https://doi.org/10.17221/208/2020-CJAS

Use of SNPs from Illumina BovineSNP50K BeadChip v3 for imputation of microsatellite alleles for parentage verification and QTL reporting

Michaela Přibáňová, Daniela Schroffelová, David Lipovský, Josef Kučera, Vladimír Šteiger, Jarmila Hromádková, Lucie Němcová*

Czech-Moravian Breeders' Corporation, a.s., Hradištko, Czech Republic

**Corresponding author: priban@cmsch.cz*

Citation: Přibáňová M., Schroffelová D., Lipovský D., Kučera J., Šteiger V., Hromádková J., Němcová L. (2020): Use of SNPs from Illumina BovineSNP50K BeadChip v3 for imputation of microsatellite alleles for parentage verification and QTL reporting. Czech J. Anim. Sci., 65: 482–490.

Abstract: The present study aimed to test the possibility of avoiding expensive retesting of the whole parental generation for single nucleotide polymorphisms (SNPs), to provide additional analysis of microsatellites in offspring in the transitional period and to analyse the likelihood of imputation of the International Society for Animal Genetics-recommended microsatellite markers from selected SNPs. The imputation and pedigree verification of 9 520 animals (representing 84 dairy bulls, 285 dairy cows, 3 202 beef bulls and 5 949 beef cows) were analysed by the method using 9 410 SNP haplotypes (incorporating an average of 73 SNPs per haplotype). The imputation method was confirmed to allow the parentage verification of up to 87% of the analysed animals without the need for retesting. The most problematic locus was *TGLA53*, with only 78% successful imputation. Seven loci (*BM2113*, *ETH225*, *TGLA227*, *BM1824*, *SPS115*, *TGLA122* and *TGLA126*) had more than 90% imputing accuracy and success of imputation. The success of imputation also depends on the breed and the call rate of the test results. The highest imputation accuracy was found for the Holstein breed; the other six breeds had over 90% successful imputation rates, four breeds had imputation rates between 85.0 and 89.9%, and ten breeds (rarely bred in the Czech Republic) had imputation rates below 85.0%. A call rate of SNP tests lower than 90% indicates problems with haplotype construction and thus deterioration in the success of imputation. The analysis of a possibility of using all possible information from Illumina BovineSNP50K BeadChip v3 showed 109 SNPs encoding 51 quantitative trait loci markers. Haplotypes were designed for interpretation of the most important markers for diseases, exterior and performance. The most important markers for Holstein breeders were chosen as kappa- (variants A, B and E) and beta-casein (variants A1, A2), Holstein haplotypes affecting fertility (HH1, HH3, and HH4) and loci causing genetic defects, bovine leukocyte adhesion deficiency and deficiency of uridine monophosphate synthetase. The results estimated from bovine bead chips corresponded to the expected distribution of the incidence of these traits in the population and were verified by PCR-RFLP tests.

Keywords: kappa-casein; beta-casein; BLAD; DUMPS; Holstein fertility haplotypes; single nucleotide polymorphism; quantitative trait locus

Supported by the Ministry of Agriculture of the Czech Republic (Project No. QK1810253).

https://doi.org/10.17221/208/2020-CJAS

Parentage verification (PV) is an integral part of animal breeding. Methods used for parentage verification have changed in the last twenty years from the practice of immunology (e.g., comparing blood groups) to routine molecular genetics.

From the molecular genetics point of view, two main methods for parentage verification can be compared: single nucleotide polymorphism (SNP) and microsatellite markers.

As mentioned by McClure et al. (2012), microsatellite markers have successfully been used for parentage verification in multiple livestock species over the past few decades. The impact of performing PV with microsatellite markers is profound. Especially from a genetic evaluation point of view, this approach means guaranteeing accurate pedigree information (Davis and Denise 1998; Gomez-Raya et al. 2008). Currently, the most preferred method is SNP genotyping given the lower costs, higher speed and better accuracy, which are the main factors ensuring excellent routine laboratory performance (McClure et al. 2013). Rincon et al. (2011) mentioned that genotyping based on SNPs has a greater than 99% call rate, whereas individual microsatellites have a 1–5% error rate (Baruch and Weller 2008). Next-generation sequencing methods are becoming a standard part of animal breeding and the calculation of genomic breeding values. A high density of bovine bead chips allows the addition of SNP markers for quantitative trait loci, health and exterior and the selection of a set of markers with sufficient variability for parentage verification.

Some studies have aimed at comparing microsatellites and SNPs as sources for parentage verification. A study by Fernandez et al. (2013) showed that a Brazilian inbred Angus herd needed only 24 SNPs to obtain the equivalent matching probability for parental verification of 18 microsatellites. Allen et al. (2010) used 43 SNPs for PV in six breeds raised in Ireland. They provided a higher matching probability than 11 microsatellites.

The problem of incompatibility of the results of two different methods must be addressed with the implementation of new methods. When the testing of blood groups changed to the testing of microsatellites, all parental generations had to be retested, and it was necessary to perform both methods for many years.

