Molecular screening of bovine \( \beta \)-casein (CSN2) A2 variant in Indonesian Holstein dairy cattle as attempt to produce digestive friendly milk

A M P Nuhriawangsa\textsuperscript{1, 2}, T Mulyani\textsuperscript{1}, G Pambuko\textsuperscript{3}, R Vanessa\textsuperscript{4}, Purwadi\textsuperscript{5}, N Widyas\textsuperscript{1, 6} and S Prastowo\textsuperscript{1, 6,*}

\textsuperscript{1}Department of Animal Science, Faculty of Agriculture, Universitas Sebelas Maret, Surakarta, Indonesia
\textsuperscript{2}Animal Food Science and Technology Research Group, Faculty of Agriculture, Universitas Sebelas Maret, Surakarta, Indonesia
\textsuperscript{3}Master Program of Animal Science, Faculty of Agriculture, Universitas Sebelas Maret, Surakarta, Indonesia
\textsuperscript{4}Graduate School of Bioscience, Faculty of Mathematic and Natural Science, Universitas Sebelas Maret, Surakarta, Indonesia
\textsuperscript{5}Faculty of Animal Husbandry, Universitas Boyolali, Boyolali, Indonesia
\textsuperscript{6}Animal Breeding, Reproduction, and Biostatistics Research Group, Faculty of Agriculture, Universitas Sebelas Maret, Surakarta, Indonesia

Corresponding author: prastowo@staff.uns.ac.id

Abstract. Cow’s milk-intolerance is a digestive problem on people who not able to digest milk. This problem may relate to the variant of \( \beta \)-casein (CSN2), especially A1, suggested due to \( \beta \)-casomorphins (BCM-7) formation during enzymatic digestion, for that selecting cattle free BCM-7 become a concern to produce digestive friendly milk. This study aimed to differentiate A1 and A2 allele variant of CSN2 gene in selected population of Indonesian Holstein cattle. In total 70 cows DNA were collected, and fragment of CSN2 exon 7 which contain Single Nucleotide Polymorphism (SNP) rs43703011 and rs43703013 were amplified. Variant analysis was done by mutation site analysis using PCR-Restriction Fragment Length Polymorphism (PCR-RFLP) with \textit{MspI} restriction enzyme and DNA sequence for confirmation. Result shows 2 allele variants of \( \beta \)-casein that are B type, representing A1 family variant, and A2 in mutation site rs43703013C>G. We found A2 allele in the studied population is superior in frequency than A1 (0.916 vs. 0.084). Of that, 8.6% cattle were heterozygotes that is BA2 and 91.4% were homozygotes A2A2. Cattle which carry A1 allele variant should be excluded from dairy cattle breeding program for further milk production free of BCM-7.

1. Introduction
Cow milk intolerance is a gastrointestinal problem which commonly occur in people who are not able to digest cow’s milk or dairy products. It is a complex problem of importance to public and individual health. One of the elements in milk intolerance is lactose malabsorption or intolerance [1]. Additionally, in some cases, the milk intolerance is also accompanied with other potential health risk problem [2]. Previous review shows that milk intolerance may be associated with BCM-7 during digestion in human
digestive tract [1]. BCM-7 is a polypeptide known as the result of A1 digestion but not A2 β-casein which affects gastrointestinal motility and inflammation [3]. In here milk with A2 protein, believed to be relatively friendly to be consumed compared to the A1 milk.

β-casein or CSN2 is one of main milk protein which has 13 variants, the most common is A1 and A2 variant protein [4]. Milk with A1 protein produce 4 times more of BCM-7 compared to A2 milk [5]. Earlier study reporting that BCM-7 has potential to be associated with health risk problem such as the risk of diabetes type-1, heart diseases, arteriosclerosis, infant dead, autism, schizophrenia, and others [6,7]. The A1 β-casein milk led to significantly higher stool consistency values compared with the A2 β-casein milk. There was also a positive association between abdominal pain and stool consistency on A1 diet, but not A2 diet. This shows the differences in gastrointestinal responses in adult humans who consuming milk containing β-casein. It appears that BCM-7 may function as an immunosuppressant and impair tolerance to dietary antigens in the gut immune system [8]. In vitro studies indicate that BCM-7 can be produced from A1 and B milk during typical digestive processes; however, BCM-7 is not a product of A2 digestion.

