Genomic prediction of residual feed intake in US Holstein dairy cattle

B. Li,[1]*** P. M. VanRaden,**[1] ⓘ **E. Guduk,**[2] ⓘ **J. R. O'Connell,**[3] **D. J. Null,**[1] **E. E. Connor,**[1]† ⓘ **M. J. VandeHaar,**[4] ⓘ
R. J. Tempelman,[4] ⓘ **K. A. Weigel,**[5] ⓘ **and J. B. Cole**[1]‡ ⓘ
[1]USDA, Agricultural Research Service, Animal Genomics and Improvement Laboratory, Beltsville, MD 20705-2350
[2]Council on Dairy Cattle Breeding, Bowie, MD 20716
[3]University of Maryland School of Medicine, Baltimore, MD 21201
[4]Department of Animal Science, Michigan State University, East Lansing 48824
[5]Department of Dairy Science, University of Wisconsin, Madison 53706

ABSTRACT

Genomic selection is an important tool to introduce feed efficiency into dairy cattle breeding. The goals of the current research are to estimate genomic breeding values of residual feed intake (RFI) and to assess the prediction reliability for RFI in the US Holstein population. The RFI data were collected from 4,823 lactations of 3,947 Holstein cows in 9 research herds in the United States, and were pre-adjusted to remove phenotypic correlations with milk energy, metabolic body weight, body weight change, and for several environmental effects. In the current analyses, genomic predicted transmitting abilities of milk energy and of body weight composite were included into the RFI model to further remove the genetic correlations that remained between RFI and these energy sinks. In the first part of the analyses, a national genomic evaluation for RFI was conducted for all the Holsteins in the national database using a standard multi-step genomic evaluation method and 60,671 SNP list. In the second part of the study, a single-step genomic prediction method was applied to estimate genomic breeding values of RFI for all cows with phenotypes, 5,252 elite young bulls, 4,029 young heifers, as well as their ancestors in the pedigree, using a high-density genotype chip. Theoretical prediction reliabilities were calculated for all the studied animals in the single-step genomic prediction by direct inversion of the mixed model equations. In the results, breeding values were estimated for 1.6 million genotyped Holsteins and 60 million ungenotyped Holsteins, The genomic predicted transmitting ability correlations between RFI and other traits in the index (e.g., fertility) are generally low, indicating minor correlated responses on other index traits when selecting for RFI. Genomic prediction reliabilities for RFI averaged 34% for all phenotyped animals and 13% for all 1.6 million genotyped animals. Including genomic information increased the prediction reliabilities for RFI compared with using only pedigree information. All bulls had low reliabilities, and averaged to only 16% for the top 100 net merit progeny-tested bulls. Analyses using single-step genomic prediction and high-density genotypes gave similar results to those obtained from the national evaluation. The average theoretical reliability for RFI was 18% among the elite young bulls under 5 yr old, being lower in the younger generations of elite bulls compared with older bulls. To conclude, the size of the reference population and its relationship to the predicted population remain as the limiting factors in the genomic prediction for RFI. Continued collection of feed intake data is necessary so that reliabilities can be maintained due to close relationships of phenotyped animals with breeding stock. Considering the currently low prediction reliability and high cost of data collection, focusing RFI data collection on relatives of elite bulls that will have the greatest genetic contribution to the next generation will give more gains and profit

Key words: genomic prediction, feed efficiency, dairy cow, residual feed intake

INTRODUCTION

Improving feed efficiency (**FE**) of dairy cattle is of great interest to dairy farmers because feed accounts for the largest part of operating costs in dairy production (European Commission, 2018; USDA, 2018). However, the available data for FE are very limited due to the high cost and difficulty in collecting individual feed intake records. Since genomic selection is well suited for

difficult-to-measure traits, genomic selection has been used as an important tool to introduce FE into dairy cattle breeding (Pryce et al., 2015). The application of genomic selection for FE has been widely developed over the years (Veerkamp et al., 2014; Pryce et al., 2015; VanRaden et al., 2018). The size of the reference population and its relationship to the predicted population remain as the limiting factors for achieving a high prediction reliability for FE. However, in the near future, a well-designed and well-recorded reference population could allow larger achievement in the genomic prediction accuracy for FE (Pryce et al., 2015; Wallén et al., 2017; VanRaden et al., 2018).

Residual feed intake (**RFI**), as one proposed FE definition trait, is defined as the difference between an animal's actual feed intake and its predicted feed intake calculated from various energy sinks [e.g., milk production, metabolic body weight (**MBW**), and change of body weight (**ΔBW**)] (Berry and Crowley, 2013; Tempelman et al., 2015a; VandeHaar et al., 2016). In the United States, RFI could receive 16% of the total emphasis in lifetime net merit (**NM$**) if it were added to the selection index (VanRaden et al., 2018). Due to low prediction reliability, RFI might only contribute about 1% extra genetic progress for lifetime profit compared with the current index (VanRaden et al., 2018). However, a 1% improvement in the rate of genetic gain is worth $4.5 million to the US dairy industry (VanRaden et al., 2018). Considering the importance of FE in dairy profits, a RFI data set has been built up across several universities and research organizations in the United States, and several studies have been recently conducted to develop trait definitions (Lu et al., 2015), estimate genetic parameters (Tempelman et al., 2015b; Manzanilla-Pech et al., 2016), and determine the genetic architecture of FE (Hardie et al., 2017; Lu et al., 2018). The development of national genomic evaluations for FE has been one of the top priorities in US dairy cattle breeding.

The goals of the current research are to estimate genomic breeding values of RFI and to assess the prediction reliabilities for RFI in the US Holstein population. In the first part of the study, a national genomic evaluation for RFI was carried out for all the Holsteins in the US population using a standard multi-step genomic evaluation method (VanRaden et al., 2009) and 60,671 SNP markers (Wiggans et al., 2016). In the second part of the study, a single-step genomic prediction method was applied to the same RFI data set, using a high-density (**HD**) genotype chip. Genomic breeding values and prediction reliabilities were calculated in both parts of the study for several animal groups.

