SIRLENE FERNANDES LÁZARO

BAYESIAN MODELS FOR GROWTH CURVES, CENSORED DATA AND VISUAL SCORES IN ANIMAL BREEDING

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-graduação em Zootecnia, para obtenção do título de *Doctor Scientiae*.

VIÇOSA
MINAS GERAIS – BRASIL
2017

Ficha catalográfica preparada pela Biblioteca Central da Universidade Federal de Viçosa - Campus Viçosa

T

L431b
2017

Lazaro, Sirlene Fernandes, 1986-
Bayesian models for growth curves, censored data and visual scores in animal breeding / Sirlene Fernandes Lazaro. - Viçosa, MG, 2017.
xi, 79f. : il. ; 29 cm.

Orientador: Paulo Sávio Lopes.
Tese (doutorado) - Universidade Federal de Viçosa.
Inclui bibliografia.

1. Suínos - Melhoramento genético. 2. Suínos - Crescimento. I. Universidade Federal de Viçosa. Departamento de Zootecnia. Programa de Pós-graduação em Zootecnia. II. Título.

CDD 22 ed. 636.40821

SIRLENE FERNANDES LÁZARO

BAYESIAN MODELS FOR GROWTH CURVES, CENSORED DATA AND VISUAL SCORES IN ANIMAL BREEDING

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-graduação em Zootecnia, para obtenção do título de *Doctor Scientiae*.

APROVADA: 22 de fevereiro de 2017.

Luis Varona Aguado

Camila Ferreira Azevedo

Henrique Torres Ventura
(Coorientador)

Fabyano Fonseca e Silva
(Coorientador)

Paulo Sávio Lopes
(Orientador)

AGRADECIMENTOS

A Deus, pelas minhas vitórias e pela fé que me dá força para superar todos os obstáculos que surgem ao longo da minha vida.

À minha mãe, Maria da Consolação, pelo exemplo de vida, dedicação, compreensão e carinho recebido, e que em certa época da vida me ensinou o sentido literal da palavra lutar, e me ensina diariamente o sentido real da palavra amar.

À Universidade Federal de Viçosa, em especial ao Departamento e ao Programa de Pós-Graduação em Zootecnia, pela oportunidade de realização deste curso.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico, pela cessão da bolsa de estudo.

À Universidade de Zaragoza pela oportunidade de realização do doutorado sanduíche.

Ao professor Luis Varona que foi meu orientedor de doutorado sanduíche na Universidade de Zaragoza e me recebeu com tanto carinho e atenção.

À pesquisadora Noelia Ibáñez-Escriche do Instituto de Pesquisa e Tecnologia de Alimentos de Lérida-Espanha (IRTA), pela excelente recepção e ensinamentos durante o períododo doutorado sanduíche.

À Associação Brasileira dos Criadores de Zebu (ABCZ) pela disponibilização dos dados para a realização deste trabalho.

Ao meu orientador, Prof. Paulo Sávio Lopes, que com sua sabedoria inigualável me guiou de forma excepcional durante a minha jornada na pós-graduação.

Aos meus coorientadores, Prof. Fabyano Fonseca e Silva e Dr. Henrique Torres Ventura, pelo engrandecimento pessoal e profissional.

À Profª. Camila Ferreira Azevedo, pela disponibilidade em participar da banca examinadora e sugestões que contribuíram para aprimorar este trabalho.

A todos os meus mestres do ensino fundamental ao superior pelo ensinamento e preparo profissional, principalmente os professores Ernane Ronie Martins e Joana Ribeiro da Gloria, a “culpada” que me conduziu ao mundo do melhoramento genético animal.

A todos os amigos da “Salinha” (Laboratório de Informática – Melhoramento Animal/UFV) e do Melhoramento Genético pelas horas de estudo e companheirismo. Vocês sempre estarão em minha memória e terão um lugar em meu coração.

A todas as pessoas que tive o (des) prazer em conhecer durante a minha jornada, vocês foram imprescindíveis na consolidação do meu caráter.

Enfim, a todos que, mesmo não citados, foram importantes pela passagem nessa importante etapa da minha vida. Que Deus os abençoe e, esteja sempre ao lado de cada um, proporcionando, a todos, momentos de alegrias.

