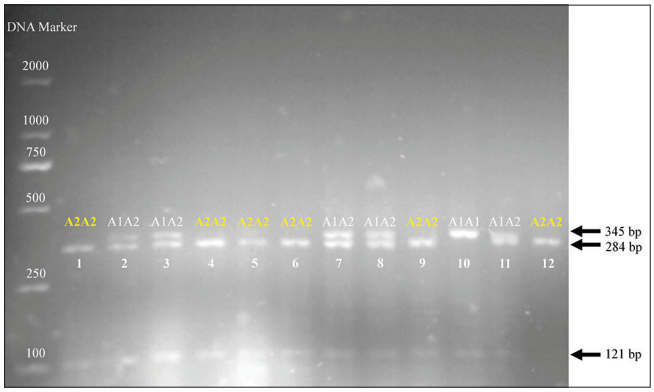

Fig 3. Representative electrophoretic profiles of PCR-RFLP products of three different genotypes (homozygous A1A1 genotype, homozygous A2A2 genotype, and homozygous A1A2 genotype). The three types of the band are associated with CSN2 allele polymorphisms. The *DdeI* restriction enzyme digested the fragment C^TNAG, demonstrated a single band in 284 bp and a 121-bp outer product, recognized it as homozygote A2A2 genotype (sample numbers 1, 4, 5, 6, 9, and 12). Conversely, when the enzyme failed to digest the amplified fragment, homozygote A1A1 genotype produced a single band in 345 bp and a 121-bp outer product (10), and heterozygote A1A2 produced a double band in 345 bp and 284 bp, with a 121-bp outer product (2, 3, 7, 8, and 11).

Herein, we have also shown that the joint application of PCR-RFLP and sequencing assays targeting the same tag can technically increase genotypic concordance. For molecular laboratories unable or unwilling to implement a high-throughput genotyping approach, the most cost-effective and robust method for precise genotyping of β-CN A1 and A2 alleles, which reduces false positives at the outlay of some false negatives, would continue to apply a *DdeI*-based assay to each DNA specimen in an observed population. The PCR-RFLP assay with various restriction enzymes has been applied in several recent β-CN A1 and A2 genotyping in dairy cattle worldwide, for example, using *DdeI* in Greek and Sri Lankan studies, *TaqI* in an Indian study, and *MspI* in an Indonesian study (Antonopoulos et al., 2021; Kumar et al., 2019; Nuhriawangsa et al., 2021; Rupasinghe et al., 2020; Sodhi et al., 2018). In this present work, we did not detect the rare CSN2 gene variants B, C, I, and L previously reported in Holstein (Yamada et al., 2021), German Yellow (Gallinat et al., 2012), Deutsches Schwarzbuntes Niederungsrind (DSN) (Meier et al., 2019), and West African taurine crossbreeds (Vanvanhossou et al., 2021), respectively. According to allele frequency analysis, the β-CN A2 rs43703011C allele, was more dominant than the A1 rs43703011A allele in both Simmental and Angus cattle breeds, which were 0.642 and 0.835, respectively. The β-CN A1A2 genotype was most prevalent in Simmental, compared to A2A2 and A1A1 genotypes, which were 0.448, 0.418, and 0.134, respectively. In contrast, in Angus, the maximum β-CN genotype frequency was revealed by A2A2 (0.706), followed by A1A2 (0.259) and A1A1 (0.035).

Table 1: Description (gene, variant, exon, BTA: bp, SNP ID, and protein position) of the four milk casein genes (CSN1S1, CSN1S2, CSN2, CSN3) causing amino acid sequence polymorphisms.

Gene	Protein	Variant	Exon	BTA: bp	Reference SNP ID	Allele	Amino acid	Protein sequence position
CSN1S1	α-S1	CSN1S1B		6:87157262	rs43703010	A/G	Glu	207 (192)
CSN1S2	α-S2	CSN1S2A		6:87266177	rs441966828	C/T	Ser	23 (8)
CSN2	β	CSN2A1		6:87181619	rs43703011	T/G	His	82 (67)
		CSN2A2		6:87181619	rs43703011	T/G	Pro	82 (67)
		CSN2I		6:87181542	rs109299401	T/G	Leu	108 (93)
		CSN3A		6:87390576	rs43703015	T/C	Thr	157 (136)
				6:87390612	rs43703016	C/A	Ala	169 (148)
				6:87390632	rs43703017	A/G	Ser	176 (155)
CSN3	κ	CSN3B		6:87390576	rs43703015	T/C	Ile	157 (136)
				6:87390612	rs43703016	C/A	Asp	169 (148)
		CSN3E		6:87390632	rs43703017	A/G	Gly	176 (155)

Bovine milk protein variants were identified following previous nomenclature (Ahmed *et al*., 2016; Caroli *et al*., 2009; Tetens *et al*., 2014). The presence of *CSN2* gene variants A1 (*CSN2A1*) and A2 (*CSN2A2*) observed in this present study is due to an SNP at the same BTA position of 6:87181619. The amino acid position was observed based on the reference protein sequence obtained from Ensembl genome database release v93 ARS-UCD1.2 UMD3.1 (GCA_002263795.2) assembly (http://www.ensembl.org/Bos_taurus/Info/Index), accessed on December 14, 2021. BTA CSN1S2A, CSN1S1B, CSN2A2, and CSN3 were accessed from Ensembl ID: ENSBTAG00000007695, ENSBTAG00000005005, ENSBTAG00000002632, and ENSBTAG00000039787, respectively. Mature protein positions are presented in parentheses and according to the protein sequence reference accessed from Ensembl.

The χ^2 values of the observed cattle population were significantly smaller than the probability value of 5.991 (d.f. = 2; $p<0.05$) and rejected the null hypotheses, showing that the allele frequencies of both cattle breeds fit with the HWE proportion (Table 2). The results of right-tailed probability analysis of a chi-squared distribution analysis exhibited *p*-values for both cattle breeds that were higher than $\alpha = 0.05$ (0.934 for Simmental and 0.861 for Angus), indicating that observed allele frequencies are significantly not associated with expected ones.

