Genomic heritabilities and genomic estimated breeding values for methane traits in Angus cattle

B. J. Hayes,*‡ K. A. Donoghue,‡ C. M. Reich,*
B. A. Mason,* T. Bird-Gardiner,‡ R. M. Herd,§ and P. F. Arthur#

ABSTRACT: Enteric methane emissions from beef cattle are a significant component of total greenhouse gas emissions from agriculture. The variation between beef cattle in methane emissions is partly genetic, whether measured as methane production, methane yield (methane production/DMI), or residual methane production (observed methane production – expected methane production), with heritabilities ranging from 0.19 to 0.29. This suggests methane emissions could be reduced by selection. Given the high cost of measuring methane production from individual beef cattle, genomic selection is the most feasible approach to achieve this reduction in emissions. We derived genomic EBV (GEBV) for methane traits from a reference set of 747 Angus animals phenotyped for methane traits and genotyped for 630,000 SNP. The accuracy of GEBV was tested in a validation set of 273 Angus animals phenotyped for the same traits. Accuracies of GEBV ranged from 0.29 ± 0.06 for methane yield and 0.35 ± 0.06 for residual methane production. Selection on GEBV using the genomic prediction equations derived here could reduce emissions for Angus cattle by roughly 5% over 10 yr.

Key words: beef cattle, genomic estimated breeding values, methane

INTRODUCTION

Enteric ruminant methane accounts for 9 to 11% of global greenhouse gas emissions, contributing approximately half of all agricultural emissions, and cattle are the single largest source of enteric methane from agriculture (Opio et al., 2013).

Methane emission levels in cattle, whether measured as methane production, methane yield (MY; methane production/DMI), or residual methane production (RMP; the difference between observed methane production and predicted methane production) are heritable traits, with heritabilities in the range of 0.19 to 0.28 (Herd et al., 2014; Donoghue et al., 2015). Selection for reduced emissions could, therefore, result in small annual but cumulative and permanent changes in emission levels. Residual methane production or MY are more attractive targets for selection than methane production rate (MPR), as they are not unfavorably correlated with production traits (Donoghue et al., 2015).

Given the current cost and difficulty of measuring these traits on individual beef cattle, it is unlikely that either MY or RMP could be measured on the scale that would be necessary to calculate traditional EBV on an ongoing basis for the beef industry. An alternative is to use genomic selection for these traits (Meuwissen et al., 2001). Genomic selection entails measuring a large reference population for MY or RMP, genotyping the reference population for a large
number of SNP markers, and then using the information to derive a genomic prediction equation, which then can be used to calculate genomic EBV (GEBV) for any selection candidate that is genotyped.

In multipurpose sheep, GEBV for gross methane production and MY have been reported, with moderate accuracies of 0.37 and 0.43, respectively (accuracies estimated by cross-validation; Rowe et al., 2014). The same authors also discovered individual markers with significant associations with these traits in a genomewide association study.

In beef cattle, accuracies of GEBV for methane traits have not been previously reported. Here we use a large group of Angus animals measured for methane emission levels (as described by Donoghue et al., 2015) and with real or imputed genotypes for 632,003 SNP to derive GEBV for MPR, MY, and RMP. The accuracies of the GEBV were moderate, enabling selection for reduced methane emission levels for Angus cattle.

**MATERIALS AND METHODS**

The project was approved by the New South Wales (NSW) Department of Primary Industries and the University of New England Animal Ethics Committees. All animals in the project were managed according to the Australian Code for the Care and Use of Animals for Scientific Purposes (NHMRC, 2013).

**Phenotypes**

The animals in the experiment and methane phenotypes are fully described in Donoghue et al. (2015). A brief summary of the animals and phenotypes is given here. There were 1,043 Angus animals phenotyped in the experiment, from the registered herds at the New South Wales Department of Primary Industries Agricultural Research Centre at Trangie in Australia. The cattle were born in 2009, 2011, 2012, and 2013 and were raised as calves by their dams on pasture until weaning at approximately 8 mo of age. The weaned bulls and heifers remained on pasture throughout their lives except for the period of methane measurement. Methane production was measured in respiration chambers over 2 consecutive 24-h periods. Methane measurements were taken at approximately yearling age (339 d of age) for the 2011-, 2012-, and 2013-born cattle, whereas the 2009-born cattle were measured for methane at approximately 2 yr of age (738 d of age). The SD of age at measurement was 167 d.