MATERIALS AND METHODS

Phenotypes

Residual feed intake data were collected from 4,823 lactations of 3,947 Holstein cows in 9 research herds in the central and eastern United States, including Iowa State University (Ames), the University of Wisconsin (Madison), the USDA Animal Genomics and Improvement Laboratory (Beltsville, MD), the University of Florida (Gainesville), the US Dairy Forage Research Center (Madison, WI), Michigan State University (East Lansing), the Purina Animal Nutrition Center (Gray Summit, MO), the Virginia Polytechnic Institute and State University (Blacksburg), and the Dairy Research Facility at the Miner Institute (Chazy, NY). Cows were born between 1999 and 2013. The experimental designs and contents of diets for the cows included in this study were described previously (Ferraretto et al., 2011, 2012; He et al., 2012; Spurlock et al., 2012; Connor et al., 2013; Yao et al., 2013; Manzanilla-Pech et al., 2016). The RFI phenotypes were calculated as the residuals of an energy sink model, where cows' DMI were adjusted for milk energy (**MilkE**), MBW, ΔBW, and several systematic effects including parity (primiparous vs. multiparous), DIM, experiment-specific rations, and test week (Tempelman et al., 2015). Nearly all RFI records were from 6-wk trials conducted during the first 50 to 250 DIM, but 202 records were from 4-wk trials. Those records from 4-wk trials were given less weight [weight $= (0.96)^2 = 0.92$] in the genetic analyses because the standard deviation was higher in 4-wk trials than 6-wk trials (1.75 vs. 1.68 kg/d) and the phenotypic correlation of 4- and 6-wk trials was 0.96.

Genotypes

Of the 3,947 cows with phenotypes, 3,505 cows were genotyped (502 cows with 777k SNP chip, 1,341 cows with 77k or 140k SNP chip, 1,251 cows with 50k SNP chip, and 411 cows with low-density chip). The genotypes were imputed to the standard national genomic evaluation set of 60,671 SNP markers (Wiggans et al., 2016), and were further imputed to a HD genotype panel of 312,614 SNP as part of a larger study that included 2,394 HD genotypes and a total of 592,757 genotyped Holsteins (VanRaden et al., 2017). Therefore, 2 types of genotype panels were available for the genotyped animals in this study, the officially used set of 60,671 SNP markers (denoted as 60k genotypes), and the HD genotype of 312,614 SNP markers. The quality control for 60k genotypes followed the procedures in (Wiggans

et al., 2010). High-density genotypes of 312,614 markers were available for 23,867 animals in the pedigree file. Animals and SNP with call rates <0.90, SNP with minor allele frequencies <0.05, monomorphic SNP, SNP deviating from Hardy-Weinberg equilibrium expectation, and animals with parent-progeny Mendelian conflicts were omitted from the HD genotypes using the preGSf90 program (version 1.10) (Misztal et al., 2018). After data filtering, a total of 278,524 SNP remained in the HD genotypes for 22,729 animals in the pedigree file.

National Genomic Evaluation for RFI

The genetic model for RFI was described previously in (VanRaden et al., 2018), where the model included the phenotype of RFI records, fixed effects of a few environmental effects, regression on inbreeding coefficients (from pedigree analyses), regression on genomic predicted transmitting ability (**GPTA**) for MilkE, regression on GPTA for body weight composite (**BWC**), and random effects of additive genetics, permanent environment, and random residual. The regressions on GPTA for MilkE and GPTA for BWC removed the remaining genetic correlations between RFI and MilkE/BWC that were not completely removed by the phenotypic regressions in the calculation of RFI (VanRaden et al., 2018).

Variance components were estimated for RFI phenotypes using MMAP release 2017_08_18 (O'Connell, 2017) with a repeatability model using either pedigree or genomic relationships among the cows. The estimated variance components were applied to further genomic evaluation for RFI.

A multi-step genomic evaluation method was used in the national genomic evaluation for RFI (VanRaden et al., 2009). A traditional evaluation was first carried out to estimate breeding values of RFI for more than 60 million Holsteins using pedigree relationships. The EBV from the traditional evaluation were then de-regressed and used as pseudo-phenotypes in the genomic evaluation to predict direct genomic values of RFI for 1.6 million genotyped Holsteins, using 60,671 SNP markers. In addition, trait deviations for RFI analogous to yield deviations were computed as RFI minus all nongenetic effects in the model and considered for use as the dependent variable instead of de-regressed EBV, but those gave slightly lower accuracies and lower correlations with single-step evaluations (results not shown).

Allele substitution effects for the SNP were estimated from de-regressed traditional EBV using an infinitesi-mal allele model with a heavy-tailed prior, in which smaller effects are regressed further toward 0 and markers with larger effects are regressed less to account for a non-normal distribution of marker effects (VanRaden, 2008). Final genomic breeding values (**GEBV**) combined 3 terms by selection index: (1) direct genomic value prediction, (2) parent average computed from the subset of genotyped ancestors using traditional relationships, and (3) parent average computed from all ancestors (VanRaden et al., 2009).

Reliability of prediction was estimated by different methods, including approximating reliabilities for evaluation (VanRaden et al., 2009; Liu et al., 2017), 5-way cross-validation (VanRaden and Hutchison, 2018), and using SCS data from the research cows to mimic RFI and estimate reliability using PTA correlations between research herd and national data for SCS since SCS has a similar heritability to RFI, but many more historic records in the database (VanRaden et al., 2018). The approximated genomic reliabilities for RFI were discounted by a factor of 0.3 to adjust the expected reliability to the observed reliability based on RFI validation studies using correlations from a 5-way cross-validation (VanRaden and Hutchison, 2018) and using SCS to mimic RFI (VanRaden et al., 2018). In addition, prediction reliabilities for RFI were also calculated by direct inversion of mixed model equations in the next part of the study on single-step genomic evaluation for RFI.