BIOGRAFIA

SIRLENE FERNANDES LÁZARO, solteira, filha de José Antônio Lázaro e Maria da Consolação Lázaro, nasceu em 08 de agosto de 1986, em Congonhas, Minas Gerais. Em março de 2006, iniciou o curso de Zootecnia na Universidade Federal de Minas Gerais – MG. Em abril de 2010 recebeu o prêmio de Destaque Estudantil Do Norte de Minas Gerais do ano de 2009. Em janeiro de 2011, recebeu o diploma de Zootecnista e em março do mesmo ano iniciou o curso de Mestrado em Zootecnia pela Universidade Federal de Minas Gerais. Em fevereiro de 2013, obteve o título de *Magister Scientiae* em Zootecnia, sob orientação do prof. Martinho de Almeida e Silva. Em abril do mesmo ano, iniciou o curso de Doutorado em Zootecnia na Universidade Federal de Viçosa com período *sandwich* pela *Universidad de Zaragoza* sob supervisão do prof. Luis Varona Aguado. Em fevereiro de 2017, submeteu-se aos exames finais de defesa de tese para obtenção do título de *Doctor Scientiae* em Zootecnia, sob orientação do prof. Paulo Sávio Lopes e coorientação do prof. Fabyano Fonseca e Silva e do Dr. Henrique Torres Ventura.

SUMÁRIO

ABSTRACT

LÁZARO, Sirlene Fernandes, D.Sc., Universidade Federal de Viçosa, February, 2017. **Bayesian models for growth curves, censored data and visual scores in animal breeding.** Advisor: Paulo Sávio Lopes. Co-advisors: Fabyano Fonseca e Silva and Henrique Torres Ventura.

In the first chapter, we proposed a genome association study for pig growth curves based on Bayesian hierarchical framework. A panel of 237 SNPs markers with the pedigree were used jointly to identify possible chromosomal regions that affect growth curve parameters (weight-age data) of 345 animals (F2 population from the Piau *vs*. commercial). Under the proposed hierarchical approach, individual growth trajectories were modeled by the nonlinear Gompertz function, so that the parameter estimates were considered to be affected by systematic, additive polygenic and SNP markers effects. The model assuming jointly pedigree and SNP markers presented the best fit based on Deviance Information Criterion. Heritability estimates ranged from 0.53 to 0.56 and from 0.55 to 0.57, respectively, for the parameters mature weight (a) and maturing rate (k). Additionally, we found high and positive genetic correlation (0.78) between “a” and "k". The percentages of the genetic variances explained by each SNP allowed identifying the most relevant chromosome regions for each phenotype (growth curve parameters). We identified three relevant SNPs (55840514 bp at SSC17, 55814469 bp at SSC17 and 76475804 bp at SSC X) affecting "a" and "k" simultaneously, and three SNPs affecting only "a" (292758 bp at SSC1, 67319 bp at SSC8 and 50290193 bp at SSC17), that are located in regions not previously described as QTL for growth traits in pigs. The modeling used was effective, and resulted in the identification of SNPs located in specific chromosomal regions that have the potential to be explored in breeding programs by marker-assisted selection. In the second chapter, we compared different methods for handling censored data of age at first calving (AFC) in Brahman cattle by Bayesian

models. Data were provided by Brazilian Association of Zebu Cattle Breeders (ABCZ). Censored records were defined as AFC records outside the interval from 731 to 1824 days. Data containing 53,703 AFC records were analyzed using four different methods: conventional linear method (LM), simulation method (SM), penalty method (PM) and a bitrait threshold-linear model considering (TLcens). The additive genetic variance components estimated from LM and PM were similar. Heritability estimates for AFC ranged from 0.09 (TLcens) to 0.20 (LM). In general, genetic breeding values correlations from different methods and the percentage of selected animals in common indicated moderate reranking, ranging from 0.82 (LM x SM) to 0.97 (LM x PM) and 32.70 (SM x TLcens) to 89.12 (LM x PM), respectively. Comparisons based on cross-validation analyses, indicated LM as a suitable alternative for predicting breeding values for AFC in this Brahman population. In the third chapter, we estimated genetic parameters for visual scores of body structure (S), precocity (P), muscularity (M) and reproductive (age at first calving - AFC) traits in Brahman cattle by using Bayesian bitrait and full multitrait models. The heritability estimates obtained using bitrait model were 0.59 (S), 0.44 (P), 0.38 (M), and 0.20 (AFC) and those obtained by full multitrait model were 0.60 (S), 0.44 (P), 0.40 (M) and 0.20 (AFC). Genetic correlations were 0.57 between body structure and precocity, 0.56 between body structure and muscularity and 0.82 between precocity and muscularity (by full multitrait model). Genetic correlations between visual scores and AFC were negatives and moderate magnitude (-0.29, -0.24 and -0.31 to S, P and M by bitrait model) and (-0.29, -0.22 and -0.29 to S, P and M by full multitrait model). These results suggest that visual scores can be used as selection criteria in Brahman cattle breeding programs and that these traits present favorable genetic correlation with age at first calving.

