Discovery of hidden pedigree errors combining genomic information with the Genomic Relationship Matrix in Texel sheep

Kaseja, K; Mucha, S; Yates, John; Smith, Ed; Banos, G; Conington, JE

Published in:
Animal

DOI:
10.1016/j.animal.2022.100468

Print publication: 01/03/2022

Document Version
Publisher's PDF, also known as Version of record

Animal 16 (2022) 100468

Contents lists available at ScienceDirect

Animal
The international journal of animal biosciences

Discovery of hidden pedigree errors combining genomic information with the genomic relationship matrix in Texel sheep

K. Kaseja [a,*], S. Mucha [a], J. Yates [b], E. Smith [b], G. Banos [a], J. Conington [a]

[a] SRUC Easter Bush, Roslin Institute Building, Edinburgh EH25 9RG, UK
[b] The British Texel Sheep Society, Stoneleigh Park, Warwickshire CV8 2LG, UK

ARTICLE INFO

Article history:
Received 5 October 2021
Revised 14 January 2022
Accepted 17 January 2022

Keywords:
Genotyping
Pedigree information
Pedigree verification
Quality control
Relationship coefficient

ABSTRACT

Genomic variants such as Single Nucleotide Polymorphisms and animal pedigree are now used widely in routine genetic evaluations of livestock in many countries. The use of genomic information not only can be used to enhance the accuracy of prediction but also to verify pedigrees for animals that are extensively managed using natural mating and enabling multiple-sire mating groups to be used. By so doing, the rate of genetic gain is enhanced, and any bias associated with incorrect pedigrees is removed. This study used a set of 8 764 sheep genotypes to verify the pedigree based on both the conventional opposing homozygote method as well as a novel method when combined with the inclusion of the genomic relationship matrix (**GRM**). The genomic relationship coefficients between verified pairs of animals showed on average a relationship of 0.50 with parent, 0.25 with grandparent, 0.13 with great grandparent, 0.50 with full-sibling and 0.27 with half-sibling. Minimum obtained values from these verified pairs were then used as thresholds to determine the pedigree for unverified pairs of animals, to detect potential errors in the pedigree. Using a case study from a population partially genotyped UK sheep, the results from this study illustrate a powerful way to resolve parentage inconsistencies, when combining the conventional 'opposing homozygote' method using genomic information together with GRM for pedigree checking. In this way, previously undetected pedigree errors can be resolved.

Implications

It is important to ensure that quality control measures are put in place to reduce any bias in results from the use of genomic data in genomic evaluations for livestock, thereby enhancing the rate of genetic gain. The procedure described here was applied to correct animal pedigrees using Ovine Single Nucleotide Polymorphism DNA arrays together with the pedigree data used routinely for genetic analyses. The results indicate that with sufficient background information, pedigree errors can be captured at the parent-progeny level, even if the genotype of the parent is not known. Moreover, looking at the connections among genotypes can lead to the identification of potentially contaminated batches of samples, which also would reduce the likelihood of introducing bias into genetic evaluations.

Introduction

The difficulty in identifying correct parentage in sheep production systems is one of the key issues that reduce the efficacy of genetic improvement programmes. The reliance of on-farm, single-sire mating groups to identify sires still can lead to errors in the pedigree and if left unchecked leads to incorrect selection of candidates for breeding. Even though single-sire mating strategies are widely used, some species and breeds still use the multiple-sire mating strategies (i.e., relying on coloured crayons to identify the sire) due to logistical and management constraints which often makes it relatively easy for pedigree errors to occur.

With the availability of genotyping, correcting pedigree errors is now possible and better fits management strategies that use multiple-sire mating groups. However, due to the potential high cost, sometimes it is impossible to genotype all breeding animals, meaning that not all pedigrees can be fully verified. Incomplete or incorrect pedigrees affect the rate of genetic response, genetic gain, as well as the accuracy of genetic predictions (Israel and Weller, 2000; Long, 1990; Nwogwugwu et al., 2020). Using key genomic information from a proportion of the population in tandem with complete or incomplete pedigree data could be beneficial in reducing errors, enhancing accuracy and accelerating genetic gain (Berry et al., 2014).