Cohorts of up to 40 cattle in 4 groups of 10 were formed and prepared for measurement. Progeny of individual sires were stratified across groups and cohorts. The cohort of 40 animals were weighed and then fed in their groups of 10. Animals were fed an amount calculated using the Australian feeding standards formulas to provide 1.2 times their estimated energy requirement for maintenance. After 10 d of feeding, the animals were weighed again, with this weight used as their test-period weight (TWT). The animals were then transported to the methane measurement facility at the University of New England (Armidale, NSW, Australia). Each animal was fed in an individual pen (1.8 by 3 m) for 2 d at 1.2 times estimated maintenance based on the TWT of the animal before being moved into the respiration chambers. Hence, each animal had a minimum of 16 d (10 d at the research station and 6 d at the methane testing facility) on the roughage ration before entry into the respiration chambers. Specifically, the ration was a commercial alfalfa and oaten hay chaff purchased from the same supplier over the duration of the experiment (Manuka “Blue Ribbon” chaff; Manuka Chaff Pty. Ltd., Quirindi, NSW, Australia), and the mean nutritional values of subsamples of the test ration over 3 yr were 88% DM, 14% CP (DM basis), 67% DM digestibility, and ME content of 9 MJ/kg DM (Herd et al., 2014). Details on the design of the respiration chambers and methane measurement protocols have been published earlier by Herd et al. (2014). The 1,043 animals with methane phenotypes were the progeny of 73 sires (average of 14 progeny per sire, ranging from 1 to 30 progeny).

Trait abbreviations and definitions are provided in Table 1. Dry matter intake during the methane measurement period and MPR were used to calculate MY as MPR per unit DMI. Four RMP traits were defined to target methane production independent of feed intake. Residual methane production is a measure of actual MPR minus expected MPR (expMPR). For the first 3 forms of RMP (RMP$_B$, RMP$_J$, and RMP$_I$ [RMP$_B$ from Blaxter and Clapperton [1965], RMP$_J$ from Johnson et al. [1995], and RMP$_I$ from the formula of the Intergovernmental Panel on Climate Change [2006]). For the last form of RMP (RMP$_R$), residuals were derived from a simple regression of MPR on DMI, using the test data from the study, with cohort fitted as a class effect. These residuals are equivalent to actual MPR minus expMPR. Additional information on the computation of RMP traits has been reported earlier by Herd et al. (2014).

Out of the 1,043 Angus cattle with methane phenotypes, 23 failed genotyping, leaving 1,020 with genotypes, and only data for the 1,020 cattle with genotypes were used in this paper. The trait averages, SD, and maximums and minimums are presented in Table 2.
Genotypes and Quality Control

The cattle were genotyped with either the 777,000 SNP Illumina (San Diego, CA) Bovine HD Array (847 animals) or the Bovine 54,000 SNP50 array (173 animals). The SNP positions used were from bovine genome assembly UMD 3.1 (University of Maryland, College Park, MD). Stringent quality control procedures were applied to the data for both genotype densities. Monomorphic SNP and SNP with less than 5 copies of the rare allele were removed. Then, genotype calls with a GenTrain (Illumina, San Diego, California) score (GenCall) > 0.6 were considered high quality and used; below this value, they were excluded. For the animals genotyped with the high density (HD) array, there were 650,934 SNP genotyped at GenCall > 0.6. Furthermore, 343 mitochondrial SNP, 1,124 Y chromosome SNP, and 1,735 unmapped SNP were excluded. Single nucleotide polymorphisms with duplicate positions or dubious positions given linkage disequilibrium with surrounding SNP were also removed, and those with poor imputation accuracies were also removed (as described by Erbe et al., 2012). Six hundred thirty-two thousand three SNP remained for HD SNP after quality control and 36,655 SNP for the Bovine SNP50 array genotyped animals.

Samples (animals) were checked for excess heterozygosity (>0.4 is a sign of sample contamination) and had to have more than 90% of SNP with GenCall scores > 0.6. All 1,020 samples passed these quality control criteria, and 97.9% of SNP were genotyped at GenCall > 0.6.