Single-Step Genomic Evaluation for RFI Using HD Genotypes

In the second part of the study, a single-step GBLUP method (**ssGBLUP**) was applied to estimate GEBV and to assess prediction reliabilities for RFI, using the same RFI data set but using HD genotypes for prediction. The ssGBLUP analysis for RFI was carried out using the program BLUPF90 (version 1.58) (Misztal et al., 2018), using a blended **H** matrix to combine pedigree and genomic information (Aguilar et al., 2010; Christensen and Lund, 2010). Pedigrees used in this ssGBLUP analysis traced back as many generations as possible for all cows with phenotypes and 5,252 elite young bulls under 5 yr old (based on the national evaluation in April 2018) and 4,029 heifers without lactation records born after January 2010 in the US research herds in which feed intake data were collected. The elite young bulls in the national database were included in the pedigree so that we would be able to estimate the prediction reliabilities for the young generations of top bulls. The young heifers in the research herds in which

intake data were collected should have closer relationships on average with the phenotyped cows, so it would be interesting to assess their prediction reliabilities. Pedigree included 94,900 animals, among which 22,729 animals have HD genotype information for the following single-step genomic evaluation for RFI.

The RFI model was based on VanRaden et al. (2018), where GPTA for milk net energy and GPTA for BWC were applied to genetic analyses for RFI to remove the remaining genetic correlations between RFI and MilkE/BWC. The GEBV for RFI estimated from the model were genetically and phenotypically uncorrelated with MilkE and BWC. The environmental effects that were removed from DMI to obtain RFI in Tempelman et al. (2015) were not included in the RFI model. The RFI model used in the ssGBLUP analyses was

$$y = \text{Age-parity-grp} + b_1 \times (\text{GPTA}_{\text{milk net energy}})$$
$$+ b_2 \times (\text{GPTA}_{\text{BWC}}) + a + pe + e,$$

where y is the RFI phenotype that has been adjusted for energy sinks of MilkE, MBW, ΔBW by phenotypic regressions, and several environmental effects, according to Tempelman et al. (2015a); Age-parity-grp is the fixed effect of the age and parity group; $\text{GPTA}_{\text{milk net energy}}$ is the cow's GPTA for milk net energy ($\text{GPTA}_{\text{milk net energy}}$ = $\text{GPTA}_{\text{Milk}} \times 0.327 + \text{GPTA}_{\text{Fat}} \times 12.95 + \text{GPTA}_{\text{Protein}} \times 7.65$) obtained from the US national genomic evaluation database for milk production traits; GPTA_{BWC} is the cow's GPTA for BWC obtained from the US national genomic evaluation database; $\text{GPTA}_{\text{milk net energy}}$ and GPTA_{BWC} were included to remove the remaining genetic correlations of RFI with milk production and BW traits that were not completely removed by the phenotypic regressions; b_1 is the regression coefficient of cow's RFI on GPTA of milk net energy, and b_2 is the regression coefficient of cow's RFI on GPTA of BWC; a is the random additive genetic effect with var(a) $\sim N\left(\mathbf{0}, \mathbf{H}\sigma_a^2\right)$, where σ_a^2 is the additive genetic variance and $\mathbf{H}$ is the relationship matrix incorporating pedigree and genomic information as defined in Legarra et al. (2009); pe is the random permanent environmental effect to account for repeated accounts from an animal, with var(pe) $\sim N\left(\mathbf{0}, \mathbf{I}\sigma_{pe}^2\right)$, where σ_{pe}^2 is the permanent environmental variance and $\mathbf{I}$ is the identity matrix; and e is the random residual with var(e) $\sim N\left(\mathbf{0}, \mathbf{R}\sigma_e^2\right)$, where σ_e^2 is the residual error variance and $\mathbf{R}$ is a diagonal matrix to adjust for the residual variance of each record based on their weights. Nearly all RFI records were from 6-wk trials but 202 records were from 4-wk trials. Those records from 4-wk trials were given

less emphasis (weight = 0.92), as described earlier in this paper. Therefore, the diagonal elements of the $\mathbf{R}$ matrix are mostly 1 (for records from 6-wk trials) and are 1/0.92 for records from 4-wk trials.

The variance components for RFI were estimated by an average information-restricted maximum likelihood algorithm using pedigree information implemented in the AIREMLF90 program (version 1.134; Misztal et al., 2018). The estimated variance components were then applied to single-step genomic prediction to estimate GEBV, implemented by the program BLUPF90 (version 1.58; Misztal et al., 2018). Prediction error variance (PEV) of the EBV for each animal was obtained by inverting the coefficient matrix of the mixed model equation. Theoretical reliability (REL) for each animal was calculated as

$$\text{REL}_i = 1 - \left(\frac{\text{PEV}_i}{G_i}\right),$$

where $G_i = (1 + \text{genomic inbreeding coefficient}) \times \sigma_a^2$ for animal i, and σ_a^2 is the additive genetic variance for RFI obtained from the previous variance components estimation.

In addition, variance of GEBV divided by the genetic variance was also calculated for RFI [i.e., var(GEBV)/σ_a^2, where var(GEBV) is the variance of the GEBV of the studied animals and σ_a^2 is the genetic variance for RFI obtained from the previous variance component estimation].

RESULTS AND DISCUSSION

National Genomic Evaluation for RFI

Heritability. The estimated heritability for RFI was 0.14 with either a pedigree or genomic model. This heritability agrees with the previous US study by (Lu et al., 2015) reporting the heritability for RFI as 0.14 when correlated traits were removed by multiple-trait modeling. Applying GPTA of MilkE and BWC to the RFI model could be one way to make RFI genetically uncorrelated with milk production traits and BWC in the selection index, to allow easy incorporation of RFI into NM$ in the US national evaluation. Without these regressions in the model, some of the top bulls for RFI (i.e., bulls with the lowest GPTA for RFI) were the top bulls for MilkE. Similar math to simultaneously model genetic and environmental correlations for RFI in dairy cattle was proposed by Lu et al. (2015, 2017). In the current study, GPTA for MilkE and BWC used estimates from full national data, which should estimate genetic

effects more precisely than using only the research cow data for those traits. Predictions of the DMI correlated with production and BW are much more accurate using the large national reference populations for those traits than using only the research cows. Furthermore, RFI is more independent from (has a lower correlation with) important index traits and is easier to assign an economic value. These were the main reasons to evaluate RFI rather than DMI in this study.

