Livestock Science 183 (2016) 4–11

Contents lists available at ScienceDirect

Livestock Science

journal homepage: www.elsevier.com/locate/livsci

Genome-wide association study of growth and body composition traits in Brangus beef cattle

Ziqing Weng [a], Hailin Su [a], Mahdi Saatchi [b], Jungjae Lee [a], Milton G. Thomas [c], Jenelle R. Dunkelberger [a], Dorian J. Garrick [a,*]

[a] Department of Animal Science, Iowa State University, Ames, IA 50010, USA
[b] Genus plc, Hendersonville, TN 37075, USA
[c] Department of Animal Sciences, Colorado State University, Fort Collins, CO 80523, USA

ARTICLE INFO

Article history:
Received 24 June 2015
Received in revised form
8 November 2015
Accepted 11 November 2015

Keywords:
Deregressed estimated breeding value
GWAS
QTL

ABSTRACT

The availability of high-density single nucleotide polymorphism (SNP) genotypes, such as BoveineHD770K, provides opportunities to identify genomic regions associated with traits in cattle. The objective of this study was to identify quantitative trait loci (QTL) associated with growth and body composition traits in Brangus beef cattle using actual and imputed 770 K SNP genotypes. A total of 1537 Brangus beef cattle were genotyped with the Bovine50K, GGPHD77K, or BovineHD770K SNP chip and deregressed estimated breeding values were derived and fitted as observations in analyses. BayesB approach was used to map QTL for each trait, and significant windows and SNPs were identified. A total of 18 QTL were identified, in which 7 were associated with more than one trait, while the remained 11 QTL were trait-specific. One pleiotropic QTL of particularly large-effect was identified on chromosome 6 at 38 Mb, which influences direct birth weight, weaning weight, and yearling weight, and harbors growth-related genes *NCAPG* and *LCORL*. Biological pathways of pleiotropic QTL were also performed using gene ontology term enrichment analysis. The QTL mapping results obtained from this study will aid in better understanding the biological processes accounting for variation in growth and body composition traits in Brangus cattle.

1. Introduction

Genome-wide association studies (GWAS) are widely used to detect chromosomal regions that are responsible for genetic variation for traits of interest (Weller, 2009). Increasing marker density, enlarging the sample size, and use of markers in strong linkage disequilibrium (LD) with quantitative trait loci (QTL), are ways to improve the accuracy of GWAS. Bayesian regression methods are useful for GWAS (Zou and Zeng, 2008) because they account for uncertainty in parameters required to construct posterior distributions for QTL inference. Recently, several regions of the bovine genome associated with growth traits have been reported in beef cattle. One region is on *Bos tauras* autosome (BTA) 6, which harbors candidate genes including *NCAPG* and *LCORL* (Nishimura et al., 2012; Saatchi et al., 2014; Santana et al., 2014; Snelling et al., 2010). Another region, located on BTA 14, contains *PLAG1, CHCHD7,* and other related genes, all of which affect growth in cattle (Nishimura et al., 2012; Pausch et al., 2011; Saatchi et al., 2014) and humans (Pryce et al., 2011; Utsunomiya et al., 2013).

Brangus is a composite breed comprising Brahman and Angus (3/8 *Bos indicus* × 5/8 *Bos taurus*). QTL mapping using Brangus might identify QTL that are segregating in both Angus and Brahman. Peters et al. (2012) and Saatchi et al. (2014) mapped QTL for growth traits using the lower density BovineSNP50 in two subsets (835 heifers and 1328 cattle) of this same dataset of U.S. Brangus cattle, respectively. However, GWAS of Brangus using high density of SNP genotypes have not been fully addressed. The objective of this study was to identify QTL associated with growth and body composition traits in a Brangus population using actual and imputed BovineHD770K SNP genotypes, and compare the results with previous studies.

2. Materials and methods

2.1. Phenotypes and genotypes

Estimated breeding values (EBVs), their reliability, and

* Corresponding author.
E-mail addresses: zweng@iastate.edu (Z. Weng), hailins@iastate.edu (H. Su), Mahdi.Saatchi@genusplc.com (M. Saatchi), leejj@iastate.edu (J. Lee), milt.thomas@colostate.edu (M.G. Thomas), jenelled@iastate.edu (J.R. Dunkelberger), dorian@iastate.edu (D.J. Garrick).

http://dx.doi.org/10.1016/j.livsci.2015.11.011
1871-1413/

Table 1
Heritability (h^2), number of genotyped animals with deregressed estimated breeding values (DEBVs), and mean reliabilities (standard error) of DEBV for the evaluated traits in Brangus cattle.

Trait	h^2[a]	Reliability	N[b]
Direct birth weight	0.42	0.58 (0.14)	1401
Maternal birth weight	0.15	0.40 (0.16)	864
Back fat thickness	0.35	0.47 (0.19)	1065
Intramuscular fat thickness	0.23	0.34 (0.19)	1152
Rib eye muscle area	0.39	0.49 (0.19)	1160
Rump fat thickness	0.38	0.49 (0.20)	860
Mature weight	0.27	0.39 (0.27)	1176
Weaning weight	0.24	0.46 (0.14)	1426
Yearling weight	0.27	0.45 (0.15)	1332

[a] Heritabilities reported by International Brangus Breeders Association. Parent average contributions have been removed to calculate the reliabilities.
[b] The number of DEBVs which had reliabilities >0.8 h^2.

heritability estimates (Table 1) of growth and body composition traits in 1537 Brangus beef cattle (983 heifers and 554 bulls from 168 farms) were obtained from the International Brangus Breeders Association. Traits studied included direct birth weight, maternal birth weight, back fat thickness, intramuscular fat thickness, rib eye muscle area, rump fat thickness, mature weight, weaning weight, and yearling weight.

