



**HAL**  
open science

## Unraveling genetic sensitivity of beef cattle to environmental variation under tropical conditions

Roberto Carneiro, Roy Costilla, Haroldo H. R. Neves, Lucia G. Albuquerque, Stephen Moore, Ben J. Hayes

► **To cite this version:**

Roberto Carneiro, Roy Costilla, Haroldo H. R. Neves, Lucia G. Albuquerque, Stephen Moore, et al.. Unraveling genetic sensitivity of beef cattle to environmental variation under tropical conditions. *Genetics Selection Evolution*, 2019, 51 (1), pp.29. 10.1186/s12711-019-0470-x . hal-02163077

**HAL Id: hal-02163077**

**<https://hal.science/hal-02163077>**

Submitted on 24 Jun 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.


L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

RESEARCH ARTICLE

Open Access



# Unraveling genetic sensitivity of beef cattle to environmental variation under tropical conditions

Roberto Carvalho<sup>1,2\*</sup> , Roy Costilla<sup>3,4</sup>, Haroldo H. R. Neves<sup>5</sup>, Lucia G. Albuquerque<sup>1,2</sup>, Stephen Moore<sup>4</sup> and Ben J. Hayes<sup>4</sup>

## Abstract

**Background:** Selection of cattle that are less sensitive to environmental variation in unfavorable environments and more adapted to harsh conditions is of primary importance for tropical beef cattle production systems. Understanding the genetic background of sensitivity to environmental variation is necessary for developing strategies and tools to increase efficiency and sustainability of beef production. We evaluated the degree of sensitivity of beef cattle performance to environmental variation, at the animal and molecular marker levels (412 K single nucleotide polymorphisms), by fitting and comparing the results of different reaction norm models (RNM), using a comprehensive dataset of Nelore cattle raised under diverse environmental conditions.

**Results:** Heteroscedastic RNM (with different residual variances for environmental level) provided better fit than homoscedastic RNM. In addition, spline and quadratic RNM outperformed linear RNM, which suggests the existence of a nonlinear genetic component affecting the performance of Nelore cattle. This nonlinearity indicates that within-animal sensitivity depends on the environmental gradient (EG) level and that animals may present different patterns of sensitivity according to the range of environmental variations. The spline RNM showed that sensitivity to environmental variation from harsh to average EG is lowly correlated with sensitivity from average to good EG, at both the animal and molecular marker levels. Although the genomic regions that affect sensitivity in harsher environments were not the same as those associated with less challenging environments, the candidate genes within those regions participate in common biological processes such as those related to inflammatory and immune response. Some plausible candidate genes were identified.

**Conclusions:** Sensitivity of tropical beef cattle to environmental variation is not continuous along the environmental gradient, which implies that animals that are less sensitive to harsher conditions are not necessarily less responsive to variations in better environmental conditions, and vice versa. The same pattern was observed at the molecular marker level, i.e. genomic regions and, consequently, candidate genes associated with sensitivity to harsh conditions were not the same as those associated with sensitivity to less challenging conditions.

## Background

Beef cattle production plays an important role in food security by converting forages and agricultural by-products into high-quality protein, and by optimizing land use through the occupation of grasslands that are unsuitable

for agriculture [1, 2]. A large proportion of beef cattle are raised on tropical grasslands [3]. Under this production system, animals are usually subject to periods of food scarcity due to seasonal patterns of rainfall and, consequently, seasonal variation of pasture quality and availability. Depending on the severity of the dry season, the animals can be subject to seasonal weight loss if no supplementation is provided, this being a major constraint to beef cattle production in the tropics [4, 5]. However, supplementation is costly, and a more effective strategy

\*Correspondence: roberto.carvalho@unesp.br

<sup>1</sup> School of Agricultural and Veterinarian Sciences, Sao Paulo State University (UNESP), Jaboticabal, SP 14884-900, Brazil

Full list of author information is available at the end of the article



would be to select cattle that are less sensitive to environmental variation in unfavorable environments and more adapted to harsh conditions.

Selection for reduced sensitivity to environmental variation can be addressed in different ways in animal breeding programs [6]. One alternative is to use reaction norm models (RNM), in which the response of each individual to environmental variation is modeled through a unique random regression curve that has its trajectory determined by a continuous environmental descriptor [7, 8]. For instance, RNM has been used to study heat tolerance in beef [9] and dairy cattle [10, 11] by modeling animals' performance and the underlying breeding value as a function of a temperature-humidity index. In this application, animals with a smaller decline in production during high heat load were considered heat tolerant.

The choice of the type of curve to be used in RNM is of primary importance to properly assess the sensitivity of each animal to environmental variation. Generally, a linear regression is fitted since higher order polynomial coefficients may be difficult to estimate and interpret. However, if the change in phenotypic expression as a function of environmental variation is not constant along the environmental gradient, the random slope of a linear regression may not be a good indicator of the animal's sensitivity. Therefore, testing higher order RNM is recommended [12].

Genotypic information has allowed extending RNM from the individual to the single nucleotide polymorphism (SNP) level [13–15]. Given a large dataset and a sufficient number of genotypes across the range of environments, the reaction norm of each SNP can reveal the size and direction of its effect on the trait of interest along the environmental gradient. An interesting feature of this approach would be to consider nonlinear reaction norms and investigate the possible existence of SNPs having environment-dependent sensitivity, i.e. SNPs that express more response to improvements in the environment and less response to its deterioration (and vice versa). This is particularly appealing in tropical beef production given the importance of genotype-by-environment interaction under these conditions [16].

Understanding the genetic background of sensitivity to environmental variation is critical in developing strategies to ameliorate current constraints on efficiency of beef cattle production in the tropics. This knowledge could be used, for example, in statistical models, which incorporate biological information to attain more accurate genomic predictions [17, 18]. The objective of our study was to unravel genetic sensitivity of beef cattle to environmental variation under tropical conditions. More specifically, we tested if sensitivity from harsh to average

environment is similar to sensitivity from average to good environment, by comparing the fit of linear and non-linear RNM for post-weaning weight gain (PWG) of Nellore cattle. We tested this at the individual (animal) and molecular marker levels.

## Methods

### Phenotypes

Phenotypic and pedigree data were obtained from the Alliance Nellore dataset ([www.gensys.com.br](http://www.gensys.com.br)), which has information for over one million Nellore calves born between 1984 and 2016, from approximately 500 different commercial herds widely distributed in Brazil, Bolivia and Paraguay, which are highly related through the intensive adoption of artificial insemination (AI) (>50% calves born from AI). This dataset is suitable to assess the sensitivity of beef cattle performance to environmental variation not only because of its size but also because of the diversity of management and environmental conditions in which the animals are raised. For instance, the average annual precipitation can vary from ~700 to ~3000 mm depending on the farm's location, and the dry season in some regions may last up to 7 months. As mentioned previously, this variation of precipitation affects drastically the quality of the forage and, consequently, the animal's performance since they are raised predominantly on pasture. Some farmers provide protein and mineral supplementation to the animals, especially during the dry season, but other farmers provide only urea to complement pasture.

We focused on the study of the post-weaning weight gain (PWG) trait because, after weaning, the animals are exposed to a wide range of environmental conditions, which makes PWG a suitable trait to assess sensitivity to environmental variation. For each animal, PWG was calculated as the difference between long-yearling and weaning weights, which were obtained at about 510 (310–730) and 210 (90–300) days of age, respectively. For the analysis, we considered only animals with known sires and dams, from contemporary groups (CG) with a minimum of 20 animals and PWG (adjusted for 300 days) between 30 and 250 kg. After filtering, the phenotypic dataset had individual PWG records for 421,585 animals from 9934 CG. Contemporary groups were formed by animals born in the same herd, year and season, from the same sex and raised in the same management group. The phenotyped animals were progeny of 6637 sires and 271,408 dams. The overall mean (sd) of PWG was 97.37 (33.13) kg. The distribution and other descriptive statistics of PWG are in Table S1 (see Additional file 1: Table S1) and in Figures S1 and S2 (see Additional file 2: Figures S1 and Additional file 3: Figure S2).

## Genotypes

Genotyping information for 13,806 Nellore cattle (*Bos indicus*) were used in genome-wide association (GWA) analyses. This information came from 1993 bulls, 4571 cows and 7242 progeny, which were genotyped with the Illumina BovineHD (HD) chip (~778 K SNPs; 4073 samples) or with a lower density (LD) chip (from ~28 to ~74 K SNPs; 9733 samples). A detailed list of the number of samples per chip is provided in Table S2 (see Additional file 1: Table S2). Lower density genotypes were imputed to HD using the software FImpute v.2.2 [19], with an expected imputation accuracy higher than 0.97 [20]. Only autosomal SNPs with a GenCall score, call rate and minor allele frequency higher than 0.15, 0.98 and 0.02, respectively, were used in the GWA analyses. We also discarded SNPs with a  $p$  value for Hardy–Weinberg equilibrium test lower than  $10^{-5}$ , and showing linkage disequilibrium ( $r^2$ ) greater than 0.998 with syntenic SNPs located within a window of 50 neighboring markers. After filtering, the total number of SNPs was equal to 412,456. All samples had a call rate higher than 0.90 (before imputation) and were kept for the GWA analyses.

