Characterization of the genetic polymorphism linked to the β-casein A1/A2 alleles using different molecular and biochemical methods

V. Vigolo,1 M. Franzoi,1 F. Cendron,1* G. Salvadore,1 M. Penasa,1 M. Cassandro,1,2 and M. De Marchi1

1Department of Agronomy, Food, Natural resources, Animals and Environment, University of Padova, Viale dell’Università 16, 35020 Legnaro (PD), Italy
2Federazione delle Associazioni Nazionali di Razza e Specie, Via XXIV Maggio 44, 00187 Roma, Italy

ABSTRACT

The 2 major subvariants of β-casein (A1 and A2), coded by CSN2 gene, have received great interest in the last decade both from the scientific community and the dairy sector due to their influence on milk quality. The consumption of the A1 variant, compared with the A2 variant, has a potential negative effect on human health after its digestion but, at the same time, its presence improves the milk technological properties. The aim of the present study was to compare the best method in terms of time required, costs, and technical engagement for the identification of β-casein A1 and A2 variants (homozygous and heterozygous animals) in milk to offer a reliable service for large-scale screening studies. Two allele-specific PCR procedures, namely RFLP-PCR and amplification refractory mutation system (ARMS-PCR), and one biochemical technique (HPLC) were evaluated and validated through sequencing. Manual and automated DNA extraction protocols from milk somatic cells were also compared. Automated DNA extraction provided better yield and purity. Chromatographic analysis was the most informative and the cheapest method but unsuitable for large-scale studies due to lengthy procedures (45 min per sample). Both allele-specific PCR techniques proved to be fast and reliable for differentiating between A1 and A2 variants but more expensive than HPLC analysis. Specifically, RFLP-PCR was the most expensive and labor-demanding among the evaluated techniques, whereas ARMS-PCR was the fastest while also requiring less technical expertise. Overall, automated extraction of DNA from milk matrix combined with ARMS-PCR is the most suitable technique to provide genetic characterization of the CSN2 gene on a large scale.

Key words: casein polymorphism, RFLP-PCR, ARMS-PCR, HPLC, milk

INTRODUCTION

Bovine milk contains 3.5 to 4% protein. About 80% of milk proteins consists of caseins (Holt et al., 2013). Caseins can be categorized into αS1-CN (CSN1S1, 39–46% of total caseins), αS2-CN (CSN1S2, 8–11% of total caseins), β-CN (CSN2, 25–35% of total caseins), and κ-CN (CSN3, 8–15% of total caseins; Huppertz, 2016). Heterogeneity of casein further includes genetic polymorphisms that have received great interest due to their possible association with milk composition, technological properties, and human health (Cendron et al., 2021; Daniloski et al., 2021; Vigolo et al., 2021). In particular, 12 genetic variants of β-CN have been identified in dairy cattle breeds (A1, A2, A3, B, C, D, E, F, G, H1, H2, and I; Sebastiani et al., 2020). The most interesting polymorphism of the CSN2 gene related to milk production is g.8101C >A (p.His67Pro), which leads to the expression of 2 families of β-CN protein variants: A1-like, carrying histidine in position 67, and A2-like, carrying proline in the same position. The mutation that causes the difference in residue is the result of a single nucleotide polymorphism at codon 67 in exon 7 of the gene CSN2, in which CCT (in the A2 variant, coding for proline) changed to CAT (in the A1 variant, coding for histidine; Kay et al., 2021). The presence of a particular AA affects the patterns of enzymatic cleavage during both in vitro and in vivo digestion; as the A2-like variants are unlikely to undergo enzymatic cleavage in position 67, the A1 polymorphism permits proteolytic cleavage which results in the release of different peptides including β-casomorphin-7 (BCM7) (Summer et al., 2020). The BCM7 is a µ-opioid receptor agonist and is considered a potential cause of numerous noncommunicable diseases, mainly concerning gastrointestinal discomfort (Daniloski et al., 2021; Kay et al., 2021). Despite the uncertainty regarding the hypothesis of a positive effect of β-CN A2 vari-
Several analytical techniques have been developed to characterize the most interesting polymorphisms of the β-CN gene, using several approaches including protein-based phenotyping and DNA-based techniques. The reference method for the detection of genetic variants is the direct sequencing of individual DNA samples with a focus on nucleotide mutations (Hashim and Al-Shuhaib, 2019), but this technique is not applicable to large-scale studies due to costs and time requirements. For this reason, several post-PCR genotyping techniques are available to identify mutations in the nucleic acid sequences; some of them are challenging to perform such as amplification-created restriction site PCR (Mayer et al., 2021) and bidirectional allele-specific PCR (Chessa et al., 2013), or too expensive such as real time quantitative PCR (Manga and Dvořák, 2010), and thus not appropriate for screening purposes. Restriction fragment length polymorphism PCR and amplification refractory mutation system (ARMS)-PCR are, instead, more suitable methods for large-scale genotyping (Jaiswal et al., 2014; Smiltina and Grislis, 2018).