RESUMO

LÁZARO, Sirlene Fernandes, D.Sc., Universidade Federal de Viçosa, fevereiro de 2017. **Modelos Bayesianos para curvas de crescimento, dados censurados e escores visuais no melhoramento animal.** Orientador: Paulo Sávio Lopes. Coorientadores: Fabyano Fonseca e Silva e Henrique Torres Ventura.

No primeiro capitulo, foi proposto um estudo de associação genômica para curvas de crescimento de suínos utilizando modelos hierárquicos Bayesianos. Utilizou-se um painel de 237 marcadores SNPs conjuntamente com informações de pedigree objetivando identificar possíveis regiões cromossômicas que afetam os parâmetros da curva de crescimento (dados de peso-idade) de 345 animais (população F2 proveniente do cruzamento Piau *vs* comercial). Assumiu-se uma trajetória de crescimento individual descrita pela função não linear de Gompertz, de forma que as estimativas de cada parâmetro desta função são influenciadas pelos efeitos sistemáticos, poligênicos aditivos e de marcadores SNPs. O modelo combinando informações de pedigree e marcadores apresentou o melhor ajuste com base no critério de informação da deviance (DIC). As estimativas de herdabilidade variaram de 0,53 a 0,56, e de 0,55 a 0,57 para os parâmetros peso a maturidade (a) e taxa de maturidade (k), respectivamente. A correlação genética entre os parâmetros “a” e “k” foi alta e positiva (0,78). As porcentagens das variâncias genéticas explicadas por cada SNP permitiram identificar as regiões cromossômicas mais relevantes para cada fenótipo (parâmetros da curva de crescimento). Foram identificados três SNPs relevantes (55840514 bp no SSC17, 55814469 bp no SSC17 e 76475804 bp no SSC X) que influenciaram, simultaneamente, os parâmetros “a” e “k”. Também foram reportados três SNPs afetando apenas “a” (292758 bp no SSC1, 67319 bp no SSC8 e 50290193 bp no SSC17) localizados em regiões cromossômicas que ainda não foram previamente descritos como QTL para características de crescimento em suínos. A modelagem utilizada foi efetiva, e resultou na identificação de marcadores SNPs

localizados em regiões cromossômicas específicas que apresentam potencial para serem exploradas em programas de melhoramento via seleção assistida por marcadores. No segundo capítulo, comparou-se as metodologias baseadas na utilização de dados censurados de idade ao primeiro parto (IPP) em bovinos Brahman por meio da abordagem Bayesiana. Os dados foram cedidos pela Associação Brasileira dos Criadores de Zebu (ABCZ). Registros censurados foram definidos como valores de IPP que extrapolaram o intervalo entre 731 e 1824 dias. Os registros de IPP (no total de 53.703 informações) foram analisados por meio de quatro diferentes metodologias: método linear convencional (LM); de simulação (SM); de penalidade (PM) e modelos bicaracterístico limiar-linear (TLcens). Os componentes de variância genética aditiva estimados para os métodos LM e PM foram similares. As estimativas de herdabilidade para IPP variaram de 0,09 (TLcens) à 0,20 (LM). De forma geral, as correlações entre os valores genéticos obtidos por meio das diferentes metodologias e a porcentagem de animais selecionados em comum variaram de 0,82 (LM x SM) à 0,97 (LM x PM), e de 32,70% (SM x TLcens) à 89,12% (LM x PM), respectivamente, indicando reordenamento moderado entre os animais. As comparações realizadas via validação cruzada indicaram o método LM como a melhor opção para predição dos valores genéticos dos animais para a característica IPP na população estudada. No terceiro capítulo, foram estimados os parâmetros genéticos para características de escores visuais de estrutura (S), precocidade (P), musculosidade (M) e reprodutiva (idade ao primeiro parto - IPP) em bovinos da raça Brahman utilizando modelos Bayesianos multicaracterístico completo e bicaracterísticos. As estimativas de herdabilidade utilizando o modelo bicaracterístico foram 0,59 (S), 0,44 (P), 0,38 (M) e 0,20 (IPP), e utilizando o modelo multicaracterístico completo foram 0,60 (S), 0,44 (P), 0,40 (M) e 0,20 (IPP). As correlações genéticas foram 0,57 entre estrutura e precocidade, 0,56 entre estrutura e musculosidade e 0,82 entre precocidade e musculosidade no modelo

multicarcterística completo. As correlações genéticas entre os escores visuais e IPP foram de moderada magnitude e negativas (-0,29, -0,24 e -0,31 para S, P e M utilizando o modelo de bicaracterístico) e (-0,29, -0,22 e -0,29 para S, P e M utilizando o modelo multicaracterístico completo). Os resultados indicam que os escores visuais podem ser utilizados como critérios de seleção em programas de melhoramento de bovinos Brahman e que essas características apresentam correlação genética favorável com a idade no primeiro parto.