EBV and Reliabilities. The breeding values for RFI were estimated for 1.6 million genotyped Holsteins and 60 million non-genotyped Holsteins in the national database using the standard multi-step genomic evaluation method (VanRaden et al., 2009). The average prediction reliabilities were calculated for different animal groups (Table 1). Few daughters of any bull were measured for RFI; therefore, all bulls have low traditional reliabilities. Including genomic information increased the prediction reliability for RFI compared with using only pedigree information (Table 1). However, due to the limited data size, the average genomic reliability for RFI was low, at 34% for the 3,965 phenotyped cows and 13% for the 1.6 million genotyped Holsteins on average. The current genomic reliability for NM$ in the United States averaged 75% for young bulls and 91% for progeny-tested bulls (VanRaden et al., 2018), whereas genomic reliability for RFI was much lower, at 12% for top 100 net merit young bulls and 16% for top 100 net merit progeny-tested sires (Table 1). For SCS, which has a similar heritability to RFI, the average genomic reliability for SCS from national evaluation is 76% for young bulls and 89% for proven bulls, which is much higher than those for RFI. Compared with RFI, genomic prediction for SCS includes many more records and an extensive historical database.

The size of the reference population and its relationship to the predicted population remain as the limiting factors in the genomic prediction for RFI. Increasing prediction reliability for RFI requires collecting more feed intake data or exchanging data internationally. Focusing RFI data collection on relatives of elite bulls that will have greatest genetic contribution to the next generation will lead to more gains and profit (VanRaden et al., 2018). Similar conclusions regarding RFI phenotyping were reached in Cottle and van der Werf (2017) on optimizing the proportion of selection candidates measured for feed intake in beef cattle breeding.

GPTA Correlations Between RFI and Index Traits. The correlations between GPTA for RFI and GPTA for other traits in the index were calculated using GPTA information of the young bulls born in the year of 2017. Young bulls born in 2017 represent one of the recent generations and have more complete records compared with bulls born after 2017. The correlations

Table 1. Average prediction reliabilities in different animal groups in the national genomic evaluation for residual feed intake (RFI) in US Holstein dairy cattle

Animal group	RFI reliability (%)	
	Traditional[1]	Genomic[2]
3,965 cows with RFI phenotypes	30	34
Top 10 sires with most RFI daughters	78	85
Top 100 net merit progeny-tested sires	8	16
Top 100 net merit young bulls	3	12
1.6 million genotyped Holsteins	5	13
60 million nongenotyped Holsteins	3	3

[1]Traditional evaluation for RFI using pedigree relationship.
[2]Genomic evaluation for RFI, where reliabilities were calculated with a discount factor of 0.3 to match the expected reliability to observed reliability based on earlier RFI studies using 5-way cross-validation (VanRaden and Hutchison, 2018) and using SCS data to mimic RFI in the evaluation (VanRaden et al., 2018).

of GPTA for RFI with GPTA for milk yield, milk components, and BWC are zero, which was as expected based on the current RFI model setting. The GPTA correlation between RFI and current NM$ is −0.12, indicating favorable correlated response on current NM$ when selecting for RFI. The GPTA correlations of RFI with productive life, livability, and SCS are −0.17, −0.16, −0.07, respectively, indicating small but favorable genetic correlations of RFI with productive life, livability, and SCS. The GPTA correlations of RFI with fertility traits were −0.09 for daughter pregnant rate, −0.02 for heifer conception rate, −0.14 for cow conception rate, and 0.01 for gestation length. The low genetic correlations between RFI and fertility traits indicate minor correlated responses on dairy cattle fertility when selecting for RFI. In addition, GPTA for RFI and for calving traits index (including service sire calving ease, daughter calving ease, service sire stillbirths, and daughter stillbirths) have a correlation of 0.14. It is tempting to infer that this implies a low but unfavorable genetic correlation of RFI with calving traits, but PTA correlations are not equivalent to genetic correlations and the 2 can differ substantially, particularly when reliabilities are low (e.g., Calo et al., 1973; Blanchard et al., 1983). Generally, GPTA correlations are low between RFI and other traits in the index. However, it is worth noting that the estimates of GPTA correlations between RFI and other traits can be affected by the methods used to model RFI (e.g., different energy sinks included in the RFI model).

Prediction Reliability Using Single-Step Genomic Evaluation and HD Genotypes

The heritability estimate for RFI from ssGBLUP was 0.14, consistent with the national evaluation for RFI

Table 2. Theoretical reliabilities (REL) for residual feed intake (RFI), SD of estimated breeding values (GEBV), and the variance of GEBV divided by genetic variance of RFI [i.e., var(GEBV)/σ_a^2] in single-step genomic prediction for RFI using high-density genotypes in US Holstein dairy cattle

	Reliability (%) by animal groups		
Item	3,947 cows with RFI phenotypes	4,029 young heifers[1]	5,252 elite young bulls[2]
Mean of REL	31	15	18
SD (GEBV)	33	21	22
var(GEBV)/σ_a^2	26	10	11

[1]4,029 young heifers without lactation records born after January 2010 in US research herds in which feed intake data were collected.

[2]5,252 elite young bulls less than 5 yr old (based on the national evaluation in April 2018).

and the previous RFI study using the same US data (Lu et al., 2015). The theoretical reliabilities for RFI were low due to the limitations of the current data size (Table 2). For the 3,947 cows with RFI phenotypes, the mean theoretical reliabilities for RFI were 0.31 and 0.26, respectively. The theoretical reliability for the elite young bulls was only 0.18 on average (Table 2), and the reliability was lower in the younger generations of elite bulls compared with older bulls (Figure 1). For the youngest elite bulls (i.e., bulls less than 1 yr old in April 2018), the average theoretical reliability was only 0.13 from single-step genomic evaluation for RFI. This result was consistent with the national evaluation for RFI where the average genomic prediction reliability for the top 100 net merit young bulls was 0.12 (Table 1). Based on our current findings, the prediction reliabilities for RFI in elite young bulls are low given the current RFI data size. Continued feed intake collection is necessary in the coming years to prevent the next generations from having even lower reliabilities for RFI due to more distant relationships between currently recorded cows and future animals for selection. Higher reliabilities will require more feed intake information

from more animals and especially young progeny to increase the reference population size and maintain its relationship to the predicted population. Considering the high cost of data collection, focusing RFI data collection on relatives (daughters, sibs, or dams) of elite bulls that will have greatest genetic contribution to the next generation will give more genetic gains and financial profit. However, such targeted phenotyping of daughters of high-index bulls could result in biased evaluations because genotypes would not be properly randomized across environments. This would be conceptually related to, but distinct from, preferential treatment (Kuhn et al., 1994) and preselection bias (e.g., VanRaden, 2012).