Allele frequencies revealed in this investigation were similar to the allele frequencies observed among the Simmental, Angus, Holstein-Friesian, and other minor breeds in other countries, such as Germany, Denmark, Croatia, Indonesia, Greece, Japan, Italy, China, Turkey, Iran, Poland, and the Netherlands. However, allele frequency investigations in different places showed different results, having been reported that the allele frequency of β-CN A1 in Angus ($n = 77$) in the USA was higher than the A2, and both alleles had the same frequencies in Holestein-Friesian ($n = 119$) in other Iranian studies (Table 3). The differences in β-CN A1 and A2 allele frequencies are likely due to crossbreeding programs and technical farming practices. The variation in SNPs provides extensive information about milk composition's phenotypic characteristics and generates a reference for designing sustainable genetic selection. It was also reported that the existence of β-CN A2 allele in cattle and buffalo milk indicates a higher chemical quality, in particular triglycerides and reduced serum cholesterol (Oleński et al., 2012). In addition, Miluchová et al. (2014) reported that fat and protein concentrations in milk

Table 2: Genotype and allele frequencies in CSN_2 gene at SNP site of g. 8101C>A in exon 7 in two dairy cattle breeds.

Breed	*n*	Genotype frequencies			Allele frequencies		χ^2	*p*-value *(d.f. = 2)*
		A1A1	A2A2	A1A2	A1	A2		
Simmental	201	0.134	0.418	0.448	0.358	0.642	0.138	0.934
Angus	85	0.035	0.706	0.259	0.165	0.835	0.299	0.861

n: total number of genotyped animals; d.f.: degree of freedom. A χ^2 standard Pearson's goodness-of-fit Chi-squared test was used to determine departure of the genotype and allele frequency of each cattle breed from HWE. The *P* value of each cattle breed in degree of freedom of 2 was higher than 0.05, indicating than null hypothesis of the right-tailed probability of the Chi-squared distribution of allele and genotype frequencies is fail to reject.

Table 3: Prevalence of CSN2 gene variants A1 and A2 in Simmental, Angus, Holstein-Friesian, and several minor cattle breeds worldwide (data sorted by year).

Breed	Country	Allele Frequency		*n*	Reference
		A1	A2		
Simmental	China	0.360	0.640	201	Present data
	Germany	0.145	0.446	217	Meier et al., 2019
	Denmark	0.327	0.673	2,626	De Noni et al., 2009
	Croatia	0.370	0.630	621	De Noni et al., 2009
Angus	China	0.167	0.833	85	Present data
	Germany	0.032	0.644	276	Meier et al., 2019
	USA	0.950	0.050	77	De Noni et al., 2009
Holstein-Friesian	Indonesia	0.084	0.916	70	Nuhriawangsa et al., 2021
	Greece	0.256	0.744	780	Antonopoulos et al., 2021
	Japan	0.320	0.640	311	Yamada et al., 2021
	Italy	0.371	0.546	1,226	Massella et al., 2017
	China	0.432	0.495	133	Dai et al., 2016
	Iran	0.500	0.500	119	Gholami et al., 2016
	Turkey	0.266	0.614	415	Gustavsson et al., 2014
	Turkey	0.485	0.456	49	Dinc et al., 2013
	Poland	0.350	0.650	650	Olénski et al., 2012
	Netherlands	0.280	0.500	1,929	Heck et al., 2009
Dengchuan	China	0.318	0.682	30	Unpublished data
Kermanshah	Iran	0.382	0.618	131	Rahimi et al., 2015
Karan-Fries	India	0.175	0.825	100	Jaiswal, 2012
Ladakhi	India	0.100	0.90	85	Sodhi et al., 2018
Frieswal crossbred	India	0.432	0.568	429	Kumar et al., 2019
Butana	Sudan	0.170	0.610	50	Ahmed et al., 2016
Vrachykeratiki	Greece	0.304	0.587	46	Antonopoulos et al., 2021
Katerinis	Greece	0	0.100	20	Antonopoulos et al., 2021
Sykias	Greece	0	0.100	20	Antonopoulos et al., 2021
Lagune	Germany	0.800	0.200	20	Vanvanhossou et al., 2021
Somba	Germany	0.630	0.370	20	Vanvanhossou et al., 2021

Other variants of the CSN2 gene are not included in the table.

containing homozygous β-CN A2A2 genotype were higher than those in milk containing A1A1 genotype. In the same report, β-CN A2 milk variant was observed to have the highest content of lactose, methionine, and proline, while β-CN A1 milk variant containing homozygote A1A1 genotype or heterozygote A1A2 genotype had the highest aggregate of cyclic adenosine monophosphate (cAMP), citric acid, glycine, and choline. Previous studies have also reported that β-CN A2 milk variant has received great interest in terms of human health concerns due to its potential to stimulate better digestion efficiency and promote antibacterial activity (Caroli et al., 2009; Deth et al., 2015; Vanvanhossou et al., 2021).

In silico physicochemical properties of β-CN protein variants A1 and A2

Various physicochemical properties, including monoisotopic mass, molecular weight, theoretical pI, amino acid composition, and other parameters of β-CN A1 with chemical formula $C_{1081}H_{1693}N_{271}O_{309}S_6$ and β-CN A2 with chemical formula $C_{1080}H_{1693}N_{269}O_{309}S_6$ were *in silico* analyzed based on their molecular protein characteristics. Monoisotopic mass, molecular weight, theoretical pI, net charge at pH 7, total number of atoms, computed instability index, and GRAVY of β-CN A1 were demonstrated to be higher values than β-CN A2, which were 23,607.24 Da, 23,622.32 Da, 5.36, –6.4, 3,360, 98.39, –0.362 for β-CN A1, and 23,567.23 Da, 23,582.30 Da, 5.24, –6.5, 3,357, 96.53, –0.355 for β-CN A2, respectively (Table 4). The results were in accordance with Darewicz et al. (2007) earlier reported that the pI value of β-CN A1 (4.90) was higher than β-CN A2 (4.76), which causes β-CN A2 to have less exposed hydrophobicity at pH 6.5–6.7, be more soluble than β-CN A1, and the maximum surface load associated with its emulsion formation stability is also lower compared to β-CN A1 (Jensen et al., 2012; Raynes et al., 2015). The pI is described as the pH value at which the protein carries no net charge or the sum of positively and negatively charged is equivalent (Saraswathy and Ramalingam, 2011).

According to present analyses using ProtParam and PepCalc tools, the pI value was observed to be less than 7 and pH was greater than pI for both β-CN protein variants, which indicated an excessive amount of OH^- in solution, captivated the positively charged amine group, causing an H^+ removal. The ProtParam tool was also used to evaluate amino acid composition, which showed that the percentage of amino acids comprised β-CN A1 and A2 was dissimilar, particularly in histidine (2.9% in β-CN A1 and 2.4% in β-CN A2) and proline (16.3% in β-CN A1 and 16.7% in β-CN A2) (Fig. 4). In the case of charged amino acid residues, the number of positively charged residues in both variants was slightly greater in comparison with negatively charged residues. Kaur et al. (2020) reported that the number of positively and negatively charged amino acid residues and pI would influence solubility, interactions, and subcellular localization of proteins. The amino acid contents of the secondary structure elements were influenced by flanking regions in the primary sequence (Khrustalev et al. 2014; 2018). Subsequently, the SWISS-MODEL tool was used to generate predicted 3D homology-modelling of the protein secondary structures of β-CN A1 and A2, while

Table 4: *In silico* physicochemical properties of β-CN protein variants A1 and A2.