## Environmental descriptor

Genetic sensitivity to environmental variation was assessed by reaction norm models (RNM) [7]. In this framework, the response of each animal to environmental variation is modeled as a unique curve that has its trajectory determined by a continuous environmental descriptor. In the absence of environmental descriptor information (e.g. temperature-humidity index or quality of pasture), descriptors derived from phenotypic data are commonly used (e.g. [21]). We opted for the use of the best linear unbiased estimates (BLUE) of CG effects as our environmental descriptor, since it encompasses the different management and environmental conditions to which the animals were exposed. In particular, we obtained BLUE from a regular mixed animal model [22] using PWG as the response variable, post-weaning age (linear and quadratic effects) and CG as fixed effects, and animal and residual as random effects. Estimated CG effects ranged from 31.38 to 225.16 kg, highlighting the diversity of conditions in which the animals were raised (~sevenfold difference between extremes). The mean (sd) was 93.71 (28.16) kg. More information about CG solutions is in Figures S3 and S4 (see Additional file 4: Figures S3 and Additional file 5: Figure S4).

The environmental descriptor used on RNM, hereafter referred to as environmental gradient (EG), was the BLUE of CG effects that were standardized to have a zero mean and standard deviation (sd) equal to 1. Few CG ( $n=72$ ) had EG greater than 3 sd. Post-weaning gain records from those CG ( $n=2728$ ) were discarded for the subsequent analyses to avoid prejudicing the reaction norm estimates, since

outliers can dramatically change the magnitude of regression coefficients [23]. After filtering, the average, minimum and maximum values of EG were equal to  $-2.21$ ,  $0.0$  and  $3.0$  sd, corresponding to CG effects on PWG equal to 31.5, 93.7 and 178.2 kg or, equivalently, to average daily gains (ADG) during the post-weaning period (300 days) equal to 105, 312 and 594 g/day, respectively. To attain this ADG over the full period, the animals most probably lost weight during the dry season (~½ post-weaning period) on lower EG, maintained their weight during the dry season on average EG, and gained weight even in the dry season on higher EG conditions. Thus, we will refer to lower, average and higher EG as harsh, average and good environments, respectively. However, it is important to keep in mind that the average EG consists of a challenging condition during the dry season. Progeny of genotyped sires comprised 78% of phenotypic records and, in general, they had records well distributed (within sire) across the range of EG (see Additional file 6: Figure S5). This is essential to obtain good reaction norm estimates at the marker level [24].

## Reaction norm models (RNM)

Five RNM were tested. The first applied RNM (RNM\_homo) assumed that the residuals for the EG were homogeneous and can be described by the equation:

$$y_{ij} = \mathbf{x}'_j \boldsymbol{\beta} + \vartheta_1 \hat{w}_i + b_{0j} + b_{1j} \hat{w}_i + e_{ij},$$

where  $y_{ij}$  is the phenotypic record (PWG) of animal  $j$  recorded in the level  $i$  of EG;  $\boldsymbol{\beta}$  is the vector of fixed effects (post-weaning period and CG), and  $\mathbf{x}'_j$  is its corresponding row covariable/incidence vector;  $\vartheta_1$  is the overall linear fixed regression coefficient of  $y_{ij}$  on  $\hat{w}_i$ ;  $\hat{w}_i$  is the covariate associated to the  $i$ -th EG (estimated in the previous step);  $b_{0j}$  is the random overall additive genetic effect or the intercept of animal  $j$  for an average EG ( $\hat{w} = 0$ );  $b_{1j}$  is the random additive genetic effect of the reaction norm slope of animal  $j$  on  $\hat{w}_i$  expressed as a deviation from  $\vartheta_1$ ; and  $e_{ij}$  is a random residual. RNM\_homo was fitted under the assumptions:

$$\begin{aligned} \{b_{0j}, b_{1j}\} &\sim N(0, \mathbf{A} \otimes \begin{bmatrix} \sigma_{b_0}^2 & \sigma_{b_0, b_1} \\ \sigma_{b_0, b_1} & \sigma_{b_1}^2 \end{bmatrix}), \\ \text{and } \{e_{ij}\} &\sim N(0, \mathbf{I} \sigma_e^2), \end{aligned}$$

where  $\mathbf{A}$  is the relationship matrix based on pedigree information ( $\otimes$  is the Kronecker product);  $\sigma_{b_0}^2$ ,  $\sigma_{b_1}^2$  and  $\sigma_{b_0, b_1}$  are the variances of the intercept, the slope and their covariance, respectively;  $\mathbf{I}$  is an identity matrix and  $\sigma_e^2$  is the residual variance.

The second RNM (RNM\_hete) was similar to RNM\_homo except that the residuals were modeled assuming

a specific residual variance for each EG, using a linear regression on  $\hat{w}_i$ . The heterogeneous residual variance coefficients for the intercept and slope were modelled using a log-residual function [25]. The third RNM (RNM\_quad) was similar to RNM\_hete except that it considered a polynomial quadratic regression to model the fixed curve and the reaction norm of the random effects (additive genetic and residual) instead of a linear regression. RNM\_quad can be described by the equation:

$$y_{ij} = \mathbf{x}'_j \boldsymbol{\beta} + \theta_1 \hat{w}_i + \theta_2 \hat{w}_i^2 + b_{0j} + b_{1j} \hat{w}_i + b_{2j} \hat{w}_i^2 + e_{ij},$$

where,  $\theta_2$  is the overall quadratic fixed regression coefficient of  $y_{ij}$  on  $\hat{w}_i$ ;  $\hat{w}_i^2$  is the squared value of  $\hat{w}_i$ ;  $b_{2j}$  is the quadratic effect of the random additive genetic effect of the reaction norm of animal  $j$  on  $\hat{w}_i$  expressed as a deviation from  $\theta_2$ ; and all other terms as previously specified.

The other two tested RNM also fitted heterogeneous residual variances and used a linear-linear (RNM\_l-l) and a quadratic-quadratic (RNM\_q-q) spline function to model the fixed curve and the reaction norm of the additive genetic and residual random effects. For both spline functions we used only one knot ( $k$ ) placed at the average EG ( $k=0$ ). RNM\_l-l can be described by the equation:

$$y_{ij} = \mathbf{x}'_j \boldsymbol{\beta} + \theta_1 \hat{w}_i + \theta_1^* \hat{w}_i^* + b_{0j} + b_{1j} \hat{w}_i + b_{1j}^* \hat{w}_i^* + e_{ij},$$

where  $\theta_1^*$  is the overall difference between the linear fixed regression coefficients of the first and second segments of the linear-linear spline function of  $y_{ij}$  on  $\hat{w}_i^*$  ( $\hat{w}_i^* = 0$  if  $\hat{w}_i < k$ ;  $\hat{w}_i^* = \hat{w}_i$  if  $\hat{w}_i \geq k$ );  $b_{1j}^*$  is the difference between the reaction norm slopes of the first and second segments of the linear-linear spline function of the random additive genetic effect of animal  $j$  on  $\hat{w}_i^*$  expressed as a deviation from  $\theta_1^*$ ; and all other terms as previously specified. With this model parameterization, the estimates of the slope of the second segment ( $b_{1seg2j}$ ) were obtained as:  $b_{1seg2j} = b_{1j} + b_{1j}^*$ . RNM\_q-q can be described by the equation:

$$y_{ij} = \mathbf{x}'_j \boldsymbol{\beta} + \theta_1 \hat{w}_i + \theta_2 \hat{w}_i^2 + \theta_1^* \hat{w}_i^* + \theta_2^* \hat{w}_i^{2*} + b_{0j} + b_{1j} \hat{w}_i + b_{2j} \hat{w}_i^2 + b_{1j}^* \hat{w}_i^* + b_{2j}^* \hat{w}_i^{2*} + e_{ij},$$

where  $\theta_2^*$  is the overall difference between the quadratic fixed regression coefficients of the first and second segments of the quadratic-quadratic spline function of  $y_{ij}$  on  $\hat{w}_i^{2*}$  ( $\hat{w}_i^{2*} = 0$  if  $\hat{w}_i < k$ ;  $\hat{w}_i^{2*} = \hat{w}_i^2$  if  $\hat{w}_i \geq k$ );  $b_{2j}^*$  is the difference between the reaction norm quadratic effects of the first and second segments of the quadratic-quadratic spline function of the random additive genetic effect of animal  $j$  on  $\hat{w}_i^*$  expressed as a deviation from  $\theta_2^*$ ; and all other terms as previously specified. With this model

parameterization, the quadratic effect estimates of the second segment ( $b_{2seg2j}$ ) were obtained as:  $b_{2seg2j} = b_{2j} + b_{2j}^*$ .

Quadratic RNM (RNM\_quad and RNM\_q-q) were used to evaluate any advantage compared with linear reaction norms. Spline RNM (RNM\_l-l and RNM\_q-q) were used to evaluate if sensitivity in the first segment (harsh to average EG) is correlated with sensitivity in the second segment (average to good EG). Higher order polynomials were tested but convergence was not achieved (results not shown).

Estimates of (co)variance components of all RNM were obtained by restricted maximum likelihood using the AIREMLF90 software [26]. The different RNM were compared based on Akaike (AIC) [27] and Bayesian (BIC) [28] information criteria.

