



Parameters of a dynamic mechanistic model of cattle growth retain enough biological interpretation for genotype-to-phenotype mapping

M.C. Freua¹, M.H.A. Santana¹ R.V. Ventura^{1,2} and J.B.S. Ferraz¹

¹Departamento de Medicina Veterinária,
Faculdade de Zootecnia e Engenharia de Alimentos, São Paulo, SP, Brasil

²Centre for Genetic Improvement for Livestock, University of Guelph,
Guelph, Ontario, Canada

Corresponding author: M.H.A. Santana
E-mail: miguel-has@hotmail.com

Genet. Mol. Res. 15 (4): gmr15048931

Received June 30, 2016

Accepted October 27, 2016

Published December 2, 2016

DOI <http://dx.doi.org/10.4238/gmr15048931>

Copyright © 2016 The Authors. This is an open-access article distributed under the terms of the Creative Commons Attribution ShareAlike (CC BY-SA) 4.0 License.

ABSTRACT. This study aimed to investigate the predictability of a phenotype when using a dynamic model of cattle growth. Genotypic and phenotypic information on Nelore (*Bos indicus*) cattle were used in a genome-wide association analysis designed to contrast the biological interpretation of core parameters [conversion efficiency of metabolizable energy to net energy for gain (k_g) and adjusted final shrunk body weight (AFSBW)] to their associated genomic regions and nearby quantitative trait loci (QTLs). Single nucleotide polymorphisms (SNPs) were used to develop prediction equations for k_g and AFSBW, which enter the model for simulative prediction purposes. QTLs and genes, one related to mature body weight and another to growth

efficiency, are consistent with the model equations. Significantly associated SNPs were used to compute parameters, which yielded reasonable model outcomes when compared with regular parameter computations. Our results provide evidence of the biological validity of using such parameters as component traits of higher phenotypes and the possibility of using genomic data for genotype-to-parameter mapping.

Key words: Beef cattle; Computational physiological genomics; Genome-wide association; Mechanistic modeling; Nellore

INTRODUCTION

In animal genetics, the expectation of a given phenotype is usually defined by a linear model, regardless of the complex biological hierarchy underlying the trait. Studies of nonlinearity between phenotypes and gene function have been reported (Francesconi and Lehner, 2014; Peng et al., 2015). However, when a complex set of variables interact to yield a phenotype, different approaches might provide some additional insights.

Genotype-to-phenotype mapping is not an easy task, because their direct association is difficult. This explains why the rise of genomics is still deterred by missing heritability, i.e., genotyping individuals does not sufficiently explained full heritable variance (Bogardus, 2009). Mapping from the genotype-to-phenotype space remains largely unknown because of a reaction space that bridges the gap. It is reasonable to hypothesize that dynamic models with some description of the hierarchy underlying a phenotype, bring in their structure a more stable behavior throughout this reaction space. In other words, genetic variance penetrates parameters directly and linearly (Wang et al., 2012).

Mathematical modeling is a solid discipline in the field of animal science. In particular, in animal nutrition, models have long been used to describe the nutrient requirements of the animal, nutrient supplied by the feed, and the growth and development of various livestock species. Assuming that such models correctly represent the environmental term of phenotypic variance, it is possible that the dynamic mechanistic model theory, combined with SNP models, could improve phenotype predictability and provide insight into genes and metabolic pathways underlying phenotypic variation. Approaches used toward this end have come to promising conclusions. Model parameters, or lower-level phenotypes, have the potential to identify more causative SNPs than association studies performed using complex phenotypes directly, i.e., outcomes of model behavior (Wang et al., 2012).

As a case study, we used the Cornell Cattle Value Discovery System (CVDS, Tedeschi et al., 2004), which is a dynamic mechanistic model of cattle growth that is able to simulate individuals. Our objective was to determine whether the core parameters of the model would retain enough biological interpretation to aid the genotype-to-phenotype mapping of complex phenotypes.

MATERIAL AND METHODS

Ethical statement

No statement from the local animal Ethics Committee was necessary for this research. The study described herein used a dataset obtained from experiments published elsewhere

(Santana et al., 2012; Gomes et al., 2013; Alexandre et al., 2015), which includes phenotype records and DNA samples from tests approved by the respective institutional animal Ethics Committee. All animal procedures were undertaken using common management practices in beef cattle feedlot operations and in accordance with standard veterinary protocols.

CVDS model

The dynamic, mechanistic model of cattle growth described by Tedeschi et al. (2004) was adopted for this study. The CVDS model is comprised of sequential equations that simulate cattle growth and development, and that has been commercially applied to beef feedlots in the United States. The model is based on the National Research Council (NRC, 2000) energy system to predict growth rate and body weight (BW) on a daily basis from the net energy available for maintenance (NEm) and growth (NEg). To account for the fact that the composition of gain at a particular weight is a function of mature BW, simulations were based on the equivalent shrunk BW (SBW), which is the standard reference weight (SRW), an overall mean BW at the same degree of maturity, multiplied by the ratio between the actual SBW and a BW value at 28% empty body fat (EBF). The CVDS calculates dry matter intake (DMI) and the NEm requirement to obtain energy available for growth from the remaining energy and the conversion efficiency of metabolizable energy to net energy for gain (k_g).