Physicochemical parameter	β-CN variant	
	A1	A2
Monoisotopic mass (Da)	23,607.24	23,567.23
Molecular weight (Da)	23,622.32	23,582.30
Theoretical isoelectric point	5.36	5.24
Amino acids composition		
Ala (A)	5 (2.4%)	5 (2.4%)
Arg (R)	4 (1.9%)	4 (1.9%)
Asn (N)	5 (2.4%)	5 (2.4%)
Asp (D)	4 (1.9%)	4 (1.9%)
Cys (C)	0 (0.0%)	0 (0.0%)
Gln (Q)	21 (10.0%)	21 (10.0%)
Glu (E)	18 (8.6%)	18 (8.6%)
Gly (G)	5 (2.4%)	5 (2.4%)
His (H)	6 (2.9%)	5 (2.4%)
Ile (I)	10 (4.8%)	10 (4.8%)
Leu (L)	22 (10.5%)	22 (10.5%)
Lys (K)	11 (5.3%)	11 (5.3%)
Met (M)	6 (2.9%)	6 (2.9%)
Phe (F)	9 (4.3%)	9 (4.3%)
Pro (P)	34 (16.3%)	35 (16.7%)
Ser (S)	16 (7.7%)	16 (7.7%)
Thr (T)	9 (4.3%)	9 (4.3%)
Trp (W)	1 (0.5%)	1 (0.5%)
Tyr (Y)	4 (1.9%)	4 (1.9%)
Val (V)	19 (9.1%)	19 (9.1%)
Pyl (O)	0 (0.0%)	0 (0.0%)
Sec (U)	0 (0.0%)	0 (0.0%)
Extinction coefficient	10,810 $M^{-1}cm^{-1}$	10,810 $M^{-1}cm^{-1}$
Net charge at pH 7	–6.4	–6.5
Estimated solubility	Poor water solubility	Poor water solubility
Total number of atoms	3,360	3,357
Computed instability index	98.39	96.53
Aliphatic index	88.47	88.47
Grand average of hydropathicity	–0.362	–0.355
Formula	$C_{1081}H_{1693}N_{271}O_{309}S_6$	$C_{1080}H_{1693}N_{269}O_{309}S_6$
Allergenicity probability	A probable allergen with the highest Tanimoto similarity index 0.93	A probable non-allergen with the highest Tanimoto similarity index 0.82

their conformational parameters, including the occurrence percentage of alpha helix (hh), extended strand (ee), beta turn (tt), and random coil (cc), were calculated using the SOPMA server tool (Fig. S2) (Combet et al., 2000; Studer et al., 2020; Waterhouse et al., 2018). The results demonstrated that random coil (60.77%) dominated among secondary structure parameters, followed by alpha helix (32.54%), extended strand (5.26%), and beta turn (1.44%), while 3_{10} helix, Pi helix, beta bridge, bend region, and ambiguous states were not detected herein. Homology modeling is the appropriate approach for revealing 3D protein coordinates with the integration of a web-based SWISS-MODEL modeling system.

Further, the computed instability index suggests the stability of proteins in both *in vivo* and *in vitro* conditions, where proteins whose instability index value is lower than 40 are referred to as stable and those with an instability index value of more than 40 are considered unstable (Gamage et al., 2019; Lugani and Sooch, 2017; Sahay et al., 2020). In this study, the instability index values for both β-CN variants were observed more than 40, which the instability index value of β-CN A1 (98.39) was higher than that of β-CN A2 (96.53), indicating those variants were unstable under a wide range of temperature changes and conditions. Other than the instability index, the aliphatic index is also another parameter used to determine the functional structural stability of proteins and a key factor for the increase in globular protein thermostability. Both β-CN variants observed in this analysis were estimated to have poor water solubility with the same extinction coefficient and aliphatic index, which were 10,810 $M^{-1}cm^{-1}$ and 88.47, respectively (Table 4). The aliphatic index may be described as the relative volume occupied by the aliphatic side chains of amino acids in their structure, including A (alanine), I (isoleucine), L (leucine), and V (valine), where valine and alanine are recognized as the aliphatic amino acids (Behbahani et al., 2020; Bhattacharjee et al., 2018; Elekofehinti et al., 2019). Ikai (1980) first proposed the aliphatic index and then evaluated the correlation between the aliphatic index and protein thermostability based on the average percentage of aliphatic amino acids.

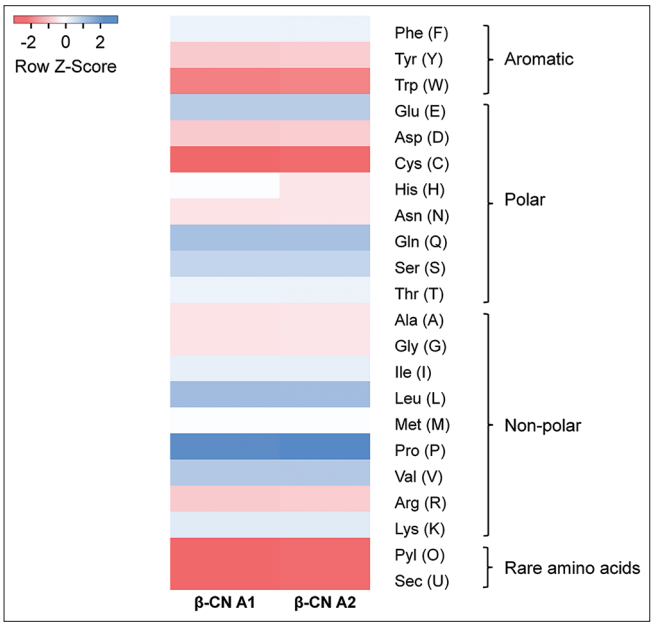

Fig 4. Heat map representation in Z-score demonstrating the comparison of amino acid composition of β-CN A2 compared with β-CN A1.