For parentage verification, 200 SNPs with demonstrable variability within ISAG rules (International Society for Animal Genetics) were selected in our study. It is not possible to compare 200 SNP profiles with 12 microsatellite profiles of parents. Nevertheless, there is a possibility of avoiding the retesting of old animals. McClure et al. (2012; 2013) published extensive studies on the imputation of microsatellite alleles using SNP haplotypes. Based on the testing of more than 8 000 animals, 39 different breeds of cattle predicted 9 410 different SNP haplotypes (with an average of 73 SNPs per haplotype), which allowed prediction of the values of all 12 microsatellites. Some rare haplotypes were breed-specific.

The parentage verification of more than 5 000 animals of different beef and dairy cattle breeds is performed in our laboratory each year. Commercial Illumina BovineSNP50K BeadChip v3 has been used for SNP testing. The output from the Genome Studio software, which is generally used for the interpretation of bovine chips on Illumina equipment, is interpreted alphabetically (i.e., in the case of a heterozygote, the output is always in AB format), and it was necessary to create an application for assembling haplotypes and comparing them with published results. Therefore, the objective of the present study was to use the imputation method for parentage verification and determine the usability of imputed microsatellite alleles in routine laboratory conditions.

MATERIAL AND METHODS

Altogether, the SNP analyses of 48 636 animals were performed for the estimation of genomic breeding values in the years 2018 and 2020 in our laboratory. The main cohort of genotyped animals was analysed by using the Illumina BovineSNP50K BeadChip v3 (USA contains 53 217 probes). Part of the beef population was genotyped on a GeneSeek GGP Bovine 150K array (USA: 138 973 probes) directly designed for beef breeds. Although it has almost three times higher density than the commercial 50K BeadChip, several key SNPs for microsatellite imputation are missing. Therefore, the use of this kind of bead chip was halted in the first year of testing. The parentage was verified in a total of 9 520 animals: 369 dairy cattle (covering four breeds) and 9 151 beef (containing 17 breeds) (see Table 1). Only results with a call rate higher than 90% were included in the subsequent analyses. Hair roots, nasal swabs and blood have been used as sources of DNA. The purification and isolation of DNA from samples were

https://doi.org/10.17221/208/2020-CJAS

Table 1. Number of tests performed by breed

Breed	Parentage verification provided
Holstein	342
Vosgienne	15
Czech red cattle (Cervinka)	10
Jersey	2
Total dairy breed animals	369
Charolais	2 132
Aberdeen Angus	2 042
Limousine	1 781
Beef Simmental	1 698
Hereford	323
Blonde d'Aquitaine	296
Highland	240
Gasconne	174
Dexter	135
Salers	126
Parthenaise	93
Piemontesse	72
Belgianblue	22
Pinzgauer	7
Bazadais	5
Uckermärker	4
Aubrac	1
Total beef breed animals	9 151

performed by magnetic binding using the commercial Omega Mag-Bind Blood&Tissue DNA HDQ Kit 96 (Omega Bio-tek, Inc., Norcross, GA, USA, for more details see Schroffelova et al. 2018).

The microsatellite markers recommended by ISAG for parentage verification of cattle, *BM1818*, *BM1824*, *BM2113*, *ETH3*, *ETH10*, *SPS115*, *ETH225*, *INRA023*, *TGLA53*, *TGLA122*, *TGLA126* and *TGLA227*, were imputed for animals meeting the established conditions (ISAG 2010). The SNPs were selected according to McClure et al. (2013) to create haplotypes for each locus. Consequently, the SNP values were converted to Illumina AB format using The Illumina GenomeStudio data analysis software, v2.0.5 (Illumina, Inc., San Diego, CA, USA) for simplicity. To compile individual haplotypes, the internal database programme iGenetika that we designed to directly manage samples of individual animals has been modified. Pedigree data, breed and date of birth are imported automatically from the linked central animal records. Genotypes were recorded in standard Illumina format as TOP alleles, *AB* alleles and CG score. A specific computer programme primarily aimed at parentage verification of SNPs was developed (and consequently used) in our laboratory through cooperation with a commercial development company. The programme compares 200 ISAG-recommended maternal and paternal SNPs with offspring SNPs. Furthermore, the programme can find potential parents by comparing more than 700 selected "discovery" SNPs from a set of all animals in the database. The BEAGLE programme (Browning and Browning 2007) may be more useful for genotype phasing, as this software can handle both bi- and multiallelic data. However, for our needs, it was necessary to develop a programme that does not require programming, works in the Windows operating system and can be incorporated as part of the iGenetika software.

Known haplotypes were entered into the programme together with the corresponding microsatellite values. The programme automatically designed all possible variants of haplotypes in any animal in the database and selected those that matched the published haplotypes by the exclusion method. If parent genotypes are in the database, variants of the haplotypes of the parents are taken into account.