GENERAL INTRODUCTION

Growth curves have long been used to describe the growth process of animals and several mathematical models have been developed to predict the growth rate of animals through various life cycle stages (Koivula et al., 2008). In general, growth curves have been studied through several nonlinear functions such as Logistic, von Bertalanffy and Gompertz (Koivula et al., 2008; Cai et al., 2012; Silva et al., 2013). These functions present a reduced number of parameters with biological interpretation (for instance, mature weight and maturing rate). In pigs, most of the genome association studies of growth assume the body weight at specific ages as phenotypes. However, it may be extended for a more general context by considering the whole weight-age data under a growth curve approach. In addition, the use of genome-wide association studies (GWAS) can be useful to search for chromosomal regions that can help to explain the genetic architecture of complex trait.

In general, genetic analysis for growth curves have been based on a two-step procedure. In the first step, a growth curve is fitted separately to the data of each individual animal, afterwards, a mixed model analysis is applied to obtain (co)variance components. In this second step the estimates of production function parameters from the previous step are taken as records (Varona et al., 1999). However, Varona et al. (1997, 1998) described a Bayesian procedure which allows the particular parameters of any production function to be estimated jointly and the (co)variance components between them. Under this approach, adjustment errors are discarded and all the available information is then used for the genetic prediction of individual growth curves (Varona et al., 1997; Blasco et al., 2003; Forni et al., 2009).

In beef cattle, the reproductive traits are important for the production system because herds which has high fertility will have greater availability of animals, whether

for sale or for selection, allowing more selective intensity and, consequently, higher genetic progress and increased profitability. In this context, the need arises of pregnant female earlier and which have lower calving interval (Boligon and Albuquerque, 2010). Heritability estimates of reproductive traits in beef cattle described in the literature are considered of low magnitude (Forni and Albuquerque, 2006; Boligon et al., 2007; Baldi et al., 2008). However, moderate heritability estimates have been found for age at first calving (Mercadante et al., 2000; Azevêdo et al., 2006; Faria et al., 2007).

For Brahman cattle, reports on heritability estimates for different traits are scarce, particularly in Brazil (Faria et al., 2011). Furthermore, *Bos indicus* cattle, such as Brahman cattle, are reportedly older at puberty when compared with most *Bos taurus* breeds (Lunstra and Cundiff, 2003; Lopez et al., 2006). These cattle have been widely used in Australia and also have been used in Brazil, due to traits such as resistance, fertility and calving facility, weight gain, weaning, longevity, finishing and crosses for meat, and they have been one of the Zebu breed more evenly distributed throughout the world being qualified as one of the best selection options for beef cattle (Lunstra and Cundiff, 2003; Lopez et al., 2006; Johnston et al., 2009; Faria et al., 2011; Bertipaglia et al., 2012; Fortes et al., 2012). The knowledge of genetic parameters and correlated responses of beef traits is important for designing specific breeding programs and conduct mating plans. It is crucial to choose selection criteria genetically and favorably related with the selection objectives of the herds and identify genetically superior animals.

When considering the traits measured in females, the age at first calving (AFC) is the most used to evaluate fertility, since it is observed relatively early, it can be easily obtained and it is expressed in the majority of the breeding females (Boligon and Albuquerque, 2010). However, selecting females directly for the lowest AFC is not simple, since reproductive traits generally have low heritability, between 0.14 and 0.19

(Pereira et al., 2002; Boligon et al., 2007). In addition, some farmers delay the entry of females into reproduction by determining age or weight for the beginning reproductive life, which makes it difficult to identify sexually precocious females. In this sense, it is necessary to verify if the traits indicating fertility and sexual precocity are associated with the visual scores (body estructure, precocity and muscularity) that are currently considered in the selection programs (Boligon and Alguquerque, 2010). Since, the advantage of including visual scores in breeding programs is that a large number of animals can be evaluated without being subjected to the stress of measurements, a fact that makes the process faster and more economically feasible (Jorge Júnior et al., 2001, 2004). However, studies that correlate visual scores with reproductive performance in zebu animals are few in the literature (Faria et al., 2009; Bertipaglia et al., 2012).

Furthermore, female fertility had been neglected in breeding programs for decades (Garrick and Ruvinsky, 1999). In Zebu breeds, the inclusion of reproductive traits on selection criteria is fundamental due to predominant poor fertility, characterized by a long postpartum anestrous period (Nava-Trujillo et al., 2010). The attempt is to select for sexual precocity in one of the most important fertility trait, the age at first calving (AFC). Lower AFC values are associated with heifer precocity, high lifetime productivity, high number of calves in a same time period and allows higher genetic progress rate. Despite easiness of routine recording, AFC data is not always appropriate to be used in genetic evaluation because of recording mistakes and non-occurrence or delay in communication of the calving at the moment of genetic evaluation and animals without AFC phenotype are ignored in routine genetic evaluation. However, their records can be reconsidered as censored observations (Tarrés et al., 2006). The analysis of censored traits requires non-usual methodologies to be implemented in current genetic evaluation programs.