Particularly, the RFLP-PCR is the most common technique for the genotyping of allelic variants and has been successfully applied in numerous genetic studies (Ratna Kumari et al., 2008; Smiltina and Grislis, 2018). In this technique, a PCR primer pair is designed in such a way it amplifies a region containing a polymorphism which renders one variant accessible for restriction enzyme digestion and another not. These can then be differentiated and visualized using electrophoretic separation of amplicons from different alleles (Panneerchelvan and Norazmi, 2003). The ARMS-PCR uses 2 universal outer primers and 2 allele-specific inner primers to alternatively amplify the 2 different alleles in products of different lengths in a single PCR reaction. It is widely used in clinical approaches due to the high quality and accurate results attainable (Little, 2001). Nonetheless, milk used to manufacture dairy products is milk pooled from several individuals. Accordingly, quantitative methods are necessary to authenticate A2 milk [e.g., biochemical techniques such as reverse-phase (RP) HPLC or isoelectric focusing; Mayer et al., 2021].

The present study investigated the polymorphism linked to the A1 and A2 variants of the CSN2 gene (GenBank M55158; g8101C > A) within exon 7, using 3 different techniques: (1) RFLP-PCR, (2) ARMS-PCR, and (3) RP-HPLC. Contextually, manual and automated DNA extraction methods were evaluated due to their potential for high-quality DNA yields in large-scale screening. The aim was to assess performance (time of analysis, expertise needed, and labor requirements) and operative cost of such techniques in differentiating the genetic variants A1 and A2 of β-CN starting from milk somatic cells.

**MATERIALS AND METHODS**

Ethical approval was not required for the present study because cows belonged to commercial herds and milk was collected during official monthly test-day milk recording scheme by the authorized personnel of the Breeders Association of Veneto Region (Vicenza, Italy).

**Samples**

Individual milk samples (50 mL) of Holstein-Friesian, Simmental, Jersey, and Rendena cattle breeds were collected between November and December 2020 by the Breeders Associations of Veneto Region (ARAV, Vicenza, Italy) during the official monthly test-day milk recording scheme of 39 herds. Samples were kept at 4°C, transported refrigerated to the Department of Agronomy, Food, Natural resources, Animals and Environment of the University of Padova (Legnaro, Italy) within 24 h from collection, and stored at −20°C until analysis.

**DNA Extraction**

To compare manual and automated protocols, 2 different methods were used to extract genomic DNA from somatic cells of milk samples. Manual extraction of genomic DNA was performed on 568 samples by adapting the DNA Purification from a Buccal Brush using the Gentra Puregene Buccal Cell Kit (Qiagen Gmbh), including an additional washing step. In particular, milk was first defatted by centrifugation at 5,000 × g for 4 min at 4°C. Then, the pellet was resuspended with 1 mL of washing solution (15 mM Tris-HCl, 25 mM NaCl, 15 mM Na2HPO4, 2.5 mM EDTA, pH 7.5) and further centrifuged at 5,000 × g for 5 min at 4°C. Finally, DNA extraction was carried out as reported by manufacturer. Automated DNA extraction was conducted on 568 samples by adopting the GenePure Pro fully automatic nucleic acid purification system (Bior) through the use of MegaBio plus Virus DNA/RNA purification Kit II (Bior).

The concentration and purity of DNA were assessed through a Multiscan Sky spectrophotometer (Thermo
Fisher Scientific) using the ratio of UV absorbance at 260 and 230 nm and at 260 and 280 nm. The means and standard deviation (SD) were calculated for all spectrophotometric measurements and compared through a 2-tailed paired t-test (P < 0.05). Statistical analysis was performed using SAS software (SAS Institute Inc.).

Evaluation of methods in terms of difficulty of performance (e.g., operators’ technical skills), time required, and costs were carried out on 96 samples extracted simultaneously. The required time was calculated as the sum of the time needed for each of the steps in the protocol, including the centrifugation and incubation. The cost per sample was calculated considering the costs of the kit and the reagents per sample. Labor was not considered.

