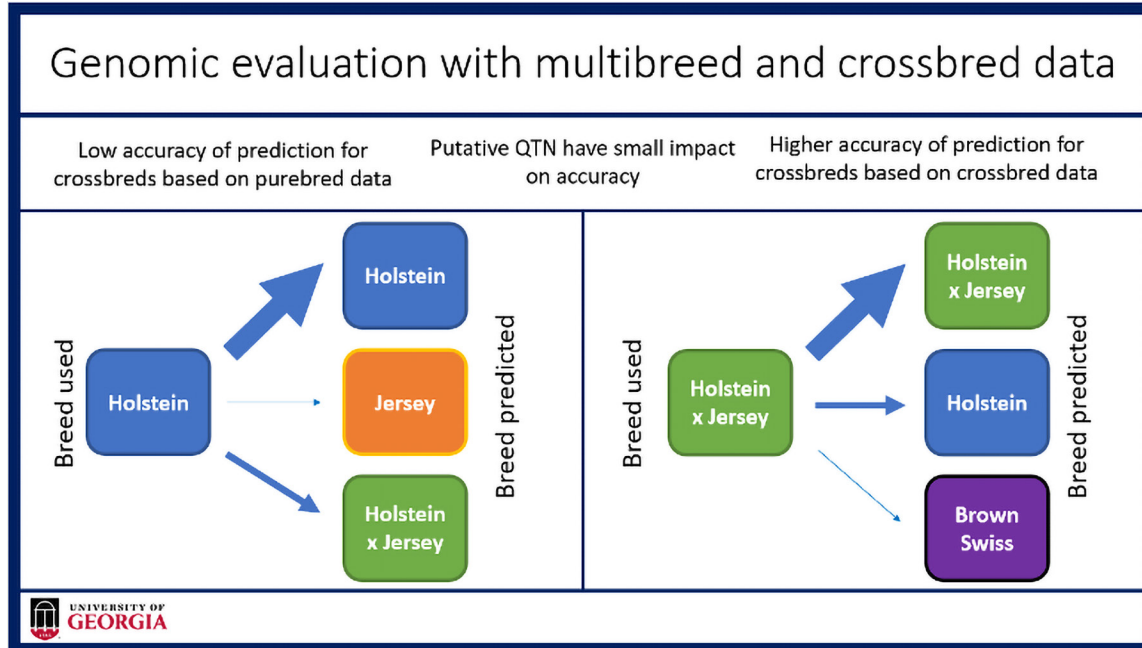


Genomic evaluation with multibreed and crossbred data*

I. Misztal,† Y. Steyn, and D. A. L. Lourenco

Graphical Abstract

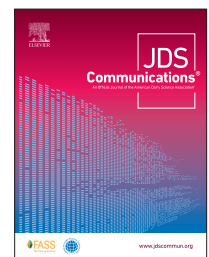


Summary

This paper reviewed studies in genomic selection of multibreed and crossbred populations. Across-breed predictions are poor. Use of putative sequence data seems to have a small impact on genomic accuracy. An accurate evaluation of any breed type requires a reference population that includes that breed type.

Highlights

- We found low accuracy of genomic evaluation of crossbreds based on purebred data.
- We found higher accuracy for crossbreds based on crossbred data.
- Use of putative sequence variants had a small impact on genomic accuracy.



*Presented as part of the Breeding and Genetics Symposium: Crossbreeding at the ADSA Annual Meeting, July 2021. Department of Animal and Dairy Science, University of Georgia, 425 River Road, Athens 30602. †Corresponding author: ignacy@uga.edu. © 2022, The Authors. Published by Elsevier Inc. and FASS Inc. on behalf of the American Dairy Science Association®. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>). Received October 08, 2021. Accepted November 21, 2021.

Genomic evaluation with multibreed and crossbred data*

I. Misztal,† , Y. Steyn, , and D. A. L. Lourenco 

Abstract: Several types of multibreed genomic evaluation are in use. These include evaluation of crossbreds based on purebred SNP effects, joint evaluation of all purebreds and crossbreds with a single additive effect, and treating each purebred and crossbred group as a separate trait. Additionally, putative quantitative trait nucleotides can be exploited to increase the accuracy of prediction. Existing studies indicate that the prediction of crossbreds based on purebred data has low accuracy, that a joint evaluation can potentially provide accurate evaluations for crossbreds but could lower accuracy for purebreds compared with single-breed evaluations, and that the use of putative quantitative trait nucleotides only marginally increases the accuracy. One hypothesis is that genomic selection is based on estimation of clusters of independent chromosome segments. Subsequently, predicting a particular group type would require a reference population of the same type, and crosses with same breed percentage but different type (F_1 vs. F_2) would, at best, use separate reference populations. The genomic selection of multibreed population is still an active research topic.