Aside from predicting protein molecular weight, composition, allergenicity probability, and stability, the hydrophobicity or hydrophilicity degree of β-CN A1 and A2 was also calculated through the GRAVY approach using the PepCalc tool. The calculation of GRAVY index is an effective method to measure the interaction properties between an amino acid's side chain and water. The GRAVY index, first proposed by Kyte and Doolittle (1982), is measured by dividing the total hydropathy values of all amino acids that appear in the observed protein with the total number of protein residues. Its score ranges between –2 to +2, where a negative score indicates the degree of hydrophilicity with good solubility in nature and a positive score means the degree of hydrophobicity, vice-versa (Di Rienzo et al., 2021). The calculation results predicted that β-CN A1 (–0.362) had a higher GRAVY index as compared with β-CN A2 (–0.355), suggesting that those protein variants were more hydrophilic in nature, globular, and difficult to identify on 2-D gels (Table 4).

According to the allergenicity probability results predicted using the AllergenFP v.1.0 tool, β-CN A1 showed potential as a probable allergenic, as it exhibited a high Tanimoto structural similarity index of 0.93 and a probable non-allergenic with the highest Tanimoto similarity index of 0.82 for β-CN A2. AllergenFP is an online free-access alignment workspace for predicting the allergenicity probability of proteins by converting amino acid main properties, including size, abundance, hydrophobicity, and the behavior of hydrogen bonding, into fingerprints that are occupied to estimate allergenicity probability through Tanimoto structural similarity matches with the allergenicity profile of previously observed proteins reported in the literature (Ji et al., 2019). Remarkably, based on research over the last 25 years, β-CN A1 protein in cattle's milk is reported to be harder to digest for some people compared to β-CN A2, which has been hypothesized that β-CN A1 is the main reason for milk protein allergies or intolerances (Park and Haenlein, 2021; Rashidinejad et al., 2017).

Digestion of β-CN A1 milk that releases BCM-7 peptide stimulates μ-opioid receptors, which can trigger allergic

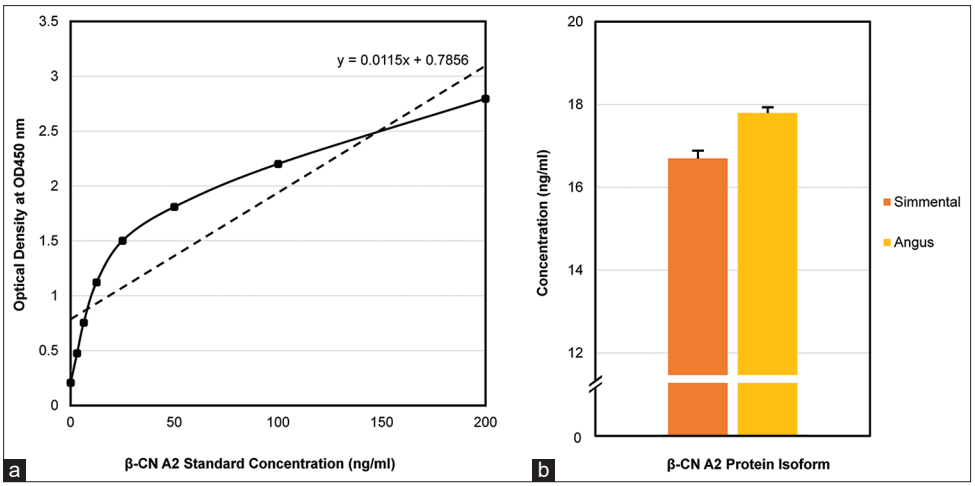

Fig 5. Standard curve of β-CN A2 at OD_{450} (a) and comparison of the average concentration of β-CN A2 protein isoform in Simmental and Angus cow's milk samples (b).

reactions or intolerance symptoms in the human body and cause potential food safety concerns, particularly in children (Jianqin et al., 2015; Mayer et al., 2021; Pal et al., 2015; Rangel et al., 2016; Suchy et al., 2010; Ul Haq et al., 2014a; Ul Haq et al., 2014b; Ul Haq, 2020; Woodford, 2009). Yadav et al. (2020) demonstrated the correlation between cattle's milk β-CN A1 and A2 feeding in mice and allergic airway disease symptoms. Mice fed with milk containing the A1A1 genotype showed elevated airway hyperresponsiveness with a concentration growth of bronchoconstrictor, which was absent in mice fed with milk containing the A2A2 genotype. The increase of interleukin 4 and 5 levels and the development of IgE and IgG levels, leading to peribronchial inflammation, were also observed in mice fed milk containing the A1A1 genotype, while an intermediate phenotype was present in mice fed milk containing the A2A2 genotype. These studies, together with the *in-silico* analysis results about allergenicity prediction presented in this present study, support the increasing public concerns about potential allergenicity caused by the consumption of cattle's milk containing β-CN A1 and its by-products. Further food safety studies on the potential milk allergenicity of β-CN A1 and A2 using animal trials are needed to establish the reproducibility of our findings reported herein.

β-CN A2 protein isoform concentration analysis in milk

The ability to measure β-CN protein, especially for A2 isoform concentration, is critical to the dairy industry. Apart from reducing the negative effects of non-β-CN A2 milk consumption on human gastrointestinal function, using β-CN A2 milk as a raw material in the production of dairy by-products would increase both product quality and technological properties. In an attempt to *in-vitro* quantify β-CN A2 protein isoform concentration in order to obtain insights into its bioactive potential, the sandwich ELISA assay at optical density OD_{450} wavelength was effectively performed in this study and did not cross-react with β-CN A1 protein. The detection using the sandwich ELISA assay allows a more direct reference of the presence of protein residues (Koppelman et al., 2021). In this assay, β-CN A2 protein isoform can be determined by the colorimetric reaction after the antigen binding with the particular enzyme-labeled antibody. The sandwich ELISA assay starts with a capture specific antibody that has been thoroughly coated onto the wells of a 96-well plate. The β-CN A2 from fresh cattle's milk as an antigen of interest was then added to the wells to bind the antibody. The primary detection antibody was added to the wells, followed by the secondary enzyme-conjugated antibody for signal amplification. After incubation was completed, the substrate developed a color change and was interpreted using a microplate reader (Alhajj and Farhana, 2022; Engvall, 2010; Kohl and Ascoli, 2017; Shah and Maghsoudlou, 2016). Thus, ELISA is specifically suitable for the quantification of β-CN A2 protein isoform independent of β-CN A1. The β-CN A2 standard curve at OD_{450}, which refers to typical optical densities in eight diluted levels of β-CN A2 concentration standard (Table S2), and the average concentration of β-CN A2 protein isoform are presented in Fig. 5.