The objective of the present study was to define the extent of pedigree errors and illustrate the use of the genomic relationship matrix (**GRM**) as a tool for pedigree verification in the case where

* Corresponding author.
E-mail address: Karolina.Kaseja@sruc.ac.uk (K. Kaseja).

https://doi.org/10.1016/j.animal.2022.100468
1751-7311/

not all animals are genotyped. Data from the UK Texel sheep population were used for this study, including historical pedigree information combined with genotypes obtained from commercially available Single Nucleotide Polymorphism (**SNP**) arrays.

Material and methods

A dataset of 8 764 genotypes from UK Texel sheep derived from four different Ovine Single Nucleotide Polymorphism DNA arrays ('chips') was used for this study provided by the British Texel Sheep Society. These included the Illumina OvineHD BeadChip with 606 006 SNPs (**HD**), Illumina OvineSNP50 with 54 241 SNPs (**50K**), Illumina OvineLD BeadChip with 15 000 SNPs (**LDv1**) and Illumina OvineLD BeadChip with 16 560 SNPs (**LDv2**). Animals were genotyped between January 2015 and October 2018 within previously run genomic research projects. Genotyped animals were raised in 1 438 different flocks that participate in the UK national sheep genetic evaluation.

Quality control

A quality control (**QC**) procedure was established to reject genotypes that did not meet criteria suitable for the derivation of genomic breeding values. Typical QC threshold that is used for the official UK National Genomic Evaluation for Beef and Dairy cattle removes genotypes with a call rate lower than 89.4%. The remaining genotypes were checked for possible duplicates. Firstly, all genotypes within the same chip type were examined, comparing each SNP position for which no SNPs were missing. Secondly, a check for duplicates between different chip types was performed. In order to facilitate this, a subset of 8 474 common SNPs across all four chip types was identified. SNPs across all chip arrays were rejected if minor allelic frequency <0.05, were not in Hardy–Weinberg Equilibrium (P-value > 0.01) or were with a call rate under 89.4% at the SNP level, leaving 8 119 common SNPs across all chip types. This subset of common SNPs was then used to perform the across-chip duplicate check as well as parentage check in the later stage.

In the case of confirmed duplicates, only one sample per animal was included in further analyses and the decision was made based on the density of the available genotype, with the higher density sample being included and lower density sample being discarded. For duplicated samples with different sample identification numbers, both samples were rejected as it was impossible to know which animal the sample belonged to.

Moreover, all genotypes were further examined for the level of homozygosity. Three genotypes were rejected from further analysis as being contaminated as determined by Laurie et al. (2010).

The QC procedures removed 482 genotypes, leaving 8 282 genotypes used for this study, with further genotypes being removed if they failed on the parentage verification procedure described below.

Parentage verification

There were 8 282 animal genotypes after final QC of which 938, 2 760, 2 396, and 2 188 were obtained from HD, 50K, LDv1, and LDv2 chip arrays, respectively. Among them, 4 219 belong to ewes and 4 062 belong to rams. For the genotyped animals, a five-generation pedigree was built (n = 42 587 animals), based on the information from the wider pedigree data held in the 'iTexel' database, pertaining to the British Texel Sheep Society. The panel of previously obtained common SNPs across all the available DNA arrays (n = 8 119) was used in order to verify the parentage of genotyped animals using the 'opposing homozygote' method (Hayes, 2011) for all genotyped parent-offspring pairs. These parent-offspring pairs were compared across all the common SNPs and the prospective parents were considered to fail parentage verification if the number of inconsistencies exceeded 1%, whereby both considered parent and offspring were homozygous for different alleles, as defined by Strucken et al. (2016).

Imputation

Following the QC for parentage verification, and after removing animals that failed on parentage, the remaining 7 995 animal genotypes were imputed up to a core set of 40 170 SNPs, which were the subset of most informative SNPs from OvineSNP50 chip after QC. Imputation was implemented using the software Findhap V3 (VanRaden et al., 2011). From the imputed genotypes, a genomic relationship matrix was constructed using the first method as described by (VanRaden, 2008).

Population structure and genomic relationship matrix analysis

The imputed genotypes were used to assess the population structure using Principal Component Analysis (Macciotta et al., 2010; Mucha et al., 2015) in R software (R Core Team, 2021), which was performed in order to define potential outliers which might not be connected strongly enough to the core population, indicating population stratification.