### Genome-wide association (GWA) analyses

GWA analyses were performed using the weighted single-step genomic BLUP method [29], which first predicts genomic estimated breeding values (GEBV) and then back-solve GEBV to SNP effects using equivalent models [30]. Two analyses were run using models similar to RNM\_hete and RNM\_l-l, but this time assuming that (co)variance components are known by using the corresponding estimates obtained in the previous step, and also replacing  $\mathbf{A}$  by  $\mathbf{H}$ , a matrix which combines pedigree ( $\mathbf{A}$ ) and genomic ( $\mathbf{G}$ ) relationships [31]. Matrix  $\mathbf{G}$  was computed as the first method proposed by [32]. Association analyses were not run for RNM\_quad and RNM\_q-q because SNP effects of the quadratic coefficient are difficult to interpret and also because sensitivity assessed as the slopes of RNM\_l-l was highly correlated (0.97-0.99) with sensitivity measured as the first derivatives of RNM\_quad and RNM\_q-q. Association analysis was not run also for RNM\_homo because it was the worst fitting model (as will be shown later).

For each GWA model (RNM\_hete and RNM\_l-l), SNP effects were iteratively recomputed weighting them proportionally to the genetic variance they explained in the previous iteration, resulting in an increased shrinkage of the SNPs explaining a lower variance and more pronounced effects of the SNPs explaining a higher proportion of genetic variance, compared to the previous iteration. Following [29], the same weight (equal to 1) was given to all SNPs in the first iteration and two additional iterations were performed. In addition to SNP effect estimates, the percentage of genetic variance explained by the segment of five adjacent SNPs was also computed for each SNP. GWA analyses were performed using BLUPF90 family programs [26].



### Functional enrichment analyses

Enrichment analyses were performed for the slopes of models RNM\_hete (b1) and RNM\_l-l (b1.seg1 and b1.seg2). For each slope, SNPs that explained at least 0.5% of the genetic variance were identified and genes that were located within 200 kb were annotated using Ensembl BioMart (Ensembl 93) [33] and UMD3.1 bovine reference genome [34]. Since a high proportion of genes within significant regions had human orthologs, functional enrichment analyses were performed with the GENE2FUNC process of the integrative web-based platform FUMA v1.3.3 [35], using all the human genes with bovine orthologs as background, totaling 18,425 genes with unique Entrez ID retrieved from the Ensembl genes 93 database. FUMA overrepresentation analyses were performed using hypergeometric tests and Benjamini–Hochberg multiple testing correction. Biological functions of genes with an adjusted enrichment  $p$  value  $\leq 0.05$  were reported. GENE2FUNC also provided enrichment of tissue-specific gene expression based on GTEx v6 RNA-seq data [36] using Bonferroni multiple test correction.

### Results

#### Reaction norm models

Additive genetic variability was observed for the slope of RNM, which indicated the presence of genotype-by-environment interaction ( $G \times E$ ) on PWG (Table 1). Evidence of  $G \times E$  was also observed in a preliminary multi-trait analyses, assuming PWG from three categories of EG ( $EG < -1$ ;  $-1 \leq EG \leq 1$ ;  $EG > 1$ ) as different traits. For instance, the genetic correlation between PWG from harsh ( $EG < -1$ ) and good ( $EG > 1$ ) environments was equal to 0.79, in the multi-trait analyses (results not shown).

Based on AIC and BIC criteria, all heteroscedastic RNM outperformed the homoscedastic model (Table 1). Among the heteroscedastic models, the quadratic (RNM\_quad) and the spline models (RNM\_l-l and RNM\_q-q) outperformed the linear model (RNM\_hete), which suggests the existence of a nonlinear component in the reaction norm of PWG in Nellore cattle. Model RNM\_q-q had the best fit among the tested models according to both criteria, AIC and BIC.

Models RNM\_homo and RNM\_hete showed a high correlation (0.96 and 0.86, respectively) between intercept and slope (Table 1), which indicates that animals

**Table 1** Estimates of variance (block-diagonals), covariance (upper triangular blocks; in italic) and correlation (lower triangular blocks; in bold) among coefficients of reaction norm models (RNM) for the additive genetic effect of post-weaning weight gain (kg) in Nellore cattle, along with residual variance estimates and Akaike (AIC) and Bayesian (BIC) information criteria

Model <sup>a</sup>	Coefficient <sup>b</sup>	b0	b1	b2	b3	b4	e <sup>c</sup>	np <sup>d</sup>	AIC <sup>e</sup>	BIC <sup>5</sup>
RNM_homo	b0 (int)	129.73	<i>48.64</i>				223.18	4	3,560,790	3,560,838
	b1 (slp)	<b>0.96</b>	19.79							
RNM_hete	b0 (int)	89.80	<i>16.33</i>				5.55	5	3,559,272 (–1518)	3,559,332 (–1506)
	b1 (slp)	<b>0.86</b>	3.99				0.23			
RNM_quad	b0 (int)	95.90	<i>11.53</i>	<i>–6.75</i>			5.58	9	3,558,581 (–2209)	3,558,689 (–2149)
	b1 (slp)	<b>0.40</b>	8.72	<i>2.13</i>			0.24			
	b2 (qdr)	<b>–0.50</b>	<b>0.52</b>	1.92			–0.06			
RNM_l-l	b0 (int)	102.86	<i>24.97</i>	<i>–1.00</i>			5.60	9	3,558,765 (–2025)	3,558,873 (–1965)
	b1 (slp1)	<b>0.89</b>	7.69	<i>0.71</i>			0.32			
	b2 (slp2)	<b>–0.02</b>	<b>0.06</b>	16.71			0.16			
RNM_q-q	b0 (int)	98.03	<i>24.51</i>	<i>–3.27</i>	<i>–1.55</i>	<i>0.82</i>	5.55	20	3,558,372 (–2418)	3,558,611 (–2227)
	b1 (slp1)	<b>0.31</b>	64.15	<i>25.80</i>	<i>–33.82</i>	<i>17.93</i>	0.06			
	b2 (qdr1)	<b>–0.09</b>	<b>0.91</b>	12.50	<i>–14.09</i>	<i>7.59</i>	–0.18			
	b3 (slp2)	<b>–0.02</b>	<b>–0.48</b>	<b>–0.45</b>	78.93	<i>–28.23</i>	0.20			
	b4 (qdr2)	<b>0.02</b>	<b>0.66</b>	<b>0.63</b>	<b>–0.93</b>	11.63	–0.01			

<sup>a</sup> RNM\_homo: linear homoscedastic; RNM\_hete: linear heteroscedastic; RNM\_quad: quadratic heteroscedastic; RNM\_l-l: spline linear–linear heteroscedastic; RNM\_q-q: spline quadratic–quadratic heteroscedastic

<sup>b</sup> b0–b4 coefficients of the RNM for the additive genetic random effect [int: intercept; slp: slope; qdr: quadratic; slp1(2): slope segment 1(2); qdr1(2): quadratic segment 1(2)]

<sup>c</sup> Residual variance (RNM\_homo) or residual coefficients associated with parameters of heteroscedastic RNM that were modeled using a log-residual function [25]

<sup>d</sup> Number of estimated parameters

<sup>e</sup> Numbers in parenthesis refer to difference in comparison with RNM\_homo

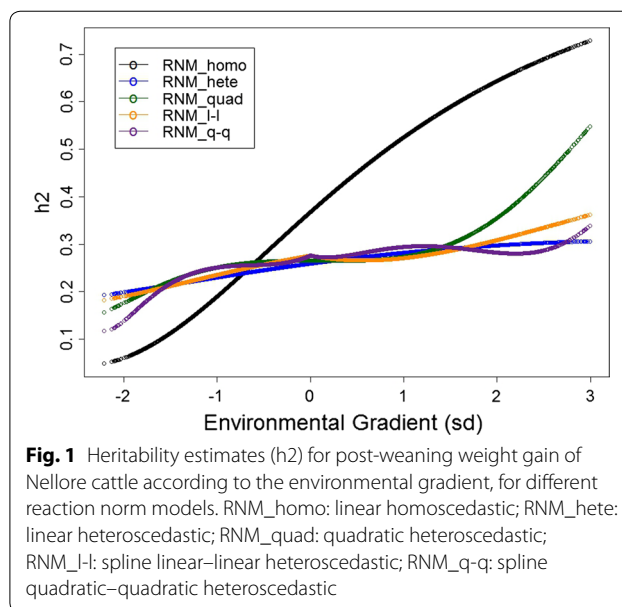
with better EBV for the average environment tended also to present higher sensitivity to environmental variation. However, correlation estimates showed by RNM\_1-1 suggest that this association is strong only for the first segment (harsh to average EG). In fact, the correlation was close to 0 for RNM\_1-1 between intercept and slope of the second segment ( $-0.02$ ) and between slopes of the first and second segments ( $0.06$ ), whereas the correlation between intercept and slope of the first segment was  $0.89$  (Table 1).

Increasing additive genetic variance (see Additional file 7: Figure S6) and heritability (Fig. 1) estimates from lower to higher EG were obtained for all models. Because homoscedastic residuals were assumed, RNM\_homo overestimated heritability in good environments and underestimated it in harsh environments, compared to heteroscedastic models. All heteroscedastic models showed similar heritability estimates for intermediate EG ( $-1.5 < EG < 1.5$ ). Model RNM\_quad presented higher heritability estimates for  $EG \geq 2$  and RNM\_q-q showed lower heritability estimates for  $EG \leq -2$ , in comparison with the other heteroscedastic models (Fig. 1).