The results indicated that the average concentration of β-CN A2 protein isoform in Angus cattle's milk was higher than that of Simmental, which was 17.8 ng/ml and 16.7 ng/ml, respectively. The optical density values of the samples detected from a sandwich ELISA assay are related to the positive interaction between β-CN A2 protein isoform and the ELISA's enzyme-labeled antibodies, as well as the relative comparison of this interaction to the protein concentration in the standard curve. Other studies have found that the sandwich ELISA assay is the most reliable test for IgE-mediated allergenic substances in food, such as

casein, albumin, globulin, gliadin, gluten, and clam proteins, with high precision, sensitivity, and relative ease of use (Bernard et al., 2021; Bosman et al., 2021; Koppelman et al., 2021; Madrid et al., 2021). Herein, the higher concentration of β-CN A2 protein isoform in Angus is likely linked with the larger plasma concentration of proline residue present in milk.

Changes in cattle's milk protein content have been associated with variations in nutrition, feeding system management, breeds, and β-CN variants (Fu et al., 2021; Franzoi et al., 2019; Amalfitano et al., 2019). An additional proline residue present in the β-CN A2 protein has been observed to support the increase in propensity for poly-L-proline helix type 2 (PPII) secondary structure formation, which is important in cellular signaling, transforming protein structural dynamics and self-assembly behavior, and might contribute to its isoform concentration (Brown and Zondlo, 2012; Raynes et al., 2015). Interestingly, Lv et al. (2020) found that there were relatively abundant metabolites in β-CN A2 variant in cattle's milk, the majority including proline, methionine, α-lactose, and acetoacetic acid, detected using nuclear magnetic resonance spectroscopy (NMR) and liquid-chromatography mass spectrometry (LC-MS)-based metabolomics approaches.

CONCLUSIONS

The DNA-based genotyping using the PCR-RFLP technique coupled with the sequencing method used in this study demonstrated rapid and promising performance in evaluating nucleotide transversion of specific-allele polymorphism in the CSN2 gene at SNP site of g.8101C>A, particularly for genotyping β-CN A1 and A2 alleles. The most frequent β-CN allele among the Simmental and Angus cattle breeds was A2, with over 64% prevalence. In Simmental, the A1A2 genotype was the most dominant, while the A2A2 genotype was mostly found in Angus. According to *in silico* physicochemical properties prediction, monoisotopic mass, molecular weight, theoretical pI, net charge at pH 7, total number of atoms, instability index, and GRAVY of β-CN A2 showed lower values in comparison with A1. Moreover, the sandwich ELISA method for quantifying β-CN A2 protein isoform concentration was effective, specific, sensitive, and had the potential to provide the dairy industry with a useful assay to double-check the robustness of DNA-based genotyping through the evaluation of fresh cattle's milk. These findings are useful in monitoring favorable allele frequency and controlling the presence of undesired non-A2A2 genotypes in order to produce a high-quality β-CN A2 milk variant. However, the effects of β-CN A2 milk variant on human health are still unclear and require further investigation through food safety studies, quantitative label-free proteomics, and metabolomics approaches.

Acknowledgments

We would like to express our heartfelt appreciation to the anonymous reviewers and colleagues for their constructive remarks and advices that helped enhance the manuscript's content quality. Financial support for the research was provided by the Leading Talent Grant Program of Yunling Industrial Technology (No. 2014-1782) and Dairy Processing and Food Safety Program of Yunnan Modern Agricultural Cattle Milk Industrial Technology System (No. 2019KJTX0014). We also gratefully thank Professor Huaming Mao in the Laboratory of Animal Genetics and Biotechnology, College of Animal Science and Technology, Yunnan Agricultural University and Beijing Tsingke Biological Technology Co., Ltd. Branch Kunming, China for experimental support.

Authors' contributions

Aixiang Huang and Xianwei Shi designed, developed, financed, supervised the experiment, and approved the final draft. Adhita Sri Prabakusuma performed the experiments, analyzed the experimental data, drafted the final report, interpretation, documentation, prepared tables and/or figures, and authored drafts of the manuscripts. Hamzah Aleryani performed the experiments, analyzed the experimental data, and reviewed drafts of the manuscripts. Xiaoyan Kong prepared samples, reagents, equipment, and approved the final draft.

Conflict of interest statement

The authors declare there is no potential conflict of interest in this research.

Supplementary information

The online version contains supplementary material for this article (Fig. S1, S2, S3, Table S1, and S2) can be accessed at https://doi.org/10.6084/m9.figshare.17297786.v2.

REFERENCES

Alhajj, M. and A. Farhana. 2022. Enzyme Linked Immunosorbent Assay. Treasure Island, FL: StatPearls Publishing.

Alzohairy, A. M. 2011. BioEdit: An important software for molecular biology. Gerf. Bull. Biosci. 2: 60-61.

Amalfitano, N., C. Cipolat-Gotet, A. Cecchinato, M. Malacarne, A. Summer and G. Bittante. 2019. Milk protein fractions strongly affect the patterns of coagulation, curd firming, and syneresis. J. Dairy. Sci. 102(4): 2903-2917.

Antonopoulos, D., D. Vougiouklaki, G. P. Laliotis, T. Tsironi, I. Valasi, A. Chatzilazarou, P. Halvatsiotis and D. Houhoula. 2021 Identification of polymorphisms of the CSN2 gene encoding β-casein in Greek local breeds of cattle. Vet. Sci. 8: 257.

Behbahani, M., P. Rabiei and H. Mohabatkar. 2020. A comparative

analysis of allergen proteins between plants and animals using several computational tools and Chou's PseAAC concept. Int. Arch. Allergy Immunol. 181: 813-821.

Bernard, H., S. Hazebrouck, N. Gaiani and K. Adel-Patient. 2021. Allergen risk assessment for specific allergy to small ruminant's milk: development of sensitive immunoassays to detect goat's and sheep's milk contaminations in dairy food matrices. Front. Allergy. 2: 733875.

Bhattacharya, M., A. Hota, A. Kar, D. S. Chini, R. C. Malick, B. C. Patra and B. K. Das. 2018. *In silico* structural and functional modelling of antifreeze protein (AFP) sequences of ocean pout (*Zoarces americanus*, Bloch and Schneider 1801). J. Genet. Eng. Biotechnol. 16: 721-730.

Bosman, G. P., S. Oliveira, P. J. Simons, T. J. Sastre, G. W. Somsen, L. M. J. Knippels, R. Haselberg, R. J. Pieters, J. Garssen and K. Knipping. 2021. Limited lactosylation of beta-lactoglobulin from cow's milk exerts strong influence on antigenicity and degranulation of mast cells. Nutrients. 13: 2041.