The accuracy of the imputed microsatellites has been verified in two ways: indirectly by comparison with the known genotypes of the parental pair during the parentage verification itself and by testing of microsatellite markers and comparison with imputed values. Microsatellite tests validated approximately one-third of imputations.

Significant quantitative trait loci (QTLs) were also analysed for interpretation from BeadChip. Evaluated QTLs included kappa- and beta-casein, Holstein haplotypes affecting fertility (HH1, HH3, HH4), QTLs causing genetic defects (*ITGB2*), QTLs causing bovine leukocyte adhesion deficiency (BLAD) and *UMPS* causing deficiency of uridine monophosphate synthetase (DUMPS).

Design of haplotypes was necessary to analyse kappa- and beta-casein. The list of SNPs used for interpretation is summarised in Table 2. At least nine alleles have been identified for kappa-casein (*A*, *B*, and *E* alleles are most common). Milk from cows with genotype *BB* produces approximately 10% more cheese and clots faster than milk from cows with *AA*. *E* is the allele associated with milk that does not clot to make cheese. The correctness of the SNP results was verified by the PCR-RFLP method. DNA primers were described by Barroso et al. (1998).

https://doi.org/10.17221/208/2020-CJAS

Table 2. List of interpreted quantitative trait loci (QTL) and gene/single nucleotide polymorphism (SNP) names

QTL name	Gene/SNP name
HH1 apoptotic protease-activating factor 1	*APAF1*
HH3 structural maintenance of chromosomes 2	*SMC2*
HH4 phosphoribosylglycinamide formyltransferase, phosphoribosylglycinamide synthetase, phosphoribosylaminoimidazole synthetase	*GART*
Leukocyte adhesion deficiency	*ITGB2*
Deficiency of uridine monophosphate synthase	*UMPS*
Beta-casein	*CSN2_1*
Beta-casein	*CSN2_2*
Beta-casein	*CSN2_3*
Beta-casein	*CSN2_4*
Beta-casein	*CSN2_5*
Beta-casein	*CSN2_6*
Beta-casein	*CSN2_7*
Beta-casein	*CSN2_8*
Beta-casein	*CSN2_X14711_6562*
Beta-casein	*CSN2_X14711_6687*
Beta-casein	*CSN2_X14711_6690*
Beta-casein	*CSN2_X14711_8101*
Beta-casein	*CSN2_X14711_8115*
Beta-casein	*CSN2_X14711_8163*
Beta-casein	*CSN2_X14711_8178*
Beta-casein	*CSN2_X14711_8219*
Beta-casein	*CSN2_X14711_8267*
Beta-casein	*CSN2_X14711_8356*
Kappa-casein	*CSN3_AY380228_12690*
Kappa-casein	*CSN3_AY380228_12940*
Kappa-casein	*CSN3_AY380228_12950*
Kappa-casein	*CSN3_AY380228_12951*
Kappa-casein	*CSN3_AY380228_12971*
Kappa-casein	*CSN3_AY380228_13065*
Kappa-casein	*CSN3_AY380228_13068*
Kappa-casein	*CSN3_AY380228_13096*
Kappa-casein	*CSN3_AY380228_13104_1*
Kappa-casein	*CSN3_AY380228_13111*
Kappa-casein	*CSN3_AY380228_13124*
Kappa-casein	*CSN3_AY380228_13165*

HH = Holstein haplotype

There are 13 known variants of beta-casein (A1, A2, A3, A4, B, C, D, E, F, H1, H2, I, and G). The most common variants are A1 and A2. In contrast, the least common are A3 and B. A2 represents the original variant of the allele, while A1 is a consequence variant of mutation during domestication. This variant could be a risk factor for the development of human diseases such as type 1 (insulin-dependent) diabetes mellitus (DM-1), ischaemic heart disease (IHD), sudden infant death syndrome and autism and schizophrenia. Two methods verified the correctness of the SNP results. In the PCR-RFLP test, DNA primers described by McLachlan (2006) were used for PCR amplification (Figure 1). The multiplex tetra-primer amplification (T-ARMS-PCR) method was used for retesting the other 155 samples. DNA primers described by Jaiswal 2012 were used for PCR amplification, and PCR products were analysed by using fragment analysis methods with an Applied Biosystems 3130 DNA Analyzer (Thermo Fisher Scientific, Waltham, MA, USA).