Missing genotypes for animals genotyped with the Bovine HD array were imputed using Beagle3 (Browning and Browning, 2009), and the same program was used to impute the animals genotyped for the Bovine HD array to 632,003 genotypes.

Statistical Analysis – Genomic Heritabilities and Genomic EBV

The models fitted for each trait were the same as described in Donoghue et al. (2015), except that genomic relationships were used to describe relationships between animals:

\[ y = Xb + Zg + e, \]

in which \( y \) is a vector of trait records (TWT, DMI, CH4, MY, RMPB, RPMJ, RPMI, or RPMR); \( b \) is a vector of fixed effects including contemporary group, age, and dam age; \( X \) is a design matrix allocating records to fixed effects; \( g \) is a vector of GEBV; \( Z \) is a design matrix allocating records to breeding values; and \( e \) is a vector of random residuals assumed distributed \( N(0, \sigma^2_e) \), in which \( \sigma^2_e \) is the error variance. The \( g \) were assumed distributed \( N(0, G \sigma^2_{gen}) \), in which \( \sigma^2_{gen} \) is the additive genetic variance and \( G \) is the genomic relationship matrix constructed from the 632,003 SNP marker genotypes, following Yang et al. (2010). Variance components were estimated on the full data.

Table 1. Definition of traits

<table>
<thead>
<tr>
<th>Trait name</th>
<th>Abbreviation</th>
<th>Units</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test-period weight</td>
<td>TWT</td>
<td>kg</td>
<td>Pretest weight</td>
</tr>
<tr>
<td>Dry matter intake</td>
<td>DMI</td>
<td>kg/d</td>
<td>Dry matter intake during methane measurement</td>
</tr>
<tr>
<td>Methane production rate</td>
<td>MPR</td>
<td>g/d</td>
<td>Methane produced</td>
</tr>
<tr>
<td>Methane yield</td>
<td>MY</td>
<td>g/kg</td>
<td>Methane production rate per unit DMI (MPR/DMI)</td>
</tr>
<tr>
<td>Residual methane production_B</td>
<td>RMP_B</td>
<td>g/d</td>
<td>Methane production rate net of expected MPR (expMPR) from the DMI, with expMPR from Blaxter and Clapperton (1965)</td>
</tr>
<tr>
<td>Residual methane production_J</td>
<td>RMP_J</td>
<td>g/d</td>
<td>Methane production rate net of expMPR from the DMI, with expMPR obtained by the formula of Johnson et al. (1995)</td>
</tr>
<tr>
<td>Residual methane production_I</td>
<td>RMP_I</td>
<td>g/d</td>
<td>Methane production rate net of expMPR from the DMI, with expMPR obtained by the formula of the Intergovernmental Panel on Climate Change (2006)</td>
</tr>
<tr>
<td>Residual methane production_R</td>
<td>RMP_R</td>
<td>g/d</td>
<td>Methane production rate net of the expMPR from the DMI, with expMPR obtained by regression of MPR on DMI</td>
</tr>
</tbody>
</table>

Table 2. Descriptive statistics for methane and growth traits for 1,020 Angus cattle with both phenotypes and genotypes

<table>
<thead>
<tr>
<th>Trait(^1)</th>
<th>Average (SD)</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>TWT, kg</td>
<td>357.7 (89.5)</td>
<td>156.0</td>
<td>640.0</td>
</tr>
<tr>
<td>DMI, kg/d</td>
<td>6.1 (1.3)</td>
<td>3.6</td>
<td>9.4</td>
</tr>
<tr>
<td>MPR, g/d</td>
<td>132.6 (25.5)</td>
<td>78.9</td>
<td>250.9</td>
</tr>
<tr>
<td>MY, g/kg DMI</td>
<td>22.0 (2.3)</td>
<td>13.1</td>
<td>29.5</td>
</tr>
<tr>
<td>RMP_B, g/d</td>
<td>−19.5 (18.4)</td>
<td>−95.6</td>
<td>26.1</td>
</tr>
<tr>
<td>RMP_J, g/d</td>
<td>10.7 (15.1)</td>
<td>−55.9</td>
<td>70.7</td>
</tr>
<tr>
<td>RMP_I, g/d</td>
<td>0.6 (16.0)</td>
<td>−69.1</td>
<td>55.9</td>
</tr>
<tr>
<td>RMP_R, g/d</td>
<td>0.0 (9.6)</td>
<td>−39.3</td>
<td>64.2</td>
</tr>
</tbody>
</table>

