Animal (2018), **12:7**, pp 1358–1362 © The Animal Consortium 2017
doi:10.1017/S1751731117003068

Genome-wide association study for growth traits in Nelore cattle

A. P. N. Terakado[1], R. B. Costa[2†], G. M. F. de Camargo[2], N. Irano[1], T. Bresolin[1], L. Takada[1], C. V. D. Carvalho[2], H. N. Oliveira[1], R. Carvalheiro[1], F. Baldi[1] and L. G. de Albuquerque[1]

[1]*Departamento de Zootecnia, Universidade Estadual Paulista (Unesp), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal-SP, Brazil;* [2]*Escola de Medicina Veterinária e Zootecnia, Universidade Federal da Bahia (UFBA), Salvador-BA, Brazil*

(Received 25 March 2017; Accepted 16 October 2017; First published online 16 November 2017)

The objective of this study was to investigate the association of single nucleotide polymorphisms (SNPs) with birth weight, weight gain from birth to weaning and from weaning to yearling, yearling height and cow weight in Nelore cattle. Data from 5064 animals participating in the DeltaGen and PAINT breeding programs were used. The animals were genotyped with a panel of 777 962 SNPs (Illumina BovineHD BeadChip) and 412 993 SNPs remained after quality control analysis of the genomic data. A genome-wide association study was performed using a single-step methodology. The analyses were processed with the BLUPF90 family of programs. When applied to a genome-wide association studies, the single-step GBLUP methodology is an iterative process that estimates weights for the SNPs. The weights of SNPs were included in all analyses by iteratively applying the single-step GBLUP methodology and repeated twice so that the effect of the SNP and the effect of the animal were recalculated in order to increase the weight of SNPs with large effects and to reduce the weight of those with small effects. The genome-wide association results are reported based on the proportion of variance explained by windows of 50 adjacent SNPs. Considering the two iterations, only windows with an additive genetic variance >1.5% were presented in the results. Associations were observed with birth weight on BTA 14, with weight gain from birth to weaning on BTA 5 and 29, with weight gain from weaning to yearling on BTA 11, and with yearling height on BTA 8, showing the genes TMEM68 *(transmembrane protein 8B) associated with birth weight and yearling height,* XKR4 *(XK, Kell blood group complex subunit-related family, member 4) associated with birth weight,* NPR2 *(natriuretic peptide receptor B) associated with yearling height, and* REG3G *(regenerating islet-derived 3-gamma) associated with weight gain from weaning to yearling. These genes play an important role in feed intake, weight gain and the regulation of skeletal growth.*

Keywords: beef cattle, height, single-step, SNP, weight

Implications

The search for genes affecting economically important traits is an useful tool to prospect major genes and better understand the genetic architecture and the physiology behind the expression of the phenotypes. The current study shows the genes that affect growth traits in Nelore beef cattle, which could lead to comparisons across breeds and future breeding strategies.

Introduction

In Brazil, beef cattle farming is undergoing a process of modernization because of the need to increase the production efficiency of the herds due to expansion of the consumer market and competition with other types of meat. One approach to ensure competitiveness is genetic selection (Ferraz Filho *et al.*, 2002). Growth traits such as body weight, which are measured from the early stage of development of the animal until older ages, are important for determining the economic efficiency of any production system. These measures are traditionally included in the selection criteria of beef cattle breeding programs as they are directly related to meat production and exhibit heritability of moderate to high magnitude, indicating that selection for these traits is effective (Frizzas *et al.*, 2009). The genetic correlation between growth traits varies from moderate to high (Boligon *et al.*, 2009; Baldi *et al.*, 2010). High genetic associations are caused by linkage disequilibrium and pleiotropic effects of genes. However, the genes that influence the traits and the magnitude of their effects are still unknown.

The use of markers permits the identification of chromosome regions with great influence on a trait. According to Peters *et al.* (2012), there are practical applications such as the understanding of the genetic structure of populations and subsequent application to genomic selection.