Table 3. The percent success rate of the imputation of microsatellite markers by analysed breed

Breed	Percent of successfully imputed microsatellites
Holstein	98.2
Aberdeen Angus	97.2
Limousine	95.4
Salers	94.8
Belgian blue	93.8
Beef Simmental	93.6
Charolais	92.7
Vosgienne	89.7
Hereford	86.7
Blonde d'Aquitaine	86.4
Jersey	84.2
Parthenaise	83.7
Aubrac	72.7
Bazadais	68.2
Piedmontesse	65.5
Gasconne	41.8
Dexter	39.8
Pinzgauer	33.6
Uckermärker	31.6
Highland	22.7

RESULTS AND DISCUSSION

The results associated with imputation are summarised in Table 3. There are differences in the obtained values between analysed breeds. In general,

https://doi.org/10.17221/208/2020-CJAS

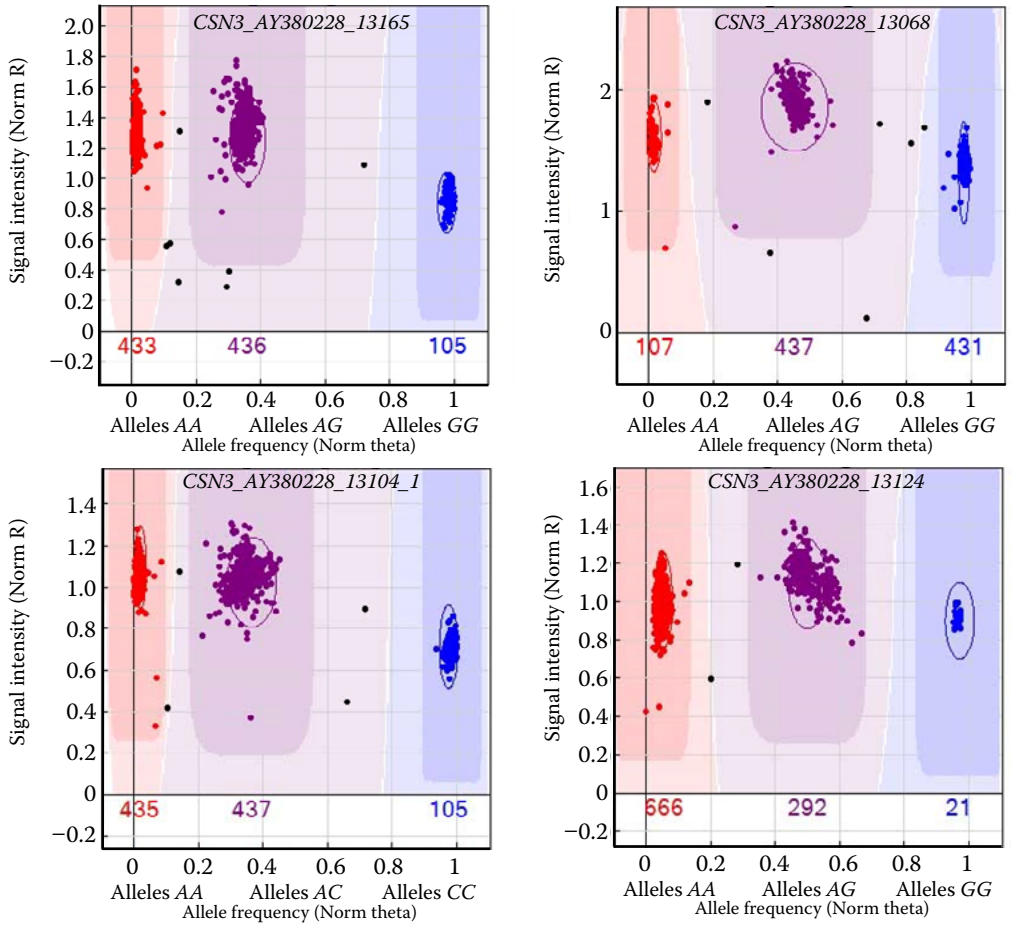

Figure 1. Allele distribution of single nucleotide polymorphisms *CSN3_AY380228_13068* (allele *A* represents A or E kappa-casein variant of milk protein, allele *G* represents kappa-casein B variant), *CSN3_AY380228_13104_1* (allele *A* represents kappa-casein A or E variant of milk protein, allele *C* represents kappa-casein B variant), *CSN3_AY380228_13124* (allele *A* represents kappa-casein A or B variant of milk protein, allele *G* represents kappa-casein E variant) and *CSN3_AY380228_13165* (allele *A* represents kappa-casein A or E variant of milk protein, allele *G* represents kappa-casein B variant)

imputation can be considered successful and used for parentage verification.

We set an 85% success rate for imputation as a limit criterion for parentage verification. The success of imputation was very low in some breeds. This was more evident in breeds with a low number of animals in our analyses and in those rarely bred in this country. The reason is probably a higher incidence of breed-specific haplotypes. In the original group tested by McClure et al. (2013), there was only a small percentage of individuals of the breeds Bazadaise (0.66%), Piedmontese (0.30%) and Gasconne (1.76%). The Dexter, Pinzgauer, Uckermärker and Highland breeds did not appear in the original group at all.