Genotyping in the US dairy industry is now widespread. Although the initial purpose of genotyping was for genetic selection, the cost has decreased enough for many farmers to use genotyping for management purposes. As of September 2020, genotyped dairy cattle in the United States exceeded 4.5 million, of which 86% are Holsteins, 12% are Jerseys, and the remainder comprises other small breeds and crossbred animals (CDCB, 2020). Approximately 2% of dairy genotypes were of crossbred animals (VanRaden et al., 2020). Due to the increased demand to accommodate the financial investment already made by genotyping crossbred animals, they were included in national genomic evaluations in 2019 (VanRaden et al., 2020). However, the genomic evaluation is separate for each purebred and for all the crossbreds, although the nongenomic evaluation that provides input for the genomic evaluation is joint for all animals.

Instead of running separate evaluations, a comprehensive evaluation including all the purebreds and crossbreds could provide many benefits. The data on each purebred could improve evaluations of other purebreds and crossbreds. The data on crossbreds could improve the evaluations of crossbreds and, if the number of genotyped crossbreds increases, could even improve the evaluation of purebreds. The requirements of the comprehensive evaluation include biases not greater and reliability not smaller than those from separate evaluations. The purpose of this study was to review current studies in multibreed evaluations, evaluate methods in use, state which assumptions they use, and determine which methods show the most promise. Although accurate imputation may be a barrier to accurate crossbred genomic prediction, especially in databases holding only one of the pure parental breeds, the imputation is not part of this review.

Several general methods were used for the genomic evaluation of multibreed and crossbred populations; see Lund et al. (2014) for an earlier review. The general methods include weighting purebred SNP effects by breed proportions for crossbred animals (VanRaden

et al., 2020; Karaman et al., 2021; Steyn et al., 2021), the use of a single genomic relationship matrix containing all breeds and breed groups (Olson et al., 2012; Hozé et al., 2014; Steyn et al., 2019), and treating different breeds as different traits in a multitrait model (Olson et al., 2012; Calus et al., 2018; van den Berg et al., 2020). Many modifications were applied to those general methods including adjusting for breed-wise allele frequencies and accounting for pedigree-based breed proportions (Makgahlela et al., 2014), taking breed-specific effects or breed origin of alleles into account (Lopes et al., 2017; Karaman et al., 2021), accounting for linkage disequilibrium (LD; Zhou et al., 2014; Rahimi et al., 2020), and using 2 genomic matrices, one of which contains the most important markers and the second the remaining markers (Khansefid et al., 2014; Raymond et al., 2020). In particular, the use of quantitative trait nucleotides (QTN) derived from sequence data could yield nearly 100% accuracy across breeds (Goddard, 2017). Some of these methods are described in greater detail below.

If phenotypes of crossbreds are relatively scarce, one option is the indirect prediction for crossbreds based on purebred data by weighting the SNP effect based on the breed proportion of each pure breed (VanRaden et al., 2020; Karaman et al., 2021; Steyn et al., 2021). For example, consider a crossbred animal composed of 50% Holstein, 25% Jersey, and 25% Brown Swiss. The indirect prediction is the sum of the SNP effect based on Holstein weighted by 0.5, the SNP effect based on Jersey weighted by 0.25, and the SNP effect based on Brown Swiss weighted by 0.25. This method indirectly assumes that SNP are markers to QTL, and the value of markers for crossbreds is a weighted average of those of the founder pure breeds.