The same samples were sent to a sequencing service (BMR Genomic sequencing service, Padova, Italy) to validate the genotypes identified by the proposed techniques.

**ARMS-PCR**

Primers used in the present study for the amplification were described by Ali et al. (unpublished data; W. R. Ali, I. Amin, M. Asif, and S. Mansoor, National Institute for Biotechnology and Genetic Engineering, Faisalabad, Pakistan) and synthesized by Thermo Fisher Scientific. The DNA amplification was performed in a total volume of 25 µL of PCR mixture containing 0.150 µL of DreamTaq Polymerase (5 U/µL), 2.5 µL of Green Buffer (1×), 2.5 µL of MgCl (20 mM), 5 µL of dNTP (2 mM; Thermo Fisher Scientific), 1 µL of Primer FW (10 mM), 1.85 µL of H2O, 0.150 µL of di DreamTaq Polymerase (5 U/µL; Thermo Fisher Scientific), and 150 ng of genomic DNA. Primers used were FA2 (forward inner) 5′-TAGTCTATCCCTCTGCCATCC-3′; RA1 (reverse inner) 5′-AGGGATGTCTTGTGGGAGGCCTG-3′; NFO (forward outer) 5′-TCTGCAAGAGCTTCCATCGA-3′; NRO (reverse outer) 5′-AAGACTGGAGCAGGCGAGGAAGGT-3′. The amount of inner and outer primer was in a ratio of 1:10 as reported by Rincon and Medrano (2003). The PCR reaction was performed in the SimplyAmp thermocycler (Thermo Fisher Scientific) and run under the following optimized conditions: initial denaturation at 94°C for 5 min; 40 cycles at 94°C for 60 s, annealing for 30 s (testing 60, 62, or 64°C), extension at 72°C for 60 s, and a final extension step at 72°C for 10 min which was not cycled. Primers used were described by Ciesielska et al. (2019) and reported in materials and methods.

The amplification products were digested using the Mph1103I restriction enzyme (NsiI, Thermo Fisher Scientific); the restriction mix consisted of 20 µL of PCR product, 8 µL of H2O, 2 µL of Buffer R (10×; Thermo Fisher Scientific), and 2 µL of Mph1103I enzyme. After enzymatic digestion at 37°C for 3 h, the obtained restriction fragments were analyzed through electrophoresis to visualize the presence or absence of the SNP that allows distinction of the β-CN variants. Non-A2 variants would yield bands demonstrating the 2 restriction fragments of 285 and 37 bp, whereas A2 variants would yield a single band of 322 bp indicating the β-CN A2 variant, and a restriction fragment of 384 bp indicating the β-CN A1 variant. The amplification products were resolved by electrophoresis in 1.5% (wt/vol) agarose gel in 600 mL of TBE (89 mM Tris, 89 mM boric acid, 2 mM EDTA; Thermo Fisher Scientific). Electrophoresis was performed at a constant voltage (200 V) for 1 h at room temperature. The costs of genotyping per sample were calculated by considering the cost of reagents and consumables.

**RFLP-PCR**

The DNA amplification was performed in a total volume of 25-µL PCR mixture containing 5 µL of dNTP (2 mM), 2.5 µL of Green Buffer (1×), 2.5 µL of MgCl (20 mM), 1 µL of Primer FW (10 mM), 1 µL of Primer RV (10 mM), 1.85 µL of H2O, 0.150 µL of di DreamTaq Polymerase (5 U/µL; Thermo Fisher Scientific), and 150 ng of genomic DNA. Primers used to perform PCR were CASB-forward, 5′-GCAGAAATTC-TAGTCTATCCCTCCCTGGAGACCATGC-3′ and CASB-reverse, 5′-ACGGACTGAGGAGGAAACAT-GACAGTTGGAGGAAG-3′. The amplification products were resolved using agarose gel electrophoresis and subsequently digested with the Mph1103I restriction enzyme. The DNA amplification was performed in the SimplyAmp thermocycler (Thermo Fisher Scientific) and run under the following optimized conditions: initial denaturation at 94°C for 5 min; 40 cycles at 94°C for 60 s, annealing for 30 s (testing 60, 62, or 64°C), extension at 72°C for 60 s, and a final extension step at 72°C for 10 min which was not cycled. Primers used were described by Ciesielska et al. (2019) and reported in materials and methods.