Heritability estimates of RNM\_homo resulted in over-shrinkage of EBV in lower EG, the opposite occurring with EBV in higher EG, compared to EBV of heteroscedastic models (Fig. 2). In accordance with (co)variance component estimates and model comparison results, there was further evidence from the comparison of EBV from the RNM\_quad, RNM\_1-1 and RNM\_q-q models that the nonlinear component in the reaction norm of PWG in Nellore cattle was important (Fig. 2c–e). The EBV from RNM\_1-1 illustrate that sensitivity to environmental variation from harsh to average EG tends to not be genetically correlated with sensitivity from average to good EG, in agreement with the genetic correlation estimates. Some animals were robust (i.e. had a flatter slope) for  $EG < 0$  but were not robust for  $EG > 0$ , and vice versa (Fig. 2d). The same pattern was observed when PWG data from the two EG classes ( $EG < 0$  and  $EG > 0$ ) were analyzed independently, using two separate heteroscedastic linear regression RNM (see Additional file 8: Figure S7), which reinforces the plausibility of the RNM\_1-1 results. Figure 2f illustrates that the different models can provide quite different EBV rankings between animals, especially at higher EG levels.

#### Genome-wide association and functional enrichment

Estimates of SNP effects on the intercept and slope of RNM\_hete were highly correlated ( $0.95$ ), suggesting that average performance and sensitivity of PWG would have a similar genetic background. However, estimates of SNP effects obtained with RNM\_1-1 indicate that this

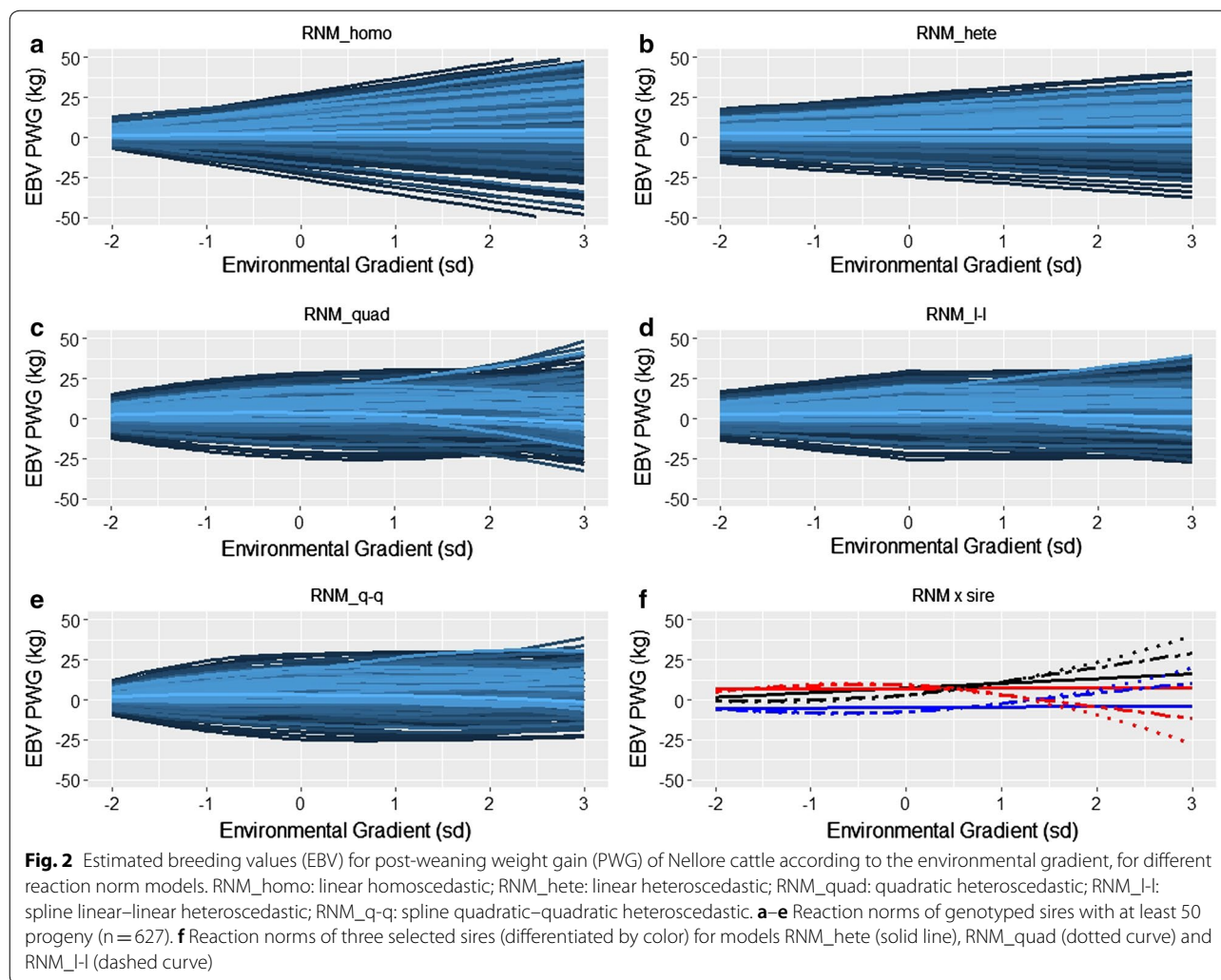


**Fig. 1** Heritability estimates ( $h^2$ ) for post-weaning weight gain of Nellore cattle according to the environmental gradient, for different reaction norm models. RNM\_homo: linear homoscedastic; RNM\_hete: linear heteroscedastic; RNM\_quad: quadratic heteroscedastic; RNM\_1-1: spline linear-linear heteroscedastic; RNM\_q-q: spline quadratic-quadratic heteroscedastic

similarity holds only for sensitivity from harsh to average environmental variation ( $r=0.97$ ), whereas a low correlation ( $0.14$ ) was observed between estimates of SNP effects of the intercept and slope of the second segment (Fig. 3).

Twenty-three genomic regions, from 16 chromosomes, were identified containing segment(s) of five adjacent SNPs that explained 0.5% or more of the genetic variance for at least one of the coefficients of RNM\_hete and RNM\_1-1 (Table 2) (see Additional file 9: Figure S8). These regions, with a size ranging from 400.0 to 436.6 kb (200 kb upstream and downstream the significant SNP(s)), include 54 annotated genes in total. The numbers of regions (genes) associated with the intercept and slope of RNM\_hete were equal to 7 (19) and 6 (13), respectively, and the numbers of regions (genes) associated with the intercept and the slopes of the first and second segments of RNM\_1-1 were equal to 7 (12), 7 (10) and 11 (30), respectively. In accordance with the results of correlation between effects of SNPs, there were overlapping candidate regions for the intercept and slope of RNM\_hete (5) and for the intercept and slope of the first segment of RNM\_1-1 (6). No overlapping candidate regions were observed between slopes of RNM\_1-1 or between its intercept and slope of the second segment (Table 2).

From the 49 genes associated with the slopes of models RNM\_hete or RNM\_1-1, 37 (75.5%) have human orthologs (Table 2) and were used for functional enrichment analyses. Seven from the eight genes with human orthologs associated with the slope of RNM\_hete (b1) had 48 significantly enriched biological processes (BP) (see Additional file 10: Figure S9). The BP associated with this set of genes (*ADCY2*, *DCLK2*,



*HAS2*, *MTRR*, *REG3A*, *REG3G* and *SLC8A1*) were predominantly related to epithelial, inflammatory, homeostatic, and cell proliferation and differentiation processes. For the slope of the first segment of RNM\_l-l (b1seg1), five genes (*HAS2*, *REG3A*, *REG3G*, *FZD1* and *SH3BP4*), out of six with human orthologs, showed 35 significantly enriched BP, predominantly related to epithelial, inflammatory, nitrogen compound, and cell proliferation and differentiation processes (see Additional file 11: Figure S10). For the slope of the second segment of RNM\_l-l (b1seg2), 16 genes (*AP3D1*, *ASTN2*, *DMRTA2*, *FAF1*, *GLUL*, *IL4*, *IL13*, *IQGAP2*, *KIF3A*, *MKNK2*, *PAPPA*, *PRCP*, *RAD50*, *SV2C*, *TRIM32* and *ZNF703*), out of 26 with human orthologs, showed 104 significantly enriched BP (the majority associated with *IL4* and *IL13*), predominantly related to immune response, metabolic, and cell death, proliferation and differentiation processes (see Additional file 12: Figure S11).