A joint multibreed evaluation uses a single genetic effect for all the genotyped animals, either an additive animal effect where all genotyped animals share a single genomic relationship matrix or the same SNP effects across all breeds (Olson et al., 2012; Hozé et al., 2014; Steyn et al., 2019). The joint evaluation indirectly relies

*Presented as part of the Breeding and Genetics Symposium: Crossbreeding at the ADSA Annual Meeting, July 2021. Department of Animal and Dairy Science, University of Georgia, 425 River Road, Athens 30602. †Corresponding author: ignacy@uga.edu. © 2022, The Authors. Published by Elsevier Inc. and Fass Inc. on behalf of the American Dairy Science Association®. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>). Received October 08, 2021. Accepted November 21, 2021.

on 2 potentially exclusive assumptions: (1) that QTL effects—and by extension SNP effects—are similar for all breeds, and (2) that a medium-density SNP chip (e.g., 50k) is large enough to account for all breeds including the crossbreds. The joint model should include effects that account for breed differences and, in the case of crossbreds, for heterosis and recombination loss. Accounting for different allele frequencies for breeds could be used for compatibility of genomic and pedigree relationships (Makgahlela et al., 2014).

Another option in the multibreed evaluation is treating each breed as a different trait (Olson et al., 2012; Calus et al., 2018; van den Berg et al., 2020). A multibreed model allows for an arbitrary correlation structure among the breed types including null correlations among the purebreds and nonzero correlations between the crossbreds and their parental breeds. The multibreed model is common in the pig and poultry industry, where crossbreds are limited to F_1 (Christensen et al., 2014) and may be applicable in dairy production systems when crosses beyond F_1 are not of interest.

A modification to any of the evaluation methods would be identification and use of QTN, usually derived from sequence data. It assumes indirectly that while many QTN are similar in location and magnitude by breed, markers to QTN vary in phase per breed, resulting in different estimates of SNP effects (van den Berg et al., 2016). Thus, if a large number of QTN are identified and their effects well estimated, the genomic predictions can be accurate for other purebreds and crossbreds. The success with the use of QTN depends not only on their identification but also on assigning a higher variance to each QTN to avoid shrinkage (Brøndum et al., 2015; Fragomeni et al., 2017).

Many studies have been carried out using the methods above, some with fairly large data sets. Their results allow us to evaluate the assumptions in the methods and subsequently give a better picture of how genomic selection works, especially in the multibreed context. For example, accounting for different allele frequencies was not found to be effective but accounting for compatibility of pedigree and genomic information was crucial (Lourenco et al., 2016). Perhaps the most obvious result is that genomic prediction across breeds is not effective. For example, Olson et al. (2012) looked at the efficiency of genomic prediction for Holsteins and Jerseys based on different reference populations. They concluded that the prediction reliabilities for one breed based on the reference population of another breed are low, and combining reference populations does not increase reliabilities over a single breed. Similar conclusions were reached in beef (Kachman et al., 2013), sheep (Moghaddar et al., 2014, 2019), and other dairy populations, even when using highly specialized techniques (Raymond et al., 2018; Karaman et al., 2021; Meuwissen et al., 2021).

Some insight on why across-breed predictions are poor can be gained from a study by Steyn et al. (2019), who simulated 5 different populations. Despite each population sharing identical QTN locations and substitution values, the across-breed accuracy was low. Their study showed that genomic prediction in populations with a small effective population size, such as dairy, is primarily by mechanisms other than by LD to QTN.

The method of breed proportions was applied to the US dairy population (VanRaden et al., 2020). For prediction of SNP effects per breed, only genotypes of purebreds were used, as the use of genotypes of crossbreds reduced reliabilities for the purebreds. In general, the measure of reliability for the crossbreds was low,

exceeding the parent average by at most 0.05, and the measures were ≤ 0.02 for SCS and daughter pregnancy rate. Karaman et al. (2021) showed that this method was less accurate than a simple joint relationship matrix containing all animals or including all animals while accounting for breed origin of alleles. Low reliability of the method suggests that either SNP effects of the crossbreds are not simply averages of their respective purebred ancestors or that prediction is mainly by factors other than markers to QTN.

Although the data on crossbreds were not directly used in predictions by breed proportion, all data (genotypes and phenotypes) are used in the joint multibreed evaluations. Winkelman et al. (2015) used such a method for the evaluation of the New Zealand dairy population consisting of Holsteins, Jerseys, and their crosses. The accuracies of prediction were high for purebreds and crossbreds and, for some traits, the accuracies were higher for the crossbreds than for the purebreds. The study suggests that high accuracy for crossbreds would be achieved only with crossbred data in the reference population. Similar conclusions were reached by Khansefid et al. (2020). In a study by Steyn et al. (2021), the population included 90k Holsteins, 9k Jerseys, and <2k crosses. Predictivity for one purebred was high if that breed was included in the reference population, it was low when it was not, and using a combined reference of 2 purebreds reduced the predictivity compared with using only 1. For crossbreds, predictivity was similar whether the reference population was any single purebred, both purebreds, or only a cross. Because accuracy with the cross population was the same as with much larger purebred populations, it seems that an accurate prediction of a breed type requires the reference population of that breed type.