The amplification products were digested using the Mph1103I restriction enzyme (NsiI, Thermo Fisher Scientific); the restriction mix consisted of 20 µL of PCR product, 8 µL of H2O, 2 µL of Buffer R (10×; Thermo Fisher Scientific), and 2 µL of Mph1103I enzyme. After enzymatic digestion at 37°C for 3 h, the obtained restriction fragments were analyzed through electrophoresis to visualize the presence or absence of the SNP that allows distinction of the β-CN variants. Non-A2 variants would yield bands demonstrating the 2 restriction fragments of 285 and 37 bp, whereas A2 variants would yield a single band of 322 bp demonstrating its inability to be digested by Mph1103I. In particular, electrophoretic separation was performed in 2.5% (wt/vol) agarose gel in 600 mL of TBE buffer (89 mM Tris, 89 mM boric acid, 2 mM EDTA; Thermo Fisher Scientific). Electrophoresis was performed at a constant voltage (200 V) for 1 h at room temperature. The costs of genotyping per sample were obtained by considering the cost of reagents and consumables.
**Table 1. Purity and concentration of the DNA obtained with different extraction methods**

<table>
<thead>
<tr>
<th>Extraction method</th>
<th>Commercial kit</th>
<th>A260/A280</th>
<th>A260/A230</th>
<th>V (mL)</th>
<th>c (µg/mL)</th>
<th>c_n (µg/mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manual extraction</td>
<td>Gentra Puregene Buccal Cell Kit (Qiagen)</td>
<td>1.12 ± 0.29^a</td>
<td>0.32 ± 0.27^b</td>
<td>15.00</td>
<td>30.37 ± 33.00^a</td>
<td>2.46 ± 2.30^a</td>
</tr>
<tr>
<td>Automated extraction</td>
<td>MegaBio plus Virus DNA/RNA purification Kit II (Bioer)</td>
<td>1.53 ± 0.27^a</td>
<td>1.55 ± 0.86^b</td>
<td>0.30</td>
<td>16.61 ± 15.10^b</td>
<td>55.60 ± 50.34^b</td>
</tr>
</tbody>
</table>

^a,bWithin a column, means with different superscript letters are significantly different (P < 0.05).

^1Values are given as mean ± SD. A260/A280 = ratio of absorbances measured at 260 and 280 nm; A260/A230 = ratio of absorbance measured at 260 and 230 nm; V = required initial milk volume; c = concentration of DNA solution; c_n = normalized DNA concentration.

**RP-HPLC**

Protein-based phenotyping was carried out through RP-HPLC analysis (Agilent 1260 Infinity II LC system; Agilent Technologies). Milk samples were prepared according to the method of Bobe et al. (1998). Briefly, 500 µL of milk was added to an aqueous solution of guanidine (Gdn) HCl (6 M GdnHCl, 0.1 M BisTris buffer, 5.37 mM sodium citrate, and 19.5 mM dithiothreitol) in a 1:1 ratio (vol/vol). Each sample was vortexed for 10 s, incubated at room temperature for 1 h to promote protein solubilization, and thereafter centrifuged for 10 min at room temperature at 13,000 × g to promote the separation of fat. The remaining soluble phase was added to a solution containing 4.5 GdnHCl diluted in a solvent consisting of water, acetonitrile, and trifluoroacetic acid (100:900:1; vol/vol/vol), in a 1:3 ratio (vol/vol). The separation of milk protein fractions and genetic variants was conducted following the method proposed by Bonfatti et al. (2008): gradient elution was carried out with a mixture of solvent A (0.1% trifluoroacetic acid in water) and solvent B (0.1% trifluoroacetic acid in acetonitrile). Separations were performed with the following gradients: linear gradient from 33 to 35% B in 5 min, from 35 to 37% B in 4 min, from 37 to 40% B in 9 min, from 40 to 41% B in 4 min, isocratic elution at 41% B for 5.5 min, linear gradient from 41 to 43% B in 0.5 min, and from 43 to 45% B in 8 min. A reversed-phase analytical column C8 (Zorbax 300SB-C8 RP, Agilent Technologies) with a silica-based packing (3.5 µm, 300 Å, 150 × 4.6 i.d.) preceded by a Security Guard Cartridge System used as a pre-column (300SB-C8 Guard Cartridges 4.6 × 12.5 mm, 4/PK, Agilent Technologies), was used for separation. Before the injection of the following sample, the column was re-equilibrated at 33% B for 8 min. The total analysis time per sample was 44 min. The flow rate was 0.5 mL/min, the column temperature was kept at 45°C, the detection was made at a wavelength of 214 nm, and the injection volume was 5 µL. In addition to β-CN identification, the analysis allowed the identification and quantification of all different milk protein fractions through the use of external standards of α-CN, β-CN, κ-CN, α-LA, β-LG A, and β-LG B (Merck KGaA) of the highest purity available. Agilent OpenLab 2 CDS software (Agilent Technologies) was used for data acquisition and analysis. The costs of phenotyping per sample were the sum of costs for solvents and consumables.