The average theoretical reliability was 0.15 for young heifers without lactation records born after January 2010 in US research herds in which feed intake data were collected (Table 2). The young heifers from the research herds could be expected to have higher prediction reliabilities than young heifers in other herds because they might have closer relationships with the phenotyped cows in the research herds. Another related question is whether RFI of lactating cows and RFI of growing heifers are the same trait. A previous study on DMI showed a genetic correlation of 0.67 (SE = 0.24) between growing heifer DMI and that of lactating cows at 70 d postcalving (Berry et al., 2014). For RFI, Connor et al. (2019) recently reported that the phenotypic correlation is 0.37 for RFI in growing heifers with RFI during the first 100 d of lactation. The genetic correlation between RFI in growing heifers (n = 417) and RFI in first-lactation cows (n = 360) in early to mid-lactation was reported to be 0.58 (Nieuwhof et al., 1992). More accurate genetic correlation estimates for RFI are required between growing heifers and lactating cows of differing ages. Obtaining those may require collecting less-dense data for FE but at more time points across different life stages. If the genetic correlation is moderate to strong, RFI information from growing heifers could be used as a predictor of genetic
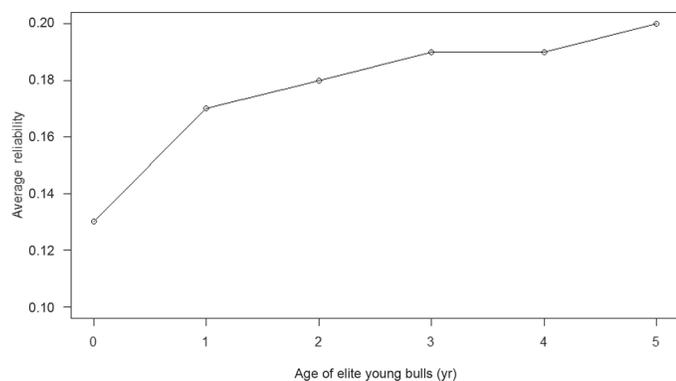
Figure 1. Theoretical reliabilities by age groups for elite young bulls less than 5 yr old (based on the national evaluation in April 2018) from single-step genomic prediction for residual feed intake.

merit for RFI in lactating cows and could be used to improve genomic evaluation for RFI in cows. Connor et al. (2019) estimated a phenotypic correlation of 0.37 for intake in growing heifers with intake in lactating cows, which suggests that heifer data could be used to increase prediction accuracy, but insufficient data were available for predicting genetic correlations.

Using HD genotypes in genomic prediction for RFI tended to give similar EBV and prediction reliability to the prediction using the standard 60,671 SNP set. In this study, the prediction reliability for RFI using the single-step method and HD genotypes was similar to that obtained from the national evaluation using the multi-step method with 60,671 genotypes. An extra study (unpublished data) on RFI has been done to compare predictions using 60,671 SNP set and HD genotypes, by applying the same RFI data set and the single-step genomic prediction method. In the results, analyses using HD genotypes and 60,671 genotypes gave similar EBV (correlation = 0.985) and prediction reliabilities (correlation = 0.999) for genotyped animals in the pedigree. Further studies are required to investigate the potential benefit of using HD genotypes in genomic prediction for RFI.

Publication of GEBV for Feed Efficiency

Genetic rankings for FE could be published in several ways based on our current results from RFI. First, RFI can be published as RFI$, representing the dollar value of RFI per lifetime. However, prediction reliabilities for RFI were generally low based on current data, so RFI could be combined with other FE-related traits in the definition of FE. Feed saved combines the reductions in feed eaten associated with RFI and lower maintenance predicted from BW (Pryce et al., 2015). The heritability for RFI is low to moderate at around 0.1 to 0.2 (Manzanilla-Pech et al., 2016), and the heritability for BW is moderate to high at around 0.5 to 0.6 (Li et al., 2018). Feed saved should have a higher heritability than RFI, and hence leads to a higher reliability compared with RFI. Based on the current US data set, if we assume prediction reliabilities average 12% for RFI and 72% for BWC in young animals, the combined average reliability for feed saved will be about 18% for young animals (VanRaden et al., 2018). In addition to Feed Saved, the Holstein Association USA (Brattleboro, VT) has published an economic estimate of feed efficiency (FE_{HA}) that combines the net income expected from higher milk yield and net loss expected from larger BWC (i.e., FE_{HA} = dollar value of milk produced − feed costs of extra milk − extra maintenance costs), but without using RFI information (Holstein Association USA, 2017). A more complete measure of feed efficiency (FeedEff) can be further subtracting the cost of RFI (RFI$) from FE_{HA} to account for the actual feed consumed (i.e., FeedEff = FE_{HA} − RFI$) (VanRaden et al., 2018).