Brito, L. F., N. Bedere, F. Douhard, H. R. Oliveira, M. Arnal, F. Peñagaricano, A. P. Schinckel, C. F. Baes and F. Miglior. 2021. Review: Genetic selection of high-yielding dairy cattle toward sustainable farming systems in a rapidly changing world. Animal. 2021: 100292.

Brooke-Taylor, S., K. Dwyer, K. Woodford and N. Kost. 2017. Systematic review of the gastrointestinal effects of A1 compared with A2 β-casein. Adv. Nutr. 8: 739-748.

Brown, A. M. and N. J. Zondlo. 2012. A Propensity scale for Type II polyproline helices (PPII): Aromatic amino acids in proline-rich sequences strongly disfavor PPII due to proline-aromatic interactions. Biochemistry. 51: 5041-5051.

Burland, T. G. 2000. DNASTAR's lasergene sequence analysis software. Methods Mol. Biol. 132: 71-91.

Caroli, A. M., S. Chessa and G. J. Erhardt. 2009. Invited review: Milk protein polymorphisms in cattle: Effect on animal breeding and human nutrition. J. Dairy. Sci. 92: 5335-5352.

Cieślińska, A., E. Fiedorowicz, G. Zwierzchowski and N. Kordulewska. 2019. Genetic polymorphism of β-casein gene in polish red cattle-preliminary study of A1 and A2 frequency in genetic conservation herd. Animals. 9: 377.

Cieślińska, A., E. Kostyra, H. Kostyra, K. Oleński, E. Fiedorowicz and S. Kamiński. 2012. Milk from cows of different β-casein genotypes as a source of β-casomorphin-7. Int. J. Food Sci. Nutr. 63: 426-430.

Čítek, J., L. Hanusová, L. Lískovcová, E. Samková, O. Hanuš, L. Hasoňová, Z. Křížová and L. Večerek. 2019. Polymorphisms in CSN3, CSN2 and LGB genes and their relation to milk production in dairy cattle in the Czech Republic. Acta. Univ. Agric. Silvic. Mendelianae. Brun. 67: 19-24.

Clegg, M. E., R. A. Tarrado, R. Reynolds, K. Kliem and S. Stergiadis. 2021. A comparative assessment of the nutritional composition of dairy and plant-based dairy alternatives available for sale in the UK and the implications for consumers' dietary intakes. Food Res. Int. 148: 110586.

Combet, C., C. Blanchet, C. Geourjon and G. Deléage. 2000. NPS@: Network protein sequence analysis. Trends. Biochem. Sci. 25: 147-150.

Dai, R., Y. Fang, W. Zhao, S. Liu, J. Ding, K. Xu, L. Yang, C. He, F. Ding and H. Meng. 2016. Identification of alleles and genotypes of beta-casein with DNA sequencing analysis in Chinese Holstein cow. J. Dairy Res. 83: 312-316.

Darewicz, M. and J. Dziuba. 2007. Formation and stabilization of emulsion with A1, A2 and B β-casein genetic variants. Eur. Food Res. Technol. 226: 147-152.

de Vries, J. H., D. Kling, A. Vidaki, P. Arp, V. Kalamara, M. M. P. Verbiest, D. Piniewska-Róg, T. J. Parsons, A. G. Uitterlinden and M. Kayser. 2022. Impact of SNP microarray analysis of compromised DNA on kinship classification success in the context of investigative genetic genealogy. Forensic Sci. Int. Genet. 56: 102625.

Demirel, A. F. and B. Çak. 2018. Discussions of effect A1 and A2 milk beta-casein gene on health. Approach. Poult. Dairy Vet. Sci. 3: 1-6.Deth, R., A. Clarke, J. Ni and M. Trivedi. 2015. Clinical evaluation of glutathione concentrations after consumption of milk containing different subtypes of β-casein: Results from a randomized, cross-over clinical trial. Nutr. J. 15: 1-6.

Di Rienzo, L., M. Miotto, L. Bò, G. Ruocco, D. Raimondo and E. Milanetti. 2021. Characterizing hydropathy of amino acid side chain in a protein environment by investigating the structural changes of water molecules network. Front. Mol. Biosci. 8: 626837.

Elekofehinti, O. O., Y. V. Aladenika, Y. R. Alli-Smith, O. C. Ejelonu and A. O. Lawal. 2019. Molecular modeling, dynamics simulation and characterization of human inositol hexakisphosphate kinase 1 (IP6K1) related to diabetes. J. Appl. Sci. Environ. Manage. 23: 461-467.

Engvall, E. 2010. The ELISA, enzyme-linked immunosorbent assay. Clin. Chem. 56: 319-20.

Franzoi, M., G. Niero, G. Visentin, M. Penasa, M. Cassandro and M. De Marchi. 2019. Variation of detailed protein composition of cow milk predicted from a large database of mid-infrared spectra. Animals. 9: 176.

Fu, L., L. Zhang, L. Liu, H. Yang, P. Zhou, F. Song, G. Dong, J. Chen, G. Wang and X. Dong. 2021. Effect of heat stress on bovine mammary cellular metabolites and gene transcription related to amino acid metabolism, amino acid transportation and mammalian target of rapamycin (mTOR) signaling. Animals. 11: 3153.

Fuerer, C., R. Jenni, L. Cardinaux, F. Andetsion, S. Wagnière, J. Moulin and M. Affolter. 2019. Protein fingerprinting and quantification of β-casein variants by ultraperformance liquid chromatography-high-resolution mass spectrometry. 2019. J. Dairy Sci. 103: 1193-1207.

Gai, N., T. Uniacke-Lowe, J. O'Regan, H. Faulkner and A. L. Kelly. 2021. Effect of protein genotypes on physicochemical properties and protein functionality of bovine milk: A review. Foods. 10: 2409.

Gallinat, J. L., S. Qanbari, C. Drögemüller, E. C. G. Pimentel, G. Thaller and J. Tetens. 2012. DNA-based identification of novel bovine casein gene variants. J. Dairy Sci. 96: 699-709.

Gamage, D. G., A. Gunaratne, G. R. Periyannan and T. G. Russell. 2019. Applicability of instability index for *in vitro* protein stability prediction. Protein. Pept. Lett. 26: 339-347.

Goodstadt, L. and C. P. Ponting. 2001. CHROMA: Consensus-based colouring of multiple alignments for publication. Bioinform. Appl. Note. 17: 845-846.