The success of imputation with individual microsatellites varied in our study. Similarly like in Sharma et al. (2018), locus *TGLA53* is the most problematic microsatellite for imputation. We were unable to predict its value in almost 17% of cases. Sharma et al. (2018) mentioned that TGLA53 had ~40% missing genotypes. The imputed value did not correspond to the actual value of the genotype in 27% of observations. Sharma et al. (2018) also found that the highest accuracy was recorded for the loci *TGLA122* and *TGLA227*. In our observation, locus *TGLA122* had the highest success of imputation (97.36%) but the second lowest imputation accuracy. The imputed value of this locus was incorrectly 143 in several cases. In our study the real value was 151 or higher. Locus *TGLA227* also had the highest imputation accuracy in our research (Table 4). Generally, we observed higher accuracies compared to the study of Sharma et al. (2018). On the other hand, the numbers of SNPs

https://doi.org/10.17221/208/2020-CJAS

Table 4. The success of imputation and imputation accuracy of individual loci

Locus	The success of imputation (%)	Imputing accuracy (%)
BM2113	88.49	97.26
ETH225	88.73	98.21
TGLA227	87.29	99.27
ETH3	79.38	98.54
ETH10	75.78	97.48
BM1824	91.61	99.23
SPS115	92.32	95.12
TGLA122	97.36	93.29
TGLA126	93.76	98.33
TGLA53	83.21	72.89
INRA23	79.62	97.89
BM1818	72.66	98.27
Arithmetic mean	85.85	95.48

used for imputation of microsatellites were lower and ranged only from nine to 24.

It is necessary to take into account the breed of an animal when the results of microsatellite imputation are analysed. McClure et al. (2013) draw attention to one phenomenon. The accuracy of microsatellite determination is affected by the fact that 83.16% of haplotypes are associated with one microsatellite allele only. Other haplotypes are associated with two or more alleles. However, this is mostly breed-specific. We also found two new microsatellite values in known haplotypes: haplotype *BABABAABBABBBBAABAAAAABAAABBABBAAAABAABABABABAAAAAABAABABAABABBAABBABA* has been interpreted as microsatellite *BM1824* with the value 188. It was observed only in one Charolais animal (McClure et al. 2013). By a microsatellite test, its interpretation as value 182 was confirmed, also in one Charolais animal. Haplotype *ABBBBBAABBBBABBAABAAAABBABABAABABBAABBBB* was also mentioned in McClure's study as equal to value 91 in microsatellite *TGLA227* in the breed Pie Rouge des Plaines, observed only once. We found it in the breed Salers, but it had the confirmed value 89.

Our imputation results indicate that Holstein is the best imputable breed. This is probably due to the lower occurrence of rare microsatellite alleles in general and because the Illumina BovineSNP50K BeadChip v3 is designed primarily for this breed.

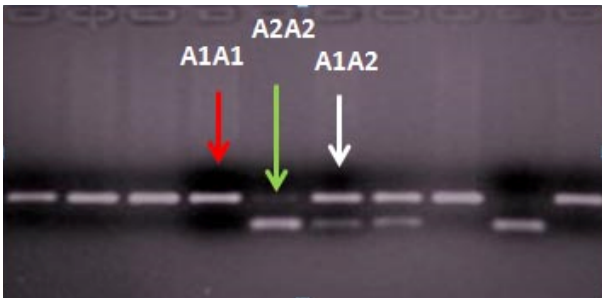

Figure 2. Fragments of three different genotypes of beta-casein

For breeds where the imputation rate is below 85%, the microsatellite alleles that require higher accuracies to design new SNP haplotypes will be predicted. In this regard, it will be necessary to test multiple generations of ancestor genotypes along with pedigree information.

Higher imputation accuracy with increased marker density was shown by Sharma et al. (2018) and McClure et al. (2012; 2013). The results are also affected by the call rate, i.e., the number of determined SNPs from the total number of probes on the chip. If the call rate is lower than 90%, the haplotype construction is problematic, and the percentage of successfully imputed microsatellites decreases proportionally.

Table 5. Allele distribution for all analysed SNPs

SNP name	Alleles	AA (%)	AB/BA (%)	BB (%)
CSN2_1	*G/A*	0.05	0.15	99.80
CSN2_2	*A*	100.00	0.00	0.00
CSN2_3	*C/G*	94.16	5.74	0.10
CSN2_4	*A/C*	86.47	13.08	0.45
CSN2_5	*A/C*	0.00	0.02	99.95
CSN2_6	*C/G*	99.90	0.10	0.00
CSN2_7	*A/C*	15.38	48.13	36.49
CSN2_8	*A/G*	0.05	0.15	99.80
CSN2_X14711_6562	*A/G*	0.15	0.25	99.60
CSN2_X14711_6687	*A/G*	0.05	0.20	99.75
CSN2_X14711_6690	*G*	0.00	0.00	100.00
CSN2_X14711_8101	*A/C*	15.38	48.08	36.54
CSN2_X14711_8115	*G*	0.00	0.00	100.00
CSN2_X14711_8163	*C*	0.00	0.00	100.00
CSN2_X14711_8178	*A/C*	86.49	13.11	0.40
CSN2_X14711_8219	*A/C*	0.00	1.00	99.00
CSN2_X14711_8267	*C/G*	94.16	5.74	0.01
CSN2_X14711_8356	*A/G*	0.05	0.10	99.85

https://doi.org/10.17221/208/2020-CJAS

A microsatellite test was automatically performed for animals where the imputed microsatellites raised doubts regarding incorrect parentage in our study. Consequently, incorrect parentage was confirmed for 72% of cases. Furthermore, microsatellites were also tested for observations where more than two microsatellites were missing after imputation, for animals from harem breeding and in cases where the semen of multiple bulls was used simultaneously for insemination. The correct father can be assigned according to an imputed genotype in more than 50% of cases. However, this possibility decreases with the number of bulls used in the harem and depending on their relationship.