The CVDS model was generated in Visual Basic 6 and is available at <http://nutritionmodels.com/cvds.html>. All simulations described herein were performed without the environmental submodel and with the exponential decay adjustment for k_g , as recommended by Tedeschi et al. (2004). Statistical analyses were conducted in R v.3.1.3 (R Core Team, 2015).

Polymorphic parameters

The SBW adjusted to 28% EBF (adjusted final shrunk BW, AFSBW) is the core parameter that represents phenotypic variability when comparing CVDS predictions among contemporary individuals fed the same diet in the same feedlot. Because the dataset used herein was comprised of *Bos indicus* animals, we performed simulations to determine the optimal SRW. Values of SRW close to 310 kg empty BW (EBW) at 22% EBF, yielded the best prediction, consistent with values reported by Tedeschi et al. (2002). To compute AFSBW, an equation from Baker et al. (2006) for ultrasound information was used:

$$FSBW = \frac{EBW + 14.26 \times (22 - pEBF)}{0.895} \quad (\text{Equation 1})$$

where *EBW* is empty body weight (kg) and *pEBF* is the predicted empty body fat percent from ultrasound information. The constant value of 14.26, as reported by Guiroy et al. (2001), represents the relationship between *pEBF* and *EBW*. The value 0.895 is assumed by Marcondes et al. (2010) to be a better conversion factor between *EBW* and *SBW* for *B. indicus* animals than the value of 0.891 assumed by the NRC (2000).

Equations described by Chizzotti et al. (2008) to compute *EBW*, and to predict body composition from ultrasound information, were used to obtain the *pEBF*:

$$EBW = -15.6 + 0.928 \times SBW$$

$$pEBF = \frac{-47.26 + 2.82 \times BFT + 0.2993 \times EBW}{EBW} \times 100 \quad (\text{Equation 2})$$

where EBW is empty body weight (kg), and BFT is body fat thickness (mm).

k_g was adopted as a second parameter in this study. Although simulated by CVDS, k_g requires an initial condition that is obtained from dietary information, and it represents individual variability as the exponential decay function calculated to retain energy as protein, which is used to update k_g , and therefore daily gain, for each time step. For statistical analysis, a mean value of k_g obtained from simulations over the entire feed days was used.

Global sensitivity analysis

Changes in model parameters represent changes in outcome variables, usually in the phenotypes of interest. To assess the parameter-to-phenotype map, we conducted a global sensitivity analysis by sampling AFSBW and k_g values with the Latin-hypercube sampling (LHS) method and then running the CVDS model to obtain predicted variables. The impact of parameters on mean DMI, average daily gain (ADG), and final EBF was quantified by linear regression of those variables on parameter values. This procedure was repeated 100 times using a Monte Carlo simulation approach. The coefficient of determination was retained for each iteration, to obtain distributions. For the LHS, AFSBW and k_g were assumed to be correlated ($r = 0.58$, $P < 0.001$). For adequate simulations, initial BW was also sampled with LHS, and correlations were assumed for both AFSBW ($r = 0.87$, $P < 0.001$) and k_g ($r = 0.88$, $P < 0.001$).

Animals and phenotypes

Individual measurements ($N = 1435$) of BW and back-fat thickness were obtained by ultrasound for Nellore (*B. indicus*) young bulls and steers from 16 feed efficiency trials conducted in feedlot operations in South (Santana et al., 2012), Southeast (Gomes et al., 2013; Alexandre et al., 2015), and Central-West Brazil (Santana et al., 2014). Data were used to compute lower-level phenotypes, i.e., individual polymorphic parameters, and to perform simulations with CVDS. When dietary information was not available, average diet metabolizable energy (ME) was assumed to be 3.0 Mcal/kg. In accordance with the methods used by Marcondes et al. (2010), initial conditions for NEm and NEg were set at 67 and 44% ME, respectively. Phenotypic data were tested for normality with the Shapiro-Wilk test, and individuals that were more than twice the interquartile range above the third quartile or below the first quartile were excluded.

Genotyping and imputation

Genomic DNA of 893 animals was extracted from either blood samples or hairs. Samples were obtained from those used in previous trials (Santana et al., 2012; Gomes et al., 2013; Alexandre et al., 2015), and DNA was prepared as previously described. Genotyping was performed with Illumina BovineHD[®] BeadChip (777,962 SNPs), Affymetrix Axiom[®] Genome-Wide BOS1 Array (648,874 SNPs), GGP Indicus Neogen HD[®] (74,677 SNPs), and

Illumina BovineSNP50[®] BeadChip (54,609 SNPs) platforms according to the manufacturer protocols. Only genotype calls (standard cluster quality) greater than 0.70 and samples with a call rate higher than 90% were used.

Imputation from these panels to a super-dense panel (SDP) was performed in two steps. First, a dense panel (1,261,128 SNPs) was created that combined genotypic information from Illumina BovineHD and Affymetrix BOS1 of 279 animals genotyped with both panels. Next, 893 animals with genotypic and phenotypic records were imputed to SDP level. Imputation accuracy was determined by cross-validation analysis within each panel in a parallel investigation, and the concordance rate between the imputed and true genotypes across all scenarios was higher than 97.51%. All imputation procedures were performed in the FImpute 2.2 program (Sargolzaei et al., 2014). Genotypic data were filtered and only autosomal SNPs with a minor allele frequency >2% and deviation from Hardy-Weinberg equilibrium (χ^2 -test, 1 d.f., $P > 1 \times 10^{-5}$) were retained for use in the association test.