The β-casein variant A1, A2, and B are commonly found in European dairy cattle breed and received much attention due to their influence on milk technology properties and on human health [9]. The two major variant A1 and A2 can be considered as β-casein “families”, which include at least 10 sub-variants. Those within the A1 family are B, C, D, F and G, while the A2 family are A3, E, H1, H2 and I [5]. The different of sub-variants within A1 family still unproved to have any effect on the release of BCM-7. However, there is a solid evidence that the B variant resulted in a particularly high release of BCM-7 compared to A1 variant [10]. Considering the food safety reason to human health, this study has objective to differentiate dairy cattle which carry different CSN2 allele variant namely A1 and A2 using a molecular approach. Further, the result could be applied in breeding dairy cattle free of BCM-7 to produce friendly digestive milk for human consumption.

2. Material and methods

2.1. Sample collection dan DNA extraction
In total 70 cows of Indonesian Holstein dairy cattle were randomly sampled from two locations. The location is in Baturaden (-7.303969142727606, 109.24800625409786) and Boyolali (7.580411827901281, 110.58255720101133) Indonesia. Those two locations were known to have large population of Indonesian Friesian Holstein dairy cattle.

Fresh blood for genomic DNA (gDNA) source were collected through coccygeal vein using 21G × 1.5 venoject needle attached to collection tube with EDTA contained. The procedure during blood collection was supervised by a veterinary medicine for animal ethically approval. Samples then stored in 5°C and transported to laboratory for long-term storage until DNA extraction. The gDNA were extracted using Wizard® Genomic DNA Purification Kit (Promega, USA) according to the protocol supplied by the company. Following extraction, gDNA then stored in -20°C for genotyping process.

2.2. β-casein genotyping
In this study, we differentiate β-casein allele variant of B and A2 on SNP site at g.8267C>G. According to Ensembl database (http://asia.ensembl.org/index.html), this site of mutation is recorded as rs43703013 (ENSBTAT0000003409.6: c.516G>C). Polymorphisms of CSN2 were detected by employing PCR-RFLP method then confirmed with DNA sequencing analysis. Gene fragment of CSN2 at exon 7 was amplified using primer listed in Table 1, position of amplicon as well as the mutation site was illustrated in Figure 1.

In every PCR reaction a positive control using GAPDH gene, and negative control by replacing the gDNA to ddH2O were performed to ensure both reactions were working and no contamination available. We use PCR reactions which consist of 1 µl gDNA sample, 1 µl each for forward and reverse primer (IDT, Singapore), 5 µl GoTaq® Master Mixes (Promega, USA), and 12 µl sterile ddH2O. All the PCR reactions were set at 35 cycles and visualized in electrophoresis reaction using 2% agarose gel.
Table 1. Primer list

| Gene Name | Sequence (5'-3') | Accession number (NCBI) | Product Size (bp) | Tm (°C) | Source |
|-----------|-----------------|------------------------|-------------------|---------|--------|
| CSN2      | F: CCAGACACAGTCCTAGTCTATCCCT<br>R: CAAACATGAGTGCAGTCCGGTCTCC | X14731.1<br>X14711.1 | 233<br>257 | 59.1<br>55 | [11]<br>[12] |
| GAPDH     | F: CCAGGGCTGCTTTTTAATTCT<br>R: ATGGCCTTTTCCATTGATGAC | NC_037332.1<br>NC_037332.1 | 257 | 55 | [12] |

Resulted PCR product then digested using \textit{MspI} restriction enzyme (Thermo Scientific, USA) which cut the DNA at restriction site 5’…C^CGG…3’. In total 5µl of PCR product were incubated at 37°C for 40 minutes along with 5 unit of \textit{MspI}. Next, digestion results then visualized in 2% agarose gel. β-casein allele types were determined by observing the resulted DNA band, A2 allele is indicated by digested PCR product into 208 and 25 bp. Meanwhile B allele was not digested. Part of PCR product simultaneously sent for sequencing using Sanger Methods. Resulted sequence then analyzed for amplified target confirmation.

Figure 1. The primer and SNP position in CSN2 exon 7 (accession number X14711.1)

2.3. Data Analysis

2.3.1. Similarity and sequencing analysis. The DNA sequencing in this study was interpreted using Unipro Ugene v. 39 software [13]. Subsequently, the result was confirmed by BLASTN on the NCBI website (https://blast.ncbi.nlm.nih.gov/Blast.cgi) [14].