† E-mail: raphaelbcosta@gmail.com

An alternative to the traditional approach of genome-wide association study (GWAS) has recently been proposed by Wang *et al*. (2012) in which all genotypes, phenotypes and pedigree information are considered in a single step (ssGBLUP), thus permitting the use of different models and all relationships simultaneously. Using this approach, all SNPs are considered at the same time together with all phenotypes of genotyped and non-genotyped animals.The objective of this study was to identify regions in the cattle genome that most influence birth weight, weight gain from birth to weaning and from weaning to yearling and yearling height and cow weight in Nelore animals genotyped with a high-density SNP panel.

Material and methods

Phenotypic data

Data from 296 871 Nelore animals (*Bos taurus indicus*) from herds belonging to farms in different regions of Brazil that participate in the DeltaGen and PAINT breeding programs were used. The following traits were evaluated: birth weight (BW), weight gain from birth to weaning (GBW) and from weaning to yearling (GWY), yearling height (YH), and cow weight (CW) Average daily gains were calculated as the difference between the subsequent and previous weight divided by the number of days comprising the period between two weight recordings.

Contemporary groups (CG) for BW were formed by the concatenation of the effects of sex, month, year, farm and management group at birth. For GBW, the farm and management group at weaning were also included. The CG for GWY and YH also included the farm and management group at yearling. Records outside the interval given by the mean of the CG plus or minus three standard deviations were excluded. In addition, CG with fewer than three observations were eliminated. Table 1 shows the overall structure of the data set analyzed. The relationship matrix of genotyped animals contained 5064 animals.

Genomic data cleaning

The animals were genotyped using a high-density panel containing 777 962 SNPs (High-Density Illumina Bovine BeadChip). Quality control of the genotype data was performed in an iterative process, in which first SNPs and then samples were excluded at each iteration until none of remaining SNPs or samples was excluded. The following criteria were used for the exclusion of SNPs: non-autosomal regions; mapped at the same position; *P*-value for Hardy–Weinberg equilibrium $<10^{-5}$; minor allele frequency <2%; GenCall score <70%; and genotyping efficiency of each SNP (call frequency) less than 95%. Samples with a call rate <92% and/or duplicate samples were excluded. The procedures were done using R software (R Core Team, 2014).

Table 1 *Number of observations* (n), *mean and respective standard deviation (SD), minimum, maximum and number of contemporary groups (CG) for birth weight (BW), weight gain from birth to weaning (GBW), weight gain from weaning to yearling (GWY) and yearling height (YH)*

Trait	*n*	Mean	SD	Minimum	Maximum	CG
BW (kg)	296 871	31.07	3.81	20.0	52.0	6922
GBW (kg)	274 137	0.71	0.14	0.16	1.36	6440
GWY (kg)	213 878	0.32	0.11	0.10	0.93	13 226
YH (cm)	127 535	131.66	6.62	104.0	166.0	7147

Genome-wide association study

The single-step methodology (ssGBLUP) proposed by Wang *et al*. (2012) was used for GWAS. The following single-trait model was applied:

$$Y = Xb + Za + e$$

where Y is the vector of phenotypic observations; X the incidence matrix that relates the phenotypes to the fixed effects; b the vector of fixed effects of the CG and covariates (linear and quadratic effects of age of the cow were included in model for BW, GBW, GWY and YH. In addition, the linear and quadratic effect of the age of the animal at measurement were used as covariates for GBW, GWY, YH and CW). Z the incidence matrix that relates the phenotypes to the animals; a the vector of effects of the animals, and e the vector of residual effects.

The variances of a and e are given by:

$$Var\begin{bmatrix} a \\ e \end{bmatrix} = \begin{bmatrix} H\sigma_a^2 & 0 \\ 0 & I\sigma_e^2 \end{bmatrix}$$

where $\sigma^2{}_a$ is the additive genetic variance and $\sigma^2{}_e$ the residual variance. H is the matrix that combines the relationship matrix and genomic information matrix as described by Aguilar *et al*. (2010), and I an identity matrix. The inverse of matrix H is:

$$H^{-1} = A^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & G^{-1} - A_{22}^{-1} \end{bmatrix}$$

where A is the relationship matrix for all animals; A_{22} is the matrix for genotyped animals, and G is the matrix of genomic information, which was calculated as described by Vanraden (2008).