**RESULTS**

**DNA Extraction Protocols**

Manual and automated extraction protocols were evaluated in terms of quantity and quality of the extracted nucleic acids, costs, time required, and difficulty of implementation (i.e., manual operation requiring technical experience). The evaluation of yield and purity of DNA extracted was based on spectrophotometric measurement ratios (A260/A280 and A260/A230, respectively) and is summarized in Table 1. Specifically, the mean values of A260/280 ratio measured on the manual and the automated extracts were 1.12 ± 0.29 (range = 0.54 to 3.25) and 1.53 ± 0.27 (range = 0.41 to 3.95), respectively. Mean values of A260/230 ratio indicate that DNA samples extracted with both methods had a purity level below optimum (manual: 0.32 ± 0.27; automated: 1.55 ± 0.86). Overall, the 2 spectrophotometric ratios demonstrate that DNA isolated through the automated protocol was of higher quality. Consequently, PCR success rate performed with DNA extracted with manual protocol approached 0%, whereas PCR performed with DNA extracted with automated protocol had a success rate of almost 80%.

The measured concentrations of DNA (c) were normalized (c_n) within each extraction method for the starting volume of the milk sample (1 mL) and the final volume of extracted DNA solution (100 µL). Results of normalized concentration (Table 1) indicate that the amount of DNA isolated using the MegaBio plus Virus DNA/RNA purification Kit II in the automated protocol was greater than that isolated in the manual protocol (55.6 and 2.46 µg/mL, respectively; P < 0.05). The time required for DNA extraction from milk differed between the automated and the manual
Table 2. Labor, time, and material costs required for DNA extraction with different extraction methods, and for genotyping or phenotyping with different techniques

<table>
<thead>
<tr>
<th>Extraction method</th>
<th>Labor requirement</th>
<th>Time for 96 samples (h)</th>
<th>Cost per sample (€)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manual protocol</td>
<td>++++</td>
<td>8.0</td>
<td>3.00</td>
</tr>
<tr>
<td>Automated protocol</td>
<td>+</td>
<td>2.0</td>
<td>4.70</td>
</tr>
<tr>
<td>Genotyping</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARMS-PCR</td>
<td>+++</td>
<td>3.2</td>
<td>0.70</td>
</tr>
<tr>
<td>RFLP-PCR</td>
<td>+++++</td>
<td>6.5</td>
<td>1.40</td>
</tr>
<tr>
<td>Phenotyping</td>
<td>++</td>
<td>75.0</td>
<td>1.50</td>
</tr>
</tbody>
</table>

1ARMS = amplification refractory mutation system; RP = reverse phase.
2(+) = very easy; (++) = easy; (+++) = difficult; (++++) = very difficult.

extraction protocol, which required 2 and 8 h, respectively. Regarding the technical engagement, the automated protocol was not labor-demanding because the commercial kit had an easy-to-follow protocol. On the other hand, the manual extraction protocol demanded a higher handling skill and attention (Table 2).

**Genotyping Using ARMS-PCR**

The ARMS methodology uses 2 different primer pairs to amplify the 2 allelic variants caused by the SNP in a single PCR reaction. In this method, 2 allele-specific amplifications occur in opposite directions by the use of 2 outer primers and 2 allele-specific inner primers which together amplify the region of the SNP. To improve the efficiency of the technique, it was necessary to optimize the correct dilution of primers. Indeed, the ratio of inner and outer primer of 1:1 proposed by Ali et al. (unpublished data; W. R. Ali, I. Amin, M. Asif, and S. Mansoor, National Institute for Biotechnology and Genetic Engineering, Faisalabad, Pakistan) provided unsatisfactory results. Adding more internal control primers (outer primer) to the reaction, following the protocol proposed by Rincon and Medrano (2003), ensured that internal primers (inner primer) did not compete with the amplification reaction. The need to optimize the technique was considered in the overall evaluation of performance.