Variants A1 and A2 of beta-casein (Figure 2) were identified in the group of 18 SNPs linked to *CSN2* loci for the Holstein breed and for animals tested in 2018. Only four SNPs showed applicable heterozygosity, another eight showed very low heterozygosity, and the rest were homozygous. The other SNPs probably represent other versions of beta-casein. Two SNPs, *CSN2_7* and *CSN2_X14711_8101*, showed a similar distribution of heterozygosity like the *A1* and *A2* alleles in the Holstein population (see Table 5 and Figure 3, respectively). This finding follows the results of Kaminski et al. (2007). Allelic frequencies were 48.25% of *A1* and 51.75% of *A2* allele, 15.38% animals had genotype *A1A1*, 48.13% *A1A2* and 36.49% *A2A2*. Kaminski et al. (2007) mentioned the frequency of the *A1* allele between 0.402 and 0.472 and the *A2* allele between 0.470 and 0.598. Additionally, Manga and Dvorak (2010)

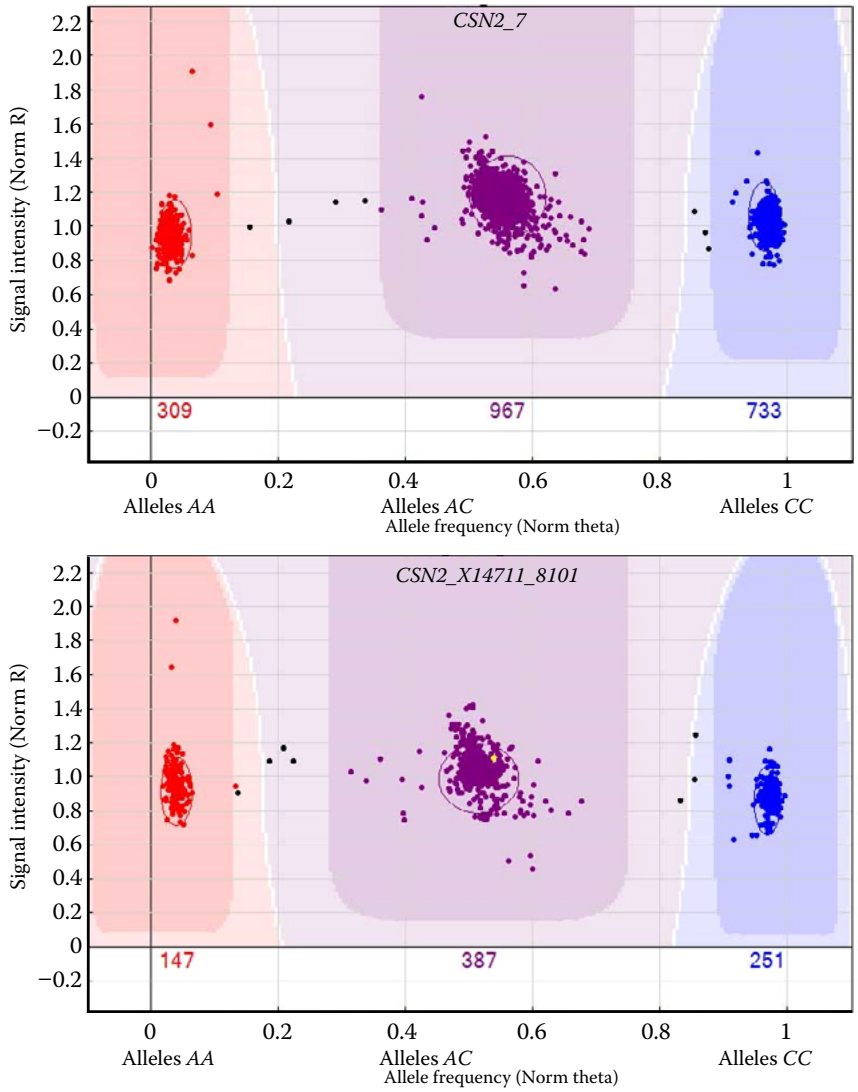

Figure 3. Allele distribution of SNPs *CSN2_7* and *CSN2_X14711_8101*. Allele *A* represents A1 variant of milk protein, and allele *C* represents A2 variant of milk protein

https://doi.org/10.17221/208/2020-CJAS

reported *A1* and *A2* allele frequencies of 0.45 and 0.55, respectively.