Association analysis

A two-step method, GRAMMAR-Gamma, which accounts for relatedness and population substructure, and uses mixed-effect modeling and regression to estimate the SNPs effects (Svishcheva et al., 2012), was used to perform genome-wide association studies (GWAS) on the polymorphic parameters. Briefly, in the first step, the fixed effects for the contemporary group and age were estimated for the phenotypic data using a mixed model. A Gamma correction factor was calculated from the computed variance-covariance matrix to correct for population substructure (i.e., genomic relationship matrix). In the second step, the transformed phenotypic data were regressed to genotypic information, and the estimated SNP effect was corrected by the Gamma factor. Those computations were performed using the GenABEL v1.7-6 package for R (Aulchenko et al., 2007). The genome-wide threshold used to declare significant SNPs was a modified Bonferroni correction for multiple testing with a 5% significant level (Gao et al., 2008).

The 1-Mb regions surrounding significant SNPs were investigated to determine whether they mapped against any previously described QTL deposited in the cattle QTLdb database (Hu et al., 2013). The Gene Ontology Annotation Database, the Kyoto encyclopedia of Genes and Genomes, and Reactome were used to identify potential functions and biological pathways of annotated genes.

Genotype to parameter simulation

After GWAS was performed, significant SNPs related to AFSBW and k_g were assumed to be true causative SNPs of the population. To generate a genotype-to-parameter map, the dataset was divided into two different subsets using a cross-validation approach. The training group was used to obtain predictive equations of parameter values given SNP genotypes. The AFSBW and k_g of individuals from the second subset (i.e., the testing group) were then predicted and entered into the CVDS for simulative prediction purposes. At this stage, genetically predicted k_g was not considered as an input that guaranteed simulations equal to those performed with the computed parameters, i.e., with initial k_g from dietary information and the exponential decay adjustment to compute k_g . However, k_g was retained in the study to assess the possibility of prediction with genomic data.

This procedure was performed using Monte Carlo simulation for 100 iterations to assess the possibility of running the model with parameter values obtained from genomic information. At each iteration, DMI, ADG, and final EBF were compared to values calculated from simulations with computed parameters by regressing them on the newly simulated outcomes to obtain distributions of the coefficient of determination.

Equations used to predict AFSBW and k_g were obtained using a non-weighted allelic profile approach (Aulchenko et al., 2009), and thus, the allele substitution effect estimated with GWAS was not included in the model. The sum of the major alleles in the genotype of an individual is considered in a joint estimation additive linear model, as follows:

$$y = \beta_i BW + \beta_j age + \beta_{Bk} (PAB_k + 2PBB_k) \quad (\text{Equation 3})$$

where y is either AFSBW or k_g , age is age at the beginning of the feeding period (days), PAB and PBB are the estimated probabilities of the AB and BB genotypes of the k^{th} SNP marker (k from 1 to k^{th} most significant SNP), and b_n represents regression coefficients. Initial BW and age at the beginning of the feedlot period were included in the model because they are usually available at a commercial plant, and thus, may be good predictor variables.

RESULTS AND DISCUSSION

Our approach is similar to causally cohesive genotype-phenotype (cGP) models (Wang et al., 2012) for computation of physiological genomics; however, the stage of model development adopted here does not permit direct measurement of parameters as phenotypes. Therefore, these parameters were estimated properly. Polymorphic parameters were calculated for the 1435 animals in the dataset (Table 1), and estimates of mean and standard deviations were used for the LHS in the sensitivity analysis (Table 2).

Table 1. Descriptive statistics of the phenotypes.

| Phenotype ^a | Minimum | Mean | Maximum | SD ^b |
|------------------------|---------|-------|---------|-----------------|
| AFSBW | 349.5 | 469.2 | 679 | 50.83 |
| k_g | 0.45 | 0.53 | 0.58 | 0.02 |
| DOF | 35 | 73.2 | 90 | 10.64 |
| iBW | 216 | 357 | 545.5 | 60.34 |
| fBW | 245 | 450.3 | 705.5 | 78.58 |
| iBF | 0.0 | 1.53 | 6.35 | 1.36 |
| fBF | 1.0 | 4.23 | 15.6 | 3.9 |
| DMI | 5.98 | 9.03 | 18.3 | 1.37 |
| ADG | 1.3 | 1.88 | 2.56 | 0.27 |
| fEBF | 9.4 | 14.77 | 17.89 | 1.66 |

^aAFSBW, adjusted final shrunk body weight (kg); k_g , efficiency of conversion of dietary metabolizable energy to net energy for growth (dimensionless); DOF, days on feed (days); iBW, initial body weight (kg); fBW, final body weight (kg); iBF, initial back-fat thickness (mm); fBT, final back-fat thickness (mm); DMI, mean dry matter intake for the DOF as simulated by CVDS (kg); ADG, average daily gain for the DOF as simulated by CVDS (kg); fEBF, final empty body fat as simulated by CVDS (% of empty body weight). ^bSD, standard deviation.