Genomic relationship coefficients ($\boldsymbol{r_{ij}}$) from GRM were used to examine the relationships between parent-offspring, grandparent-offspring, great-grandparent-offspring and great great-grandparent-offspring, as well as those among full- and half-siblings. In the first instance, these family connections were only examined for pairs of animals that were considered as having confirmed parentage using the opposing homozygote method. These GRM values were then used to determine population-specific 'thresholds' against which the verification of unconfirmed relatives was performed. In the second stage of the analysis, all suspiciously high or low (i.e., outside the estimated range) r_{ij} were examined along with the provided pedigree, to identify pairs of animals for which r_{ij} might be erroneous. A detailed pedigree was built for each animal with suspicious r_{ij} and examined further by looking at the r_{ij} between the animal itself and its close relatives (such as parents or offspring), focusing on the so-far verified (based on the opposing homozygote method) relationships. Seeing still some unexpected r_{ij} values, further examination was performed by looking into the wider pedigree and checking if there were verified grandparents or great grandparents, along with grandchildren and great grandchildren. If no close relatives were genotyped, or if any confusion was apparent, coming from the side of the non-genotyped parent, the examination focused on aunts, uncles and cousins, in order to identify the connection in the pedigree which might confirm or reject the status of the pedigree. This exercise led to a case study entailing closer examination of pedigree and genomic relationships, where the observed r_{ij} values indicated that there might be some inconsistencies in the pedigree provided by the breeder. In the case study, the sires of three potential maternal half-sibs with very high r_{ij} had not been genotyped; therefore, it was not possible to capture this error at an earlier stage with the use of the conventional opposing homozygote method.

Results and discussion

Three different 'types' of duplicated samples that emerged from the QC are summarised in Table 1. Firstly, there were multiple genotypes belonging to the same animal, which had been genotyped more than once. The similarity of the genotypes belonging

Table 1
Types of duplicated samples within available sheep genotypes.

Duplicate type	ID reported	Genotypes	% similarity	Number	Action
Genuine duplicated sample	The same ID	Exact	>99.3%	127 genotypes (reported against 61 IDs)	Best genotype per animal chosen
Mislabelled duplicated sample	The same ID	Different	36–79%	12 genotypes (reported against 6 IDs)	All genotypes rejected
Unknown duplicates	Different ID	Exact	>99.1%	18 genotypes (reported against 18 IDs)	All genotypes rejected

ID = identity number.

to the same animal was >99.3%, and a single genotype was chosen to be kept for the particular animal in the ensuing analyses. Secondly, there were also duplicated samples belonging to different animals but labelled as corresponding to a single animal. This is often the result of human error and requires checking and removing before any genomic analysis takes place. The similarity of these genotypes in this population was between 36 and 79%, clearly indicating that they could not be of the same animal. These findings confirm the importance of the proper QC before the genomic analysis, as human errors are sometimes unavoidable. Thirdly, there were unknown duplicates pertaining to genotypes that are the same (similarity over 99.1%) but had been reported against different animal IDs. This would also require special attention as including them in the dataset might cause errors during parentage verification. Moreover, unknown duplicates were cross-checked with the information available in the reported pedigree and data such as flock of origin, reported parents and date of birth to exclude potential monozygotic twins.

Table 2
Number of parent-offspring pairs verified within available sheep genotypes.

Relationship type	*N* available pairs	*N* failed[1]	*N* confirmed[2]
Dam-offspring	2 091	108 (5.2%)	1 983 (94.8%)
Sire-offspring	3 306	195 (5.9%)	3 111 (94.1%)
Dam and sire-offspring[3]	1 152	15 (1.3%)	1 037 (90.0%)

[1] Number of pairs that failed at the level of >1% (in reality >6.5%) inconsistent Single Nucleotide Polymorphisms (SNPs) between the genotype of parent and offspring.
[2] Number of pairs that passed the parentage check with the level of inconsistent SNPs not exceeding 0.1%.
[3] For pairs where both parents were checked.