Some methods have been proposed to deal with censored traits in genetic evaluations. One is based on simulation of censored records from positive truncated normal distributions taking into account the estimated effects of the model (Donoghue et al., 2004; Korsgaard et al., 2003). Another one is the penalty methodology proposed by Johnston and Bunter (1996), which consists to impute information by adding a constant (number of days) to real data. For AFC, 21 days are often included based on the assumption that the heifer should be fertile in the subsequent estrous cycle. The linear-threshold bivariate analysis considers the censoring status (threshold binary trait) as an additional trait to improve the accuracy of genetic parameter estimates.

We proposed a genome association study for pig growth curves based on Bayesian hierarchical framework considering different sets of SNP markers and pedigree and to identify possible chromosome regions affecting the growth curve parameters. And we aimed to apply and compare different methodologies that deal with censored AFC records and to estimate genetic parameters between AFC and visual scores by linear bitrait and full multitrait models that were previously determined in the previous step, under a Bayesian framework in Brazilian Brahman cattle by accessing predictive performance via cross-validation.

In the first chapter, we proposed a genome association study for pig growth curves based on Bayesian hierarchical framework considering different sets of SNP markers and pedigree and we aimed also to identify possible chromosome regions affecting the growth curve parameters. In the second chapter, we aimed to compare the mentioned methods under a Bayesian framework for genetic evaluation of AFC in Brazilian Brahman cattle. And in the last chapter, we aimed to estimate genetic parameters between age at first calving and visual scores (body structure, precocity and muscularity) by using linear bitrait and full multitrait Bayesian models in Brazilian Brahman cattle.

LITERATURE CITED

AZEVÊDO, D.M.M.R.; FILHO, R.M.; LÔBO, R. N.B.; et al. Reproductive performance of Nelore cows in the North and Northeast Regions of Brazil. **Revista Brasileira de Zootecnia**, v.35, n.3, p.988-996, 2006.

BALDI, F.; ALENCAR, M. M.; FREITAS, A. R.; et al. Genetic parameters for body size, condition score, reproductive and longevity traits in females of the Canchim breed. **Revista Brasileira de Zootecnia**, v.37, n.2, p.247-253, 2008.

BERTIPAGLIA, T.S.; CARREÑO, L.O.D.; MACHADO, C.H.C.; et al. Estimates of genetic parameters for visual scores and their correlation with production and reproductive traits in Brahman cattle. **Revista Brasileira de Zootecnia**, v.41, n.6, p.1407-1411, 2012.

BLASCO, A.; PILES, M.; VARONA, L. A Bayesian analysis of the effect of selection for growth rate on growth curves in rabbits. **Genetics Selection Evolution**, 35, 21-41, 2003.

BOLIGON, A.A.; RORATO, P.R.N.; ALBUQUERQUE, L.G. Genetic correlations between male scrotal circumference and female productive and reproductive traits in Nellore cattle. **Revista Brasileira de Zootecnia**, v.36, n.3, p.565-571, 2007.

BOLIGON, A.A.; ALBUQUERQUE, L.G. Genetic correlations between visual scores and reproductive traits in Nelore cattle using Bayesian inference. **Pesquisa Agropecuária Brasileira**, v.45, n.12, p.1412-1418, 2010.

CAI, W., KAISER, M.S., DEKKERS, J.C.M. Bayesian analysis of the effect of selection for residual feed intake on growth and feed intake curves in Yorkshire swine. **Journal of Animal Science**, 90, 127–141, 2012.

DONOGHUE, K.A.; REKAYA, R.; BERTRAND, J.K. Comparison of methods for handling censored records in beef fertility data: Simulation study. **Journal Animal Science**, v.82, p.351-356, 2004.

FARIA, C.U.; MAGNABOSCO, C.U.; REYES, A.; et al. Bayesian inference on field data for genetics parameters for some reproductive traits and related traits of Nellore cattle. **Genetics and Molecular Biology**, v.30, n.2, p.343-348, 2007.

FARIA, C.U.; MAGNABOSCO, C.U.; ALBUQUERQUE, L.G.; et al. Bayesian analysis in the estimation of genetic correlations between visual scores and reproductive traits in Nelore cattle using linear-threshold models. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v.61, n.4, p.949-958, 2009.

FARIA, L.C.; QUEIROZ, S.A.; VOZZI, P.A.; et al. Quantitative genetic study on growth traits of Brahman cattle in Brazil. **ARS Veterinaria**, v.27, n.1, p.030-035, 2011.