2.3.2. Allele and genotype frequency. Allele and genotype frequency were analyzed using formula from Nei and Kumar [15], as described in formulas (1) and (2).

\[ x_i = \frac{2n_{ii} + \sum n_{ij}}{2N} \]  \hspace{1cm} (1)

Where: \( x_i \) is allele frequency \( i \), \( n_{ii} \) is the sum of genotype \( ii \) individual, \( n_{ij} \) is the sum of heterozygote genotype \( ij \), \( N \) is total individual.

\[ x_{ii} = \frac{n_{ii}}{N} \]  \hspace{1cm} (2)
Where: \( x_{ii} \) is genotype frequency \( ii \), \( n_{ii} \) is the sum of homozygote genotype \( ii \), \( N \) is total individual. Later, the genotype distribution in population was determined by Hardy-Weinberg Equilibrium (HWE) and statistically analyzed by chi-square test (\( \chi^2 \)) [15].

3. Results and discussion

Part of exon 7 of CSN2 gene was successfully amplified in this study as shown in Figure 2. The PCR reaction was able to amplify 233 bp of targeted DNA by primers designed in the earlier study [11]. Control positive (+) using GAPDH confirming that reaction was running well, and there is no contamination occurred as confirmed no DNA band in control negative (-) reaction. Additionally, the sequencing result of PCR product samples confirming the amplicon is part of CSN2 exon 7 with similarity 98.97% (Table 2).

![Figure 2. CSN2 gene visualization at 233 bp PCR product. M = ladder marker 1000 bp, 1-10: sample K+: control positive, K-: control negative](image)

| Sample ID | Similarity  | Sequence name                                          |
|-----------|-------------|--------------------------------------------------------|
| 1         | 98.97%      | Bos taurus \( \beta \)-casein (CSN2) gene, exon 7 and partial cds |
| 2         | 94.00%      | Bos taurus \( \beta \)-casein (CSN2) gene, exon 7 and partial cds |

Following amplification, genotyping was done by restriction enzyme digestion. As the result, 2 types of alleles namely B and A2 variant were revealed. In genotype variation, we found 2 genotypes that are BA2, shown by two DNA bands, and A2A2 which only single band (Figure 3). Furthermore, in our studied population, we didn’t find any cattle carry BB genotype. The sequence analysis to confirm that result is presented in chromatogram (Figure 4), shows the SNP different in the mutation site from G (allele B) to C (allele A2). Analysis of allele show that allele A2 was higher in frequency compared to
B variant. This also applied to the frequency of A2A2 genotype which higher compared to BA2 and BB genotypes (Table 3). HWE analysis shows no significant (P>0.05) different between genotype, which means the studied population is normally distributed.

**Table 3. Genotype and allele frequency of CSN2 variant**

| Allele frequency | Genotype frequency (n) |
|------------------|------------------------|
| B                | 0.084                  |
| A2               | 0.916                  |
| BB               | 0 (0)                  |
| BA2              | 0.086 (6)              |
| A2A2             | 0.914 (64)             |

Commonly CSN2 A1 variant has only found in European dairy cattle (Bos taurus) origin [9]. However, the variation of CSN2 allele now can be found in Bos indicus or other different dairy cattle worldwide. Purebred Asian and African cattle produce milk containing only the A2 beta-casein type, although some cattle produce A1 beta-casein as a consequence of crossbred ancestry. The A1 and A2 variant was reported in all dairy cattle both in Bos taurus and Bos indicus. For example the I, C, and B variant was found in Bos taurus and Bos indicus [16,17], while the E variant was reported only in Bos taurus [9]. Interestingly CSN2 B variant was reported to be existed in Bos javanicus which known to be Indonesian native cattle [18]. Notably, the mutations of CSN2 have been spreading and inherited since long time ago as part of evolutionary process of cattle domestication.

The result of higher frequency of A2 allele in this study compared to B allele, is in alignment with the previous report [19]. It is reported that A2 allele variant of CSN2 in Holstein cattle population studied in China, Rumania, and Italy have higher frequency compared to allele A1, B, I, A3, and C. Though the area of those studied populations are different but seem there are the same trend to produce digestive friendly milk is being a concern. This can be assumed due to the growing evidence of the negative effect of BCM-7 to human health risk factors.

CSN2 B is the variant of CSN2 which resulting highest BCM-7 compared to other variant during digestion [1,20]. In here, B variant particularly high release of BCM-7 compared to A1 variant [10]. It is explained that BCM-7 associated with lactose intolerance by inhibiting the lactase enzyme activity then affecting intestinal condition, thus lead to the problem of digestive system [1].