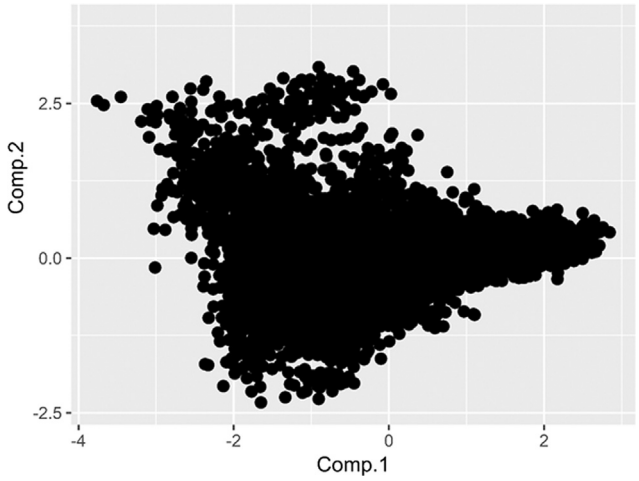

Fig. 1. Plot of first (Comp.1) and second (Comp.2) principal components of the genomix relationship matrix. Research population (sheep) clustering based on the principal component analysis, plotting first and second principal components of the genomic relationship matrix, with 14.8 and 4.7% genetic variation explained by first and second component, respectively.

The indicative overall level of pedigree error was calculated to be just over 5% for dam-offspring and just under than 6% for sire-offspring pair (Table 2), which is in line with reports on other UK livestock species (Moore et al., 2018; Visscher et al., 2002) and slightly lower than for the Irish sheep population (Berry et al., 2014).

Fig. 1 illustrates the breed composition of the studied population by plotting first and second principal components of the genomic relationship matrix, indicating no major outlier groups. This result indicates this population is mainly homogeneous; hence, genomic analyses can be performed for all the animals included in this study. Over 35% of the variation can be explained by the first ten principal components, with 14.8 and 4.7% explained by first and second components, respectively.

The genomic relationship coefficients (r_{ij}) from GRM for this population based on the animals that passed pedigree verification based on the opposing homozygote method are summarised in Table 3. Results indicate that, on average, the relationship coefficients are almost exactly as theoretically expected (Falconer and Mackay, 1996). The minimum reported r_{ij} values were used for the verification of potential relationships among animals (namely animal-parent, animal-grandparent, animal-great grandparent or siblings). The minimum r_{ij} for animal and great grandparent (0.01) was considered as being too low for verification purposes. Two pairs of maternal half-siblings were reported to have r_{ij} higher than 0.70, which is considerably higher than the calculated threshold of 0.25 (Fig. 2). The dam was verified for each animal, however, neither of their reported sires was genotyped, hence, this relation-

Table 3
Genomic relationship coefficients among sheep according to pedigree relationship type.

Relationship Type	Theoretical relationship	*N*[1]	Avg[1]	SD	Min[1]	Max[1]
Parent – offspring	0.5	4 994	0.50	0.04	0.36	0.95
Grandparent – offspring	0.25	1 829	0.25	0.05	0.11	0.54
Great grandparent – offspring	0.125	361	0.13	0.05	0.03	0.40
Great great grandparent – offspring	0.06	37	0.07	0.05	0.01	0.23
Full-siblings	0.5	238	0.50	0.05	0.34	0.64
Half-siblings	0.25	27 729	0.27	0.05	0.13	0.60

[1] *N* is the number of relationship pairs contributing to the category; Avg is the average genomic relationship coefficient; Min is the minimum genomic relationship coefficient; Max is the maximum genomic relationship coefficient.