\(^1\)TWT = test-period weight; MPR = methane production rate; MY = methane yield; RMP_B = residual methane production_B; RMP_J = residual methane production_J; RMP_I = residual methane production_I; RMP_R = residual methane production_R, as described in the text.
set (1,020 records) using ASReml (Gilmour et al., 2002). Genomic heritabilities were then calculated as

$$h^2 = \frac{\hat{\sigma}^2_{\text{gen}}}{\hat{\sigma}^2_{\text{gen}} + \hat{\sigma}^2_{\text{u}}}.$$ 

We also attempted to estimate the proportion of variance explained by the SNP that was not a result of capturing close pedigree relationships. This is the variation that could be exploited when predicting animals unrelated, or at least less related, to the reference set (but of the same breed). We did this by extending the model above to jointly fit relationships from pedigree and genomic relationships:

$$y = Xb + Zg + Zu + e,$$

in which terms are as defined above, and $u$ is a polygenic breeding value distributed $N(0, A\sigma^2_u)$, in which $A$ is the relationship matrix calculated from the pedigree. The proportion of variance captured by the SNP when the joint model was fitted was then

$$\frac{\hat{\sigma}^2_{\text{gen}}}{\hat{\sigma}^2_{\text{gen}} + \hat{\sigma}^2_{\text{u}}}.$$

The accuracy of GEBV was evaluated by predicting the youngest cohort of animals, those screened in 2014 (273). The reference population was then all the other animals (747). There were no half-sib groups split across the reference and validation sets; that is, sires had all progeny either in the reference or in the validation set. The average genomic relationship between reference and validation animals was 0.09 (on a scale where the genomic relationship between full sibs is 0.5).

The accuracy of prediction was calculated for the animals in the validation set, the correlation of their GEBV, and their phenotypes (corrected for age, dam age, and contemporary group) divided by the heritability of the trait: $r(\text{GEBV}, y^*)/(h^2)^{1/2}$. Standard errors of the accuracies were calculated as $1/(n)^{1/2}$, in which $n$ is the number of animals in the validation set (273).

For comparison with the validation with the youngest cohort of animals, a 5-fold cross-validation was also performed where 20% of the animals were randomly chosen as a validation set and the other animals were in the reference set. The accuracy of genomic prediction for each fold of the 5-fold cross-validation was calculated as described above.

Two methods were used to predict GEBV – genomic best linear unbiased prediction (BLUP) and BayesR. Genomic BLUP is the method described above to estimate genomic heritabilities, except in this case, phenotypes for validation animals were excluded from the analysis.

BayesR uses a Gibbs sampling approach to estimate variant effects that are modeled as a mixture distribution of 4 normal distributions including a null distribution, $N(0, 0.0\sigma^2_g)$, and 3 others, $N(0, 0.0001\sigma^2_g)$, $N(0, 0.001\sigma^2_g)$, $N(0, 0.01\sigma^2_g)$, in which $\sigma^2_g$ is the additive genetic variance for the trait. This allows many uninformative variants to be dropped from the model and permits remaining variants to have moderate to large effects (unlike GBLUP, where all effects are dramatically shrunk). BayesR analytical methodology was described by Erbe et al. (2012), with further detail and additions in Kemper et al. (2015). Our implementation followed Kemper et al. (2015), but briefly, the model fitted to the reference set for each trait was