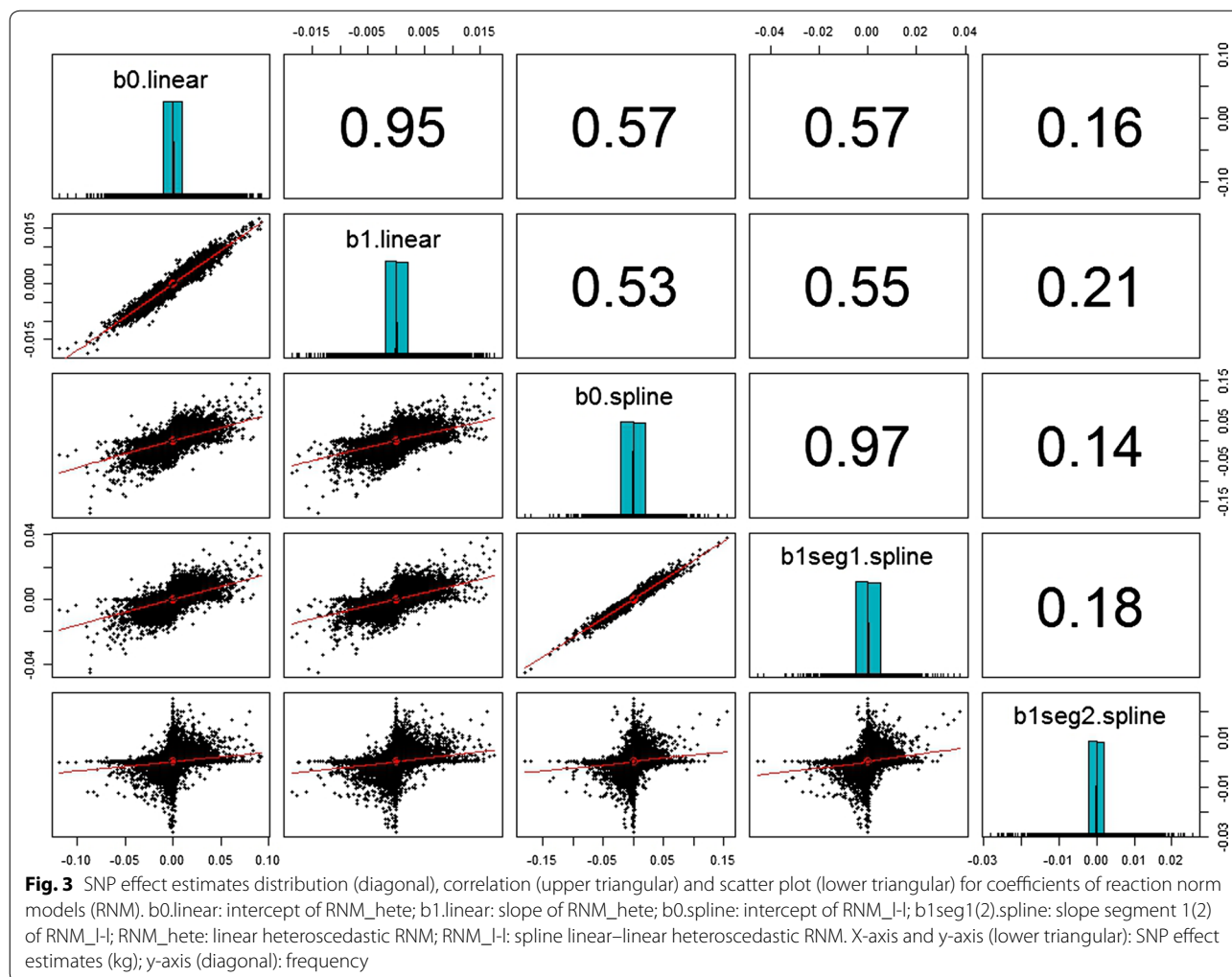
GENE2FUNC gene expression results evidenced that the human orthologs of the bovine genes associated with sensitivity assessed by the slope of RNM\_hete (b1) tend to be more differently expressed in humans in the small intestine and adrenal gland (see Additional file 13: Figure S12). For the genes associated with the slope of the first segment of RNM\_l-l, GENE2FUNC identified significant enrichment of up-regulated genes in the human pancreas (see Additional file 14: Figure S13). For sensitivity assessed by the slope of the second segment of RNM\_l-l, the set of candidate genes tended to be more differently expressed in humans in the brain (see Additional file 15: Figure S14).

## Discussion

### Reaction norm models

When running genetic evaluations, genetic sensitivity to environmental variation can be addressed in different ways in an animal breeding context. In this study, we used





RNM to model the response to environmental variation of the trait PWG in Nellore cattle raised under tropical pasture conditions. The existence of additive genetic variability for the slope of RNM indicated the presence of  $G \times E$  on PWG. It offers the possibility of using the individual random slopes as a selection criterion for robustness, since animals with a flatter slope are less sensitive to environmental variation. Other studies have also identified, through RNM, the importance of  $G \times E$  for different traits in Nellore cattle [37–39].

An interesting modelling alternative to our environmental descriptor is a Bayesian hierarchical RNM (BHRNM). This model has been recommended when the environmental descriptor needs to be derived from phenotypic data [40, 41]. We chose not to use BHRNM because it would have drastically increased the computational time required for the analyses. Moreover, there is also evidence in the literature that “two-step” RNM

(using BLUE previously estimated as environmental covariates) achieves a similar fit to BHRNM [42].

As in [37], we also observed that heteroscedastic RNM provided a better fit than homoscedastic RNM. In addition, we found that spline and quadratic RNM outperformed linear RNM, which suggests the existence of a nonlinear component in the reaction norm of PWG in Nellore cattle. This nonlinearity indicates that within-animal sensitivity depends on the EG level and that animals may present different patterns of sensitivity depending on the degree of EG. For instance, some animals may be less sensitive to variation in harsher environmental gradients and more sensitive to variation in not so challenging environmental gradients, and vice versa. This nonlinearity in the reaction norm seems to be less significant in cattle that are raised in more intensive systems, at least in dairy cattle [43], possibly because environmental factors that influence intensive systems tend to be more

**Table 2 Genomic regions associated with coefficients of reaction norm models for post-weaning weight gain in Nellore cattle, and annotated genes within those regions**

Region	%Var <sup>a</sup>					Chr <sup>b</sup>	Window start <sup>b</sup> (bp)	Window end <sup>b</sup> (bp)	Gene stable ID <sup>c</sup>	Gene symbol <sup>c</sup>
	b0	b1	b0s	b1s1	b1s2					
1			0.55	0.55		1	84,867,382	85,267,382	–	–
2					0.58	3	35,940,081	36,344,427	ENSBTAG00000014530	NTNG1
3					0.78	3	96,120,576	96,520,576	ENSBTAG00000014482	FAF1
									ENSBTAG00000045402	RF00322
									ENSBTAG00000010414	DMRTA2
4				0.55		3	114,849,177	115,249,177	ENSBTAG00000013869	SH3BP4
5	0.56		1.52	1.61		4	8,329,351	8,741,255	ENSBTAG00000002107	FZD1
									ENSBTAG000000027319	
6	1.00	0.88	1.25	1.14		5	84,442,697	84,878,158	ENSBTAG00000039906	
									ENSBTAG00000046108	
									ENSBTAG00000000749	LMNTD1
7					0.74	7	22,728,136	23,164,748	ENSBTAG00000009034	AP3D1
									ENSBTAG00000025448	IZUMO4
									ENSBTAG00000014526	MOB3A
									ENSBTAG00000018049	MKNK2
									ENSBTAG00000025445	
									ENSBTAG00000008609	SEPT8
									ENSBTAG00000031246	CCNI2
									ENSBTAG00000025477	KIF3A
									ENSBTAG00000015957	IL4
									ENSBTAG00000015953	IL13
									ENSBTAG00000011252	RAD50
8			0.76	0.75		7	95,572,442	95,972,442	–	–
9					0.82	8	107,325,318	107,747,332	ENSBTAG00000004010	PAPPA
									ENSBTAG00000047511	
									ENSBTAG00000025667	
									ENSBTAG00000017155	TRIM32
10					0.53	9	58,226,166	58,643,182	–	–
11					0.50	10	7,290,264	7,690,264	ENSBTAG00000017507	SV2C
									ENSBTAG00000000897	IQGAP2
12	0.79	0.63				11	22,788,685	23,203,999	ENSBTAG00000013861	SLC8A1
									ENSBTAG00000037153	RF00001
13	0.60	0.63	0.68	0.55		11	56,219,017	56,652,372	ENSBTAG00000007135	REG3G
									ENSBTAG00000011314	REG3A
14	0.56	0.64	0.77	0.74		14	19,482,365	19,901,806	ENSBTAG00000044730	RF00026
									ENSBTAG00000019892	HAS2
15					0.50	14	72,964,805	73,364,805	–	–
16					0.83	16	64,552,728	64,958,425	ENSBTAG00000046243	ZNF648
									ENSBTAG00000027426	
									ENSBTAG00000013631	GLUL
17		0.50				17	7,729,955	8,129,955	ENSBTAG00000008098	DCLK2
18			0.51			17	11,993,902	12,393,902	ENSBTAG00000034522	SLC10A7
									ENSBTAG00000048020	
									ENSBTAG00000014060	LSM6

**Table 2 (continued)**

Region	%Var <sup>a</sup>					Chr <sup>b</sup>	Window start <sup>b</sup> (bp)	Window end <sup>b</sup> (bp)	Gene stable ID <sup>c</sup>	Gene symbol <sup>c</sup>
	b0	b1	b0s	b1s1	b1s2					
19	0.54					18	59,112,039	59,512,039	ENSBTAG00000045571 ENSBTAG00000037440 ENSBTAG00000047712 ENSBTAG00000030454 ENSBTAG00000011052	
20	0.84	0.71				20	65,401,804	65,809,687	<i>ENSBTAG00000009401</i> <i>ENSBTAG00000009400</i> <i>ENSBTAG00000019210</i>	<i>MTRR</i> <i>FASTKD3</i> <i>ADCY2</i>
21				0.63		27	11,730,389	12,130,389	<i>ENSBTAG00000047427</i>	<i>RF00100</i>
22				0.52		27	32,310,351	32,710,351	<i>ENSBTAG00000033381</i> <i>ENSBTAG00000013943</i>	<i>ZNF703</i>
23				0.62		29	12,876,785	13,278,418	<i>ENSBTAG00000033731</i> <i>ENSBTAG00000006812</i> <i>ENSBTAG00000044832</i>	<i>PRCP</i> <i>FAM181B</i>

<sup>a</sup> Percentage of genetic variance (%Var) explained by the leading segment (within region) of five adjacent SNPs, for each coefficient. b0: intercept of a heteroscedastic linear reaction norm model (RNM\_hete); b1: slope of RNM\_hete; b0s: intercept of a heteroscedastic spline linear-linear reaction norm model (RNM\_l-l); b1s1: slope of segment 1 of RNM\_l-l; b1s2: slope of segment 2 of RNM\_l-l. Only %Var  $\geq$  0.5% are presented

<sup>b</sup> Chromosome (Chr) and bp position (bpp) according to the UMD3.1 assembly

<sup>c</sup> Retrieved from the BioMart Ensembl genes 93 database; genes with human orthologs, used for enrichment analyses, are presented in italics

controlled, compared to pasture-based systems, resulting in a smaller environmental variation.