The determination of BW at a given body composition, as represented by AFSBW, is regarded as a key parameter that shapes growth models and influences the accuracy and precision of model estimates (Arnold and Bennetti, 1991).

Table 2. Coefficient of determination (R^2) from the global sensitivity analysis.

| Phenotypes ^a | AFSBW ^b | | | k_g ^c | | |
|-------------------------|--------------------|-----------------------|-----------------------|--------------------|-------|-------|
| | Mean | 2.5% | 97.5% | Mean | 2.5% | 97.5% |
| DMI | 0.455 | 0.4 | 0.504 | 0.109 | 0.079 | 0.14 |
| ADG | 0.642 | 0.592 | 0.684 | 0.016 | 0.004 | 0.03 |
| EBF | 0.003 | 2.34×10^{-5} | 1.13×10^{-2} | 0.039 | 0.012 | 0.071 |

^aDMI, dry matter intake; ADG, average daily gain; and EBF, empty body fat. ^bAFSBW, adjusted final shrunk body weight. Mean and quartiles for Monte Carlo runs. ^c k_g , efficiency of conversion of dietary metabolizable energy to net energy for growth. Mean and quartiles for Monte Carlo runs.

The AFSBW is responsible for distinguishing individuals with respect to DMI and ADG, as shown by our sensitivity analysis. On average, AFSBW explains 46 and 64% of the phenotypic variation in DMI and ADG, respectively. One of the main sources of error in the CVDS model is the computation of AFSBW when carcass traits or ultrasound measurements are not available, or when it has to be computed from biometric measurements (Fox et al., 1988) in order to obtain growth projections at the beginning of the feeding period. This explains why the genotype-to-phenotype mapping approach proposed herein using certain parameters is appealing to support the discovery of important genomic regions associated with higher level phenotypes. In addition, the use of genomic information as a predictor is ideal for application in feedlots, particularly to organize individuals into more homogenous pens, which typically leads to increased profitability.

The k_g parameter is also responsible for some of the variation in DMI and ADG, which may be due to the moderate correlation between AFSBW and k_g . However, a change in k_g also implies differences in body composition (i.e., final EBF), although this is rather small (only 7% by the sensitivity analysis). In fact, higher levels of fat deposition are energetically more efficient than protein deposition, which means that animals that were fatter are expected to have a greater k_g as imposed by the CVDS structure. Animals that are able to convert ME from the diet to NE_g are expected to be more energetically efficient. This is because for a given fed energy concentration, a higher conversion would be expressed as greater ADG. It is important to consider the composition of gain because greater ADG does not necessarily equate to greater protein deposition, but fat deposition. Therefore, the fact that k_g only explains a fraction of the EBF indicates that k_g relates to energy efficiency, but very few with fat thickness. Therefore, k_g may be an interesting phenotype to further explore as a selection criterion for feed efficiency. The variability of EBF reported by the CVDS model does not fully related to AFSBW and k_g , which means that other biological processes still not modeled, contribute to EBF.

We expected a higher influence of AFSBW on body composition due to its potential to influence growth trajectory. Animals with higher AFSBW should have a longer fattening period to obtain a targeted body composition when compared to individuals with lower AFSBW values. However, this develops with the same initial body weight and days of feeding, and the variation encountered in the dataset might have contributed to the decreased association between AFSBW and simulated final EBF.

Genome-wide association

We report the 10 most significant SNPs found to be associated with AFSBW and k_g (Tables 3 and 4) as determined from a GWAS performed using 893 animals. Many of these SNPs are intron variants of specific genes, and the regions surrounding them harbor

QTLs for several production traits. As expected, SNPs significant for AFSBW are found in regions previously associated with BW at different growth stages (birth, weaning, yearling, slaughter, and mature). Residual feed intake and post-weaning ADG, QTLs that are related to feed efficiency and growth trajectory, also confirm the biological interpretation expected for AFSBW. This is also the case for k_g , where QTLs for BW and feed conversion ratio were found. Milk production traits such as milk, protein, and fat yield found for AFSBW and k_g , may have some association with energy efficiency, which is particularly expected for k_g . It is interesting that reproductive traits were also significant QTLs found for both parameters. This may reveal that there is a relationship between reproduction and the animal's growth potential.

Table 3. Estimated statistics and parameters for the 10 most significant single-nucleotide polymorphisms (SNPs) for adjusted final shrunk body weight (AFSBW).

| SNP ID | Chr | Position (bp) | β^a | SE | P value | Gene ^b |
|-------------|-----|---------------|-----------|-------|-----------------------|-------------------|
| rs207966751 | 4 | 35,125,282 | 11.433 | 2.272 | 4.85×10^{-7} | - |
| rs135020999 | 4 | 35,086,804 | 14.2 | 2.938 | 1.34×10^{-6} | - |
| rs521816871 | 17 | 4,982,198 | 14.677 | 3.18 | 3.92×10^{-6} | <i>TMEM154</i> |
| rs109782726 | 16 | 34,462,941 | -8.126 | 1.765 | 4.13×10^{-6} | <i>SDCCAG8</i> |
| rs42735715 | 16 | 34,460,742 | 7.737 | 1.691 | 4.76×10^{-6} | <i>SDCCAG8</i> |
| rs136058533 | 24 | 1,124,251 | -14.404 | 3.155 | 4.97×10^{-6} | <i>ATP9B</i> |
| rs132812243 | 18 | 18,597,882 | 9.05 | 1.987 | 5.23×10^{-6} | <i>HEATR3</i> |
| rs137449097 | 24 | 1,147,611 | -13.866 | 3.061 | 5.92×10^{-6} | <i>ATP9B</i> |
| rs137299855 | 21 | 70,307,103 | -7.282 | 1.618 | 6.75×10^{-6} | - |
| rs517269657 | 4 | 77,569,249 | 9.018 | 2.014 | 7.57×10^{-6} | - |