$$y = Xb + Za + Wv + e,$$

in which $y$ is the vector of $n$ phenotypes for animals in the reference; $b$ is the vector of fixed effects as described above; $a$ is the vector of $q$ polygenic breeding values, distributed $N(0, A\sigma^2_a)$; $v$ is the vector of $m$ variant effects, distributed as a mixture of 4 distributions, $N(0, 0.0\sigma^2_g)$, $N(0, 0.0001\sigma^2_g)$, $N(0, 0.001\sigma^2_g)$, and $N(0, 0.01\sigma^2_g)$, in which $\sigma^2_g$ is the additive genetic variance for each trait; $e$ is the vector of $n$ residual errors, distributed $N(0, 16\sigma^2_e)$; $X$ is an $n \times p$ design matrix allocating phenotypes to fixed effects; $Z$ is an $n \times q$ design matrix allocating phenotypes to polygenic breeding values; $W$ is an $n \times m$ design matrix of variant genotypes that are centered and standardized following Yang et al. (2010); $A$ is the numerator relationship matrix calculated from sire and dam pedigree records; $\sigma^2_a$ is the additive genetic variance not explained by the variants, and $\sigma^2_e$ is the error variance. As in Erbe et al. (2012), the proportion of SNP in each of the 4 ($k$) distributions ($prk$) was updated each iteration by sampling from a Dirichlet posterior distribution $prk \sim Dir(\alpha_k + \beta_k)$, in which the starting value was $\alpha = [1, 1, 1, 1]$ and $\beta_k$ was the current number of SNP with effects sampled from distribution $k$. We excluded variants from the analysis with minor allele frequency < 0.01. In each BayesR analyses, we ran 2 replicate chains of the Gibbs sampler, and each chain ran for 20,000 iterations with 10,000 iterations. Final parameter estimates were derived from the means of the sampled effects in the post-burn-in iterations, obtained separately for each of the 5 chains.

As BayesR gives the posterior probability of each SNP being in the model (e.g., not in the zero distribution) and the posterior probability of being in the large distribution (the distribution with the largest variance), it can also be used for QTL mapping (e.g., Moser et al., 2015). We plotted the posterior probability for each SNP of being in the large distribution against genomic location.
Table 3. Heritability estimates (SE) derived from analyses using either pedigree or genomic information to construct relationships between animals

<table>
<thead>
<tr>
<th>Trait</th>
<th>$h^2_{\text{pedigree}}$ (SE)</th>
<th>$h^2_{\text{genomic}}$ (SE)</th>
<th>Proportion of genetic variance captured by SNP in joint model</th>
</tr>
</thead>
<tbody>
<tr>
<td>TWT</td>
<td>0.43 (0.08)</td>
<td>0.42 (0.07)</td>
<td>0.66</td>
</tr>
<tr>
<td>DMI</td>
<td>0.44 (0.08)</td>
<td>0.37 (0.07)</td>
<td>0.49</td>
</tr>
<tr>
<td>MPR</td>
<td>0.27 (0.06)</td>
<td>0.28 (0.06)</td>
<td>0.67</td>
</tr>
<tr>
<td>MY</td>
<td>0.22 (0.06)</td>
<td>0.20 (0.05)</td>
<td>0.59</td>
</tr>
<tr>
<td>RMP$_B$</td>
<td>0.19 (0.06)</td>
<td>0.18 (0.05)</td>
<td>0.41</td>
</tr>
<tr>
<td>RMP$_I$</td>
<td>0.19 (0.05)</td>
<td>0.18 (0.05)</td>
<td>0.49</td>
</tr>
<tr>
<td>RMP$_R$</td>
<td>0.19 (0.05)</td>
<td>0.18 (0.05)</td>
<td>0.46</td>
</tr>
</tbody>
</table>

1 $h^2_{\text{pedigree}}$ refers to the proportion of variance due to additive genetic effects, estimated from the correlation of relatives ages (either phenotypically or genetically, or a combination of both). $h^2_{\text{genomic}}$ refers to the proportion of variance due to additive genetic effects, estimated from the correlation of relatives ages (either phenotypically or genetically, or a combination of both). The proportion of additive genetic variance captured by the SNP when the joint model was fitted ranged from 0.67 to 0.41 (Table 3). This was an encouraging result, indicating the SNP were picking up substantial genetic variation for most of the traits, even when close relationships in the data are accounted for (the proportion of genetic variation explained by the SNP sets an upper limit on the accuracy of GEBV that can be achieved; Garrick et al., 2009).

RESULTS AND DISCUSSION

Estimates of genomic heritabilities were similar to those calculated using pedigree with the same phenotypes (Donoghue et al., 2015) for most methane traits and were within 1 SE for all traits. The proportion of the additive genetic variance captured by the SNP when the joint model was fitted ranged from 0.67 to 0.41 (Table 3). This was an encouraging result, indicating the SNP were picking up substantial genetic variation for most of the traits, even when close relationships in the data are accounted for (the proportion of genetic variation explained by the SNP sets an upper limit on the accuracy of GEBV that can be achieved; Garrick et al., 2009).