With many breeds, the number of possible types of crossbreds can be very high, and there is a temptation to include all breeds and crossbreds in a joint multibreed evaluation. In beef, such an evaluation for the Irish breeds includes 47 breeds plus crossbreds (Berry et al., 2016; Mäntysaari et al., 2017) and, for Igenity, it includes at least 10 breeds plus crossbreds (Saatchi et al., 2014). While the first evaluation used a ~50k SNP chip, the second used <3k selected SNP. Assuming that predictivity from breed to breed is low, which means nearly separate SNP solutions for each breed, a joint evaluation would result in effectively using a reduced number of SNP per breed, potentially reducing accuracy compared with purebred evaluations. Steyn et al. (2019) found that the reliability is lower compared with a single breed evaluation when a joint evaluation uses many breeds with small number of randomly selected SNP.

Recently, Cesarani et al. (2021) ran a 5-breed single-step evaluation using complete US data with over 3 million genotyped animals. The breeds were of greatly unbalanced sizes, with Holstein comprising 90% of the data and Ayrshire comprising 0.1% of the data. With appropriate options [core animals in the algorithm for proven and young (APY) having enough animals from each breed], the reliabilities were as good as those from single-breed evaluations. It is possible that no reduction in accuracy despite many breeds was aided by the APY algorithm, where the unneeded impact of one breed on the other was limited by the reduction in dimensionality. A subsequent study will determine whether including the crossbred data with such a method will provide acceptable evaluations for the crossbreds without negatively affecting the purebreds.

Using putative sequence information had a positive but still very small impact on across-breed reliabilities (Meuwissen et al.,

2021). In the latter study, relatively large populations of Australian cattle were used to identify putative QTN and evaluate their effect on reliability within and across breeds. Methods included BayesR, which can assign a different variance to each SNP, reducing shrinkage to properly identified QTN, and an improved method to BayesR named BayesGC. Reliabilities of prediction for single breeds were high with genomic BLUP (**GBLUP**; 0.71 for Holsteins and 0.64 for Jersey), which mostly ignores the putative QTN, and increased by, at most, 0.01 with BayesR and 0.04 with BayesGC, which potentially account for putative QTN. Reliabilities for an unrelated breed (Australian Red Cow) were low (0.23 for GBLUP and 0.27 for BayesR), indicating the small value of putative QTN on prediction across breeds. A study by Khansefid et al. (2020) looked at the accuracy of genomic selection for various purebreds and crossbreds, with 3 types of SNP information and analyses either by GBLUP or by a Bayesian model. Predicted accuracies for crossbreds were the highest with the crossbred data. The Bayesian model improved prediction accuracy for the crossbreds based on the purebred data but less so with purebred and crossbred data; the improvements were smaller with SNP information that included putative QTN variants.

If accuracies are calculated for all breeds or all types of crossbreds together, they may be inflated if the model does not account for the breed type (Bermann et al., 2021). The breed type is defined here as either a purebred or a specific type of crossbred, where F_1 and F_2 are different breed types. Assume that accuracies are calculated by predictivity, which is a correlation of genomic EBV and phenotypes adjusted for fixed effects (Legarra et al., 2008). If the fixed effects do not include the breed effect, the predictivity mostly predicts the missing breed effect. In the extreme case, the accuracy calculated as the ratio of predictivity and a square root of heritability may be >1 . Accounting for breed types of crossbreds may be hard because their number can be high (e.g., F_1 , F_2 , various reciprocal crosses, 3-way crosses). Without accounting for breed types, genomic EBV would predict the breed type well, but prediction within a breed type may be poor.

The cited studies indicate that (1) there is little predictivity from breed to breed even if QTL are identical; (2) the accuracy of crossbreds based on purebreds is low; (3) using crossbred reference populations to predict crossbreds seems to be successful; and (4) using putative QTN only marginally improved accuracy. This can be explained by the nature of genomic selection in populations with a small effective population size. Genomic selection is based primarily on predicting clusters of the most common chromosome segments in each population (Daetwyler et al., 2010; Pocrnic et al., 2019), and less on LD to QTN. Additionally, the number of detectable QTN with estimable effect and without pleiotropy for traits under selection is likely small (Georges et al., 2019; Misztal et al., 2020). Thus, it is important for all breeds and types of crossbreds to be present in the reference population. Extending the idea, accurate prediction of crosses requires a reference population that includes the relevant cross types. Genomic multibreed evaluation is still an active research topic.