^a β , allele substitution effect estimated with GWAS. ^bGene where SNP is located. *TMEM154*, *Bos taurus* transmembrane protein 154; *SDCCAG8*, serologically-defined colon cancer antigen 8; *ATP9B*, ATPase, class II, type 9B; *HEATR3*, HEAT repeat containing 3.

Table 4. Estimated statistics and parameters for the 10 most significant single-nucleotide polymorphisms (SNPs) for k_g (efficiency of dietary metabolizable energy to net energy for gain).

| SNP ID | Chr | Position (bp) | β^a | SE | P value | Gene ^b |
|-------------|-----|---------------|-----------|-------|-----------------------|-------------------|
| rs132652367 | 10 | 56,197,171 | 0.011 | 0.002 | 1.32×10^{-7} | <i>UNC13C</i> |
| rs719659528 | 7 | 59,706,678 | -0.005 | 0.001 | 1.06×10^{-6} | - |
| rs134411628 | 24 | 32,067,866 | -0.004 | 0.001 | 4.23×10^{-6} | - |
| rs109334860 | 4 | 8,389,419 | 0.003 | 0.001 | 5.63×10^{-6} | - |
| rs109346688 | 25 | 34,770,404 | 0.004 | 0.001 | 8.24×10^{-6} | <i>MDH2</i> |
| rs137152396 | 4 | 82,460,519 | 0.003 | 0.001 | 8.64×10^{-6} | <i>POU6F2</i> |
| rs133459104 | 17 | 29,212,172 | 0.009 | 0.002 | 8.93×10^{-6} | <i>SCLT1</i> |
| rs137006070 | 10 | 34,301,459 | 0.003 | 0.001 | 1.04×10^{-5} | - |
| rs132862266 | 13 | 59,129,338 | 0.005 | 0.001 | 1.35×10^{-5} | - |
| rs43634421 | 10 | 55,970,723 | -0.004 | 0.001 | 1.38×10^{-5} | <i>UNC13C</i> |

^a β , allele substitution effect estimated with GWAS. ^bGene where SNP is located. *UNC13C*, unc-13 homolog; *MDH2*, malate dehydrogenase 2, NAD (mitochondrial); *POU6F2*, POU class 6 homeobox 2; *SCLT1*, sodium channel and clathrin linker 1.

The biological interpretation of AFSBW and k_g is also supported by the annotated genes. For instance, for k_g , the *UNC13C* gene is related to the synaptic vesicle cycle (GO:0007268). Although an association between energy balance and the nervous system is not straightforward, a feedback loop between plasticity of adipose tissue, which in some cases modulate feed intake, and the nervous system has previously been reported (Sutton et al., 2014). The CVDS model seems to be consistent with this metabolic pathway as k_g influences both EBF and DMI (Table 2).

MDH2 encodes a malate dehydrogenase localized to the mitochondria, which is an enzyme acting within the tricarboxylic acid cycle, an aerobic pathway that yields energy (GO:0030060). Kim et al. (2013) showed that ectopic expression of *MHD2* induced the acceleration of adipogenic differentiation, indicating that the acetylation of *MDH2* is very important for adipogenesis. Zhou et al. (2012) reported that leaner swine breeds exhibit lower expression of *MDH2*, and that, regardless of the breed, females express the gene at higher levels than males. This suggests that fat metabolism may underlie the conversion efficiency of dietary energy to energy available for BW gain, as is expected for k_g . In addition, the pattern of *MDH2* expression differs depending on the tissue (Zhou et al., 2012). *MDH2* was expressed at higher levels in subcutaneous adipose tissue than in visceral adipose tissue. In fact, the calculation of k_g by the CVDS model considers not only the energy retained by the animal, but mainly, the composition of gain. Higher levels of fat deposition mean higher k_g values, which matches the association between k_g and *MDH2*. This suggested association between k_g and fat metabolism, as shown by the genomic analysis, is interesting because its concept is adopted by the standard energy system of the NRC (2000), which drives the development of many model and experiments in ruminant nutrition.

Information available for the *POU6F2* and *SCLT1* genes, which were also significant for k_g , still does not allow any kind of association to be established with cattle growth traits. *POU6F2* and *SCLT1* remain uncharacterized for *Bos taurus*; however, in *Equus caballus* and *Ovis aries*, *PU6F2* has DNA-binding functions that regulate transcription (GO:0003700).