In the present study, the ARMS-PCR technique was successfully applied to identify individual SNPs in *CSN2* of 568 individual milk samples; the allele-specific segments differed sufficiently in size to be distinguished by agarose gel electrophoresis. The ARMS-PCR produced a fragment of 586 bp as internal control, a fragment of 384 bp for the A1 allele, and a fragment of 256 bp for the A2 allele. Figure 1 (right side) depicts the electrophoretic output after amplification of *CSN2* allelic variants. It shows the difference between the length of amplicons between A2 homozygous individuals (which yield the control fragment and the 256 bp fragment), the A1 individuals (which yield control fragment and the 384 bp fragment), and the A2A1 heterozygotes.

![Figure 1](image_url)  # Figure 1. Electropherogram of β-CN A1/A2 locus. Observed (A) and theoretical (B) patterns of the PCR amplicons from RFLP and amplification refractory mutation system (ARMS), respectively. The dashed red line represents the reading area of the genotypes (top: genotypes; bottom: nonspecific bands). The asterisks indicate not confirmable genotypes due to no control-band (586 bp) present.
with the control fragments and both the 256 and 384 bp amplicons (e.g., lane h).

**Genotyping Using RFLP-PCR**

The RFLP-PCR technique was successfully applied for the genotyping of 230 individual milk samples. Performing a PCR using primers suggested by Cieślińska et al. (2019) resulted in amplicons of 322 bp. The following digestion of the amplicons with the enzyme Mph1103I (NsiI, Thermo Fisher Scientific), a restriction enzyme that recognizes the ATGCAT sequence as the restriction site, produced a restriction pattern that clearly differentiated homozygote from heterozygote cows. This is because an allele-specific SNP lies within the restriction site of Mph1103I rendering the enzyme incapable of cleaving the amplicon deriving from the A2 variant. The amplicon from the A1 variant, however, retains its restriction site due to the lack of the SNP, thereby yielding 2 fragments of 285 and 37 bp in length following digestion. This step is the most delicate and greatly affected the evaluation of the expertise needed to perform the technique; indeed, time and temperature must be extremely precise as errors at this stage could lead to false negatives. Thus, the amount of enzyme and the incubation time were optimized following the manufacturer’s protocol and performing several digestion-time trials. The maximum digestion activity was obtained in 3 h with a digested PCR product of 100%.

Again, the produced fragments differed sufficiently in size to be distinguished by agarose gel electrophoresis thereby allowing the differentiation between A2 homozygotes (e.g., lane 2; Figure 1, left side), non-A2 cows (e.g., lane 3; Figure 1, left side), and A2X heterozygotes (e.g., lane 1; Figure 1, left side).

**Phenotyping Using RP-HPLC**

Chromatography analysis allowed for the acquisition of detailed protein composition and quantification of 568 individual milk samples. Regarding β-CN, the most common genetic variants of β-CN (i.e., A1, A2, and B) were identified by comparing elution times of the chromatographic peaks with those reported in literature (Bonfatti et al., 2008). Retention time for β-CN B, A1, and A2 variants was 23.1, 25.7, and 26.7 min, respectively (Figure 2). For the genetic variants of β-CN, the resolution between peaks exceeded 1 min, which is considered satisfactory in chromatographic separation. Extraction and chromatographic injection are limited to a trivial laboratory task; particular attention should be paid during soluble phase collection, to avoid sample contamination, which may lead to invalid results.

**DISCUSSION**

**Yield and Quality of DNA Prepared with Different Extraction Methods**

The setup of robust extraction protocols is a key part in the development of fast and reliable services to be offered to a large number of farmers. Even though milking is an easy and convenient way of acquiring samples, obtaining a good DNA extraction for subsequent molecular analysis from milk samples is quite challenging (Volk et al., 2014). Indeed, several components of extraction mixtures and milk constituents such as fat, carbohydrates, and enzymes are known as PCR interferers. For this reason, several commercial kits are available, and a wide number of DNA extraction methods have been proposed (Psifidi et al., 2010; Volk et al., 2014). The DNA is extracted from somatic cells which are present in milk in varying concentrations depending on the health status of the mammary gland (Franzoi et al., 2020), but nonetheless always in a sufficient quantity to provide enough isolated DNA for PCR applications. In the present study, the first goal was to investigate the quality of DNA extracted from milk using an automatic and a manual DNA extraction protocol, with a focus also on time and technical requirements. The extraction methods were chosen due to their popularity in laboratories for their potential to retrieve high-quality DNA yield and for the ability to process a large number of samples. The manual DNA extraction protocol produced a yield and purity inferior to the automatic protocol method, but it has a cost almost 40% lower (Table 2). Spectrophotometric measurements ratios indicated poor quality of the extracted DNA with ratios below the optimum: pure DNA solutions have an 260/280 ratio of ≥1.8 while lower ratios indicate contamination (Held, 2001). Similarly, the desired level for 260/230 ratios is >1.5 to 1.8 and samples that present lower values reveal a significant presence of contaminants (Pfaffl, 2004). Specifically low ratios reported for manual extraction in Table 1 probably reflected contamination of organic inhibitors present in the milk matrix such as protein and carbohydrates, but also contamination of trace components (e.g., isopropanol, EDTA). Instead, values reported for the automated protocol indicated possible contamination during the washing step with buffer containing guanidine salts. Despite the low ratios, values are comparable to those reported by Volk et al. (2014) who evaluated different methods for DNA extraction starting from the same matrix.