The accuracies of GEBV from GBLUP were moderate and quite similar across traits (Table 4) and validation method (youngest animals or 5-fold cross-validation). The accuracies were all significantly different than 0; the SE of the correlation between GEBV and phenotypes (which, divided by square root of heritability, gives the accuracy) was 0.06, and for all traits, this correlation was positive and at least twice this SE. Accuracies from BayesR were similar to those from GBLUP. Improvements in accuracy from BayesR compared with GBLUP has been observed for many traits in dairy cattle although with much larger reference populations than was available here (Kemper et al., 2015). This is because with small reference sizes, the advantage of BayesR over GBLUP is diminished, as the power to discriminate whether a SNP of moderate effect size belongs in the BayesR distribution with very small, small, or moderate variance is limited, unless there are causal mutations with large effects (e.g., the accuracy of genomic prediction with BayesR is substantially higher than GBLUP for fat percent from dairy cattle, even with small reference sizes, due to the large effect of the DGAT1 mutation).

Table 4. Accuracy of genomic EBV (GEBV), where GEBV were calculated with either genic BLUP (GBLUP) or BayesR. Accuracy was $r(GEBV, y^*)/\sqrt{h^2}$, in which GEBV are the GEBV for 273 youngest animals used for validation or 5-fold cross-validation with animals in folds selected at random (GBLUP only). The $y^*$ are the phenotypes of these animals corrected for contemporary group and age effects, and $h^2$ is the genomic heritability of the trait from Table 3.

<table>
<thead>
<tr>
<th>Trait</th>
<th>GBLUP</th>
<th>BayesR</th>
<th>GBLUP</th>
</tr>
</thead>
<tbody>
<tr>
<td>TWT</td>
<td>0.37</td>
<td>0.38</td>
<td>0.26 ± 0.08</td>
</tr>
<tr>
<td>DMI</td>
<td>0.35</td>
<td>0.36</td>
<td>0.39 ± 0.09</td>
</tr>
<tr>
<td>MPR</td>
<td>0.35</td>
<td>0.38</td>
<td>0.32 ± 0.04</td>
</tr>
<tr>
<td>MY</td>
<td>0.29</td>
<td>0.36</td>
<td>0.37 ± 0.09</td>
</tr>
<tr>
<td>RMP$_B$</td>
<td>0.30</td>
<td>0.30</td>
<td>0.38 ± 0.08</td>
</tr>
<tr>
<td>RMP$_I$</td>
<td>0.34</td>
<td>0.33</td>
<td>0.35 ± 0.08</td>
</tr>
<tr>
<td>RMP$_R$</td>
<td>0.34</td>
<td>0.31</td>
<td>0.35 ± 0.08</td>
</tr>
</tbody>
</table>

1 TWT = test-period weight; MPR = methane production rate; MY = methane yield; RMP$_B$ = residual methane production$_B$; RMP$_I$ = residual methane production$_I$; RMP$_R$ = residual methane production$_R$. As defined in the text.

2 From Donoghue et al. (2015).

In fact, there was little evidence for individual genes with large effects on any of the methane traits. The posterior probability from BayesR for any SNP to have a large effect (that is, coming from the distribution with the largest variance) was 8%, which is for a SNP on chromosome 12 affecting MPR (Fig. 1A), and for a SNP on chromosome 20, for which the posterior probability was 4%, affecting MY. For RMP$_I$, the maximum posterior probability for any SNP was only 1.3% (Fig. 1B), and results were similar across all the residual methane traits. The results suggest the methane traits are affected by a large number of genes with small individual effects on each trait.

The accuracy of GEBV for MPR (0.38) is similar to that reported from a multibreed sheep population (0.37) by Rowe et al. (2014). The higher accuracy of GEBV for MY in their paper (0.40 versus 0.29) could reflect the larger reference population in their study (1,872 total animals in their study vs. 1,020 in this experiment). Also, greater genetic diversity (a multibreed sheep population) in their study compared with the single-breeds Angus population used here may be contributing to their higher accuracy for this trait. The accuracies of GEBV for methane traits we report here are also quite similar to those reported by Pickering et al. (2015), using a methane phenotype in dairy cattle, that were methane emissions predicted from DMI and BW (e.g., no actual methane production measured). The authors reported an accuracy of GEBV of 0.30 for this trait (1,726 cows in the reference set).