FORNI, S. and ALBUQUERQUE, L.G. Environmental effects and genetic parameters estimates for days to calving in Nelore cattle. **Revista Brasileira de Zootecnia**, v.35, n.4, p.1239-1335, 2006.

FORNI, S.; PILES, M.; BLASCO, A.; et al. Comparison of different nonlinear functions to describe Nelore cattle growth. **Journal of Animal Science**, 87, 496–506, 2009.

FORTES, M.R.S.; LEHNERT, S.A.; BOLORMAA, S.; et al. Finding genes for economically important traits: Brahman cattle puberty. **Animal Production Science**, v.52, p.143-150, 2012.

GARRICK, D.; RUVINSKY, A. **The genetics of cattle.** 2nd edition, 641 pg., CABI: Oxfordshire, UK, 1999.

JOHNSTON, D. J.; BUNTER, K. L. Days to calving in Angus cattle: Genetic and environmental effects, and covariances with other traits. **Livestock Production Science**, v.45, n.1, p.13-22, 1996.

JOHNSTON, D. J.; BARWICK, S.A.; CORBET, N.J.; et al. Genetics of heifer puberty in two tropical beef genotypes in northern Australia and associations with heiferand steer-production traits. **Animal Production Science**, v.49, p.399-412, 2009.

JORGE JÚNIOR, J.; DIAS, L.T.; ALBUQUERQUE, L.G. Fatores de correção de escores visuais de conformação, precocidade e musculatura, à desmama, para idade da vaca ao parto, data juliana de nascimento e idade à desmama em bovinos da raça Nelore. **Revista Brasileira de Zootecnia**, v.33, n.6, p.2044-2053, 2004 (supl. 2).

JORGE JÚNIOR, J.; PITA, F.V.C.; FRIES, L.A. et al. Influência de alguns fatores de ambiente sobre os escores de conformação, precocidade e musculatura à desmama em um rebanho da raça Nelore. **Revista Brasileira de Zootecnia**, v.30, n.6, p.1697-1703, 2001.

KOIVULA, M., SEVÓN-AIMONEN, M.-L., STRANDÉN, I.; et al. Genetic (co)variances and breeding value estimation of Gompertz growth curve parameters in Finnish Yorkshire boars, gilts and barrows**. Journal of Animal Breeding and Genetics**, 125, 168–175, 2008.

KORSGAARD, I.R.; LUND, M.S.; SORENSEN, D.; et al. Multivariate Bayesian analysis of Gaussian, right censored Gaussian, ordered categorical and binary traits using Gibbs sampling. **Genetics Selection Evolution**, v.35, p.159-183, 2003.

LOPEZ, R.; THOMAS, M.G.; HALLFORD, P.D.M.; et al. Case study: Metabolic hormone profiles and evaluation of associations of metabolic hormones with body fat and reproductive characteristics of Angus, Brangus, and Brahman heifers. **The Professional Animal Scientist**, v.22, p.273-282, 2006.

LUNSTRA, D.D.; CUNDIFF, L.V. Growth and pubertal development in Brahman-, Boran-, Tuli-,Belgian Blue-, Hereford- and Angus-sired F1 bulls. **Journal Animal Science**, v.81, p.1414-1426, 2003.

MERCADANTE, M.E.Z.; LÔBO, R.B.; OLIVEIRA, H.N. Estimates of (Co)Variances among Reproductive and Growth Traits in Female Nellore Cattle. **Revista Brasileira de Zootecnia**, v.29, n.4, p.997-1004, 2000.

NAVA-TRUJILLO, H.; SOTO-BELLOSO, E.; HOET, A.E. Effects of clinical mastitis from calving to first service on reproductive performance in dual-purpose cows. **Animal Reproduction Science**, v.121, n.1, p.12-16, 2010.

PEREIRA, E.; ELER, J.P.; FERRAZ, J.B.S. Genetic analysis of reproductive traits in Nelore cattle. **Pesquisa Agropecuária Brasileira**, v.37, p.703-708, 2002.

SILVA, F.F., RESENDE, M.D.V., ROCHA, G.S.; et al. Genomic growth curves of an outbred pig population. **Genetics and Molecular Biology**, 36, 4, 520-527, 2013.

TARRÉS, J.; PIEDRAFITA, J.; DUCROCQ, V. Validation of an approximate approach to compute genetic correlations between longevity and linear traits. **Genetics Selection Evolution**, v.38, p.65-83, 2006.

VARONA, L., MORENO, C., GARCÍA CORTÉS, L.A. et al. Two-step versus joint analysis of Von Bertalanffy function. **Journal of Animal Breeding and Genetics**, 116, 331–338, 1999.