The *TMEM154* gene encodes a transmembrane protein that was found to be associated with AFSBW. The function of this protein has not yet been determined in any livestock species; however, in humans, variations in *TMEM154* have been found to have a negative effect on beta cell functioning, and thus, on the level of insulin resistance and glucose. As reported by Harder et al. (2015), high *TMEM154* mRNA expression is found in B lymphocytes, and immunohistochemical analysis of the human gastrointestinal tract revealed high expression of this protein in glandular cells of the digestive tract. As AFSBW is calculated from outcomes of animal performance, such as body composition at slaughter, it can be suggested that *TMEM154* may have some relationship to nutrient partitioning, which in turn, would result in AFSBW variation. Following the pathways of the antagonizing effect of the growth hormone on insulin action known for cattle (Lucy, 2008), depression of beta cells, as depicted by *TMEM154*, would benefit the development of lean tissue.

The *SDCCAG8* gene encodes a centrosome-associated protein, which is involved in organizing the centrosome during interphase and mitosis (Reactome:5834576). Pathways within the mitotic cell cycle are expected to be associated with AFSBW. In humans, *SDCCAG8* seems to be relevant for BW regulation with high transcript abundance observed in the hypothalamus, pituitary, and adrenals (Scherag et al., 2012), a hormonal axis that also plays an important role in cattle growth traits (Perkins et al., 2014). An intronic variant of *SDCCAG8* has been associated with lower levels of weight loss in overweight children and adolescents (Scherag et al., 2012). If this also occurred in beef cattle, AFSBW could be related to the animal's ability for BW regulation, which is consistent with its interpretation as stated by the CVDS model.

The *ATP9B* gene encodes a P4-ATPase, which is important for translocating phospholipids from the exoplasmic to the cytoplasmic leaflet of lipid bilayers (Takatsu et al., 2011). Information on *ATP9B* and *HEATR3*, which were also associated with AFSBW, do not allow for an interpretation of their effects on cattle growth traits.

Model performance with genetically predicted phenotypes

To draw the genotype-to-parameter map, and develop equations to predict AFSBW and k_g from genotypic data, we used the 10 most significant SNPs; however, SNPs close to each other (within the 1-Mb range) were treated as one to account for the possibility of linkage disequilibrium. Prediction of AFSBW and k_g were of good precision. The joint model was able to explain 76% of the variation in AFSBW and 73% of that in k_g (Table 5). This result suggests that it may be possible to derive values of mechanistic model parameters using genomic information. Interestingly, the 30% lack of ability of the predictive model to compute AFSBW did not spread over the simulations. CVDS runs with AFSBW predicted by genotypes yield good predictions of the outcome variables DMI, ADG, and final EBF (Table 5).

Table 5. Statistics of the coefficient of determination (R^2) from the genotype-to-parameter simulations.

| Phenotypes ^a | Mean ^b | Quartiles ^b | |
|-------------------------|-------------------|------------------------|-------|
| | | 2.5% | 97.5% |
| AFSBW | 0.757 | 0.727 | 0.775 |
| k_g | 0.734 | 0.709 | 0.760 |
| DMI | 0.973 | 0.962 | 0.980 |
| ADG | 0.850 | 0.783 | 0.887 |
| EBF | 0.984 | 0.972 | 0.989 |

^aAFSBW, adjusted final shrunk body weight; k_g , efficiency of conversion of dietary metabolizable energy to net energy for growth; DMI, dry matter intake; ADG, average daily gain; and EBF, empty body fat. ^bMean and quartiles for Monte Carlo runs.

Although high-level phenotypes when compared to parameters usually adopted in cGP models, AFSBW and k_g can be understood as component traits of more complex phenotypes. The biological interpretation of AFSBW and k_g , as imposed by the CVDS model, may be confirmed genetically, as shown by the results of the GWAS study and by the surrounding QTLs (Tables 6 and 7). SNPs found to be significant for the parameters revealed many associated QTLs and annotated genes that could be further explored to better describe biological pathways responsible for cattle growth and energetic efficiency. Specifically, our analyses have highlighted that k_g may be related to fat metabolism. To this end, we also suggest that CVDS could be improved by more detailed modeling of the intersection among fat pools, energy balance, and growth dynamics. As expected, AFSBW, a parameter close to the concept of BW at maturity, seems to be associated with mechanisms of cell cycle and mitosis.

Table 6. Quantitative trait loci (QTLs) associated with production traits within the 1-Mb region surrounding the most significant single nucleotide polymorphisms (SNPs) for AFSBW (adjusted final shrunk body weight).

| SNP ID | QTLs ^a |
|-------------|--|
| rs207966751 | BW slaughter (15723), BW weaning (10708), milk fat % (5055), milk protein yield (10277), calving ease (4655) |
| rs521816871 | R Rump angle (3448), BW birth (11038), BW yearling (11039), RFI (4448), post-weaning ADG (4484), milk yield (4743) |
| rs109782726 | BW birth (11025), ADG (7101), BW weaning (4482), pre-weaning ADG (4486), BW weaning (11026), calving ease (1700) |
| rs136058533 | RFI (5307), BW yearling (11184), calving ease (11181) |
| rs132812243 | RFI (5293, 5294), BW mature (11061) |

^aBW, body weight; RFI, residual feed intake; ADG, average daily gain; number in parenthesis refers to the QTL ID at the Cattle QTL Database (Hu et al., 2013).