The high genetic correlation estimate between the intercept and slope of RNM\_hete suggested little opportunity of joint selection for increased performance and reduced sensitivity. A similar result was also observed by [41], in a study on PWG of Angus cattle raised in Brazil. However, the results of RNM\_l-l revealed that this high correlation holds only for harsher environments and that there are good opportunities to select for increased PWG and reduced sensitivity in not so challenging environmental conditions.

An alternative to select for increased PWG and reduced sensitivity jointly is to identify and select the best genotype for specific environmental conditions. However, this strategy may provide suboptimal genetic gain especially for tropical pasture systems where environmental conditions can vary substantially among years, even within farms. This reinforces the relevance of using sensitivity as a selection criterion.

#### Genome-wide association and functional enrichment

There is strong evidence that the mechanisms associated with sensitivity depend on the degree of EG. GWAS results of RNM\_l-l revealed that important genomic regions associated with sensitivity from harsh to average environmental conditions differed from those associated with sensitivity from average to good environmental conditions. This result is not surprising given the diversity

of environmental conditions in which the animals were raised and the complexity of livestock response to stressful conditions [44].

In harsher environments, genes associated with acute inflammatory response and keratinocyte proliferation and differentiation processes seem to play an important role in sensitivity of beef cattle that is assessed as the slope of the first segment of RNM\_l-l (b1seg1). For instance, the *regenerating islet-derived 3 alpha (REG3A)* gene is associated (in humans and mice) with wound repair after skin injury and with homeostasis of the skin, thus contributing to immune defense [45]. Another candidate gene of the REG3 family (*REG3G*), also associated with sensitivity to harsher environments (b1seg1), was found to be related with antimicrobial defense of the mammalian intestine [46] and with intestinal strategies for maintaining symbiotic host-microbial relationships [47]. The barrier function of gastrointestinal epithelium in host protection is well documented in the literature [48, 49] and severe feed restriction has been shown to compromise the total tract barrier function in beef cattle [50]. These results reinforce the plausibility of *REG3A* and *REG3G* being associated with sensitivity of beef cattle to harsh environments. Another candidate gene potentially associated with sensitivity to harsh environmental conditions is the *SH3 domain binding protein 4 (SH3BP4)*. This gene was identified in a region with copy number variation in South African Nguni cattle, which are recognized for their ability to sustain harsh

environmental conditions and resistance to parasites and disease [51].

In not so harsh environments, candidate genes associated with inflammatory and immune response seem also to play an important role in sensitivity of beef cattle that is assessed as the slope of the second segment of RNM\_1-1 (b1seg2). Among them, the *interleukin 4* (*IL4*) and *interleukin 13* (*IL13*) are perhaps the most plausible candidates. Both genes are known to share a wide range of activities on monocytes, epithelial cells and B cells, and thus play an important role in host defense [52–55]. Although interleukins and other cytokines are mainly reported to be involved in the regulation of immune response, there is also evidence about their role in regulating protein metabolism and muscle function [56], among other metabolic functions as, for example, energy homeostasis [57, 58].

Members of the heat shock protein (HSP) family have been considered as candidate genes for sensitivity or stress response in different studies [59–62]. No HSP genes were observed in our candidate regions. However, for sensitivity in not so harsh environments (b1seg2), the *glutamate-ammonia ligase* (*GLUL*) was identified as a candidate gene. *GLUL* is involved in glutamine synthesis, which, in turn, has been shown to be a powerful effector of HSP expression [63]. *GLUL* has also been reported to be associated with cellular response to starvation [64].

Other important candidate genes associated with b1seg2 are the: *adaptor related protein complex 3 subunit delta 1* (*AP3D1*), *kinesin family member 3A* (*KIF3A*) and *RAD50 double strand break repair protein* (*RAD50*). *AP3D1* is involved in the regulation of the sequestering of zinc ion, and zinc has been shown to have multiple impacts on the immune function [65]. *KIF3A* is upregulated in heat-shocked Holstein oocytes [66] and differently expressed in a bovine tick-resistance study [67]. *RAD50* is a key gene for the immune system [68] and was identified as candidate gene for harsh environmental adaptation in cattle [69].

Eleven out of the 30 genes associated with sensitivity to not so harsh environments (b1seg2) were from the same region between 22.7 and 23.2 Mb on chromosome 17. The genes within this region are highly conserved in different species and multiple organisms have orthologs with those genes (NCBI databases; [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)). In humans, they are located in two regions, according to the reference assembly GRCh38.p12, one on chromosome 5 between 132.6 and 132.8 Mb (*SEPT8*, *CCNI2*, *KIF3A*, *IL4*, *IL13*, *RAD50*) and another on chromosome 19 between 2.0 and 2.2 Mb (*AP3D1*, *IZUMO4*, *MOB3A*, *MKNK2*). In Nellore cattle, the markers within this region are in high linkage disequilibrium (see Additional file 16: Figure S15). It is important to note that this

candidate region would not have been identified, if we had performed GWAS using only RNM\_hete.

Enrichment results from FUMA showed that besides the differences in candidate regions associated with b1seg1 and b1seg2, the genes in those regions also differ in their gene expression profile. This result cannot be extrapolated to cattle since it is based on human data, but can serve as a guide for future gene expression studies in cattle that aim at identifying differentially expressed genes associated with adaptation or sensitivity. Based on the results of the FUMA analysis, in harsher environments, sampling of gastrointestinal and pancreas tissues is recommended. In not so harsh but still challenging environments, sampling of brain and skin tissues is also recommended.

## Conclusions

Our results reveal important  $G \times E$  interactions that affect the performance of beef cattle raised under tropical conditions or, more specifically, that  $G \times E$  interactions affect PWG in Nellore cattle. Sensitivity to environmental variation at the individual level is not linear along the environmental gradient, since animals that are less sensitive to environmental changes under harsher conditions are not necessarily less responsive to not so challenging conditions, and vice versa. The same pattern was observed at the molecular marker level, i.e. genomic regions associated with sensitivity to harsh conditions are not the same as those associated with sensitivity to not so harsh (but still challenging) conditions, although the genes within those regions share common biological processes, such as those related with inflammatory and immune response. These results highlight the diversity and complexity of mechanisms that are involved in cattle response to stressful/unfavorable conditions, and should serve as a basis for future studies aiming at unraveling the genetic control of sensitivity.

## Additional files

**Additional file 1: Table S1** Descriptive statistics for post-weaning weight gain (kg), adjusted for 300 days, of Nellore cattle. **Table S2** Number of genotyped samples per SNP chip and animal category.

**Additional file 2: Figure S1.** Number of animals (counts) with own records for post-weaning weight gain, by year of birth and sex.

**Additional file 3: Figure S2.** Smoothed density plots of post-weaning weight gain (PWG), adjusted for a period of 300 days, of Nellore cattle.

**Additional file 4: Figure S3.** Smoothed density plots of contemporary group (CG) solutions (best linear unbiased estimates) for post-weaning weight gain (kg) of Nellore cattle.

**Additional file 5: Figure S4.** Boxplot of contemporary group (CG) solutions (best linear unbiased estimates) for post-weaning weight gain (kg) of Nellore cattle, by year of birth and sex.



**Additional file 6: Figure S5.** Scatterplot and histograms of progeny's environmental gradient (EG) standard deviation (sd) and range (max(EG)-min(EG)), by genotyped sire with at least 5 progeny (n = 1384).

**Additional file 7: Figure S6.** Additive (left) and residual (right) variance component estimates for post-weaning weight gain of Nelore cattle according to the environmental gradient, for different reaction norm models. RNM\_homo: linear homoscedastic; RNM\_hete: linear heteroscedastic; RNM\_quad: quadratic heteroscedastic; RNM\_l-l: spline linear-linear heteroscedastic; RNM\_q-q: spline quadratic-quadratic heteroscedastic.

**Additional file 8: Figure S7.** Estimated breeding values (EBV) for post-weaning weight gain (PWG) of Nelore cattle according to the environmental gradient (EG), obtained with a heteroscedastic linear reaction norm model (RNM\_hete). To generate the results of this plot, data of the two segments (EG < 0 and EG > 0) were analyzed separately, running two independent RNM\_hete analyses. The correlation estimates between the intercept (b0) and slope of the two segments (b1\_seg1 and b1\_seg2) are provided. Reaction norms of genotyped sires with at least 50 progeny and in common among the two EG segments (n = 621) are presented.

