Parentage Verification Using Imputed Microsatellite and SNP Data in Slovenian Brown Swiss Population

J. Obsteter[1], B. Logar[1], A. Opara[1] and J. Jenko[1,2]
[1]Agricultural Institute of Slovenia, Hacquetova ulica 17, 1000 Ljubljana, Slovenia
[2]The Roslin Institute and Royal (Dick) School of Veterinary Studies, The University of Edinburgh, Easter Bush, Midlothian, Scotland, UK
E-mail: jana.obsteter@kis.si, betka.logar@kis.si

Abstract

Parentage verification for cattle in Slovenia is routinely done using microsatellite markers (MS). With the incoming SNP data there is currently a discordance between genotypic data available for different animals, which is not suitable for parentage verification. To overcome this problem imputation of MS from SNP data was implemented for Slovenian Brown-Swiss animals. The imputation was performed adopting McClure *et al*. (2013) methodology and haplotype reference. MS imputation was performed utilising Beagle software. The overall microsatellite imputation accuracy was 91.7 %. Additionally, already confirmed parentages were re-tested using imputed MS. Out of 65 cases of two-parents testing 15.4 % and 44.6 % had 0 and 1 MS misconcordances, respectively, and the parentage was confirmed. For 40.0 % cases with ≥ 2 MS misconcordances parentage was rejected. Next, parentage verification using 800 SNPs was implemented. Out of 43 cases of one-parent testing 90.7 % and 9.3 % had 0 and 1 SNP misconcordances, respectively, therefore verifying all tested parentages. When replacing the real parent with a half-sibling or grandparent, the number of SNP misconcordances was >25 in all cases. The results suggest that MS imputation needs additional optimisation to reach required accuracies, possibly by using a haplotype reference consisting of animals that are genetically more similar to the studied Brown-Swiss population. On the other hand, verification using SNPs has proven as a reliable tool for routine use.

Key words: parentage verification, cattle, microsatellite imputation, SNP haplotype, Slovenia

Introduction

Parentage testing for cattle in Slovenia has been routinely performed using 12 microsatellite markers (MS) proposed by the International Society for Animal Genetics (ISAG). Parentage verification using MS is performed for pure-breed bull candidates, breeding material, i.e. semen, embryos and ovary cells, and for random parentage testing. Recently, we have also started to receive SNP-genotypes for some animals due to the introduction of genomic selection to Slovenian cattle breeding programmes. Genomic selection in Slovenia has first been implemented for Brown-Swiss in 2014 (Rigler *et al*., 2016) by virtue of participation in Interbull's project InterGenomics (Santus, 2011). Animals being SNP-genotyped are candidate bull calves and some other breeding animals. The transition from MS to SNP genotyping results in a discordance in the type of genotypic data available for parents and offspring, since some have MS and some SNP genotypic data, which does not allow parentage testing. MS imputation presents a feasible solution for this problem (McClure *et al*., 2013). Therefore, to avoid re-genotyping SNP-genotyped animals for MS genotypes, this study aimed to evaluate the suitability of MS imputation from SNPs for Slovenian Brown Swiss population. Additionally, parentage verification using only SNP markers was evaluated.

Materials and Methods

Data

The study population consisted of 214 Slovenian Brown-Swiss animals, 115 males and 99 females. The animals were genotyped on six different SNP-chips of various densities (Table 1). Not all of the SNP-chips contained the minimum set of the specified 880 SNPs required for optimal MS-

imputation (McClure *et al*., 2013). Therefore, SNP imputation onto the chip containing the maximum number of the minimum set SNPs was implemented. The chosen SNP-chip was GeneSeek Genomic Profiler version 4 (GGPv04) containing 30,105 SNPs in total and 878 out of the specified 880 SNPs. The SNP-imputation was performed using FIMPUTE software (Sargolzaei *et al*., 2011). The accuracy of SNP-imputation was 94.8 % reported as allelic concordance.

MS imputation and parentage verification

Microsatellites of 214 animals were imputed using haplotype reference provided by McClure *et al*. (2013). MS imputation was performed utilising Beagle 3.3.2 software (Browning and Browning, 2007). Beagle uses localised haplotype clustering for genotype phasing and imputation. The MS-imputation was run with different number of iterations (10, 100 and 1000) and the accuracy of each scenario was assessed. Afterwards, the accuracy of parentage verification using 12 imputed MS was assessed. MS data for 65 trios with parentage already confirmed based on genotyped MS data was collected and offspring's genotyped MS data was replaced with imputed MS data. Twelve MS were used for parentage verification allowing ≤ 1 MS misconcordance when testing both parents to account for possible genotyping errors.

Table 1.