References

- Bermann, M., D. Lourenco, V. Breen, R. Hawken, F. Brito Lopes, and I. Misztal. 2021. Modeling genetic differences of combined broiler chicken populations in single-step GBLUP. *J. Anim. Sci.* 99:1–7. <https://doi.org/10.1093/jas/skab056>.
- Berry, D. P., F. Kearney, R. Evans, E. Wall, and A. Cromie. 2016. Genomic evaluations in dairy cattle, beef cattle, and sheep in Ireland. *J. Anim. Sci.* 94(Suppl. 4):8–9. <https://doi.org/10.2527/jas2016.94supplement48a>.
- Brøndum, R. F., G. Su, L. Janss, G. Sahana, B. Gulbrandsen, D. Boichard, and M. S. Lund. 2015. Quantitative trait loci markers derived from whole genome sequence data increases the reliability of genomic prediction. *J. Dairy Sci.* 98:4107–4116. <https://doi.org/10.3168/jds.2014-9005>.
- Calus, M. P., M. Goddard, Y. Wientjes, P. Bowman, and B. Hayes. 2018. Multibreed genomic prediction using multitrait genomic residual maximum likelihood and multitask Bayesian variable selection. *J. Dairy Sci.* 101:4279–4294. <https://doi.org/10.3168/jds.2017-13366>.
- CDCB. 2020. Council on Dairy Cattle Breeding Activity Report Oct 19/ Sep 2020. Accessed Dec. 2, 2020. https://www.uscdcb.com/wp-content/uploads/2020/10/2020-CDCB-Activity-Report_103020_lowres.pdf.
- Cesarani, A., D. Lourenco, Y. Masuda, S. Tsuruta, A. Legarra, E. Nicolazzi, P. Vanraden, and I. Misztal. 2021. Multi-breed genomic evaluation for dairy cattle in the US using single-step GBLUP. Page 207 in Book of Abstracts of the 72nd EAAP Annu. Mtg. EAAP.
- Christensen, O. F., P. Madsen, B. Nielsen, and G. Su. 2014. Genomic evaluation of both purebred and crossbred performances. *Genet. Sel. Evol.* 46:23. <https://doi.org/10.1186/1297-9686-46-23>.
- Daetwyler, H. D., R. Pong-Wong, B. Villanueva, and J. A. Woolliams. 2010. The impact of genetic architecture on genome-wide evaluation methods. *Genetics* 185:1021–1031. <https://doi.org/10.1534/genetics.110.116855>.
- Fragomeni, B. O., D. A. L. Lourenco, Y. Masuda, A. Legarra, and I. Misztal. 2017. Incorporation of causative quantitative trait nucleotides in single-step GBLUP. *Genet. Sel. Evol.* 49:59. <https://doi.org/10.1186/s12711-017-0335-0>.
- Georges, M., C. Charlier, and B. Hayes. 2019. Harnessing genomic information for livestock improvement. *Nat. Rev. Genet.* 20:135–156. <https://doi.org/10.1038/s41576-018-0082-2>.
- Goddard, M. E. 2017. Can we make genomic selection 100% accurate? *J. Anim. Breed. Genet.* 134:287–288. <https://doi.org/10.1111/jbg.12281>.