VARONA, L.; MORENO, C.; GARCIA CORTES, L.A.; et al. Bayesian Analysis of Wood's Lactation Curve for Spanish Dairy Cows. **Journal of Dairy Science**, 81, 1469–1478, 1998.

VARONA, L.; MORENO, C.; GARCIA CORTES, L.A.; et al. Multiple trait genetic analysis of underlying biological variables of production functions. **Livestock Production Science**, 47, 201-209, 1997.

Chapter 1

Bayesian analysis of pig growth curves combining pedigree and genomic information

ABSTRACT: We proposed a genome association study for pig growth curves based on Bayesian hierarchical framework considering different sets of SNP markers and pedigree. Additionally, we aimed also to identify possible chromosome regions affecting the growth curve parameters using empirical weight-age data from an outbred F2 (Brazilian Piau vs commercial) pig population. Under the proposed hierarchical approach, individual growth trajectories were modeled by the nonlinear Gompertz function, so that the parameter estimates were considered to be affected by additive polygenic, systematic and SNP markers effects. The model assuming jointly pedigree and SNP markers presented the best fit based on Deviance Information Criterion. Heritability estimates ranged from 0.53 to 0.56 and from 0.55 to 0.57, respectively for the parameters mature weight (a) and maturing rate (k). Additionally, we found high and positive genetic correlation (0.78) between “a” and "k". The percentages of the genetic variances explained by each SNP allowed identifying the most relevant chromosome regions for each phenotype (growth curve parameters). The majority of these regions were closed to QTL regions previously reported for growth traits. However, we identified three relevant SNPs (55840514 bp at SSC17, 55814469 at SSC17 and 76475804 at SSC X) affecting "a" and "k" simultaneously, and three SNPs affecting only "a" (292758 bp at SSC1, 67319 bp at SSC8 and 50290193 bp at SSC17), that are located in regions not previously described as QTL for growth traits in pigs.

Keywords: Hierarchical nonlinear model, Gompertz, SNP markers.

INTRODUCTION

Most of the genome association studies of pig growth assume the body weight at specific ages as phenotypes. However, it may be extended for a more general context by considering the whole weight-age data under a growth curve approach. In general, pig growth curves have been studied through several nonlinear functions such as Logistic, von Bertalanffy and Gompertz (Koivula et al., 2008; Cai et al., 2012; Silva et al., 2013). These functions present a reduced number of parameters with biological interpretation (for instance, mature weight and maturing rate). Thus, breeding goals can be defined aiming to change the shape of the growth curves by treating these parameter estimates as phenotypic observations in statistical genetic models.

Traditionally, genetic analysis of growth curves considering only pedigree information has been performed in two distinct steps. First, the growth curve parameters are estimated for each animal; and, second, (co)variance components, genetic and environmental effects are estimated on them. This approach ignores the adjustment errors and does not allow estimating growth curve parameters for individuals with a scarce amount of records (Varona et al., 1999). In this context, hierarchical Bayesian models for growth curves were proposed by calculating joint posterior distributions for the curve parameters, (co)variance components, and systematic and genetic effects. Under this approach, adjustment errors are discarded and all the available information is then used for the genetic prediction of individual growth curves (Varona et al., 1997; Blasco et al., 2003; Forni et al., 2009).

Ibáñez-Escriche and Blasco (2011) generalized the hierarchical Bayesian models for growth curves under a genome wide selection approach considering a simulated population. These procedures provide information on location of specific genome regions affecting growth curve components, that may lead to new insights about marker assisted

selection in pig breeding approaching desirable genetic changes on growth curves. However, generalization for genome association studies have been under exploited in literature, especially for real data.

In this context, we proposed a genome association study for pig growth curves based on Bayesian hierarchical framework considering different sets of SNP markers and pedigree. Additionally, we aimed also to identify possible chromosome regions affecting the growth curve parameters.

MATERIALS AND METHODS

Experimental population and phenotypic data

The phenotypic data was obtained from the Pig Breeding Farm of the Department of Animal Science, Universidade Federal de Viçosa (UFV), MG, Brazil. A three-generation resource population was created and managed as described by Hidalgo et al. (2013) and Verardo et al. (2015). Briefly, two naturalized Piau breed grandsires were mated with 18 granddams from a commercial line composed of Large White, Landrace and Pietrain breeds, to produce the F1 generation from which 11 F1 sires and 54 F1 dams were selected. These F1 individuals were mated to produce the F2 population, of which 345 animals were weighed at birth and at 21, 42, 63, 77, 105 and 150 days of age.

DNA extraction, genotyping and SNP quality control

DNA was extracted at the Animal Biotechnology Lab from Animal Science Department of Universidade Federal de Viçosa. Genomic DNA was extracted from white cells of parental, F1 and F2 animals, more details can be found in Band et al. (2005). The low-density customized SNPChip with 384 markers was based on the Illumina Porcine SNP60 BeadChip (San Diego, CA, USA, Ramos et al., 2009).