$$G = ZDZ'\lambda$$

where Z is the coefficient matrix adjusted for allele frequency; D the weight matrix for SNPs (initially $D = I$), and λ the weighting factor based on SNP frequencies (Vanraden, 2008):

$$\lambda = \frac{\sigma_u^2}{\sigma_a^2} = \frac{1}{\Sigma_{i=1}^{M} 2p_i(1-p_i)}$$

The effects of the SNPs and weights were derived following the steps proposed by Wang *et al*. (2014). The iteration process was repeated twice, that is, the weights of SNPs were recalculated, repeating steps 2 to 6. The objective of these iterations is to increase the weight of SNPs with large effects and to reduce the weight of those with small effects (Wang *et al*., 2014). The analyses were processed with the BLUPF90 family of programs (Misztal, 2012). The GWAS results are reported as the proportion of variance explained by a window of 50 adjacent SNPs. The Map Viewer tool

for the bovine genome available at NCBI was used to report the genes within the windows: (http://www.ncbi.nlm.nih.gov/projects/mapview/map_search.cgi?taxid=9913&build=103.0). Assembly: Bos_taurus_UMD_3.1.

Results and discussion

Supplementary Material Figures S1, S2, S3, S4 and S5 show the Manhattan plots with the variances of SNP windows that explain >1.5% of the additive genetic variance for BW, GBW, GWY, YH and CW, respectively, after two iterations. Practically all traits have SNP windows that explain >1.5% of additive genetic variance. Important SNP windows were identified on BTA 14 for BW, on BTA 5 and 29 for GBW, on BTA 11 and 8 for GWY, and on BTA 12 for YH. Cow weight exhibited no SNP windows that explained >1.5% of genetic variance.

Tables 2, 3, 4 and 5 show the genes (chromosome and position) found in a window of 50 SNPs that explain >1.5% of the additive genetic variance for BW, GBW, GWY and YH, respectively. For BW, the window that explains the highest percentage of additive genetic variance in the trait contains genes already described in previous studies for other traits. Lindholm-Perry *et al*. (2011) found SNPs in the *TMEM68* gene. This gene is expressed in the rumen, abomasum, intestine and adipose tissue. It codes for an enzyme whose role is to catalyze important reactions that involve proteins, lipids or other substrates located within or near the fat layers (Lindholm-Perry *et al*., 2011). The *TMEM68* and *XKR4* genes have been associated with feed intake and average daily gain in cattle. SNPs near the *XKR4* gene have also been associated with subcutaneous fat thickness (Porto-Neto *et al*., 2012) and the gene has been indicated as a candidate for carcass traits (Ramayo-Caldas *et al*., 2014), both also in cattle.

The *LYN* gene has been identified in a candidate region for BW of Nelore cattle in other GWAS (Utsunomiya *et al*., 2013). In another nearby window on the same chromosome, the *ASPH* gene has been associated with muscle hypertrophy (Perez *et al*., 2010) and also indicated as a candidate for carcass traits in cattle (Ramayo-Caldas *et al*., 2014). Some genes are candidates for the same trait (BW), while others are candidates for traits correlated with BW, indicating the pleiotropic effect of these genes. Utsunomiya *et al*. (2013), studying BW in Nelore cattle, found SNPs located on BTA 14 that explained 4.62% of the genetic variance in the trait. This region has also been indicated by other authors to be associated with reproductive traits, highlighting the *PLAG1*, *CHCHD7*, *MOS*, *RPS20*, *LYN*, *RDHE2* (SDR16C5) and *PENK* genes that are associated with height in humans and cattle. The *LYN* gene was the only gene found in the present study (associated with BW) that has also been reported by other authors.

