Proc. Assoc. Advmt. Anim. Breed. Genet: 24: 163-166

ACROSS-COUNTRY PREDICTION OF METHANE EMISSIONS USING RUMEN MICROBIAL PROFILES

M.K. Hess[1], A. Donaldson[2], H.M. Henry[1], D.L. Robinson[2], A.S. Hess[1], J.C. McEwan[1], V.H. Oddy[2] and S.J. Rowe[1]

[1] AgResearch Ltd, Invermay, Mosgiel, New Zealand
[2] NSW Department of Primary Industries, University of New England, Armidale, NSW 2351

SUMMARY

Rumen microbial profiles have been shown to be accurate predictors of methane emissions in a variety of species, however, it can be very costly and slow to generate a dataset with a sufficient number of individuals measured for methane who also have had rumen samples collected and processed into rumen microbial profiles for these benefits to be applied in industry. We evaluated the potential of combining datasets from New Zealand and Australian sheep to improve our ability to accurately predict methane emissions in Australian sheep. Prediction of Australian sheep methane emissions using rumen microbial profiles and phenotypes from New Zealand was possible, however, it was important to closely match the diets the sheep were fed to have confidence in the predictions. Prediction accuracies of Australian sheep methane emissions were higher when training on data collected on Australian sheep than training on New Zealand sheep; however augmentation of New Zealand data collected on a similar diet enabled more complex models to be run and an improvement in prediction accuracy.

INTRODUCTION

The rumen microbiome has been shown to play an important role methane production and feed efficiency and improve prediction accuracy in these traits (Hess *et al*. Submitted-b). However, large sample numbers are typically required for accurate trait prediction. Over 3,000 New Zealand sheep rumen microbial profiles have been generated with associated methane emission phenotypes, representing a variety of breed compositions, ages and diets (Hess *et al*. Submitted-a). Robinson *et al*. (2020) describe a study in over 500 Australian merino sheep that have been measured for methane emissions with rumen samples collected during the study. This study predicted methane emissions in Australian merino sheep under two scenarios: 1) when Australian sheep had no methane data collected and models were trained using the New Zealand dataset, and 2) when some Australian sheep had methane data collected and added to the New Zealand training dataset. The models used in our study utilized genomic information, rumen microbial profiles or both.

MATERIALS AND METHODS

Australian Microbiomes. Rumen samples were collected from 502 Information Nucleus Flock follower ewes on a chaffed lucerne and cereal hay diet at 1.5-1.6 times maintenance (Robinson *et al*. 2020). Restriction Enzyme-Reduced Representation Sequencing (Hess *et al*. 2020) was used to generate Reference Free Rumen Microbial Profiles, as described in Hess *et al*. (Submitted-a).

New Zealand Microbiomes. Reference Free Rumen Microbial Profiles were generated on 3,019 rumen samples from 1,200 dual purpose composite ewes (Hess *et al*. Submitted-a; Hess *et al*. Submitted-b). Rumen Microbial Profiles were separated into 3 groups based on diet (all fed ad lib) and age: lamb on ryegrass-based pasture/grass (GL, n = 1051), adult on ryegrass-based pasture/grass (GA, n = 1010), and lambs on a lucerne pellet diet (LL, n = 958).

Methane Phenotypes. Australian sheep had methane phenotypes collected in Respiration and Portable Accumulation Chambers (Robinson *et al*. 2020) during the same experiment in which rumen samples were collected. Methane emission phenotypes for the Australian sheep used in this

study were the genetic plus permanent environmental effects for respiration chamber measurements based on the model without covariates for liveweight and feed intake of Robinson *et al*. (2020)

New Zealand methane phenotypes were the methane emission phenotypes from Portable Accumulation Chambers, adjusted for the fixed effects of birth rear rank, age of dam and birth date deviation (Hess *et al*. Submitted-b). Adjusted methane phenotypes were normalized within group, such that each group had a mean of zero and standard deviation of one to account for differences in measurement type (respiration chamber vs portable accumulation chamber), differences in methane emissions due to effects such as diet and age, and differences in the methane yield models.

Genotypes. High density genotypes were available on all New Zealand sheep and 322 of the Australian sheep. Sheep were genotyped on a variety of SNP chips, then imputed to a high density set of SNPs separately within each country. After imputation, the two datasets were combined and SNPs that were segregating in both populations (471,596 SNPs) were used to generate a genomic relationship matrix (GRM) using the first method of Van Raden (2008).

Models. Three models were run in ASReml v 4.1 (Gilmour *et al*. 2015), which explained variation in methane phenotype using genotypes, Microbial Profiles or both:

$$y = \mu + G + e;$$
$$y = \mu + M + e;$$
$$y = \mu + G + M + e$$

where y is the adjusted methane phenotype; μ is the mean; G is the random animal genetic effect with relationships between animals represented by the GRM described above; M is the random microbial effect with relationships between samples represented by the cohort-adjusted microbial relationship matrix, calculated as described in Hess *et al*. (2020); and e is the residual.