- Hozé, C., S. Fritz, F. Phocas, D. Boichard, V. Ducrocq, and P. Croiseau. 2014. Efficiency of multi-breed genomic selection for dairy cattle breeds with different sizes of reference population. *J. Dairy Sci.* 97:3918–3929. <https://doi.org/10.3168/jds.2013-7761>.
- Kachman, S. D., M. L. Spangler, G. L. Bennett, K. J. Hanford, L. A. Kuehn, W. M. Snelling, R. M. Thallman, M. Saatchi, D. J. Garrick, R. D. Schnabel, J. F. Taylor, and E. J. Pollak. 2013. Comparison of molecular breeding values based on within- and across-breed training in beef cattle. *Genet. Sel. Evol.* 45:30. <https://doi.org/10.1186/1297-9686-45-30>.
- Karaman, E., G. Su, I. Croue, and M. S. Lund. 2021. Genomic prediction using a reference population of multiple pure breeds and admixed individuals. *Genet. Sel. Evol.* 53:46. <https://doi.org/10.1186/s12711-021-00637-y>.
- Khansefid, M., M. E. Goddard, M. Haile-Mariam, K. V. Konstantinov, C. Schrooten, G. de Jong, E. G. Jewell, E. O'Connor, J. E. Pryce, H. D. Daetwyler, and I. M. MacLeod. 2020. Improving genomic prediction of crossbred and purebred dairy cattle. *Front. Genet.* 11:598580. <https://doi.org/10.3389/fgene.2020.598580>.
- Khansefid, M., J. E. Pryce, S. Bolormaa, S. P. Miller, Z. Wang, C. Li, and M. E. Goddard. 2014. Estimation of genomic breeding values for residual feed intake in a multibreed cattle population. *J. Anim. Sci.* 92:3270–3283. <https://doi.org/10.2527/jas.2014-7375>.
- Legarra, A., C. Robert-Granié, E. Manfredi, and J.-M. Elsen. 2008. Performance of genomic selection in mice. *Genetics* 180:611–618. <https://doi.org/10.1534/genetics.108.088575>.
- Lopes, M. S., H. Bovenhuis, A. M. Hidalgo, J. A. M. van Arendonk, E. F. Knol, and J. W. M. Bastiaansen. 2017. Genomic selection for crossbred performance accounting for breed-specific effects. *Genet. Sel. Evol.* 49:51. <https://doi.org/10.1186/s12711-017-0328-z>.
- Lourenco, D. A., S. Tsuruta, B. O. Fragomeni, C. Y. Chen, W. O. Herring, and I. Misztal. 2016. Crossbreed evaluations in single-step genomic best linear unbiased predictor using adjusted realized relationship matrices. *J. Anim. Sci.* 94:909–919. <https://doi.org/10.2527/jas.2015-9748>.
- Lund, M. S., G. Su, L. Janss, B. Gulbrandsen, and R. F. Brøndum. 2014. Genomic evaluation of cattle in a multi-breed context. *Livest. Sci.* 166:101–110. <https://doi.org/10.1016/j.livsci.2014.05.008>.
- Makgahlela, M. L., I. Strandén, U. Nielsen, M. Sillanpää, and E. Mäntysaari. 2014. Using the unified relationship matrix adjusted by breed-wise allele