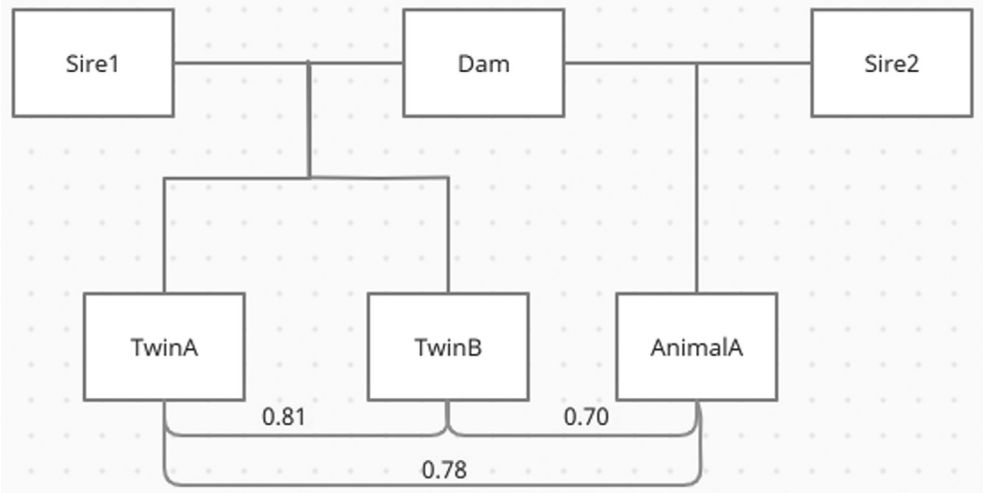

Fig. 2. Family tree as reported for sheep with suspiciously high genomic relationships (r_{ij}). Family tree for the case study, showing that the r_{ij} values between animals reported as half-siblings are much higher than expected, indication that further investigation needs to be performed.

ship could not be verified and was not captured using the conventional opposing homozygote parentage verification method. For this reason, these animals were subsequently used for the case study described below.

Case study

As shown in Fig. 2, the estimates of r_{ij} were higher than expected among three animals: these are a pair of twins (TwinA and TwinB) and their maternal half-sibling (AnimalA). All three animals were born recently on one farm, and for which both parents had been reported and registered with the breed society.

The r_{ij} was estimated to be 0.78 and 0.70 between TwinA and AnimalA, and TwinB and AnimalA, respectively, while the r_{ij} between TwinA and TwinB was 0.81 (Fig. 2). The r_{ij} values between AnimalA and the twins are much higher than the average of 0.25 for half-siblings which could be as a result of inbreeding (Moore et al., 2018) although such a big discrepancy cannot be explained by inbreeding only. Unfortunately, neither the reported sires nor the older ancestors had been genotyped, hence, it was impossible to validate the pedigree using the conventional opposing homozygote method. As the relationship coefficient r_{ij} calculated between each of the twins and AnimalA indicates a stronger relationship than that of being half-sibs, further investigation was performed to check the family connections for these animals using the GRM approach.

There were genotypes available for several progeny of both reported sires for these animals. The potential half-siblings were used to assess the level of the relationship in order to verify the sires. The reported sire for the twins (Sire1) had in total 54 genotyped offspring with valid genotypes, of which 24 passed the opposing homozygote validation for the dam. The reported sire for AnimalA (Sire2) had 34 offspring with valid genotypes of which eight had passed the validation of the mother. Only offspring with a validated dam were taken into account as at least one parent was considered to be verified for them using opposing homozygote method. In each case, the dams of these offspring were different than the dam of twins or AnimalA, meaning all the animals should be related as half-siblings with average r_{ij} ranging between 0.25 (theoretical) and 0.27 (calculated for this population) (Table 3). Table 4 summarises the average genomic relationship coefficients between TwinA, TwinB, AnimalA and the other offspring of Sire1 and Sire2. Clearly, the average relationship coefficients are consistent across TwinA, TwinB and AnimalA, and were lower for Sire1 and higher for Sire2, indicating closer relationship for all three lambs to Sire2. High SDs of r_{ij} for Sire1 indicate that there might be some animals closely related to the twins and AnimalA, but which are not necessarily half-siblings. Again, these high SDs are observable for Sire1 and all three lambs. Further examination of the relationship of other offspring of Sire1 with AnimalA and the twins indicated that among the offspring of Sire1, seven are related to AnimalA and the twins with average r_{ij} = 0.21. Looking at previously obtained thresholds for this population (Table 3), this relationship level is close to average r_{ij} between animal and grandparents. However, looking at the age of these animals, this clearly cannot be the case. However, the theoretical relationship between animal and uncle/aunt is 0.25 (Falconer and Mackay, 1996), therefore, such a relationship coefficient of 0.21 as well as the age of these animals indicated they could be aunt/uncle – animal relationships. This is in line with the provided pedigree where Sire1 was seen as a maternal grandsire of these seven animals related to AnimalA and the twins.

Table 4
Genomic relationship coefficients for sheep used in the case study.

		Avg r_{ij} (SD)[1]		
	N[1]	TwinA	TwinB	AnimalA
Sire1	26	0.05 (0.11)	0.06 (0.11)	0.05 (0.10)
Sire2	8	0.39 (0.06)	0.39 (0.06)	0.36 (0.04)

[1] N is the number of potential half-sibling with verified dam registered against each sire; Avg r_{ij} is average genomic relationship between N progeny of Sire1 or Sire2 and particular animal, followed by SD.

