730. Value of data from ram breeding flocks as an industry reference population for Australian sheep

P. Alexandri[1*], S.F. Walkom[1], A.A. Swan[1], J.H.J. van der Werf[2] and D.J. Brown[1]

[1]AGBU, a joint venture of NSW Department of Primary Industries and University of New England, 2351, Armidale, Australia; [2]School of Environmental and Rural Science, University of New England, Armidale, NSW 2351, Australia; palexan8@une.edu.au

Abstract

This study examined the value of using carcass data from seedstock ram breeding flocks to build upon an industry sheep reference population in Australia. Data from 995 lambs managed in 15 commercial ram breeder flocks were collected between 2017 and 2020 for carcass and meat quality measurements: hot carcass weight, tissue depth on the 12th rib (GR site), eye muscle depth, fat between the 12th and the 13th rib (C – site), intramuscular fat and shear force. Industry data were cross-validated with and without reference data from the MLA Resource Flock. Industry data did not bias the estimation of breeding values when used in combination with the reference population (MLA Resource Flock). Therefore, industry data can be used to expand an industry reference population if data collection is accurate and consistent with industry standards.

Introduction

The MLA Resource Flock (RF) is the current Australian sheep reference population and a succession of the Sheep CRC Information Nucleus Flock (Van der Werf *et al.* 2010). The flock provides a well-designed and ongoing sheep reference population and is mostly funded with national levy funds. This industry investment helps subsidise the high cost for progeny testing and measuring expensive traits such as shear force and intramuscular fat of lamb meat.

Historically most RF sites were well-resourced research stations and RF data has been used for research as well as in routine genetic evaluation to produce Australian Sheep Breeding Values (ASBVs). Large reference populations lead to increased accuracy of genomic predictions (Habier *et al.*, 2010) which can be further enhanced by collecting phenotypic and genotypic information from additional animals. As part of expanding the reference population, a series of projects were established where breeders funded data collection from their own flocks with co-investment support from industry funds. These projects allowed for more animals to be genotyped and measured for carcass and meat quality traits to help build an industry reference population with reduced investment per animal measured from industry funds. The quality of data collected on these projects, however, varied and procedures were not always fully comparable to data from the research stations used in the RF. Therefore, we investigated the value of the records collected on industry animals via the co-investment funding model as part of the reference population for sheep in Australia.

Materials & methods

Data. The seedstock data came from 995 lambs born between 2017 and 2020. The lambs came from 15 seedstock ram breeding flocks and were the progeny of 281 sires across 3 breeds (Poll Dorset, White Suffolk and Southdown). A second dataset with 4,027 RF animals was used for validation. The RF animals were born between 2015 and 2020 and were the progeny of 445 sires across the same three breeds (Poll Dorset, White Suffolk and Southdown). Carcass traits for industry and RF animals were measured after slaughter in commercial abattoirs. Weight (CWT) and tissue depth at the GR site (GRFAT, total tissue depth measured with a GR knife on the 12th rib) were measured on the hot carcass. After overnight chilling (3-4 °C), at a cut

R.F. Veerkamp and Y. de Haas (eds)

DOI: 10.3920/978-90-8686-940-4_730, © P. Alexandri *et al.* 2022

between the 12th and 13th ribs of each carcass, eye muscle (*M. longissimus thoracis et lumborum*, LL) depth (EMD) and fat depth at the C site (CFAT) were measured. The percentage of intramuscular fat (IMF) at the LL was determined using a near infrared procedure (NIR) as described by (Perry *et al.* 2001). Shear force (SF5) at 5 days after slaughter was measured at a section of the LL as described by Hopkins *et al.* (2010).

Statistical analyses. Estimated Breeding values (EBVs) were estimated using the LAMBPLAN genetic evaluation software OVIS (Brown *et al.* 2018). All data was pre-adjusted for birth and rear type (single, twins, triplets and quadruplets or more lambs), lamb age and age of dam. Contemporary group, defined by breed, flock, management group, sex, date of measurement and kill group, was used as a fixed effect. Hot carcass weight was included as a covariate to adjust other carcass traits to a weight constant. All models included the random effects of animal, genetic group (Swan *et al.* 2016) and sire × flock interaction. To estimate the differences in accuracy of prediction an internal cross-validation procedure was used for each data set as described by Legarra and Reverter (2018). RF data were separated into four data sets with approximately the same size as the seedstock data. RF animals were randomly assigned to groups one to four based on their sires so that half-sib families were not represented in multiple groups. All seedstock animals were assigned to group five. Numbers of records, sires and unadjusted trait means for each group are shown in Table 1.