**Additional file 9: Figure S8.** Manhattan plots of percentage of genetic variance explained by the segment of 5 adjacent SNPs, for each parameter of the reaction norm models RNM\_hete and RNM\_l-l (RNM\_hete: heteroscedastic linear; RNM\_l-l: heteroscedastic spline linear-linear). The horizontal dashed line at 0.5 represents the empirical threshold used to identify candidate genomic regions ( $\geq 0.5$ ).

**Additional file 10: Figure S9.** Significantly enriched gene ontology biological processes, identified by GENE2FUNC process of FUMA, associated with candidate genes for sensitivity to environmental variation of post-weaning weight gain in Nelore cattle, assessed by a random slope (b1) of a heteroscedastic linear random regression model (RNM\_hete).

**Additional file 11: Figure S10.** Significantly enriched gene ontology biological processes, identified by GENE2FUNC process of FUMA, associated with candidate genes for sensitivity to environmental variation of post-weaning weight gain in Nelore cattle, assessed by a random slope of the first segment (b1\_seg1) of a heteroscedastic spline linear-linear random regression model (RNM\_l-l).

**Additional file 12: Figure S11.** Significantly enriched gene ontology biological processes, identified by GENE2FUNC process of FUMA, associated with candidate genes for sensitivity to environmental variation of post-weaning weight gain in Nelore cattle, assessed by a random slope of the second segment (b1\_seg2) of a heteroscedastic spline linear-linear random regression model (RNM\_l-l).

**Additional file 13: Figure S12.** Enrichment of human tissue specific gene expression based on GTEx v6 RNA-seq data [36], obtained with GENE2FUNC process of FUMA, for candidate ortholog bovine genes associated with sensitivity to environmental variation of post-weaning weight gain in Nelore cattle, assessed by a random slope (b1) of a heteroscedastic linear random regression model (RNM\_hete).

**Additional file 14: Figure S13.** Enrichment of human tissue specific gene expression based on GTEx v6 RNA-seq data [36], obtained with GENE2FUNC process of FUMA, for candidate ortholog bovine genes associated with sensitivity to environmental variation of post-weaning weight gain in Nelore cattle, assessed by a random slope of the first segment (b1\_seg1) of a heteroscedastic spline linear-linear random regression model (RNM\_l-l). Significant enrichment at Bonferroni corrected  $P \leq 0.05$  are colored in red.

**Additional file 15: Figure S14.** Enrichment of human tissue specific gene expression based on GTEx v6 RNA-seq data [36], obtained with GENE2FUNC process of FUMA, for candidate ortholog bovine genes associated with sensitivity to environmental variation of post-weaning weight gain in Nelore cattle, assessed by a random slope of the second segment (b1\_seg2) of a heteroscedastic spline linear-linear random regression model (RNM\_l-l).

**Additional file 16: Figure S15.** Linkage disequilibrium ( $r^2$ ) between markers of the 22.7-23.2 Mb region of chromosome 7 (UMD 3.1 assembly), in Nelore cattle.

## Acknowledgements

The authors acknowledge GenSys (Porto Alegre, Brazil) and the breeding programs contributing to the Alliance Nelore dataset ([www.gensys.com.br/sumario/aliancanelore/](http://www.gensys.com.br/sumario/aliancanelore/)) for providing the data.

## Authors' contributions

RCa, HHRN and LGA conceived the study. RCa, SM and BJH designed the analyses and coordinated the study. HHRN imputed the genotypes and prepared the data files. RCa ran the statistical and GWA analyses. RCo ran the functional enrichment analyses. RCa led the manuscript preparation. All authors interpreted the results and contributed to edit the manuscript. All authors read and approved the final manuscript.

## Funding

This study was financially supported by the Sao Paulo Research Foundation (FAPESP) Brazil (grant numbers 17/18415-3 and 17/10630-2) and by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) Brazil (Finance Code 001).

## Availability of data and materials

All the data supporting the results of this study are included in the article and in the Additional files. The phenotypic and genotypic data cannot be shared because they are owned by commercial breeding programs.

## Ethics approval and consent to participate

Animal Care and Use Committee approval was not required for this study because the data were obtained from an existing dataset of commercial Nelore breeding programs.

## Consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

## Author details

<sup>1</sup> School of Agricultural and Veterinarian Sciences, Sao Paulo State University (UNESP), Jaboticabal, SP 14884-900, Brazil. <sup>2</sup> National Council for Scientific and Technological Development (CNPq), Brasília, DF 71605-001, Brazil. <sup>3</sup> Institute for Molecular Bioscience (IMB), University of Queensland, St. Lucia, QLD 4072, Australia. <sup>4</sup> Queensland Alliance for Agriculture and Food Innovation (QAAFI), Centre for Animal Science, University of Queensland, St. Lucia, QLD 4072, Australia. <sup>5</sup> GenSys Associated Consultants, Porto Alegre, RS 90680-000, Brazil.

Received: 19 November 2018 Accepted: 4 June 2019

Published online: 20 June 2019

## References

- Mottet A, De Haan C, Falcucci A, Tempio G, Opio C, Gerber P. Livestock: on our plates or eating at our table? A new analysis of the feed/food debate. *Glob Food Secur.* 2017;14:1–8.
- Layman DK. Assessing the role of cattle in sustainable food systems. *Nutr Today.* 2018;53:160–5.
- Gerber PJ, Mottet A, Opio CI, Falcucci A, Teillard F. Environmental impacts of beef production: review of challenges and perspectives for durability. *Meat Sci.* 2015;109:2–12.
- Lamy E, van Harten S, Sales-Baptista E, Guerra MM, Almeida AM. Factors influencing livestock productivity. In: Sejian V, Naqvi SMK, Ezeji T, Lakritz J, Rattan L, editors. *Environmental stress and amelioration in livestock production.* Berlin: Springer; 2012. p. 19–51.
- Poppi DP, Quigley SP, da Silva TACC, McLennan SR. Challenges of beef cattle production from tropical pastures. *R Bras Zootec.* 2018;47:e20160419.
- Carvalho R, Hayes BJ, Albuquerque LG, Moore S. Genetic sensitivity of beef cattle to environmental variation. In: Hermes S, Dominik S, editors. *Breeding focus 2018—reducing heat stress.* Armidale: Anim Genet Breed Unit; 2018. p. 17–29.