Thus, using the GRM enabled the interrogation of the correctness of the pedigree for the sire of the twins and also for AnimalA, revealing an incorrectly recorded sire for two animals (TwinA and TwinB) and revealing that Sire2 is more likely to be the true sire of the twins, even though none of the sires themselves were genotyped. This particular case has shown that having access to a large number of genotypes may help identify potential errors in the pedigree, even for non-genotyped animals.

Conclusion

The present study showed that additional quality control on potential duplicated samples is critically important and can reveal contaminated samples, which might have passed the standard

quality checks for call rate, minor allelic frequency, and Hardy–Weinberg equilibrium.

Despite the opposing homozygote method being used widely in the livestock genomic evaluations as a tool to verify reported pedigree, it is not sensitive enough in the case of missing genotypes for some critical animals. Careful analysis and correct interpretation of the unexpected values obtained for genomic relationship coefficients can indicate and potentially help to resolve pedigree inconsistencies even in situations when doubtful relatives are not genotyped. Routine investigation of GRM for dubious estimated relationships among specific individuals can help to identify animals that need to be genotyped and also suggest possible parents from the available pool when parentage verification has failed. Using GRM will help to validate the pedigree and improve the accuracy of selection, thereby speeding up genetic gain. Furthermore, when only a limited number of animals can be sampled for genotyping, the use of the GRM may help to identify and select the animals that might be the most critical ones in order to establish the correct pedigree. Careful investigation of GRM can also contribute to identifying errors before imputing missing parentage genotypes based on genotyped progeny as described by Berry et al. (2014). All this may not be achieved by using conventional opposing homozygous method alone, as the latter is not comparing pairs of relatives for which genotypic information is missing, while GRM uses data mining to check all the available genetic information on relatives.

As the current UK national evaluation for sheep is based purely on the pedigree provided using the conventional Best Linear Unbiased Prediction (**BLUP**) method, it is extremely important for the breed society to provide parentage verification service in order to avoid bias in the genetic evaluation. With time, it is expected that the UK national evaluation for Texel sheep will be based on Single-Step BLUP, where the information coming from genotypes is incorporated, reducing the potential bias coming from the unreliable pedigree. However, the official pedigree still needs to be adjusted where possible, as even if the genomic breeding values are more accurate when genomic information is included, the pedigree will still report the incorrect parent, introducing further errors and confusion to the breeders. Incorporating the check of genomic relationship values routinely may facilitate the identification and resolution of pedigree errors that were not captured using the conventional opposing homozygote method. By reporting and resolving such errors, the pedigree can be mended, which undoubtedly increases the accuracy of the estimated genomic breeding values, thus increase genetic gain and speed up the achievement of desirable breeding goal.

Arguably, some amendments of the pedigree may not be well received by breeders who have purchased lambs being erroneously registered as sired by potentially valuable sires. Nonetheless, when GRM along with the conventional parentage verification is used, there is more assurance about the authenticity of the pedigree provided. The new 'gold standard' of the genomic evaluation should be using the opposing homozygous method to verify the pedigree, followed by use of the GRM to deeply check the connections between animals in the pedigree.

Ethics approval

Not applicable.

Data and model availability statement

The data that support the study findings are private and confidential.

Authors ORCIDs

K. Kaseja: https://orcid.org/0000-0002-9092-8798
S. Mucha: https://orcid.org/0000-0002-2863-5695
J. Yates: https://orcid.org/0000-0002-9339-7275
E. Smith: https://orcid.org/0000-0003-0695-5601
G. Banos: https://orcid.org/0000-0002-7674-858X
J. Conington: https://orcid.org/0000-0002-2387-3555

Author' contributions

KK undertook the analysis and wrote the paper. JC and GB were principal investigators (PIs) for the original research project, collated the data and wrote the paper. SM, JY and ES reviewed and wrote the manuscript. ES also collected the data. All authors have read and approved the final manuscript.

Declaration of interest

None.