Future Perspectives

Current findings indicate challenges but also opportunities in genomic evaluation for RFI. Obtaining more RFI phenotypes (with a good design for data recording) remains the main issue for this new trait. Increases in genomic reliability with additional RFI data can be forecasted. For young calves, such reliability could be 12% with 5,000, 19% with 10,000, 31% with 20,000, and 52% with 50,000 cows in the reference population (VanRaden and Hutchison, 2018). A related question is the timeliness of data collection. The prediction reliability for the youngest elite bulls (Figure 1) decreased substantially because they were less genetically connected to the previously recorded reference cows. A stable or growing rate of FE data collection is required so that newly arriving records provide more information than what is lost as previous data become more remote from current selection candidates. Thus, among data sets of 5,000 total records, recent records from 1,000 cows per year for 5 yr are better than 500 cows per year for 10 yr or 250 cows per year for 20 yr. Because data collection for FE is expensive, experimental design should guide the time plan of data collection, and international collaboration can help maximize reference population size.

Several other strategies in addition to maximizing size and timeliness of the reference population can help optimize FE data collection. First, data collection could be focused on certain groups of animals (e.g., relatives of elite bulls). Obtaining RFI phenotypes from relatives (daughters, sibs, or dams) of elite bulls that have greatest genetic contribution to the next generation will produce more genetic gains and financial profit. Second, considering the high cost of data collection, sparse recording of feed intake data on more animals could be more useful to obtain large-scale RFI phenotypes and to obtain FE information from different life stages of animals (Negussie et al., 2019). Third, strategies to use new phenotypes (e.g., milk spectra data) or on-farm applications (e.g., sensors, smart collars, and imaging technology) may also provide opportunities to generate proxy phenotypes for DMI to speed up data collection for FE (Chizzotti et al., 2015; Ruuska et al., 2016; Thomasen et al., 2018).

CONCLUSIONS

Residual feed intake, a candidate FE trait in dairy cattle, can be defined to be genetically uncorrelated with major energy sink traits (e.g., milk production, BW), by including GPTA of these energy sink traits in the genetic analyses for RFI. Including RFI in the US national dairy cattle evaluation gave the EBV estimates of RFI for 1.6 million genotyped Holsteins and 60 million ungenotyped Holsteins. The GPTA correlations between RFI and other traits in the index (e.g., fertility) are generally low, indicating minor correlated responses on other index traits when selecting for RFI. Including genomic information increased the prediction reliability for RFI compared with using only pedigree information. Due to the limitation of data size, the current prediction reliability was low, 34% for phenotyped animals, 13% for 1.6 million genotyped animals, 16% for the top 100 net merit progeny-tested bulls. The theoretical reliability for the elite young bulls under 5 yr old was only 0.18 on average, and the reliability was lower in the younger generations of elite bulls compared with older bulls. Continued feed intake collection is necessary in the coming years to prevent the next generations to have lower reliabilities due to the lower relationship between currently recorded cows and future animals for selection. Considering the currently low prediction reliability and high cost of data collection, focusing RFI data collection on relatives of elite bulls that will have greatest genetic contribution to the next generation will give more gains and profit.

REFERENCES

Aguilar, I., I. Misztal, D. L. Johnson, A. Legarra, S. Tsuruta, and T. J. Lawlor. 2010. Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. J. Dairy Sci. 93:743–752. https://doi.org/10.3168/jds.2009-2730.

Berry, D. P., M. P. Coffey, J. E. Pryce, Y. de Haas, P. Løvendahl, N. Krattenmacher, J. J. Crowley, Z. Wang, D. Spurlock, K. Weigel, K. Macdonald, and R. F. Veerkamp. 2014. International genetic evaluations for feed intake in dairy cattle through the collation of data from multiple sources. J. Dairy Sci. 97:3894–3905. https://doi.org/10.3168/jds.2013-7548.

Berry, D. P., and J. J. Crowley. 2013. CELL BIOLOGY SYMPOSIUM : Genetics of feed efficiency in dairy and beef cattle. J. Anim. Sci. 91:1594–1613. https://doi.org/10.2527/jas.2012-5862.

Blanchard, P. J., R. W. Everett, and S. R. Searle. 1983. Estimation of genetic trends and correlations for Jersey cattle. J. Dairy Sci. 66:1947–1954. https://doi.org/10.3168/jds.S0022-0302(83)82033-5.

Calo, L. L., R. E. McDowell, L. D. Van Vleck, and P. D. Miller. 1973. Genetic aspects of beef production among Holstein-Friesians pedigree selected for milk production. J. Anim. Sci. 37:676–682. https://doi.org/10.2527/jas1973.373676x.

Chizzotti, M. L., F. S. Machado, E. E. L. Valente, L. G. R. Pereira, M. M. Campos, T. R. Tomich, S. G. Coelho, and M. N. Ribas. 2015. Technical note: Validation of a system for monitoring individual feeding behavior and individual feed intake in dairy cattle. J. Dairy Sci. 98:3438–3442. https://doi.org/10.3168/jds.2014-8925.

Christensen, O. F., and M. S. Lund. 2010. Genomic prediction when some animals are not genotyped. Genet. Sel. Evol. 42:2. https://doi.org/10.1186/1297-9686-42-2.

Connor, E. E., J. L. Hutchison, H. D. Norman, K. M. Olson, C. P. Van Tassell, J. M. Leith, and R. L. Baldwin VI. 2013. Use of residual feed intake in Holsteins during early lactation shows potential to improve feed efficiency through genetic selection. J. Anim. Sci. 91:3978–3988. https://doi.org/10.2527/jas.2012-5977.

Connor, E. E., J. L. Hutchison, C. P. Van Tassell, and J. B. Cole. 2019. Defining the optimal period length and stage of growth or lactation to estimate residual feed intake in dairy cows. J. Dairy Sci. 102:6131–6143. https://doi.org/10.3168/jds.2018-15407.

Cottle, D. J., and J. H. J. van der Werf. 2017. Optimising the proportion of selection candidates measured for feed intake for a beef cattle breeding objective that includes methane emissions. J. Anim. Sci. 95:1030–1041. https://doi.org/10.2527/jas.2016.1177.

European Commission. 2018. EU Dairy Farms Report Based on 2016 FADN Data. Accessed Mar. 12, 2019. http://ec.europa.eu/agriculture/rica/pdf/Dairy_report_2016.pdf.