Two methods verified the correctness of the SNP results. One hundred fourteen animals of different genotypes were tested by PCR-RFLP test, and the other 155 animals were analysed by the T-ARMS-PCR method. The results confirmed SNPs *CSN2_7* and *CSN2_X14711_8108* as informative for detection of the *A1* and *A2* alleles. *CSN2* haplotype *GACACCCGGGGCGCACG* is equal to the *A2* allele, and *CSN2* haplotype *GACACCAGGGGAGCACG* is equal to *A1*.

A similar method for interpreting kappa-casein was used, when variants were analysed in groups of 12 SNPs associated with kappa-casein mutations in a group of 980 tested animals. Only four SNPs from this group showed heterozygosity: *CSN3_AY380228_13068*, *CSN3_AY380228_13104_1*, *CSN3_AY380228_13124* and *CSN3_AY380228_13165* (Figure 1). Frequencies in SNP *CSN3_AY380228_13068* were 107 × *AA* genotype, 437 × *AG* genotype and 431 × *GG* genotype. *CSN3_AY380228_13104_1* had 435 × AA, 437 × *AC* and 105 × *CC* genotypes. Similarly, in *CSN3_AY380228_13165* there were 433 × *AA*, 436 × *AG* and 105 × *GG*. Only SNP *CSN3_AY380228_13124* showed a different allele frequency (666 times as *AA*, 292 times as *AG* and 21 times as *GG*). This SNP was chosen as the candidate for detecting the *E* allele. Bezdicek (2007) stated that its frequency in the population was 0.094 68. Our observed frequency was 0.095 9. Allele *A* has also been estimated in other SNPs (*CSN3_AY380228_13068*). Allele *C* has been estimated in SNP *CSN3_AY380228_13104_1* and allele *G* in SNP *CSN3_AY380228_13165* as candidates for the B variant of kappa-casein. A frequency of 0.218 3 was observed in our study. Variant A of kappa-casein had a frequency of 0.685 8. According to Bezdicek (2007), the frequencies of the A, B and E variants of kappa-casein reached the values of A = 0.75, B = 0.155 2 and E = 0.094 8 in 2007.

The loci *GAGGAGGAAAAA* as genotype A, *GAGGAGAACAAG* as genotype B and *GAGGAGGAAAGA* as genotype E were constructed by haplotypes in *CSN3* with the PCR-RFLP test. PCR primers were described by Barroso et al. (1998).

Furthermore, the results of BLAD, represented by the *ITGB2* locus, have been confirmed. Six heterozygous animals were found in the analysed Czech population. Variant *A* represents the wild-type allele, whereas variant *G* represents the mutated allele. The results were confirmed by direct PCR-RFLP testing, where 140 animals, including all heterozygotes, were tested.

A and *G* alleles in the *APAF1* SNP represent HH1 mutations of the Holstein haplotype affecting fertility, which was also confirmed in our study. Allele *A* was undesirable and was present in 3.85% of animals. At the *SMC2* locus, the HH3 mutations were *A* and *G* alleles, and the *G* allele was undesirable and was present in 2.39% of animals. At the *GART* locus, the HH4 mutations were *A* and *C* alleles, the *C* allele was undesirable, and it was present in 2.08% of animals. *UMPS* has *A* and *G* alleles. Genotype *GG* represents a DUMPS-negative animal (*AG* is a DUMPS carrier).

CONCLUSION

The method of imputing microsatellite markers from SNPs can be used to verify the parentage of most common cattle breeds in this country. This method also eliminates a costly process of retesting all breeding animals or applying both testing methods for offspring. For breeds such as Dexter or Highland, it will be necessary to design their own haplotypes and determine the values of their microsatellite alleles. This method is less easily applicable to finding sires in the case of harem breeding. The known value of all imputing SNPs is necessary because with each missing data point, the accuracy of assembling the correct haplotype decreases.

In practice, we recommend retesting the animals by a direct microsatellite test in the following cases: the parentage, verified with the imputed microsatellites, seems to be doubtful, or less than nine alleles are imputed, and in the case of harem breeding and automatically in breeds where the success rate of imputation is less than 80%.

There is a possibility of interpretation for 51 individual QTL markers. We chose some haplotypes that are essential for Holstein cattle breeders: fertility haplotypes HH1, HH3 and HH4, bovine leukocyte adhesion deficiency syndrome, deficiency of uridine monophosphate synthetase and variants of beta and kappa-caseins. We found possible variants in all of the selected SNPs or SNP groups and confirmed their interpretation by PCR tests. Our study has shown that the interpretation of the results of these QTLs from the chip is possible with a high probability of accuracy.

https://doi.org/10.17221/208/2020-CJAS

Conflict of interest

The authors declare no conflict of interest.