Moreover, the repeatability of the manual protocol was not satisfactory showing a great SD for both spectrophotometric measurements and DNA concentration.
(Table 1). This often led to the impediment of a successful PCR. The reasons for such negative performances could be related to the greater amount of time required for the additional washing step of the cellular pellet and to the presence of several manual operations in the protocol (e.g., centrifugations, fat removal, pellet resuspension, DNA and protein precipitation; Table 2). However, it should be considered that the high SD could be also due to the variability of DNA content in the analyzed samples: the somatic cells from which the DNA was extracted are present in milk in variable amount. Although the manual protocol was cheaper compared with the automated extraction method, the present study proved it to be unsuitable for DNA extraction on a large scale due to low consistency in terms of yield as well as the presence of contaminants postextraction. This was also the reason for an almost 100% abortion rate of the PCR amplification. The au-

Figure 2. Reverse-phase HPLC chromatograms relative to individual milk samples. Genetic variants were identified on the basis of their retention time. (A) Heterozygous A1A2; (B) homozygous A2A2; (C) homozygous A1A1; (D) homozygous BB. κ-CN (variants A, E, and B); β-CN (variants B, A1, and A2).
tomated extraction protocol, on the other hand, had several advantages. In particular, a satisfactory purity level of extracted DNA was consistently achieved and thereby an advantageous final quantity/volume ratio was readily obtained which allowed for several rounds of PCR analysis.

Even though blood is considered the best matrix for genetic studies, in the present study we observed that the quality and quantity of genetic material extracted from milk through the automated extractor was fully satisfactory. Additionally, although it is not suitable for all animals, it is important to consider that milk sampling does not require specialized personnel and it is periodically collected during milk recording schemes. It therefore allows for large-scale application and provides an opportunity to present a rapid response to farmers.

**Genotyping Using ARMS-PCR**

The ARMS-PCR technique was optimized for reagent concentration and for primer annealing temperatures. Optimization of the primer pair was carried out by performing a PCR at different temperatures: 62, 64, and 66°C. At 62°C the PCR led to the formation of a great number of nonspecific fragments, impairing the analysis of the results, though the annealing temperature of 64°C had greater specificity for the 256 bp amplicon (variant A2) compared with the 384 and 586 bp amplicons (variant A1 and control, respectively). The annealing temperature of 66°C was confirmed to be the most effective for the primer pair, providing the best results during electrophoretic separation (unpublished data; W. R. Ali, I. Amin, M. Asif, and S. Mansoor, National Institute for Biotechnology and Genetic Engineering, Faisalabad, Pakistan). To decrease the amplification abortion rate, the PCR mix was optimized by increasing the dNTP concentration and modifying the primer concentration compared with Ali et al. (unpublished data; W. R. Ali, I. Amin, M. Asif, and S. Mansoor, National Institute for Biotechnology and Genetic Engineering, Faisalabad, Pakistan).

Among the 3 methods proposed in the present study, indeed, ARMS-PCR was the technique that required the least time to be performed and also the most cost-effective. However, to obtain optimal use of the ARMS-PCR amplification, the presence of an internal control during the analysis is essential to allow accurate and reliable identification of allelic variants.