The accuracies of GEBV for methane traits (0.26–0.38) reported here are quite similar to those reported for
Genomic predictions for methane traits in cattle

residual feed intake: another hard-to-measure trait in beef cattle and with a similar heritability to MY. Chen et al. (2013) reported an accuracy of 0.29 for GEBV residual feed intake in Angus, when validation animals did not include animals very closely related to the reference set (accuracy was 0.58 when the validation set did include closely related animals). Mujibi et al. (2011) reported accuracies of GEBV of up to 0.25 for residual feed intake in 721 crossbred beef cattle. Bolormaa et al. (2014) did report a higher accuracy of GEBV of 0.58 for residual feed intake in Angus cattle closely related to the population here and a similar size reference set, although the SE on that estimate was reasonably large.

The age of the animal at measurement can substantially affect phenotypes for methane traits. An interesting question is, therefore, whether including reference animals with methane traits recorded at different ages to the validation set improves the accuracy of genomic prediction or decreases it. We tested whether removing these older animals from the reference set decreased or increased the accuracy of genomic prediction. In our case, there was a cohort of animals with average age of 738 d at time of measurement in the reference set, whereas the average age of the animals in the validation set at time of measurement was 338 d. In fact, removing the oldest cohort of animals substantially decreased the accuracy of genomic prediction for RMPB (from 0.3 with all animals to 0.22 with only animals similar in age to the validation animals). This would suggest that including these older animals is useful, provided age is included as a covariate in the model.

Given the accuracy of GEBV of 0.3 for MY, a response to selection for this trait that could be achieved per year can be approximately calculated. Genetic gain is given by

$$\Delta G = i r \sigma_{gen}/L,$$

in which $i$ is the intensity of selection (assume 1.5), $L$ is the generation interval (assume 3.5 years), $r = 0.3$ is the accuracy of GEBV, and $\sigma_{gen}$ is the genetic SD for the trait (estimates from our data were used). The selection response for MY would be 0.084 g CH$_4$/kg DMI. This is 0.4% of the mean for this trait, suggesting 10 years of selection could lead to a 4% reduction in MY, using the GEBV derived with the data set used here. This rate of gain is, for example, lower than, but not a different order of magnitude than, the rate of gain for milk yield in dairy cattle, a much easier trait to measure, where a roughly 1.5% gain per year has been achieved. Note that the calculation of genetic gain above assumes the accuracy of genomic prediction for MY is constant over time; in practice, it is likely this accuracy will reduce over time unless additional animals are measured for MY, genotyped, and added to the reference population.

The rate of genetic gain for reduction in methane emissions could be improved with more accurate GEBV. This could be achieved either by expanding the reference population of Angus animals phenotyped for methane emissions or by accumulating information across cattle breeds, when animals of other cattle breeds are phenotyped for these traits. Khansefid et al. (2014) used this second approach to improve the accuracy of GEBV for residual feed intake, combining data from beef and dairy breeds. Some studies report a relationship between residual feed intake and methane emissions (Nkrumah et al., 2006; Hegarty et al., 2007; Fitzsimons et al., 2013), although this is not always observed (Freetly and Brown-Brandl, 2013; Freetly et al., 2015). So another possibility to improve accuracy of GEBV for methane traits might be to use a multitrait approach with residual feed intake and/or methane trait phenotypes in the same analysis. Another approach would be an appropriately constructed selection index for profit including methane production level, DMI, and production traits. The advantage of this

Figure 1. Posterior probability for each SNP of being in the large effect distribution against genomic location, from BayesR analysis. Odd numbered chromosomes are colored in red and even number chromosomes are colored blue. The SNP numbers, ordered by genome location, are given on the x-axis. A) Methane production rate, B) methane yield, and C) residual methane production.
approach is that the residual traits (RFI and RMP) do not have to be calculated a priori; however, the selection index approach does require good estimates of the genetic correlations between these traits, which in some cases are available (e.g., Donoghue et al., 2015) as well as economic weights for the objective traits.

**LITERATURE CITED**