Table 7. Quantitative trait loci (QTLs) associated with production traits within the 1-Mb region surrounding the most significant single nucleotide polymorphisms (SNPs) for k_g (efficiency of dietary metabolizable energy to net energy for gain).

| SNP ID | QTLs ^a |
|-------------|---|
| rs132652367 | BW birth (10878) |
| rs719659528 | BW birth (10810), BW weaning (10808), BW yearling (10807), milk yield (10291/2448), milk protein % (2532), milk protein yield (2449), calving ease (10811) |
| rs134411628 | Feed conversion ratio (5308), BW weaning (11198), BW yearling (11197), calving ease (11196) |
| rs109346688 | BW weaning (11213), height mature (11216), rum angle (1715), weaning weight-maternal milk (11215), milk protein yield (2610/10351), milk yield (2591/1538/10345), milk fat percentage (10342), calving ease (15224) |
| rs133459104 | BW slaughter (11708), fatty acid |
| rs137006070 | BW mature (10873), ADG (22836) |

^aBW, body weight; ADG, average daily gain; number in parenthesis refers to the QTL ID at the Cattle QTL Database (Hu et al., 2013).

In conclusion, the results of our study show that the integration of genetics and nutritional models of cattle growth have potential application for phenotype predictability. It is possible to draw a genotype-to-parameter map, and then simulate the phenotype by using genomic information to compute model core parameters. The integration of these two disciplines, genomics and system dynamics, provides opportunities to advance towards the discovery of candidate genes for important phenotypes in the animal sciences, and to develop a theory of predictive modeling for decision support tools.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

The contributions of Núcleo de Criadores de Nelore do Norte do Paraná, Luciano Borges (Rancho da Matinha), and Eduardo Penteado Cardoso (Faz. Mundo Novo) are gratefully acknowledged. We would like to thank Dr. Zhong Wang for help with the *gwas.lasso* package. Research supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, #2012/02039-9, #2013/26902-0, #2013/20571-2, and #2014/07566-2) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, #473249/2013-8 and #442345/2014-3).

REFERENCES

- Alexandre PA, Kogelman LJA, Santana MHA, Passarelli D, et al. (2015). Liver transcriptomic networks reveal main biological processes associated with feed efficiency in beef cattle. *BMC Genomics* 16: 1073. <http://dx.doi.org/10.1186/s12864-015-2292-8>
- Arnold RN and Bennetti GL (1991). Evaluation of four simulation models of cattle growth and body composition: Part I Comparison and characterization of the models. *Agric. Syst.* 35: 401-432. [http://dx.doi.org/10.1016/0308-521X\(91\)90117-S](http://dx.doi.org/10.1016/0308-521X(91)90117-S)
- Aulchenko YS, Ripke S, Isaacs A and van Duijn CM (2007). GenABEL: an R library for genome-wide association analysis. *Bioinformatics* 23: 1294-1296. <http://dx.doi.org/10.1093/bioinformatics/btm108>
- Aulchenko YS, Struchalin MV, Belonogova NM, Axenovich TI, et al. (2009). Predicting human height by Victorian and genomic methods. *Eur. J. Hum. Genet.* 17: 1070-1075. <http://dx.doi.org/10.1038/ejhg.2009.5>
- Baker MJ, Tedeschi LO, Fox DG, Henning WR, et al. (2006). Using ultrasound measurements to predict body composition of yearling bulls. *J. Anim. Sci.* 84: 2666-2672. <http://dx.doi.org/10.2527/jas.2006-006>