The above models were trained using GL, GA, LL or all NZ samples, and used to predict breeding values (BV) and microbial values (MV) in the Australian dataset. For models including both G and M, the BV and MV were summed to get the combined value (GMV). Accuracies were estimated as the correlation between the phenotype and the BV, MV or GMV. The accuracy of the microbial values were calculated using all Australian samples or just the samples associated with genotyped animals, and models containing G were only run for animals with genotype information available. Accuracies were estimated for each cohort separately and the standard errors of the accuracies estimated as the standard deviation across all cohorts. There were 10 validation cohorts with 50 ± 26 Australian sheep in the full dataset and 5 of these cohorts had 64 ± 29 genotyped sheep.

The three models above were also trained using Australian samples excluding the cohort that was being predicted, as well as these samples augmented with the LL or all NZ samples. Microbial relationship matrices used for each model were generated using tags that were present in all groups found in either the training or prediction set for that model. There were 79,328 tags present in both GA and AUS groups, 69,120 tags present in both GL and AUS groups, 39,502 tags present in both LL and AUS groups, 29,456 tags present in all groups (GA, GL, LL and AUS), and 150,687 tags present in the AUS group.

RESULTS AND DISCUSSION

Across-country prediction. Our first analysis aimed to use various rumen microbial profiles from New Zealand sheep to predict methane emissions in Australian sheep. Microbial value estimates for either all Australian samples or samples associated with a genotyped animal were poor and tended to be negative when New Zealand samples were used as the training set, with the exception of the samples from lambs fed lucerne pellets (Table 1). The highest accuracy (0.23) was from BV estimated using the full NZ dataset and a model fitting both genomic and microbial effects. This model contains the most information, with up to three methane phenotypes collected on each individual (one each in GL, GA and LL), compared to one for each of the other groups.

Models trained on the LL data had low but positive accuracies (0.09-0.13) and the lowest

standard errors (Table 1). Most training individuals were represented in all three NZ groups (GA, GL and LL), so the difference in BV accuracies from the model just fitting genomics is largely driven by differences in the methane phenotype. The diet fed to the Australian sheep (chaffed lucerne and cereal hay) is more similar to the lucerne pellet diet of the LL group than the ryegrass-based pasture of the other two New Zealand groups, therefore it is likely the drivers of methane emissions in these Australian sheep are most similar to those in the New Zealand LL group.

The model fitting both genomic and microbial effects and trained on the LL dataset showed the highest GMV accuracy, but this was no higher than the accuracy of the BV in the model just fitting genomic effects with the same training data (Table 1), suggesting that incorporating microbial information doesn't always improve accuracy beyond just fitting genotypes even when the microbial profiles had some predictive ability (e.g. LL). For the model fitting genomic and microbial effects and trained on the other NZ datasets (GL, GA and all NZ), there is some evidence that including the microbial component into the model can improve BV accuracy (0.08-0.23) compared to a model fitting only the genetic effect (-0.01-0.17).

Table 1. Accuracy of predicting Australian methane emissions using Genotypes and/or Microbial Profiles from New Zealand sheep

Training set	All AUS	Genotyped AUS				
				GRM+MRM		
	MV	BV	MV	BV	MV	GMV
GL	-0.10 ± 0.14	-0.01 ± 0.18	-0.12 ± 0.09	0.08 ± 0.17	-0.12 ± 0.10	-0.11 ± 0.09
GA	-0.20 ± 0.14	0.04 ± 0.15	-0.22 ± 0.11	0.21 ± 0.21	-0.23 ± 0.11	-0.20 ± 0.11
LL	0.13 ± 0.08	0.13 ± 0.09	0.09 ± 0.07	0.12 ± 0.12	0.12 ± 0.08	0.13 ± 0.08
NZ	-0.02 ± 0.12	0.17 ± 0.14	-0.06 ± 0.04	0.23 ± 0.16	-0.01 ± 0.06	0.01 ± 0.06

All AUS = Genotyped and non-genotyped Australian sheep, Genotyped AUS = genotyped subset of All AUS
GL = Grass lamb, GA = Grass adult, LL = Lucerne pellet lamb, NZ = All NZ samples (GL + GA + LL)
MV = Microbial value, BV = Breeding value, GMV = Genetic plus Microbial value

Incorporating data from other countries. Our second analysis aimed to evaluate whether including data from another country can improve prediction accuracy. All accuracies were higher when incorporating Australian data into the training set (Table 2) compared to training on different combinations of the New Zealand dataset (Table 1). BV and MV accuracies were high when using the training set of Australian samples despite the smaller size (Table 2). The highest accuracies were observed for GMV using the AUS+LL training set, followed by the MV estimated for genotyped animals when training on just the Australian dataset.