7. Falconer D. Selection in different environments: effects on environmental sensitivity (reaction norm) and on mean performance. *Genet Res.* 1990;56:57–70.
8. Kirkpatrick M, Lofsvold D, Bulmer M. Analysis of the inheritance, selection and evolution of growth trajectories. *Genetics.* 1990;124:979–93.
9. Bradford HL, Fragomeni BO, Bertrand JK, Lourenco DAL, Misztal I. Genetic evaluations for growth heat tolerance in Angus cattle. *J Anim Sci.* 2016;94:4143–50.
10. Ravagnolo O, Misztal I. Genetic component of heat stress in dairy cattle, parameter estimation. *J Dairy Sci.* 2000;83:2126–30.
11. Nguyen TTT, Bowman PJ, Haile-Mariam M, Pryce JE, Hayes BJ. Genomic selection for tolerance to heat stress in Australian dairy cattle. *J Dairy Sci.* 2016;99:2849–62.
12. Schaeffer LR. Application of random regression models in animal breeding. *Livest Prod Sci.* 2004;86:35–45.
13. Lillehammer M, Hayes BJ, Meuwissen TH, Goddard ME. Gene by environment interactions for production traits in Australian dairy cattle. *J Dairy Sci.* 2009;92:4008–17.
14. Hayes BJ, Bowman PJ, Chamberlain AJ, Savin K, Van Tassell CP, Sonstegard TS, et al. A validated genome wide association study to breed cattle adapted to an environment altered by climate change. *PLoS One.* 2009;4:e6676.
15. Silva FF, Mulder HA, Knol EF, Lopes MS, Guimarães SEF, Lopes PS, et al. Sire evaluation for total number born in pigs using a genomic reaction norms approach. *J Anim Sci.* 2014;92:3825–34.
16. Burrow HM. Importance of adaptation and genotype x environment interactions in tropical beef breeding systems. *Animal.* 2012;6:729–40.
17. Pérez-Enciso M, Rincón JC, Legarra A. Sequence vs. chip-assisted genomic selection: accurate biological information is advised. *Genet Sel Evol.* 2015;47:43.
18. MacLeod IM, Bowman PJ, Vander Jagt CJ, Haile-Mariam M, Kemper KE, Chamberlain AJ, et al. Exploiting biological priors and sequence variants enhances QTL discovery and genomic prediction of complex traits. *BMC Genomics.* 2016;17:144.
19. Sargolzaei M, Chesnais JP, Schenkel FS. A new approach for efficient genotype imputation using information from relatives. *BMC Genomics.* 2014;15:478.
20. Carvalho R, Boison SA, Neves HHR, Sargolzaei M, Schenkel FS, Utsunomiya YT, et al. Accuracy of genotype imputation in Nelore cattle. *Genet Sel Evol.* 2014;46:69.
21. Calus MP, Groen AF, de Jong G. Genotype x environment interaction for protein yield in Dutch dairy cattle as quantified by different models. *J Dairy Sci.* 2002;85:3115–23.
22. Henderson CR. Applications of linear models in animal breeding. Guelph: University of Guelph; 1984.
23. Rousseeuw PJ, Leroy AM. Robust regression and outlier detection. Hoboken: Wiley; 2003.
24. Hayes BJ, Daetwyler HD, Goddard ME. Models for genome x environment interaction: examples in livestock. *Crop Sci.* 2016;56:2251–9.
25. Foulley JL, Quaas RL. Heterogeneous variances in Gaussian linear mixed models. *Genet Select Evol.* 1995;27:211–28.
26. Misztal I, Tsuruta S, Lourenco D, Aguilar I, Legarra A, Vitezica Z. Manual for BLUPF90 family of programs. 2016. [http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90\\_all2.pdf](http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all2.pdf). Accessed 19 Oct 2018.
27. Akaike H. A new look at the statistical model identification. *IEEE Trans Automat Control.* 1974;19:716–23.
28. Schwarz G. Estimating the dimension of a model. *Ann Stat.* 1978;6:461–4.
29. Wang H, Misztal I, Aguilar I, Legarra A, Muir WM. Genome-wide association mapping including phenotypes from relatives without genotypes. *Genet Res (Camb).* 2012;94:73–83.
30. Strandén I, Garrick DJ. Technical note: derivation of equivalent computing algorithms for genomic predictions and reliabilities of animal merit. *J Dairy Sci.* 2009;92:2971–5.
31. Aguilar I, Misztal I, Johnson DL, Legarra A, Tsuruta S, Lawlor TJ. Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *J Dairy Sci.* 2010;93:743–52.
32. VanRaden PM. Efficient methods to compute genomic predictions. *J Dairy Sci.* 2008;91:4414–23.
33. Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J, et al. Ensembl 2018. *Nucleic Acids Res.* 2018;46:D754–61.
34. Zimin AV, Delcher AL, Florea L, Kelley DR, Schatz MC, Puiu D, et al. A whole-genome assembly of the domestic cow, *Bos taurus*. *Genome Biol.* 2009;10:R42.
35. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun.* 2017;8:1826.
36. The GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science.* 2015;348:648–60.
37. Chiaia HLJ, de Lemos MVA, Venturini GC, Aboujaoude C, Berton MP, Feitosa FB, et al. Genotype x environment interaction for age at first calving, scrotal circumference, and yearling weight in Nelore cattle using reaction norms in multitrait random regression models. *J Anim Sci.* 2015;93:1503–10.
38. Santana ML Jr, Eler JP, Bignardi AB, Menendez-Buxadera A, Cardoso FF, Ferraz JBS. Multi-trait linear reaction norm model to describe the pattern of phenotypic expression of some economic traits in beef cattle across a range of environments. *J Appl Genet.* 2015;56:219–29.
39. Oliveira DP, Lourenco DAL, Tsuruta S, Misztal I, Santos DJA, de Araújo Neto FR, et al. Reaction norm for yearling weight in beef cattle using single-step genomic evaluation. *J Anim Sci.* 2018;96:27–34.
40. Su G, Madsen P, Lund MS, Sorensen D, Korsgaard IR, Jensen J. Bayesian analysis of the linear reaction norm model with unknown covariates. *J Anim Sci.* 2006;84:1651–7.
41. Cardoso FF, Tempelman RJ. Linear reaction norm models for genetic merit prediction of Angus cattle under genotype by environment interaction. *J Anim Sci.* 2012;90:2130–41.
42. Cardoso LL, Neto JB, Cardoso FF, Cobuci JA, Biassus IO, Barcellos JOJ. Hierarchical Bayesian models for genotype x environment estimates in post-weaning gain of Hereford bovine via reaction norms. *R Bras Zootec.* 2011;40:294–300.
43. Streit M, Reinhardt F, Thaller G, Bennewitz J. Reaction norms and genotype-by-environment interaction in the German Holstein dairy cattle. *J Anim Breed Genet.* 2012;129:380–9.
44. MohanKumar SMJ, Balasubramanian P, Dharmaraj M, MohanKumar PS. Neuroendocrine regulation of adaptive mechanisms in livestock. In: Sejian V, Naqvi SMK, Ezeji T, Lakritz J, Rattan L, editors. Environmental stress and amelioration in livestock production. Heidelberg: Springer; 2012. p. 263–98.
45. Lai Y, Li D, Li C, Muehleisen B, Radek KA, Park HJ, et al. The antimicrobial protein REG3A regulates keratinocyte proliferation and differentiation after skin injury. *Immunity.* 2012;37:74–84.
46. Mukherjee S, Hooper LV. Antimicrobial defense of the intestine. *Immunity.* 2015;42:28–39.
47. Cash HL, Whitham CV, Behrendt CL, Hooper LV. Symbiotic bacteria direct expression of an intestinal bactericidal lectin. *Science.* 2006;313:1126–30.
48. Steele MA, Penner GB, Chaucheyras-Durand F, Guan LL. Development and physiology of the rumen and the lower gut: targets for improving gut health. *J Dairy Sci.* 2016;99:4955–66.
49. Chung LK, Raffatellu M. G.I.pros: antimicrobial defense in the gastrointestinal tract. *Semin Cell Dev Biol.* 2018;88:129–37.
50. Zhang S, Albornoz RI, Aschenbach JR, Barreda DR, Penner GB. Short-term feed restriction impairs the absorptive function of the reticulo-rumen and total tract barrier function in beef cattle. *J Anim Sci.* 2013;91:1685–95.
51. Wang MD, Dzama K, Hefer CA, Muchadeyi FC. Genomic population structure and prevalence of copy number variations in South African Nguni cattle. *BMC Genomics.* 2015;16:894.
52. McKenzie GJ, Bancroft A, Grecis RK, McKenzie AN. A distinct role for interleukin-13 in Th2-cell-mediated immune responses. *Curr Biol.* 1998;8:339–42.
53. Buitkamp J, Jann O, Fries R. The cattle interleukin-13 gene: genomic organization, chromosomal location, and evolution of the promoter. *Immunogenetics.* 1999;49:872–8.
54. Trigona WL, Brown WC, Estes DM. Functional implications for signaling via the IL4R/IL13R complex on bovine cells. *Vet Immunol Immunopathol.* 1999;72:73–9.
55. Woods AM, Judd AM. Interleukin-4 increases cortisol release and decreases adrenal androgen release from bovine adrenal cells. *Domest Anim Endocrinol.* 2008;34:372–82.
56. Zoico E, Roubenoff R. The role of cytokines in regulating protein metabolism and muscle function. *Nutr Rev.* 2002;60:39–51.

57. Matarese G, La Cava A. The intricate interface between immune system and metabolism. *Trends Immunol.* 2004;25:193–200.
58. Odegaard JI, Chawla A. The immune system as a sensor of the metabolic state. *Immunity.* 2013;38:644–54.
59. Queitsch C, Sangster TA, Lindquist S. Hsp90 is a capacitor of phenotypic variation. *Nature.* 2002;417:618–24.
60. Sangster TA, Salathia N, Undurraga S, Milo R, Schellenberg K, Lindquist S, et al. HSP90 affects the expression of genetic variation and developmental stability in quantitative traits. *Proc Natl Acad Sci USA.* 2008;105:2963–8.
61. Bahbahani H, Tijjani A, Mukasa C, Wragg D, Almathen F, Nash O, et al. Signatures of selection for environmental adaptation and Zebu × Taurine hybrid fitness in East African Shorthorn Zebu. *Front Genet.* 2017;8:68.
62. Macciotta NPP, Biffani S, Bernabucci U, Lacetera N, Vitali A, Ajmone-Marsan P, et al. Derivation and genome-wide association study of a principal component-based measure of heat tolerance in dairy cattle. *J Dairy Sci.* 2017;100:4683–97.
63. Sanders MM, Kon C. Glutamine is a powerful effector of heat shock protein expression in *Drosophila* Kc cells. *J Cell Physiol.* 1991;146:180–90.
64. Boj SF, van Es JH, Huch M, Li VS, José A, Hatzis P, et al. Diabetes risk gene and Wnt effector Tcf7l2/TCF4 controls hepatic response to perinatal and adult metabolic demand. *Cell.* 2012;151:1595–607.
65. Haase H, Rink L. Multiple impacts of zinc on immune function. *Metalomics.* 2014;6:1175–80.
66. Ticianelli JS, Emanuelli IP, Satrapa RA, Castilho ACS, Loureiro B, Sudano MJ, et al. Gene expression profile in heat-shocked Holstein and Nelore oocytes and cumulus cells. *Reprod Fertil Dev.* 2017;29:1787–802.
67. Franzin AM, Maruyama SR, Garcia GR, Oliveira RP, Ribeiro JMC, Bishop R, et al. Immune and biochemical responses in skin differ between bovine hosts genetically susceptible and resistant to the cattle tick *Rhipicephalus microplus*. *Parasit Vectors.* 2017;10:51.
68. Roth S, Rottach A, Lotz-Havla AS, Laux V, Muschwackh A, Gersting SW, et al. Rad50-CARD9 interactions link cytosolic DNA sensing to IL-1 $\beta$  production. *Nat Immunol.* 2014;15:538–45.
69. Yurchenko AA, Daetwyler HD, Yudin N, Schnabel RD, Vander Jagt CJ, Soloshenko V, et al. Scans for signatures of selection in Russian cattle breed genomes reveal new candidate genes for environmental adaptation and acclimation. *Sci Rep.* 2018;8:12984.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