- frequencies in genomic evaluation of a multibreed population. *J. Dairy Sci.* 97:1117–1127. <https://doi.org/10.3168/jds.2013-7167>.
- Mäntysaari, E. A., R. D. Evans, and I. Strandén. 2017. Efficient single-step genomic evaluation for a multibreed beef cattle population having many genotyped animals. *J. Anim. Sci.* 95:4728–4737. <https://doi.org/10.2527/jas2017.1912>.
- Meuwissen, T., I. van den Berg, and M. Goddard. 2021. On the use of whole-genome sequence data for across-breed genomic prediction and fine-scale mapping of QTL. *Genet. Sel. Evol.* 53:19. <https://doi.org/10.1186/s12711-021-00607-4>.
- Misztal, I., D. Lourenco, and A. Legarra. 2020. Current status of genomic evaluation. *J. Anim. Sci.* 98:skaa101. <https://doi.org/10.1093/jas/skaa101>.
- Moghaddar, N., M. Khansefid, J. H. J. van der Werf, S. Bolormaa, N. Duijvesteijn, S. A. Clark, A. A. Swan, H. D. Daetwyler, and I. M. MacLeod. 2019. Genomic prediction based on selected variants from imputed whole-genome sequence data in Australian sheep populations. *Genet. Sel. Evol.* 51:72. <https://doi.org/10.1186/s12711-019-0514-2>.
- Moghaddar, N., A. A. Swan, and J. H. Van Der Werf. 2014. Comparing genomic prediction accuracy from purebred, crossbred and combined purebred and crossbred reference populations in sheep. *Genet. Sel. Evol.* 46:58. <https://doi.org/10.1186/s12711-014-0058-4>.
- Olson, K. M., P. VanRaden, and M. Tooker. 2012. Multibreed genomic evaluations using purebred Holsteins, Jerseys, and Brown Swiss. *J. Dairy Sci.* 95:5378–5383. <https://doi.org/10.3168/jds.2011-5006>.
- Pocrnic, I., D. A. L. Lourenco, Y. Masuda, and I. Misztal. 2019. Accuracy of genomic BLUP when considering a genomic relationship matrix based on the number of the largest eigenvalues: A simulation study. *Genet. Sel. Evol.* 51:75. <https://doi.org/10.1186/s12711-019-0516-0>.
- Rahimi, S. M., A. Rashidi, and H. Esfandyari. 2020. Accounting for differences in linkage disequilibrium in multi-breed genomic prediction. *Livest. Sci.* 240:104165. <https://doi.org/10.1016/j.livsci.2020.104165>.
- Raymond, B., A. C. Bouwman, C. Schrooten, J. Houwing-Duistermaat, and R. F. Veerkamp. 2018. Utility of whole-genome sequence data for across-breed genomic prediction. *Genet. Sel. Evol.* 50:27. <https://doi.org/10.1186/s12711-018-0396-8>.
- Raymond, B., Y. C. J. Wientjes, A. C. Bouwman, C. Schrooten, and R. F. Veerkamp. 2020. A deterministic equation to predict the accuracy of multi-population genomic prediction with multiple genomic relationship matrices. *Genet. Sel. Evol.* 52:21. <https://doi.org/10.1186/s12711-020-00540-y>.
- Saatchi, M., R. D. Schnabel, J. F. Taylor, and D. J. Garrick. 2014. Large-effect pleiotropic or closely linked QTL segregate within and across ten US cattle breeds. *BMC Genomics* 15:442. <https://doi.org/10.1186/1471-2164-15-442>.
- Steyn, Y., D. Gonzalez-Pena, Y. L. Bernal Rubio, N. Vukasinovic, S. K. DeNise, D. A. L. Lourenco, and I. Misztal. 2021. Indirect genomic predictions for milk yield in crossbred Holstein-Jersey dairy cattle. *J. Dairy Sci.* 104:5728–5737. <https://doi.org/10.3168/jds.2020-19451>.
- Steyn, Y., D. A. L. Lourenco, and I. Misztal. 2019. Genomic predictions in purebreds with a multibreed genomic relationship matrix. *J. Anim. Sci.* 97:4418–4427. <https://doi.org/10.1093/jas/skz296>.
- van den Berg, I., D. Boichard, and M. S. Lund. 2016. Comparing power and precision of within-breed and multibreed genome-wide association studies of production traits using whole-genome sequence data for 5 French and Danish dairy cattle breeds. *J. Dairy Sci.* 99:8932–8945. <https://doi.org/10.3168/jds.2016-11073>.
- van den Berg, I., I. MacLeod, C. Reich, E. Breen, and J. Pryce. 2020. Optimizing genomic prediction for Australian Red dairy cattle. *J. Dairy Sci.* 103:6276–6298. <https://doi.org/10.3168/jds.2019-17914>.
- VanRaden, P. M., M. E. Tooker, T. C. S. Chud, H. D. Norman, J. H. Megonigal Jr., I. W. Haagen, and G. R. Wiggans. 2020. Genomic predictions for crossbred dairy cattle. *J. Dairy Sci.* 103:1620–1631. <https://doi.org/10.3168/jds.2019-16634>.
- Winkelman, A. M., D. L. Johnson, and B. L. Harris. 2015. Application of genomic evaluation to dairy cattle in New Zealand. *J. Dairy Sci.* 98:659–675. <https://doi.org/10.3168/jds.2014-8560>.
- Zhou, L., B. Heringstad, G. Su, B. Gulbrandsen, T. H. E. Meuwissen, M. Svendsen, H. Grove, U. S. Nielsen, and M. S. Lund. 2014. Genomic predictions based on a joint reference population for the Nordic Red cattle breeds. *J. Dairy Sci.* 97:4485–4496. <https://doi.org/10.3168/jds.2013-7580>.

Notes

I. Misztal  <https://orcid.org/0000-0002-0382-1897>

Y. Steyn  <https://orcid.org/0000-0001-5467-9555>

D. A. L. Lourenco  <https://orcid.org/0000-0003-3140-1002>

This study was supported by grants from the Agriculture and Food Research Initiative of the National Institute of Food and Agriculture (Washington, DC; 2020-67015-31030) and by Zoetis Inc. (Parsippany, NJ).

We are grateful for the useful comments by the 2 anonymous reviewers.

The authors have not stated any conflicts of interest.