Following analysis of the full data set, three different validation scenarios were investigated. First, EBVs were calculated for each RF validation group, using RF data from the other three groups as the training population (RF – RF analysis, performed four times – one for each group). Second, prediction of industry animals was carried out using only RF data (groups one to four) (F – seedstock analysis, replicated four times). Third, prediction of RF animals was performed using four different replicated combinations of RF (groups one to four) and industry animals (group five) (combined analysis).

For each trait, four validation metrics were calculated and averaged across replicates. Accuracy and dispersion metrics were calculated using the LR method (Legarra and Reverter 2018) as the correlation and regression slopes between the EBVs from each of the three analyses (RF – RF, RF – industry, combined) with EBVs from the full analysis. The regression slopes between EBVs are expected to have a value close to one if there is no over or under dispersion. Accuracy and dispersion were also calculated as the correlation and regressions for EBVs on phenotypes adjusted for fixed effects, with regressions performed in ASReml (Gilmour *et al.* 2015).

Table 1. Number of animals (N), sires (Sires), sires per contemporary group (Sires/CG) and unadjusted mean record values (standard deviation) for all traits and validation groups used in the validation analysis. CWT: hot carcass weight (kg), EMD: carcase eye muscle depth (cm), CFAT: fat depth at the C-site (cm), GRFAT: tissue depth at GR-site (cm), IMF: intramuscular fat (%), SF5: shear force 5 days after slaughter (N).

Group[1]	N	Sires	Sires/CG	CWT	EMD	CFAT	GRFAT	IMF	SF5
1	1,100	112	16.7	23.6 (3.8)	32.9 (4.8)	4.3 (2.4)	14.5 (6.2)	4.4 (1.0)	39.3 (12.9)
2	1,162	110	17.7	26.3 (3.7)	35.1 (4.2)	4.8 (2.6)	18.0 (5.1)	4.4 (1.0)	35.3 (10.1)
3	953	111	15.4	26.7 (3.1)	35.4 (4.2)	5.2 (2.4)	20.7 (5.8)	4.7 (1.2)	32.9 (10.3)
4	812	112	26.4	27.5 (3.5)	34.9 (4.1)	5.7 (2.9)	22.3 (6.0)	4.7 (1.1)	29.0 (8.9)
5	995	140	9.6	26.9 (5.2)	37.0 (4.5)	4.4 (2.8)	16.4 (6.7)	3.9 (1.0)	38.3 (15.0)

[1] Data for groups 1 to 4 came from RF animals, group 5 included seedstock data.

Results

Validation results followed similar patterns for two of the three different validation scenarios. For the RF – RF and combined scenarios LR EBV dispersion metrics had values close to one for CWT and CFAT and greater than one for EMD, GRFAT, IMF and SF5 (Table 2). LR EBV correlations ranged from 0.36 (EMD, combined) to 0.52 (GRFAT for RF – RF and combined, SF5 for RF – RF). Phenotypic dispersions for the same validation scenarios (RF – RF and combined) were close to one for GRFAT (0.95 and 0.94 respectively) and CFAT (0.87 for both). However, estimates greater than one were observed for EMD, IMF and SF5 for both analyses. Correlations between EBVs and phenotypes were consistent between the two analyses ranging from 0.19 (EMD) to 0.45 (CWT).

Differences were observed in the metrics when RF data was used to predict into industry animals (scenario RF – seedstock). LR EBV dispersions were closer to one for GRFAT (0.94), CWT (0.85) and CFAT (0.73) although for the last two traits the values observed were lower than the other two validation scenarios (Table 2). LR EBV accuracies were higher for all traits ranging from 0.49 (CFAT) to 0.64 (EMD). Higher phenotypic dispersions were observed for all traits ranging from 0.67 (CWT) to 1.67 (SF5). Phenotypic accuracies were similar to other validation analyses for EMD and CFAT but higher for the rest of the traits. In general, validation patterns were very similar when using RF data to predict RF progeny (RF – RF) and combined RF and seedstock data (combined) were used to validate RF phenotypes but over-dispersion was observed when RF data was used to predict industry phenotypes (RF – seedstock).

Discussion

Accuracy of genomic predictions can benefit from larger reference populations (Habier *et al.* 2010). In this study the value of expanding the Australian sheep reference population by using data from seedstock ram breeding flocks for carcass and meat quality traits was explored. Cross validation results showed that it was possible to use data from seedstock ram breeding flocks to expand the industry reference population for traits recorded using common protocols. Comparison between observed and predicted performances in a

Table 2. Validation metrics for each validation scenario averaged across replicates. CWT: hot carcass weight, EMD: carcase eye muscle depth, CFAT: fat depth at the C-site, GRFAT: tissue depth at GR-site, IMF: intramuscular fat, SF5: shear force 5 days after slaughter.