Acknowledgements

The authors wish to thank Texel Sheep Society for providing access to their pedigree and genotype database. Genotype data were collected under InnovateUK / (BBSRC) project reference number 131791 (BB/M018377/1) and also project reference number 102646 (BB/M02833X/1).

Financial support statement

This work was supported by the InnovateUK/(BBSRC) project reference number 131791 (BB/M018377/1) and also project reference number 102646 (BB/M02833X/1).

References

Berry, D.P., McHugh, N., Randles, S., Wall, E., McDermott, K., Sargolzaei, M., O'Brien, A.C., 2014. Imputation of Ungenotyped Parental Genotypes in Dairy and Beef Cattle from Progeny Genotypes. Animal 8, 895–903.

Falconer, D.S., Mackay, T.F.C., 1996. Introduction to Quantitative Genetics. Longmans Green, Harlow, Essex, UK.

Hayes, B.J., 2011. Technical Note: Efficient Parentage Assignment and Pedigree Reconstruction with Dense Single Nucleotide Polymorphism Data. Journal of Dairy Science 94, 2114–2117.

Israel, C., Weller, J.I., 2000. Effect of Misidentification on Genetic Gain and Estimation of Breeding Value in Dairy Cattle Populations. Journal of Dairy Science 83, 181–187.

Laurie, C.C., Doheny, K.F., Mirel, D.B., Pugh, E.W., Bierut, L.J., Bhangale, T., Boehm, F., Caporaso, N.E., Cornelis, M.C., Edenberg, H.J., Gabriel, S.B., Harris, E.L., Hu, F.B., Jacobs, K.B., Kraft, P., Landi, M.T., Lumley, T., Manolio, T.A., McHugh, C., Painter, I., Paschall, J., Rice, J.P., Rice, K.M., Zheng, X., Weir, B.S.GENEVA Investigators, 2010. Quality control and quality assurance in genotypic data for genome-wide association studies. Genetic Epidemiology 34 (6), 591–602. https://doi.org/10.1002/gepi.20516. PMID: 20718045; PMCID: PMC3061487.

Long, T.E., 1990. Effects of Errors in Pedigree on Three Methods of Estimating Breeding Value for Litter Size, Backfat and Average Daily Gain in Swine. University of Nebraska, U.S. Department of Agriculture, Lincoln, NE, USA.

Macciotta, N.P.P., Gaspa, G., Steri, R., Nicolazzi, E.L., Dimauro, C., Piermati, C., Cappio-Borlino, A., 2010. Using Eigenvalues as Variance Priors in the Prediction of Genomic Breeding Values by Principal Component Analysis. Journal of Dairy Science 93, 2765–2774.

Moore, K.L., Vilela, C., Kaseja, K., Mrode, R., Coffey, M., 2018. Forensic Use of the Genomic Relationship Matrix to Validate and Discover Livestock Pedigrees. Journal of Animal Science 97, 35–42.

Mucha, S., Mrode, R., MacLaren-Lee, I., Coffey, M., Conington, J., 2015. Estimation of Genomic Breeding Values for Milk Yield in UK Dairy Goats. Journal of Dairy Science 98, 8201–8208.

Nwogwugwu, C.P., Kim, Y., Chung, Y.J., Jang, S.B., Roh, S.H., Kim, S., Lee, J.H., Choi, T.J., 2020. Effect of Errors in Pedigree on the Accuracy of Estimated Breeding Value for Carcass Traits in Korean Hanwoo Cattle. Asian-Australasian Journal of Animal Sciences 33, 1057–1067.

Strucken, E.M. et al., 2016. How Many 311 Markers Are Enough? Factors Influencing Parentage Testing in Different Livestock Populations. Journal of Animal Breeding and Genetics 133, 13–23.

R Core Team, 2021. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria https://www.R316project.org/.
VanRaden, P.M., O'Connell, J.R., Wiggans, G.R., Weigel, K.A., 2011. Genomic Evaluations with Many More Genotypes. Genetics Selection Evolution 43, 1–11.
VanRaden, P.M., 2008. Efficient Methods to Compute Genomic Predictions. Journal of Dairy Science 91, 4414–4423.
Visscher, P.M., Woolliams, J.A., Smith, D., Williams, J.L., 2002. Estimation of Pedigree Errors in the UK Dairy Population Using Microsatellite Markers and the Impact on Selection. Journal of Dairy Science 85, 2368–2375.