Table 2 *Identification of genes, chromosome and position based on the genetic variance explained by windows of 50 single nucleotide polymorphisms for birth weight*

Birth weight			
Chromosome	Position	Gene	% Variance explained
14	28 561 092 to 28 788 465 bp	*CLVS1* *ASPH*	1.97
	24 565 383 to 24 874 608 bp	*XKR4* *TMEM68* *TGS1* *LYN*	2.32

Table 3 *Identification of genes, chromosome and position based on the genetic variance explained by windows of 50 single nucleotide polymorphisms for weight gain from birth to weaning*

Weight gain from birth to weaning			
Chromosome	Position	Gene	% Variance explained
5	51 972 554 to 52 329 387 bp	*FAM19A2* *LOC101903699*	1.78
29	26 803 260 to 27 160 083 bp	*LOC506317* *OR8D4* *OR4D5* *LOC787418* *LOC790885* *LOC787455* *LOC532517* *LOC781760* *LOC781804* *LOC526286* *LOC100298846* *LOC100337086* *LOC506981* *LOC100300434* *LOC510293* *LOC522385* *OR10S1* *LOC100299320* *LOC100336852* *LOC100299628* *LOC100336852* *LOC100301071* *LOC782288* *LOC782248* *LOC782216* *LOC782366*	1.55

Table 4 *Identification of genes, chromosome and position based on the genetic variance explained by windows of 50 single nucleotide polymorphisms for weight gain from weaning to yearling*

Weight gain from weaning to yearling			
Chromosome	Position	Gene	% Variance explained
11	4 676 595 to 5 104 110 bp	*REG3G*	2.03

Table 5 *Identification of genes, chromosome and position based on the genetic variance explained by windows of 50 single nucleotide polymorphisms for yearling height*

Yearling height			
Chromosome	Position	Gene	% Variance explained
8	60 325 913 to 60 771 025 bp	*GBA2* *RGP1* *MSMP* *LOC101904249* *NPR2* *SPAG8* *HINT2* *FAM221B* *TMEM8B* *LOC529284* *LOC784594* *LOC618717* *LOC615852* *LOC784434* *LOC784343* *LOC618828* *LOC100337274* *LOC506202* *LOC784185* *LOC784108* *HRCT1* *OR2S2* *LOC100299372* *LOC523083* *LOC532403* *LOC504766* *LOC613441* *LOC783787* *RECK*	1.77
12	45 929 754 to 46 319 826 bp	–	1.88
	43 417 321 to 43 760 894 bp	–	2.33

In the analysis of GBW, the *FAM19A2* gene was identified on BTA 5. This gene is associated with body mass index and weight in humans (Comuzzie *et al*., 2012), providing evidence of its participation in performance and body growth in other species. Also for GBW, the second window that explained most of the additive genetic variance, on BTA 29, contains olfactory receptors (OR) genes (*OR8D4*, *OR4D5* and *OR10S1*) and many LOC genes corresponding to OR genes. In a study on rats, some olfactory receptor genes exhibited high levels of expression in the duodenum when the animals received diets rich in fatty acids (Primeaux *et al*., 2012). It means that OR genes may be related to fatty acids metabolism/absorption in the duodenum, establishing a clear explanation for their influence in GBW trait.

Analysis of the association with GWY identified the *REG3G* gene located on BTA 11. This gene plays a role in jejunal homeostasis and has been suggested to act on the renewal of intestinal cells in mice (Everard *et al*., 2014). More efficient renewal of the intestinal epithelium improves the absorption of nutrients and consequently increases weight gain. Buzanskas *et al*. (2014), studying growth traits such as weaning and yearling weight in Canchim cattle, found SNPs associated with weaning weight on BTA 4, 6 and 11 and with yearling weight on BTA 7, 22, 25 and 27. These regions are different from the ones found here. Divergent results can be explained by allele variations among different breeds and by differences in the methodologies used in each study, factors that can lead to over- or underestimation of the effects of QTLs and SNPs.

A total of 29 genes located on BTA 8 were associated to YH. Two SNP windows stood out on BTA 12, which explained 1.88% and 2.33% of the additive genetic variance of this trait, but there is no description of any gene in these regions. In the window associated with BTA 8, the *NPR2* gene is an important regulator of skeletal growth. In humans, the loss of function or mutations in this gene cause acromesomelic dysplasia, type Maroteaux (AMDM), a condition that results in disproportionally short stature. Humans with AMDM are born with a normal size and abnormal growth becomes apparent after birth (Olney *et al*., 2005).

The *TMEM8B* gene was associated with YH, which was also identified in the analysis of the association with BW in the present study, plays an important role in different biological processes with its role explained above. The *OR2S2* is an olfactory receptor gene as probably the LOC genes at the same window. They possibly play a role as the olfactory receptors associated with GBW. There are no genes described or associated with cow weight that explain more than 1.5% of the additive genetic variance in the trait.