	Trait						
Metric	CWT	EMD	CFAT	GRFAT	IMF	SF5	Scenario[1]
LR EBV dispersion	0.93	1.16	0.94	1.07	1.28	1.14	RF – RF
	0.85	1.56	0.73	0.94	1.13	1.07	RF – seedstock
	0.92	1.15	0.94	1.08	1.27	1.13	Combined
LR EBV correlations	0.46	0.37	0.44	0.52	0.49	0.52	RF – RF
	0.50	0.64	0.49	0.57	0.56	0.60	RF – seedstock
	0.45	0.36	0.43	0.52	0.48	0.50	Combined
Phenotypic dispersion	0.61	1.21	0.87	0.95	1.33	1.12	RF – RF
	0.67	1.61	1.25	1.01	1.06	1.67	RF – seedstock
	0.58	1.19	0.87	0.94	1.31	1.10	Combined
Phenotypic – EBV correlations	0.45	0.19	0.24	0.40	0.30	0.40	RF – RF
	0.49	0.19	0.24	0.43	0.59	0.65	RF – seedstock
	0.45	0.19	0.24	0.40	0.30	0.40	Combined

[1] RF-RF: prediction using different datasets of RF animals, RF – seedstock: using RF animals to predict into seedstock animals, Combined: using different combinations of RF animals.

cross-validation analysis is important to measure the efficiency of the application of the analysis to specific data sets (Legarra *et al.* 2008). Our results show that for reference and seedstock data, phenotypic and EBV dispersions can be similar when seedstock data are used in combination with a well recorded reference population (Analyses RF – RF and combined, Table 2). Using seedstock and reference data together does not introduce further bias to breeding values estimation. When reference data was used to predict seedstock phenotypes results depended upon the recorded traits. For example, EMD presents higher biases than other carcass traits both for EBV and phenotypic dispersions (Table 2) but also exhibits higher mean values in the seedstock animals compared to the reference data sets (Table 1). Moreover, data structure is different between RF and seedstock animals. The number of sires per contemporary group is typically lower for seedstock data (Table 1) and RF data also normally represents a bigger range of breeds (van der Werf *et al.* 2010). This highlights the importance of recorded data quality; the establishment of an industry reference population can benefit from accurate and consistent data recording.

In conclusion, data collected from seedstock ram breeding flocks can be used to complement managed progeny test sites to create an industry reference population. The effectiveness of commercial data depends on the trait measured (completeness of data and good representation of the flock's diversity) and the influence of fixed effects recorded on the flock. Co-investment into industry recorded carcass traits using levy funds was found to be beneficial to the growth of the reference population. However, these projects should be carried out with caution as there is a risk of reduced sire diversity, less consistent data collection, and reduced data quality if they are not managed properly.

Acknowledgements

This work was conducted as part of the Advanced Measurement Technologies for globally competitive Australian meat project, funded by the Australian Government Department of Agriculture and Water Resources as part of its Rural R&D for Profit programme. The authors thank the teams behind the Sheep CRC Information Nucleus Flock and MLA resource flock.

References

Brown D.J., Swan A.A., Boerner V., Li L., Gurman P.M., *et. al.* (2018) Proc. of the 11th WCGALP, Aukland, New Zealand.

Gilmour A.R., Gogel B.J., Cullis B.R., Welham S.J. & Thompson R. (2015) ASReml User Guide Release 4.1. In: *VSN International Ltd, Hemel Hempstead, UK*

Habier, D., Tetens J., Seefried F.R., Lichtner P. & Thaller G. (2010) Genet. Sel. Evol. 42:5. https://doi.org/10.1186/1297-9686-42-5.

Hopkins D.L., Toohey E.S., Warner R.D., Kerr M.J. & van de Ven R. (2010) *Anim. Prod. Sci.* 50(6):382-385. https://doi.org/10.1071/AN09162.

Legarra A. & Reverter A. (2018) *Genet. Sel. Evol.* 50:53. https://doi.org/10.1186/s12711-018-0426-6.

Legarra A., Robert-Granie C., Manfredi E. & Elsen J.M. (2008) *Genetics* 180 (1):611-618. https://doi.org/10.1534/genetics.108.088575.

Perry D., Shorthose W.R., Ferguson D.M. & Thompson J.M. (2001) *Aust. J. Exp. Agric.* 41(7):953-957. https://doi.org/10.1071/EA00092.

Swan A.A., Brown D.J. & van der Werf J.H.J. (2016) *Ani. Prod. Sci.* 56(1):87-94. https://doi.org/10.1071/AN14560.

Van der Werf J.H.J., Kinghorn B.P. & Banks R.G. (2010) *Ani. Prod. Sci.* 50(12):998-1003. https://doi.org/10.1071/AN10151.