A subset of these Brangus cattle (1121) were genotyped with Bovine50K (Illumina, San Diego, CA), while another 243 cattle were genotyped with BovineHD770K (Illumina, San Diego, CA), and the rest 173 cattle were genotyped with GGPHD77K (Gene Seek, Lincoln, NE). These Brangus cattle were born between 1976 and 2012. Distribution of their birth year was shown in Table 2. Most animals genotyped with BovineHD770K and Bovine50K were born between 2001 and 2010, while majority of animals genotyped with GGPHD77K were born after 2010. Genotypes were obtained using DNA extracted from blood, semen, or hair samples and did not require an approved animal care and use statement.

2.2. Imputation

All of the animals genotyped with the Bovine50K or GGPHD77K were imputed to BovineHD770K using FImpute (Sargolzaei et al., 2014). Before imputation, 77 1089 SNPs on BovineHD770K with call rate >0.90 (6173 SNPs were removed) and could be mapped to SNPchiMP (Nicolazzi et al., 2014) were retained as HD reference panel. SNPs aligned to HD reference panel with call rate >0.90 were remained on Bovine50K (48 740 SNPs) and GGPHD77K (73 989 SNPs), respectively. All of 1537 animals had call rate >0.90.

After imputation, 736 053 segregating markers remained after quality control and limiting markers to those that were uniquely assigned to 29 bovine autosomes and X chromosome on UMC assembly (University of Missouri). In total, 31 506 SNP with minor allele frequency <0.005, 1565 SNPs with Mendelian inconsistency rate between parent and progeny >0.05, and 1965 unmapped SNP were removed, respectively. Genotype quality control was conducted using PLINK v1.0 (Purcell et al., 2007).

Table 2
Distribution of birth year of genotyped Brangus cattle on each SNP panel.

Birth year	Total animals	Bovine HD770K	GGPHD 77K	Bovine 50K
1976–1990	19	3	0	16
1991–1995	25	7	0	18
1996–2000	96	11	4	81
2001–2005	365	80	20	265
2006–2010	878	130	44	704
2011–2012	154	12	105	37

Genotyped animals were clustered into three groups according to SNP genotype panels (Bovine50K or GGPHD77K or BovineHD770K). The average additive genetic relationship coefficient (a_{mean}) and maximum additive genetic relationship coefficient (a_{max}) were calculated between each animal and other animals from the same group or other groups using pedigree information. To qualify the relationship between different groups, the mean values of a_{mean} and a_{max} within and between groups were calculated across all animals in each group.

A simulation study was conducted to evaluate the imputation accuracy of FImpute. A sample of 10 animals was randomly selected from HD770K group. The remained 233 animals were used as reference population. Imputation from 77 K to 770 K SNP panels was performed in these selected animals based on the use of only that subset of 73 989 SNPs from the 770 K panel that were on the 77 K. The same strategy was conducted to impute genotypes from 50 K to 770 K SNP panels only using the overlapped 48 740 SNPs. Each scenario was repeated 5 times, in order to avoid sample bias. Allelic r^2 was used to qualify imputation accuracy, which depends less on SNP allele frequency compared with concordance rate (Browning and Browning, 2009). It was calculated as the squared correlation between the imputed genotypes and original genotypes on the 770 K SNP panel in the selected animals.

2.3. Deregressed EBVs

Individual EBVs combine information from the performance of the individual, any progeny, and its parents, based on pedigree relationships. In this study, deregressed estimated breeding values (DEBVs) were derived according to Garrick et al. (2009). DEBVs on each trait used as response variables to estimate SNP effects in a weighted Bayesian analysis. Weighting factors were used to account for heterogeneous variance caused by differences in reliabilities of individual DEBVs. Only the DEBVs with reliability >0.8 h^2 were retained in the analysis which eliminated animals without individual or offspring information contributing to their EBVs, causing the number of genotyped animals with DEBVs to vary between traits (Table 1).

2.4. Genome-wide association study

Mapping of QTL associated with the performance traits of interest was conducted using the BayesB method (Meuwissen et al., 2001) with weighting factors implemented in GENSEL4.4 software (Fernando and Garrick, 2013; Garrick and Fernando, 2013). According to the BayesB method, each SNP effect follows an independent, univariate t-distribution with null mean,degree of freedom v_a, and scale parameter S_a^2, which is equivalent to a univariate normal distribution with null mean and locus specific variance (Fernando and Garrick, 2013). For each trait, DEBVs were used as response variables in the following model:

$$y_i = \mu + \sum_{j=1}^{k} z_{ij} u_j + e_i \ ,$$

where y_i is the DEBV for animal i, μ is the population mean, k is the number of SNP loci, z_{ij} is the marker genotype coded (0/1/2) for SNP j in animal i, u_j is the substitution effect for SNP j with $u_j>0$ (with probability of $1-\pi$), or $u_j=0$ (with probability π), and ei is a residual effect with heterogeneous variance. Weighting factors (Garrick et al., 2009) were calculated as:

$$w_i = \frac{1-h^2}{[c+(1-r_i^2)/r_i^2]h^2}$$

where c is the proportion of the genetic variance (GV%) not explained by markers, which was assumed to be 0.40 ((Saatchi