- Bogardus C (2009). Missing heritability and GWAS utility. *Obesity* 17: 209-210. <http://dx.doi.org/10.1038/oby.2008.613>
- Chizzotti ML, Tedeschi LO and Valadares Filho SC (2008). A meta-analysis of energy and protein requirements for maintenance and growth of Nellore cattle. *J. Anim. Sci.* 86: 1588-1597. <http://dx.doi.org/10.2527/jas.2007-0309>
- Fox DG, Sniffen CJ and Connor JDO (1988). Adjusting nutrient requirements of beef cattle for animal and environmental variations. *J. Anim. Sci.* 66: 1475-1495. <http://dx.doi.org/10.2527/jas1988.6661475x>
- Francesconi M and Lehner B (2014). The effects of genetic variation on gene expression dynamics during development. *Nature* 505: 208-211. <http://dx.doi.org/10.1038/nature12772>
- Gao X, Starmer J and Martin ER (2008). A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genet. Epidemiol.* 32: 361-369. <http://dx.doi.org/10.1002/gepi.20310>
- Gomes RC, Silva SL, Carvalho ME, Rezende FM, et al. (2013). Protein synthesis and degradation gene SNPs related to feed intake, feed efficiency, growth, and ultrasound carcass traits in Nellore cattle. *Genet. Mol. Res.* 12: 2923-2936. <http://dx.doi.org/10.4238/2013.August.12.8>
- Guiroy PJ, Fox DG, Tedeschi LO, Baker MJ, et al. (2001). Predicting individual feed requirements of cattle fed in groups. *J. Anim. Sci.* 79: 1983-1995. <http://dx.doi.org/10.2527/2001.7981983x>
- Harder MN, Appel EVR, Grarup N, Gjesing AP, et al. (2015). The type 2 diabetes risk allele of TMEM154-rs6813195 associates with decreased beta cell function in a study of 6,486 Danes. *PLoS One* 10: e0120890. <http://dx.doi.org/10.1371/journal.pone.0120890>
- Hu ZL, Park CA, Wu XL and Reecy JM (2013). Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era. *Nucleic Acids Res.* 41: D871-D879. <http://dx.doi.org/10.1093/nar/gks1150>
- Kim EY, Han BS, Kim WK, Lee SC, et al. (2013). Acceleration of adipogenic differentiation via acetylation of malate dehydrogenase 2. *Biochem. Biophys. Res. Commun.* 441: 77-82. <http://dx.doi.org/10.1016/j.bbrc.2013.10.016>
- Lucy MC (2008). Functional differences in the growth hormone and insulin-like growth factor axis in cattle and pigs: implications for post-partum nutrition and reproduction. *Reprod. Domest. Anim.* 43 (Suppl 2): 31-39. <http://dx.doi.org/10.1111/j.1439-0531.2008.01140.x>
- Marcondes M, Chizzotti M, Valadares Filho S, Gionbelli M, et al. (2010). Energy requirements of Zebu beef cattle. In: Nutrient Requirements of Zebu Beef Cattle BR-CORTE (Valadares Filho S, Marcondes M, Chizzotti M and Paulino P, eds.). Universidade Federal de Viçosa, Viçosa.
- NRC (2000). Nutrient Requirements of Beef Cattle, 7th ed. National Academy Press, Washington, D.C.
- Peng H, Ma J, Bai Y, Lu J, et al. (2015). MeDiA: Mean distance association and its applications in nonlinear gene set analysis. *PLoS One* 10: e0124620. <http://dx.doi.org/10.1371/journal.pone.0124620>
- Perkins SD, Key CN, Garrett CF, Foradori CD, et al. (2014). Residual feed intake studies in Angus-sired cattle reveal a potential role for hypothalamic gene expression in regulating feed efficiency. *J. Anim. Sci.* 92: 549-560. <http://dx.doi.org/10.2527/jas.2013-7019>
- R Core Team (2015). R: A language and environment for statistical computing.
- Santana MHA, Rossi P, Almeida R and Cucco DC (2012). Feed efficiency and its correlations with carcass traits measured by ultrasound in Nellore bulls. *Livest. Sci.* 145: 252-257. <http://dx.doi.org/10.1016/j.livsci.2012.02.012>
- Santana MHA, Utsunomiya YT, Neves HHR, Gomes RC, et al. (2014). Genome-wide association study for feedlot average daily gain in Nellore cattle (*Bos indicus*). *J. Anim. Breed. Genet.* 131: 210-216. <http://dx.doi.org/10.1111/jbg.12084>
- Sargolzaei M, Chesnais JP and Schenkel FS (2014). A new approach for efficient genotype imputation using information from relatives. *BMC Genomics* 15: 478. <http://dx.doi.org/10.1186/1471-2164-15-478>
- Scherag A, Kleber M, Boes T, Kolbe AL, et al.; NUGENOB Consortium (2012). SDCCAG8 obesity alleles and reduced weight loss after a lifestyle intervention in overweight children and adolescents. *Obesity* 20: 466-470. <http://dx.doi.org/10.1038/oby.2011.339>
- Sutton AK, Pei H, Burnett KH, Myers Jr MG, et al. (2014). Control of food intake and energy expenditure by Nos1 neurons of the paraventricular hypothalamus. *J. Neurosci.* 34: 15306-15318. <http://dx.doi.org/10.1523/JNEUROSCI.0226-14.2014>
- Svishcheva GR, Axenovich TI, Belonogova NM, van Duijn CM, et al. (2012). Rapid variance components-based method for whole-genome association analysis. *Nat. Genet.* 44: 1166-1170. <http://dx.doi.org/10.1038/ng.2410>
- Takatsu H, Baba K, Shima T, Umino H, et al. (2011). ATP9B, a P4-ATPase (a putative aminophospholipid translocase), localizes to the trans-Golgi network in a CDC50 protein-independent manner. *J. Biol. Chem.* 286: 38159-38167. <http://dx.doi.org/10.1074/jbc.M111.281006>
- Tedeschi LO, Boin C, Fox DG, Leme PR, et al. (2002). Energy requirement for maintenance and growth of Nellore bulls and steers fed high-forage diets. *J. Anim. Sci.* 80: 1671-1682. <http://dx.doi.org/10.2527/2002.8061671x>
- Tedeschi LO, Fox DG and Guiroy PJ (2004). A decision support system to improve individual cattle management. 1. A mechanistic, dynamic model for animal growth. *Agric. Syst.* 79: 171-204. [http://dx.doi.org/10.1016/S0308-521X\(03\)00070-2](http://dx.doi.org/10.1016/S0308-521X(03)00070-2)
- Wang Y, Gjuvsland AB, Vik JO, Smith NP, et al. (2012). Parameters in dynamic models of complex traits are containers of missing heritability. *PLoS Comput. Biol.* 8: e1002459. <http://dx.doi.org/10.1371/journal.pcbi.1002459>
- Zhou SL, Li MZ, Li QH, Guan JQ, et al. (2012). Differential expression analysis of porcine MDH1, MDH2 and ME1 genes in adipose tissues. *Genet. Mol. Res.* 11: 1254-1259. <http://dx.doi.org/10.4238/2012.May.9.4>