REFERENCES

Allen AR, Talyor M, McKeown B, Curry AI, Lavery JF, Mitchell A, Hartshorne D, Fries R, Skuce RA. Compilation of a panel of informative single nucleotide polymorphisms for bovine identification in the Northern Irish cattle population. BMC Genet. 2010 Jan 25;11(1): [8 p.].

Barroso A, Dunner S, Canon J. Detection of bovine kappa-casein variants A, B, C, and E by means of polymerase chain reaction-single strand conformation polymorphism (PCR-SSCP). J Anim Sci. 1998 Jun;76(6):1535-8.

Baruch E, Weller JI. Estimation of the number of SNP genetic markers required for parentage verification. Anim Genet. 2008 Oct;39(5):474-9.

Bezdicek J. Allele and genotype frequencies of milk protein kappa-casein (CSN3) in artificial insemination bulls of Czech Fleckvieh and Holstein breed. Acta Univ Agric et Silvic Mendelianae Brun. 2007;5:17-22.

Browning SR, Browning BL. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. Am J Hum Genet. 2007 Nov;81(5):1084-97.

Davis GP, Denise SK. The impact of genetic markers on selection. J Anim Sci. 1998 Sep 1;76(9):2331-9.

Fernandez ME, Goszczynski DE, Liron JP, Villegas-Castagnasso EE, Carino MH, Ripoli MV, Rogberg-Munoz A, Posik DM, Garcia P, Giovambattista G. Comparison of the effectiveness of microsatellites and SNP panels for genetic identification, traceability and assessment of parentage in an inbred Angus herd. Genet Mol Biol. 2013 Sep;36(2):185-91.

Gomez-Raya L, Priest K, Rauw WM, Okomo-Adhiambo M, Thain D, Bruce B, Rink A, Torell R, Grellman L, Narayanan R, Beattie CW. The value of DNA paternity identification in beef cattle: Examples from Nevada's free-range ranches. J Anim Sci. 2008 Jan;86(1):17-24.

ISAG – International Society for Animal Genetics. Cattle molecular markers and parentage testing workshop [Internet]. Edinburgh, Scotland; 2010 July 28 [cited 2020]. Available from: http://www.isag.us/Docs/CattleMMPTest_CT.pdf.

Jaiswal KP. PCR based genotyping of cross breed cattle for A_1,A_2 beta casein types [master's thesis]. [Karnal, India]: National Dairy Research Institute; 2012. 67 p.

Kaminski S, Cieslinska A, Kostyra E. Polymorphism of bovine beta-casein and its potential effect on human health. J Appl Genet. 2007 Sep 1;48(3):189-98.

Manga I, Dvorak J. TaqMan allelic discrimination assay for A1 and A2 alleles of the bovine CSN2 gene. Czech J Anim Sci. 2010 Aug 19;55(8):307-12.

McClure M, Sonstegard T, Wiggans G, Van Tassell CP. Imputation of microsatellite alleles from dense SNP genotypes for parental verification. Front Genet. 2012 Aug 14;3: [9 p.].

McClure M, Sonstegard T, Wiggans G, Van Eenennaam A, Weber K, Penedo MC, Berry D, Flynn J, Garcia JF, Carmo A, Regitano L, Albuquerque M, Silva M, Machado M, Coffey M, Moore K, Boscher MY, Genestout L, Mazza R, Taylor J, Schnabel R, Simpson B, Marques E, McEwan J, Cromie A, Coutinho L, Kuehn L, Keele J, Piper E, Cook J, Williams R, Van Tassell C. Imputation of microsatellite alleles from dense SNP genotypes for parentage verification Gross multiple Bos taurus and Bos Indicus breed. Front Genet. 2013 Sep 18;4: [11 p.].

McLachlan CNS, inventor; A2 Corporation Limited, assignee. Breeding and milking cows for milk free of β-casein A1, United States patent US 7,094,949 B2. 2006 Aug 22.

Rincon G, Weber KL, Eenennaam AL, Golden BL, Medrano JF. Hot topic: Performance of bovine high-density genotyping platforms in Holsteins and Jerseys. J Dairy Sci. 2011 Dec;94(12):6116-21.

Schroffelova D, Nemcova L, Hromadkova J, Kucera J, Lipovsky D, Steiger V, Pribanova M. Optimalizace odberu alternativnich biologickych vzorku pro navaznou kvalitni izolaci genomicke DNA [Optimalization of alternative biologic samples collection for consequent DNA isolation] [certificated methodology]. Prague: CMSCH a.s.; 2018. 19 p. Czech.

Sharma A, Park JE, Park B, Park MN, Roh SH, Jung WY, Lee SH, Chai HH, Chang GW, Cho YM, Lim D. Accuracy of imputation of microsatellite markers from BovineSNP50 and BovineHD BeadChip in Hanwoo population of Korea. Genomics Inform. 2018 Mar;16(1):10-3.

Received: August 4, 2020
Accepted: December 1, 2020