Many studies have reported genes associated with growth traits in regions different from those found in this study (Alexander *et al*., 2007; Gutiérrez-Gil *et al*., 2009; Bolormaa *et al*. (2011); Pryce *et al*. (2011); Buzanskas *et al*., 2014). This fact might be explained by the different genetic constitution of breeds. Variations in allele frequency, linkage disequilibrium and SNP coverage of the chip in the breed can affect the results. Some factors such as methodology and the number of animals used can also be responsible for these differences. The present study used a large number of animals and an innovative methodology and we believe that the results obtained are reliable for the breed and for future comparisons between breeds.

The identification of genome regions that affect economic traits in cattle is extremely important since knowledge of these regions will likely contribute in the future to the selection of these traits in breeding programs of the Nelore breed. Genes of great contribution to the variance of traits are candidates for future studies on causal mutations to compose SNP chips of low density and to help with the process of genetic evaluation at a better cost–benefit ratio.

Conclusions

The ssGBLUP methodology permitted to identify chromosome regions that were associated with BW, GBW, GWY and

YH. The regions identified in this study help elucidate the genes that most contribute to the variance of the traits, permit a better understanding of the physiological process underlying the expression of phenotypes, and indicate candidate genes for future fine-mapping studies. Moreover, it evidences possible strong genetic differences across breeds, useful for future breeding comparison strategies.

Acknowledgments

The costs associated to genotyping of animals was supported by the thematic project 'Genomic tools for the genetic improvement of economically important traits in Nelore cattle' supported by Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP grant numbers 2009/16118-5).

Supplementary material

To view supplementary material for this article, please visit https://doi.org/10.1017/S1751731117003068

References

Aguilar I, Misztal I, Johnson DL, Legarra A, Tsuruta S and Lawlor TJ 2010. Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. Journal of Dairy Science 93, 743–752.

Alexander LJ, Geary TW, Snelling WM and MacNeil MD 2007. Quantitative trait loci with additive effects on growth and carcass traits in a Wagyu-Limousin F2 population. Animal Genetics 38, 413–416.

Baldi F, Albuquerque LG and Alencar MM 2010. Random regression models on Legendre polynomials to estimate genetic parameters for weights from birth to adult age in Canchim cattle. Journal of Animal Breeding and Genetics 127, 289–299.

Boligon AA, Albuquerque LG, Mercadante MEZ and Lôbo RB 2009. Herdabilidades e correlações entre pesos do nascimento à idade adulta em rebanhos da raça Nelore. Revista Brasileira de Zootecnia 38, 2320–2326.

Bolormaa S, Hayes BJ, Savin K, Hawken R, Barendse W, Arthur PF, Herd RM and Goddard ME 2011. Genome-wide association studies for feedlot and growth traits in cattle. Journal of Animal Science 89, 1684–1697.

Buzanskas ME, Grossi DA, Ventura RV, Schenkel FS, Sargolzaei M, Meirelles SLC, Mokry FB, Higa RH, Mudadu MA, Silva MVGB, Niciura SCM, Torres Júnior RAA, Alencar MM, Regitano LCA and Munari DP 2014. Genome-wide association for growth traits in Canchim beef cattle. PLoS One 9, e94802.

Comuzzie AG, Cole SA, Laston SL, Voruganti VS, Haack K, Gibbs RA and Butte NF 2012. Novel genetic loci identified for the pathophysiology of childhood obesity in the Hispanic population. PLoS One 7, e51954.

Everard A, Lazarevic V, Gaia N, Johansson M, Stahlman M, Backhed F, Delzenne NM, Schrenzel J, François P and Cani PD 2014. Microbiome of prebiotic-treated mice reveals novel targets involved in host response during obesity. Multidisciplinary Journal of Microbial Ecology 8, 2116–2130.

Ferraz Filho PB, Ramos AA, Silva LOC, Souza JC, Alencar MM and Malhado CHM 2002. Tendência genética dos efeitos direto e materno sobre os pesos à desmama e pós-desmama de bovinos da raça Tabapuã no Brasil. Revista Brasileira de Zootecnia 31, 635–640.