Goud, T. S., R. C. Upadhyay, A. Kumar, S. Karri, R. Choudhary, S. Ashraf, S. V. Singh, O. S. Kumar and C. Kiranmai. 2018. Novel extraction of high-quality genomic DNA from frozen bovine blood samples by using detergent method. Open. Vet. J. 8: 415-422.

Jensen, H. B., J. W. Holland, N. A. Poulsen and L. B. Larsen. 2012. Milk protein genetic variants and isoforms identified in bovine milk representing extremes in coagulation properties. J. Dairy Sci. 95: 2891-2903.

Ji, D., C. C. Udenigwe and D. Agyei. 2019. Antioxidant peptides encrypted in flaxseed proteome: An in silico assessment. Food Sci. Hum. Wellness. 8: 306-314.

Jianqin, S., X. Leiming, X. Lu, G. W. Yelland, J. Ni and A. J. Clarke. 2015. Effects of milk containing only A2 beta casein

versus milk containing both A1 and A2 beta casein proteins on gastrointestinal physiology, symptoms of discomfort, and cognitive behavior of people with self-reported intolerance to traditional cows' milk. Nutr. J. 15: 35.

Kaur, A., P. K. Pati, A. M. Pati and A. K. Nagpal. 2020. Physico-chemical characterization and topological analysis of pathogenesis-related proteins from Arabidopsis thaliana and *Oryza sativa* using *in-silico* approaches. PLoS One. 15: e0239836.

Khrustalev, V. V., E. V. Barkovsky and T. A. Khrustaleva. 2014. The influence of flanking secondary structures on amino acid content and typical lengths of 3/10 helices. Int. J. Proteomics. 2014: 360230.

Kohl, T. O. and C. A. Ascoli. 2017. Immunometric double-antibody sandwich enzyme-linked immunosorbent assay. Cold. Spring. Harb. Protoc. 6: 093724.Kolenda, M. and B. Sitkowska. 2021. The polymorphism in various milk protein genes in Polish Holstein-friesian dairy cattle. *Animals*. 11: 389.

Koppelman, S. J., A. L. Lardizabal, L. Niemann, J. L. Baumert and S. L. Taylor. 2021. Development of a sandwich enzyme-linked immunosorbent assay for detection and quantification of clam residues in food products. Biomed. Res. Int. 2021: 6685575.

Kumar, A., R. V. Singh, A. Chauhan, K. Ilayakumar, S. Kumar, A. Kumar, A. Sonwane, S. Kumar, M. Panigrahi and B. Bhushan. 2019. Genetic association analysis reveals significant effect of β-casein A1/A2 loci on production and reproduction traits in Frieswal crossbred cows. Biol. Rhythm. Res. 51: 1259-1272.

Kyte, J. and R. F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157: 105-132.

Lugani, Y. and B. S. Sooch. 2017. *In silico* characterization of cellulases from genus *Bacillus*. Int. J. Curr. Res. Rev. 9(13): 30-37.

Lv, Z., H. Liu, Y. Yang, D. Bu, C. Zang, K. Yang, X. Yu and J. Wang. 2020. Changes in metabolites from bovine milk with β-casein variants revealed by metabolomics. Animals. 10: 954.

Madrid, R. A. García-García, P. Cabrera, I. González, R. Martín and T. García. 2021. Survey of commercial food products for detection of walnut (*Juglans regia*) by two ELISA methods and real time PCR. Foods. 10: 440.

Marchitelli, C., G. Contarini, G. De Matteis, A. Crisà, L. Pariset, M. C. Scatà, G. Catillo, F. Napolitano and B. Moioli. 2013. Milk fatty acid variability: Effect of some candidate genes involved in lipid synthesis. J. Dairy Res. 80: 165-173.

Mayer, H. K., K. Lenz and E. Halbauer 2021. "A2 milk" authentication using isoelectric focusing and different PCR techniques. Food. Res. Int. 147: 110523.

Meier, S., P. Korkuć, D. Arends and G. A. Brockmann. 2019. DNA sequence variants and protein haplotypes of casein genes in German Black Pied Cattle (DSN). Front. Genet. 10: 1-9.

Miluchová, M., M. Gábor, J. Candrák, A. Trakovická and K. Candráková. 2018. Association of HindIII-polymorphism in kappa-casein gene with milk, fat and protein yield in Holstein cattle. Acta Biochim. Pol. 65: 403-407.

Miluchová, M., M. Gabor and A. Trakovicka. 2014. Analysis of genetic structure in Slovak Pinzgau cattle using five candidate genes related to milk production traits. Genetika. 46: 865-875.

Minkiewicz, P., A. Iwaniak and M. Darewicz. 2019. BIOPEP-UWM database of bioactive peptides: Current opportunities. Int. J. Mol. Sci. 20: 5978.

Nei, M. and Kumar, S. 2000. Molecular Evolution and Phylogenetics. Oxford University Press, New York.

Nuhriawangsa, A. M. P., T. Mulyani, G. Pambuko, R. Vanessa, Purwadi, N. Widya and S. Prastowo. 2021. Molecular screening of bovine β-casein (CSN2) A2 variant in Indonesian Holstein dairy cattle as attempt to produce digestive friendly milk. IOP Conf. Earth. Environ. Sci. 902: 012060.

Oleński, K. A. Cieślińska, T. Suchocki, J. Szyda and S. Kamiński. 2012. Polymorphism in coding and regulatory sequences of beta-casein gene is associated with milk production traits in Holstein-Friesian cattle. Anim. Sci. Pap. Rep. 30: 5-12.

Pal, S., K. Woodford, S. Kukuljan and S. Ho. 2015. Milk intolerance, beta-casein and lactose. Nutrients. 7: 7285-7297.

Park, Y. W. and G. F. W. Haenlein. 2021. A2 Bovine milk and caprine milk as a means of remedy for milk protein allergy. Dairy. 2: 191-201.

Pedrosa, V. B., F. S. Schenkel, S. Y. Chen, H. R. Oliveira, T. M. Casey, M. G. Melka, L. F. Brito. 2021. Genomewide association analyses of lactation persistency and milk production traits in Holstein cattle based on imputed whole-genome sequence data. Genes. 12: 1830.

Pooja, K., S. Rani and B. Prakash. 2017. *In silico* approaches towards the exploration of rice bran proteins-derived angiotensin-I-converting enzyme inhibitory peptides. Int. J. Food. Prop. 20: 2178-2191.