BV accuracy was not significantly impacted by adding LL or all NZ data to the Australian dataset (Table 2). This is likely driven by the different breed compositions between the two countries, leading to genomic relationships that were mostly negative between animals from NZ and Australia (Mean = -0.05; Range = -0.09 to 0.09); while those within Australia were mostly positive (Mean = 0.19; Range = -0.03 to 0.73).

The model fitting both genomic and microbial effects gave higher accuracies than the models fitting just genomic or just microbial relationships for the models trained on AUS+LL and AUS+NZ data (Table 2). The model fitting both genomic and microbial effects is more complex than the other two models used in our study, this led to singularity issues when used on the Australian dataset, likely driven by the smaller training set of 322 genotyped animals. Augmentation of the Australian dataset with New Zealand samples allows a more complicated model to successfully run and

produces a higher prediction accuracy than using a model that just fits genomic or microbial effects.

Table 2. Accuracy of predicting Australian methane emissions using Genotypes and/or Microbial Profiles from New Zealand and Australian sheep

Training set	All AUS	Genotyped AUS				
				GRM+MRM		
	MV	BV	MV	BV	MV	GMV
AUS	0.54 ± 0.12	0.45 ± 0.26	0.57 ± 0.16	Singularities		
AUS+LL	0.47 ± 0.13	0.46 ± 0.24	0.49 ± 0.16	0.44 ± 0.25	0.48 ± 0.15	0.60 ± 0.17
AUS+NZ	0.40 ± 0.14	0.44 ± 0.25	0.39 ± 0.17	0.43 ± 0.25	0.37 ± 0.15	0.53 ± 0.15

All AUS = Genotyped and non-genotyped Australian sheep, Genotyped AUS = genotyped subset of All AUS
AUS = Australian Samples, AUS+LL = Australian and Lucerne Lamb samples, AUS+NZ = AUS and All NZ samples; MV = Microbial value, BV = Breeding value, GMV = Genetic plus Microbial value

Factors influencing these results. Several factors will be influencing these results and their application to other datasets. The design of the Australian and New Zealand datasets were different in terms of sheep breed, the method for measuring methane (respiration chambers vs portable accumulation chambers), diet the sheep were on, and slightly different methods of rumen sample processing. Differences in rumen microbial profiles between New Zealand and Australian datasets were observed in Hess *et al*. (Submitted-a). These differences are likely largely driven by environmental factors, such as diet, but could also be partially due to differences in sample preparation. Cohort-adjusted rumen microbial profiles, as were used in this study, did not show the same differences between New Zealand and Australian samples (Hess *et al*. Submitted-a).

CONCLUSIONS

This study shows that prediction of methane emissions across country using microbial profiles is possible even when genetic linkages are not strong, however, care needs to be taken in matching the diets as closely as possible to have some confidence in the predictions. Prediction accuracies of Australian sheep methane emissions were higher when training on data collected on Australian sheep than training on New Zealand sheep. Importantly, augmentation of the Australian dataset with data collected on New Zealand sheep that were on a similar diet enabled more complex models to be run and an improvement in prediction accuracy.

REFERENCES

Gilmour A., Gogel B., Cullis B., Welham S. and Thompson R. (2015) VSN International Ltd.

Hess M.K., Hodgkinson H., Hess A.S., Zetouni L., Budel J., Henry H.M., Van Stijn T.C., Kirk M.R., Dodds K.G., Brauning R., McCulloch A.F., Hickey S.M., Johnson P.L., Jonker A., Morton N., Hendy S., Janssen P.H., McEwan J.C. and Rowe S.J. (Submitted-a) *BMC Genomics*.

Hess M.K., Johnson P.L., Brauning R., McCulloch A.F., Hickey S.M., Henry H.M., Stijn T.C.V., Dodds K.G., Hess A.S., Jonker A., Janssen P.H., McEwan J.C. and Rowe S.J. (Submitted-b) *Genet. Sel. Evol*.

Hess M.K., Rowe S.J., Van Stijn T.C., Henry H.M., Hickey S.M., Brauning R., McCulloch A.F., Hess A.S., Kirk M.R., Kumar S., Pinares-Patiño C., Kittelmann S., Wood G.R., Janssen P.H. and McEwan J.C. (2020) *PLOS One* 15, e0219882.

Robinson D., Dominik S., Donaldson A. and Oddy V. (2020) Animal Production Science 60, 880.

VanRaden P.M. (2008) *J. Dairy Sci*. **91**, 4414.