Ferraretto, L. F., R. D. Shaver, and S. J. Bertics. 2012. Effect of dietary supplementation with live-cell yeast at two dosages on lactation performance, ruminal fermentation, and total-tract nutrient digestibility in dairy cows. J. Dairy Sci. 95:4017–4028. https://doi.org/10.3168/jds.2011-5190.

Ferraretto, L. F., R. D. Shaver, M. Espineira, H. Gencoglu, and S. J. Bertics. 2011. Influence of a reduced-starch diet with or without exogenous amylase on lactation performance by dairy cows. J. Dairy Sci. 94:1490–1499. https://doi.org/10.3168/jds.2010-3736.

Hardie, L. C., M. J. VandeHaar, R. J. Tempelman, K. A. Weigel, L. E. Armentano, G. R. Wiggans, R. F. Veerkamp, Y. de Haas, M. P. Coffey, E. E. Connor, M. D. Hanigan, C. Staples, Z. Wang, J. C. M. Dekkers, and D. M. Spurlock. 2017. The genetic and biological basis of feed efficiency in mid-lactation Holstein dairy cows. J. Dairy Sci. 100:9061–9075. https://doi.org/10.3168/jds.2017-12604.

He, M., K. L. Perfield, H. B. Green, and L. E. Armentano. 2012. Effect of dietary fat blend enriched in oleic or linoleic acid and monensin supplementation on dairy cattle performance, milk fatty acid profiles, and milk fat depression. J. Dairy Sci. 95:1447–1461. https://doi.org/10.3168/jds.2011-4635.

Holstein Association USA. 2017. TPI Formula - August 2017. Accessed Aug. 31, 2017. http://www.holsteinusa.com/genetic_evaluations/ss_tpi_formula.html.

Kuhn, M. T., P. J. Boettcher, and A. E. Freeman. 1994. Potential biases in predicted transmitting abilities of females from preferential treatment. J. Dairy Sci. 77:2428–2437. https://doi.org/10.3168/jds.S0022-0302(94)77185-X.

Legarra, A., I. Aguilar, and I. Misztal. 2009. A relationship matrix including full pedigree and genomic information. J. Dairy Sci. 92:4656–4663. https://doi.org/10.3168/jds.2009-2061.

Li, B., W. F. Fikse, P. Løvendahl, J. Lassen, M. H. Lidauer, P. Mäntysaari, and B. Berglund. 2018. Genetic heterogeneity of feed intake, energy-corrected milk, and body weight across lactation in primiparous Holstein, Nordic Red, and Jersey cows. J. Dairy Sci. 101:10011–10021. https://doi.org/10.3168/jds.2018-14611.

Liu, Z., P. M. VanRaden, M. H. Lidauer, M. P. Calus, H. Benhajali, H. Jorjani, and V. Ducrocq. 2017. Approximating genomic reliabilities for national genomic evaluation. Interbull Bull. 75–85.

Lu, Y., M. J. Vandehaar, D. M. Spurlock, K. A. Weigel, L. E. Armentano, E. E. Connor, M. Coffey, R. F. Veerkamp, Y. de Haas, C. R. Staples, Z. Wang, M. D. Hanigan, and R. J. Tempelman. 2018. Genome-wide association analyses based on a multiple-trait approach for modeling feed efficiency. J. Dairy Sci. 101:3140–3154. https://doi.org/10.3168/jds.2017-13364.

Lu, Y., M. J. Vandehaar, D. M. Spurlock, K. A. Weigel, L. E. Armentano, C. R. Staples, E. E. Connor, Z. Wang, N. M. Bello, and R. J. Tempelman. 2015. An alternative approach to modeling genetic merit of feed efficiency in dairy cattle. J. Dairy Sci. 98:6535–6551. https://doi.org/10.3168/jds.2015-9414.

Lu, Y., M. J. Vandehaar, D. M. Spurlock, K. A. Weigel, L. E. Armentano, C. R. Staples, E. E. Connor, Z. Wang, M. Coffey, R. F. Veerkamp, Y. de Haas, and R. J. Tempelman. 2017. Modeling genetic and nongenetic variation of feed efficiency and its partial relationships between component traits as a function of management and environmental factors. J. Dairy Sci. 100:412–427. https://doi.org/10.3168/jds.2016-11491.

Manzanilla-Pech, C. I. V., R. F. Veerkamp, R. J. Tempelman, M. L. van Pelt, K. A. Weigel, M. VandeHaar, T. J. Lawlor, D. M. Spurlock, L. E. Armentano, C. R. Staples, M. Hanigan, and Y. De Haas. 2016. Genetic parameters between feed-intake-related traits and conformation in 2 separate dairy populations–the Netherlands and United States. J. Dairy Sci. 99:443–457. https://doi.org/10.3168/jds.2015-9727.

Misztal, I., S. Tsuruta, D. A. L. Lourenco, Y. Masuda, I. Aguilar, A. Legarra, and Z. Vitezica. 2018. Manual for BLUPF90 family of programs. Vol. 2018. Accessed Dec. 26, 2019. http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all7.pdf.

Negussie, E., T. Mehtiö, P. Mäntysaari, P. Løvendahl, E. A. Mäntysaari, and M. H. Lidauer. 2019. Reliability of breeding values for feed intake and feed efficiency traits in dairy cattle: When dry matter intake recordings are sparse under different scenarios. J. Dairy Sci. 102:7248–7262. https://doi.org/10.3168/jds.2018-16020.

Nieuwhof, G. J., J. A. M. van Arendonk, H. Vos, and S. Korver. 1992. Genetic relationships between feed intake, efficiency and production traits in growing bulls, growing heifers and lactating heifers. Livest. Prod. Sci. 32:189–202. https://doi.org/10.1016/S0301-6226(12)80001-7.

O'Connell, J. R. 2017. MMAP: Mixed Model Analysis for Pedigrees and Populations. Accessed Aug. 1, 2017. https://mmap.github.io/.

Pryce, J. E., O. Gonzalez-Recio, G. Nieuwhof, W. J. Wales, M. P. Coffey, B. J. Hayes, and M. E. Goddard. 2015. Hot topic: Definition and implementation of a breeding value for feed efficiency in dairy cows. J. Dairy Sci. 98:7340–7350. https://doi.org/10.3168/jds.2015-9621.