Rangel, A. H. N., D. C. Sales, S. A. Urbano, J. G. B. Galvãojúnior, J. C. de Andrade Neto and C. S. Macêdo. 2016. Lactose intolerance and cow's milk protein allergy. Food Sci. Technol. 36: 179-187.

Rashidinejad, A., P. Bremer, J. Birch and I. Oey. 2017. Nutrients in cheese and their effect on health and disease. In: Nutrients in Dairy and their Implications on Health and Disease. Ch. 14. Academic Press, Cambridge.

Raynes, J. K., L. Day, M. A. Augustin and J. A. Carver. 2015. Structural differences between bovine A1 and A2 β-casein alter micelle self-assembly and influence molecular chaperone activity. J. Dairy Sci. 98: 2172-2182.

Rousset, F. 2008. Genepop'007: A complete reimplementation of the Genepop software for Windows and Linux. Mol. Ecol. Resour. 8: 103-106.

Rupasinghe, R. K., N. Shanjayan, L. G. S. Lokugalappatti and S. Wickramasinghe. 2020. Genetic variants of β-casein gene in indigenous and exotic dairy cattle breeds in Sri Lanka. *Asian. J. Dairy Food* Res. 39: 217-220.

Sahay, A., A. Piprodhe and M. Pise. 2020. *In silico* analysis and homology modeling of strictosidine synthase involved in alkaloid biosynthesis in *Catharanthus roseus*. J. Genet. Eng. Biotechnol. 18: 44.

Saraswathy, N. and P. Ramalingam P. 2011. Two-dimensional gel electrophoresis of proteins. In: Biomedicine, Concepts and Techniques in Genomics and Proteomics. Ch. 11. Woodhead Publishing, Sawston.

Sebastiani, C., C. Arcangeli, M. Ciullo, M. Torricelli, G. Cinti, S. Fisichella and M. Biagetti. 2020. Frequencies evaluation of β-casein gene polymorphisms in dairy cows reared in central Italy. Animals. 10: 252.

Shah, K. and P. Maghsoudlou. 2016. Enzyme-linked immunosorbent assay (ELISA): The basics. Br. J. Hosp. Med. (Lond). 77: C98-C101.

Smouse, P. E., S. C. Banks and R. Peakall. 2017. Converting quadratic entropy to diversity: both animals and alleles are diverse, but some are more diverse than others. PLoS One. 12: e0185499.

Sodhi, M., R. Kataria, S. K. P. Niranjan, P. Verma, S. Swami, A. Sharma, V. Bharti, B. Kumar, M. Iqbal, S. Rabgais, P. Kumar, A. Giri, S. Kalia, D. Gagoi, P. Sarangi and M. Mukesh. 2018. Sequence characterisation and genotyping of allelic variants of beta casein gene establishes native cattle of Ladakh to be a natural resource for A2 milk. Def. Life Sci. J. 3: 177-181.

St-Gelais, D. and S. Haché. 2005. Effect of β-casein concentration in cheese milk on rennet coagulation properties, cheese composition and cheese ripening. Food Res. Int. 38: 523-531.

Studer, G., C. Rempfer, A. M. Waterhouse, G. Gumienny, J. Haas and T. Schwede. 2020. QMEANDisCo distance constraints applied on model quality estimation. Bioinformatics. 36: 1765-1771.

Suchy, F. J., P. M. Brannon, T. O. Carpenter, J. R. Fernandez, V. Gilsanz, J. B. Gould, K. Hall, S. Hui, J. Lupton and J. Mennella. 2010. NIH consensus development conference statement: lactose intolerance and health. NIH Consens. State. Sci. Statements. 27: 1-27. Available from: https://consensus.nih.gov/2010/docs/LI_CDC_2010_Final%20Statement.pdf [Last accessed on 2021 Dec 04].

Tetens, J. L. 2014. Genetic Variability of Bovine and Ovine Milk Protein Genes. (Doctoral dissertation), Christian-Albrechts-Universität zu Kiel, Germany. Available from: https://www.tierzucht.uni-kiel.de/de/forschung/dissertationen-1/diss_tetens_14.pdf [Last accessed on 2021 Dec 04].

Ul Haq, M. R. 2020. A1 milk and heart diseases. In: β-Casomorphins. Springer, Singapore.Ul Haq, M. R., R. Kapila, U. K. Shandilya and S. Kapila. 2014a. Impact of milk derived β-casomorphins on physiological functions and trends in research: A review. Int. J. Food Prop. 17: 1726-1741.

Ul Haq, M. R., R. Kapila, R. Sharma, V. Saliganti and S. Kapila. 2014b. Comparative evaluation of cow β-casein variants (A1/A2) consumption on Th_2-mediated inflammatory response in mouse gut. Eur. J. Nutr. 53: 1039-1049.

Untergasser, A., I. Cutcutache, T. Koressaar, J. Ye, B. C. Faircloth, M. Remm and S. G. Rozen. 2012. Primer3 new capabilities and interfaces. Nucleic Acids. Res. 40: e115.

Vanvanhossou, S. F. U., J. J. Giambra, T. Yin, K. Brügemann, L. H. Dossa and S. König. 2021. First DNA sequencing in Beninese indigenous cattle breeds captures new milk protein variants. *Genes*. 12: 1702.

Waterhouse, A., M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. T. Heer, T.A.P. de Beer, C. Rempfer, L. Bordoli, R. Lepore and T. Schwede. 2018. SWISS-MODEL: Homology modelling of protein structures and complexes. Nucleic Acids Res. 46(W1): W296-W303.

Woodford, K. 2009. Devil in the Milk: Illness, Health and the Politics of A1 and A2 Milk. Chelsea Green Publishing, Vermont.

Yadav, S., N. D. S. Yadav, A. Gheware, A. Kulshreshtha, P. Sharma and V. P. Singh. 2020. Oral feeding of cow milk containing A1 variant of β-casein induces pulmonary inflammation in male Balb/c mice. Sci. Rep. 10: 8053.

Yamada, A., M. Sugimura and T. Kuramoto. 2021. Genetic polymorphism of bovine beta-casein gene in Japanese dairy farm herds. Anim. Sci. J. 92: e13644.

Yasmin, I., R. Iqbal, A. Liaqat, W. A. Khan, M. Nadeem, A. Iqbal, M. F. J. Chughtai, S. J. Ur Rehman, S. Tehseen, T. Mehmood, S. Ahsan, S. Tanweer, S. Naz and A. Khaliq. 2020. Characterization and comparative evaluation of milk protein variants from Pakistani dairy breeds. Food Sci. Anim. Resour. 40: 689-698.