**Genotyping Using RFLP-PCR**

The RFLP-PCR relies on the presence or absence of a particular recognition site of a restriction enzyme in the target sequence. In the present study, the detection of β-CN alleles (A1, A2) was carried out using the restriction enzyme sequence as described by Cieślińska et al. (2019). The protocol was optimized for primer annealing temperature by testing 3 annealing temperatures: 62, 64, and 66°C. Contrary to Cieślińska et al. (2019), the optimal annealing temperature proved to be 64°C. Compared with the other techniques of the present study, the RFLP-PCR was the most difficult in terms of technical requirements because of the involvement of a restriction enzyme very sensitive to contaminants and temperature changes. Regarding the required time per sample, this technique is less efficient than the ARMS-PCR but still suitable for large-scale genotyping, because it takes about 4 min per sample at an affordable cost (Hashim and Al-Shuhaib, 2019). Moreover, the ability to detect other SNPs revealing different allelic variants (other than A1 and A2) is not possible because the enzymatic digestion is site-specific. For instance, to detect the B variant, it is necessary to develop a different protocol considering the SNP (C > G) at position 8,267 bp. For this reason, in the present study, B allele was not detected but instead segregated with the A1 genotype. However, these results may be insignificant if we consider that farmers are mainly interested in knowing whether their cows are homozygous A2A2, heterozygous A2X, or non-A2 animals. In contrast to ARMS-PCR in which the A1 allele is identified by a single PCR amplicon (384 bp), in RFLP-PCR the A1 variant is characterized by 2 fragments (285 and 37 bp) after amplicon digestion with the Mph1103I restriction enzyme (Figure 1). This may be of critical importance, because an amplicon that was not cut due to an error in the procedure would be classified incorrectly. In addition, the small difference in fragment length after digestion can complicate the interpretation of bands on low-resolution agarose gel.

**Phenotyping Using RP-HPLC**

Protein separation by RP-HPLC is based on the interaction of the hydrophobic groups of the different protein fractions with insoluble hydrophobic groups immobilized on the stationary phase. The substitution of the Pro67 (a nonpolar, partially hydrophobic AA) in the β-CN A2 with a charged polar His67 in variant A1, decreases the hydrophobicity of the latter variant resulting in a shorter elution time from the column compared with the A2 variant. Significant variations were observed between protein chromatograms from samples of different β-CN polymorphisms (Figure 2). The change in hydrophobicity caused by the different AA substitutions in the different variants allows for the detection of other β-CN polymorphism beyond A1 and A2 due to the resulting change in elution time.
In restriction enzyme analysis techniques, instead, no further genotyping beyond variant A1 and A2 is possible without additional methods (i.e., appropriate restriction sites have to be artificially created). Additionally, further information can be gathered from the output of the chromatographic analysis such as the quantification of each protein fraction and genetic variants of other milk proteins, namely α-CN and β-LG. The former can provide the amount of each β-CN variant present in milk, which is useful for farmers who desire to know the A2 incidence (cows) on the herds by analyzing one bulk sample. On the other hand, even if the skills required from the operator for sample preparation are low (Table 2), the analysis time required for one sample is 45 min making this method difficult to adopt for large-scale screening.

CONCLUSIONS

We investigated an easy, quick, economical, and thorough scientific protocol to provide a service to farmers on individual genetic characterization. First, the automated extraction of DNA from a complex matrix such as milk has proven to be the best method to obtain high amounts of total genomic material. Second, the HPLC approach showed the highest and most efficient way to indirectly genotype the cow. It does, however, require more time for analysis. Thus, allele-specific PCR proved to be a very reliable method (PCR-RFLP and PCR-ARMS) for the characterization of the most common β-CN variants (A1/A2). The molecular approaches required less time for the analysis. In particular, the ARMS-PCR was the cheapest, fastest, and most user-friendly method. Finally, the RFLP-PCR method, which relies on the use of a restriction enzyme, required a more complicated and costly procedure to determine the cow’s genotype.

ACKNOWLEDGMENTS

Funding was provided by the projects COVID-MILK 2020 (Breeders Association of Veneto Region; ARAV, Vicenza, Italy) and MEP CASEUS (Strategie per il miglioramento dell’efficienza produttiva e qualitativa del settore lattiero-caseario, Latteria di Soligo; Farra di Soligo, Italy). We thank ARAV (Vicenza, Italy) for providing milk samples, and the National Association of Holstein, Brown Swiss, and Jersey Breeds (ANAFIBJ, Cremona, Italy), the National Association of Italian Simmental Breed (ANAPRI, Udine, Italy), and the National Association of Rendena Breed (ANARE, Trento, Italy) for technical support. The authors have not stated any conflicts of interest.

REFERENCES


**ORCIDs**

V. Vigolo https://orcid.org/0000-0001-6413-3257
M. Franzoi https://orcid.org/0000-0001-8701-5632
F. Cendron https://orcid.org/0000-0002-8221-7566
M. Penasa https://orcid.org/0000-0001-9984-8738
M. Cassandro https://orcid.org/0000-0002-8799-2870
M. De Marchi https://orcid.org/0000-0001-7814-2525