Ruuska, S., S. Kajava, M. Mughal, N. Zehner, and J. Mononen. 2016. Validation of a pressure sensor-based system for measuring eating, rumination and drinking behaviour of dairy cattle. Appl. Anim. Behav. Sci. 174:19–23. https://doi.org/10.1016/j.applanim.2015.11.005.

Spurlock, D. M., J. C. M. Dekkers, R. Fernando, D. A. Koltes, and A. Wolc. 2012. Genetic parameters for energy balance, feed efficiency, and related traits in Holstein cattle. J. Dairy Sci. 95:5393–5402. https://doi.org/10.3168/jds.2012-5407.

Tempelman, R. J., D. M. Spurlock, M. Coffey, R. F. Veerkamp, L. E. Armentano, K. A. Weigel, Y. de Haas, C. R. Staples, E. E. Connor, Y. Lu, and M. J. VandeHaar. 2015a. Heterogeneity in genetic and nongenetic variation and energy sink relationships for residual feed intake across research stations and countries. J. Dairy Sci. 98:2013–2026. https://doi.org/10.3168/jds.2014.8510.

Thomasen, J. R., J. Lassen, G. G. B. Nielsen, C. Borggard, P. R. B. Stentebjerg, R. H. Hansen, N. W. Hansen, and S. Borchersen. 2018. Individual cow identification in a commercial herd using 3D camera technology. Proc. World Congr. Genet. Appl. Livest. Prod. 11:613.

USDA. 2018. Milk Cost-of-Production Estimates-2016 Base, Milk Cost of Production Estimates. Accessed Feb. 10, 2018. https://www.ers.usda.gov/data-products/milk-cost-of-production-estimates/milk-cost-of-production-estimates/#Milk Cost-of-Production Estimates-2016 Base.

VandeHaar, M. J., L. E. Armentano, K. Weigel, D. M. Spurlock, R. J. Tempelman, and R. Veerkamp. 2016. Harnessing the genetics of the modern dairy cow to continue improvements in feed efficiency. J. Dairy Sci. 99:4941–4954. https://doi.org/10.3168/jds.2015-10352.

VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. J. Dairy Sci. 91:4414–4423. https://doi.org/10.3168/jds.2007-0980.

VanRaden, P. M. 2012. Avoiding bias from genomic pre-selection in converting daughter information across countries. Interbull Bull. 1–5.

VanRaden, P. M., and J. L. Hutchison. 2018. Methods to compute reliabilities for genomic predictions of feed intake. J. Dairy Sci. 101 (Suppl. 2):370–371 (Abstr. 392) in American Dairy Science Association Proceedings, Knoxville, TN. American Dairy Science Association, Champaign, IL.

VanRaden, P. M., J. R. O'Connell, E. E. Connor, M. J. Vandehaar, R. J. Tempelman, and K. A. Weigel. 2018. Including feed intake data from U.S. Holsteins in genomic prediction. Page 125 in Proc. 11th World Congress on Genetics Applied to Livestock Production, Vol. Biology–Feed Intake and Efficiency 1. World Congress on Genetics Applied to Livestock Production, Auckland, New Zealand.

VanRaden, P. M., M. E. Tooker, J. R. O'Connell, J. B. Cole, and D. M. Bickhart. 2017. Selecting sequence variants to improve genomic predictions for dairy cattle. Genet. Sel. Evol. 49:32. https://doi.org/10.1186/s12711-017-0307-4.

VanRaden, P. M., C. P. Van Tassell, G. R. Wiggans, T. S. Sonstegard, R. D. Schnabel, J. F. Taylor, and F. S. Schenkel. 2009. Invited Review: Reliability of genomic predictions for North American Holstein bulls. J. Dairy Sci. 92:16–24. https://doi.org/10.3168/jds.2008-1514.

Veerkamp, R. F., M. P. L. Calus, G. De Jong, R. Van Der Linde, and Y. De Haas. 2014. Breeding value for dry matter intake for Dutch bulls based on DGV for DMI and BV for predictors. Page 115 in Proceedings of 10th World Congress of Genetics Applied to Livestock Production.

Wallén, S. E., M. Lillehammer, and T. H. E. Meuwissen. 2017. Strategies for implementing genomic selection for feed efficiency in dairy cattle breeding schemes. J. Dairy Sci. 100:6327–6336. https://doi.org/10.3168/jds.2016-11458.

Wiggans, G. R., T. Cooper, P. VanRaden, C. Van Tassell, D. Bickhart, and T. Sonstegard. 2016. Increasing the number of single nucleotide polymorphisms used in genomic evaluation of dairy cattle. J. Dairy Sci. 99:4504–4511. https://doi.org/10.3168/jds.2015-10456.

Wiggans, G. R., P. M. VanRaden, L. R. Bacheller, M. E. Tooker, J. L. Hutchison, T. A. Cooper, and T. S. Sonstegard. 2010. Selection and management of DNA markers for use in genomic evaluation. J. Dairy Sci. 93:2287–2292. https://doi.org/10.3168/jds.2009-2773.

Yao, C., D. M. Spurlock, L. E. Armentano, C. D. Page Jr., M. J. Vandehaar, and D. M. Bickhart. 2013. Random Forests approach for identifying additive and epistatic single nucleotide polymorphisms associated with residual feed intake in dairy cattle. J. Dairy Sci. 96:6716–6729. https://doi.org/10.3168/jds.2012-6237.

ORCIDS

P. M. VanRaden ⓘ https://orcid.org/0000-0002-9123-7278
E. Guduk ⓘ https://orcid.org/0000-0002-9724-0566
E. E. Connor ⓘ https://orcid.org/0000-0002-7846-7247
M. J. VandeHaar ⓘ https://orcid.org/0000-0002-8475-3493
R. J. Tempelman ⓘ https://orcid.org/0000-0002-7833-6730
K. A. Weigel ⓘ https://orcid.org/0000-0002-2391-6260
J. B. Cole ⓘ https://orcid.org/0000-0003-1242-4401