Chip name	# SNPs on the chip	# animals genotyped	# SNPs for MS imputation	# SNPs for parentage testing
GGPv02	19,720	6	607	800
GGPv03	26,151	43	682	800
GGPv04	30,105	22	878	800
HDv01	76,883	4	751	800
HDv02	138,892	11	840	775
50Kv01	54,001	128	57	799

Parental verification using SNP markers

Parentage testing using SNPs was performed for 43 available SNP-genotyped parent-offspring duos. No such trios were available. Test was performed using a list of specified 800 SNPs proposed by McClure *et al*. (2015) allowing 1% SNP misconcordance rate when testing one parent. Only genotyped SNPs were used for verification. Not all utilised SNP-chips contained all the specified 800 SNPs (Table 1). However, McClure *et al*. report that ≥ 500 of the SNPs suffices for parentage verification. Additionally, the reliability and specificity of the test was assessed by replacing the SNP data of the real parent with SNP data of a half-sibling or grand-parent and the number of SNP misconcordances was inspected.

Results and Discussion

The best accuracy for MS imputation, 91.7 % reported as allelic concordance rate, was achieved implementing 1000 iterations. However, 29.0 % of animals as well as four MS (ETH10, ETH3, TGLA53, BM1818) had imputation accuracy < 90.0 %. Following up, testing the parentage by comparing offspring's imputed MS against parental genotyped MS did not reach required sensitivity. Out of 65 cases of parentage testing 15.4 % and 44.6 % had 0 and 1 MS misconcordances, respectively, and resulted in verified parentage. However, 40.0 % of the cases had >2 SNP misconcordances and resulted in rejected parentage. The type II error of the test was therefore high since 40 % of the results were false negatives. The number of rejected parentages using imputed MS was much higher than previously reported, e.g. McClure *et al*. (2013) reported ~9 % of cases with ≥2 SNP misconcordances.

Additionally, parentage was tested using SNPs only. Out of 43 cases of one parent testing 90.7 % and 9.3 % cases had 0 and 1 SNP misconcordances, respectively, which would result in all parentages being verified. Additionally, the specificity of the test was assessed by replacing the listed parent with a

half - sibling or grandparent. In all such cases the number of SNP misconcordances was > 25 thus rejecting the parentage. This provides additional evidence for SNP parentage testing using 800 SNPs as being a reliable and accurate method.

Conclusions

To overcome the parentage testing problem of genotypic data discordance, introduced by the transition from MS to SNP data, MS imputation was implemented for Brown-Swiss in Slovenia. In this study MS imputation did not reach required accuracies for a routine use in parentage testing. The procedure needs additional optimisation by increasing the number of genotyped animals or using a different haplotype reference consisting of animals more genetically similar to the studied Brown-Swiss population. On the other hand parentage testing using the 800 SNPs has proven as a reliable tool for routine use.

References

Browning, S.R. & Browning, B.L. 2007. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am. J. Hum. Genet.* 81, 1084–1097. doi:10.1086/521987

McClure, M.C., McCarthy, J., Flynn, P., Weld, M., Keane, M., O'Connell, K., Mullen, M., Waters, S. & Kearney, J.F. 2015. SNP Selection for Nationwide Parentage Verification and Identification in Beef and Dairy Cattle, in: *ICAR Technical Series 19: Performance Recording in the Genotyped World*. Presented at the ICAR Technical Meeting, Krakow, Poland, pp. 175–181.

McClure, M.C., Sonstegard, T.S., Wiggans, G.R., Van Eenennaam, A.L., Weber, K.L., Penedo, C.T., Berry, D.P., Flynn, J., Garcia, J.F., Carmo, A.S., Regitano, L.C.A., Albuquerque, M., Silva, M.V.G.B., Machado, M.A., Coffey, M., Moore, K., Boscher, M.-Y., Genestout, L., Mazza, R., Taylor, J.F., Schnabel, R.D., Simpson, B., Marques, E., McEwan, J.C., Cromie, A., Coutinho, L.L., Kuehn, L.A., Keele, J.W., Piper, E.K., Cook, J., Williams, R. & Van Tassell, C.P. 2013. Imputation of microsatellite alleles from dense SNP genotypes for parentage verification across multiple Bos taurus and Bos indicus breeds. *Front. Genet*. 4. doi:10.3389/fgene.2013.00176

Rigler, M., Sadar, M. & Potočnik, K. 2016. Izvajanje genomske selekcije pri rjavi pasmi v Sloveniji. *Rjavo Govedo* 11, 12–13.

Santus, E. 2011. Intergenomics: business rules and transition into services. *Interbull Bulletin* 2.

Sargolzaei, M., Chesnais, J.P. & Schenkel, F.S. 2011. FImpute - An efficient imputation algorithm for dairy cattle populations. *J. Dairy Sci. 94 (E-Suppl. 1)*, 421.