The growing demand of milk product and its safety, a concern to produce safer and friendly milk product, then manifested in the dairy cattle selection program. As we know, Friesian Holstein dairy cattle is the majority breed reared in Indonesia. This breed is selected and imported from either New Zealand or Australia to establish the herd population [21]. Since a lot of claims of unfavorable effect of BCM-7 to human health, New Zealand and Australia started to select and sold milk containing A2
protein as premium brand [5]. The selection was based on the DNA technology developed to determine whether the animal carry A1 or A2 allele variant or in combination of both. This also can be a justification which explaining why the percentage A2A2 cattle we found in our study is higher compared to another genotype group (Table 3).

Selection in New Zealand and Australia have been designed to breed cattle free of BCM-7. Converting a specific herd by selective breeding to eliminate A1 β-casein from the herd can be achieved by intensive methods of animal selection that incorporated reproductive technology [22]. It can be started with the selection of A2A2 bulls for sperm source in artificial insemination program.

4. Conclusions
The β-casein gene variant at rs43703013 in the current study can be identified as B allele, A1 family variant, and A2 allele. In the studied population of Indonesian Holstein, A2 allele is superior in frequency compared to A1 variant. Further study needs to be performed in determine CSN2 allele in other SNP mutation site, followed with protein identification of A1 and A2 for confirmation in milk. Later, the attempt to produce friendly digest (healthy) milk would need to be emphasized in dairy cattle breeding program free of BCM-7.

Acknowledgement
The current study is part of research project entitled “The variation analysis of Bovine Follicle Stimulating Hormone Receptor (FSHR) and its association to Indonesian Holstein Reproductive Performance”. The project was funded by PKLP-UNS Grant at fiscal year 2018, coordinated and administrated by Sigit Prastowo. Part of data presented in this paper was based on the bachelor thesis of Tri Mulyani with title “Variasi Alel Gen Beta-Casein (CSN2) pada Sapi Perah Friesien Holstein”. She was a student in Animal Science Department, Faculty of Agriculture, Universitas Sebelas Maret, Surakarta which started her research project on October 2020 under supervision of Sigit Prastowo and Adi Magna Patriadi Nuhriawangsa. She was successfully defending her thesis on April 26th, 2021.

References
[1] Pal S, Woodford K, Kukuljan S and Ho S 2015 Nutrients 31 7285–97
[2] Bell S J, Grochoski G T and Clarke A J 2006 Crit. Rev. Food Sci. Nutr. 46 93–100
[3] Ho S, Woodford K, Kukuljan S and Pal S 2014 Eur. J. Clin. Nutr. 68 994–1000
[4] De Noni I and Cattaneo S 2010 Food Chem. 119 560–6
[5] Kamiński S, Cieślińska A and Kostyra E 2007 Eur. J. Clin. Nutr. 68 994–1000
[6] Laugesen M and Elliott R 2003 N. Z. Med. J. 124
[7] Tailford K A, Berry C L, Thomas A C and Campbell J H 2003 Atherosclerosis 170 13–9
[8] Clemens R A 2011 Nestle Nutrition Workshop Series: Pediatric Program 67 187–95
[9] Massella E, Piva S, Giacometti F, Liuzzo G, Zambrini A V Serraino A 2017 Ital. J. Food Saf. 6 131–3
[10] Noni I De 2008 Food Chem. 110 897–903
[11] Barroso A, Dunner S and Canhã J 1999 J. Anim. Sci. 77 2629–32
[12] Prastowo S et al. 2016 Reprod. Fertil. Dev. 29 890–905
[13] Okonechnikov K, Golosova O and Fursov M 2012 Bioinformatics 28 1166–7
[14] Boratyn G M et al. 2013 Nucleic Acid Res. 41 29–33
[15] Nei M and Kumar S 2000 Molecular Evolution and Phylogenetics (New York: Oxford University Press)
[16] Gallinat J L, Qanbari S, Drögemüller C, Pimentel E C G, Thaller G and Tetens J 2013 J. Dairy Sci. 96 699–709
[17] Jann O, Ceroitti G, Caroli A and Erhardt G 2002 J. Anim. Breed. Genet. 119 65–8
[18] Caroli A M, Chessa S and Erhardt G J 2009 J. Dairy Sci. 92 5335–52
[19] Dâi R, Fang Y, Zhao W, Liu S, Ding J, Xu K, Yang L, He C, Ding F and Meng H 2016 J. Dairy Res. 83 312–6
[20] Olenski K, Kami S, Szyda J and Cieslinska A 2010 Elsevier 131 137–40
[21] Pangestu M, Subagyo Y, Yuwonol P and Rustomo B 2000 Asian-Australasian J. Anim. Sci. 13 505–8
[22] Mencarini I, Woodford K and Old K 2013 Proc. New Zeal. Soc. Anim. Prod. 73 149–54