Frizzas OG, Grossi DA, Buzanskas ME, Paz CCP, Bezerra LAF, Lôbo RB, Oliveira JA and Munari DP 2009. Heritability estimates and genetic correlations for body weight and scrotal circumference adjusted to 12 and 18 months of age for male Nellore cattle. Animal 3, 347–351.

Gutiérrez-Gil B, Williams JL, Homer D, Burton D, Haley CS and Wiener P 2009. Search for quantitative trait loci affecting growth and carcass traits in a cross population of beef and dairy cattle. Journal of Animal Science 87, 24–36.

Lindholm-Perry AK, Sexten AK, Kuehn LA, Smith TPL, King DA, Shackelford SD, Wheeler TL, Ferrell CL, Jenkins TG, Snelling WM and Freetly HC 2011. Association, effects and validation of polymorphisms within the NCAPG – LCORL locus located on BTA6 with feed intake, gain, meat and carcass traits in beef cattle. BMC Genetics 12, 103.

Misztal I 2012. BLUPF90 – a flexible mixed model program in Fortran 90, University of Georgia. Retrieved on 15 December 2014 from http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90.pdf.

Olney RC, Bükülmez H, Bartels CF, Prickett TCR, Espiner EA, Potter LR and Warman ML 2005. Heterozygous mutations in natriuretic peptide receptor-B (NPR2) are associated with short stature. Journal of Clinical Endocrinology and Metabolism 91, 1229–1232.

Perez R, Cañón J and Dunner S 2010. Genes associated with long-chain omega-3 fatty acids in bovine skeletal muscle 1010. Journal of Applied Genetics 51, 479–487.

Peters SO, Kizilkaya K, Garrick DJ, Fernando RL, Reecy JM, Weaber RL, Silver GA and Thomas MG 2012. Bayesian genome wide association analyses of growth and yearling ultrasound measures of carcass traits in Brangus heifers. Journal of Animal Science 90, 3398–3409.

Porto-Neto LR, Bunch RJ, Harrison BE and Barendse W 2012. Variation in the XKR4 gene was significantly associated with subcutaneous rump fat thickness in indicine and composite cattle. Animal Genetics 43, 785–789.

Primeaux SD, Braymer HD and Bray GA 2012. High fat diet differentially regulates the expression of olfactory receptors in the duodenum of obesity-prone and obesity-resistant rats. Digestive Diseases and Sciences 58, 72–78.

Pryce JE, Hayes BJ, Bolormaa S and Goddard ME 2011. Polymorphic regions affecting human height also control stature in cattle. Genetics 187, 981–984.

Ramayo-Caldas Y, Ballester M, Fortes MRS, Esteve-Codina A, Castelló A, Noguera JL, Fernández AI, Pérez-Enciso M, Reverter A and Folch JM 2014. From SNP co-association to RNA co-expression: novel insights into gene networks for intramuscular fatty acid composition in porcine. BMC Genomics 15, 232–246.

R Core Team 2014. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Utsunomiya YT, Carmo AS, Carvalheiro R, Neves HHR, Matos MAC, Zavarez LB, O'Brien AMP, Sölkner J, McEwan JC, Cole JB, Vantassel CP, Schenkel FS, Silva MVGB, Porto-Neto LR, Sonstegard TS and Garcia JF 2013. Genome-wide association study for birth weight in Nellore cattle points to previously described orthologous genes affecting human and bovine height. BMC Genetics 14, 52.

Vanraden PM 2008. Efficient methods to compute genomic predictions. Journal of Dairy Science 91, 4414–4423.

Wang H, Misztal I, Aguilar I, Legarra A, Fernando RL, Vitezica Z, Okimoto R, Wing T, Hawken R and Muir WM 2014. Genome wide association mapping including phenotypes from relatives without genotypes in a single-step (ssGWAS) for 6 week body weight in broiler chickens. Frontiers in Genetics 5, 134.

Wang H, Misztal I, Aguilar I, Legarra A and Muir WM 2012. Genome-wide association mapping including phenotypes from relatives without genotypes. Genetic Research 94, 73–83.