These SNPs were selected according to QTL positions previously identified on this population using meta-analyses (Silva et al., 2011) and fine mapping (Hidalgo et al., 2013, Verardo et al., 2015). Thus, although a small number of markers have been used, the customized SNPchip based on previous identified QTL positions ensures an appropriate coverage of the relevant genome regions in this population. From the total of 384 markers, 66 SNPs were discarded for no amplification, and from the remaining 318 SNPs, 81 were discarded due to a minor allele frequency (MAF) < 0.05. Thus, 237 SNPs markers were used and distributed as follows: SSC1 (56), SSC4 (54), SSC7 (59), SSC8 (30), SSC17 (25) and SSCX (13), being the average distance within each chromosome, respectively, 5.17, 2.37, 2.25, 3.93, 2.68 and 11.00 Mb.

The model

A hierarchical Bayesian model was applied to analyze individual pig growth curves based on nonlinear Gompertz function, whose parameters were modeled by a multitrait linear model including additive polygenic, SNP marker and systematic effects.

In the first stage, it was considered the following Gompertz growth model:

$$y_{ij} = a_i \exp(-b_i \exp(-k_i t_{ij})) + \varepsilon_{ij}, \quad (1)$$

where y_{ij} is the observed body weight of individual i at age j, a_i represents the mature weight, b_i is a time scale parameter (it does not have biological interpretation), k_i is the maturing rate, t_{ij} is the day in which the body weight were measured, and ε_{ij} is the residual term, considered to be independent and normally distributed among individuals. The following distribution was assumed for the weight-age data in this first stage:

$$f(y_{ij} \mid a_i, b_i, k_i, \sigma^2_{\varepsilon j}) \sim N(a_i \exp(-b_i \exp(-k_i t_{ij})), \sigma^2_{\varepsilon j})$$

The standard deviation ($\sigma_{\varepsilon j}$) for the residual term in (2) was considered as a linear function of two parameters (r_a and r_b) aiming to model its trajectory over time (*i.e.*, to consider residual heterogeneity of variance):

$$\sigma_{\varepsilon j} = r_a + r_b t_{ij}. \quad (2)$$

In the second stage, additive polygenic, systematic and SNP marker effects were estimated under a multitrait linear model considering the parameter estimates from the first stage as phenotypic observations. Three alternative models, characterized by the inclusion of different genetics effects in addition to the systematic effects, were proposed. The first one assumed the additive polygenic effects (Pedigree) – M1 (3); the second one the SNP genotypes effects (Markers) – M2 (4); and the third one considered both previously mentioned effects (Pedigree and markers) – M3 (5). These models are given respectively by:

$$\boldsymbol{\theta} = \mathbf{X\beta} + \mathbf{Zu} + \mathbf{e}, \quad (3)$$

$$\boldsymbol{\theta} = \mathbf{X\beta} + \mathbf{Mc} + \mathbf{e}, \quad (4)$$

$$\boldsymbol{\theta} = \mathbf{X\beta} + \mathbf{Mc} + \mathbf{Zu} + \mathbf{e}, \quad (5)$$

where $\boldsymbol{\theta}$ is a vector containing the estimates of the parameter "a", "b", and "k" for all individuals, $\boldsymbol{\theta}' = [\mathbf{a}, \mathbf{b}, \mathbf{k}]' = [a_1, a_2, ..., a_n, b_1, b_2, ...b_n, k_1, k_2, ...k_n]'$; $\boldsymbol{\beta}$ is the vector of systematic effects (intercept and 19 contemporary groups given by the combination of sex, batch and halothane gene genotype (Band et al., 2005)), $\boldsymbol{\beta} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\beta \otimes \mathbf{I})$, being $\boldsymbol{\Sigma}_\beta$ a known diagonal matrix with values 1e+10 (large variances) to represent vague prior knowledge; $\mathbf{u} = (u_{a1}, u_{a2}, ..., u_{an}, u_{b1}, u_{b2}, ..., u_{bn}, u_{k1}, u_{k2}, ..., u_{kn})$ is the vector of additive polygenic effects, assumed as: $\mathbf{u} \mid \boldsymbol{\Sigma}_g, \mathbf{A} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_g \otimes \mathbf{A})$, n is the total number of individuals and in this study was too the number of records within individual, $\mathbf{A}$ is the additive relationship matrix among the animals and $\boldsymbol{\Sigma}_g$ is the additive genetic (co)variance matrix; $\mathbf{c} = (c_{a1}, c_{a2}, ..., c_{am}, c_{b1}, c_{b2}, ..., c_{bm}, c_{k1}, c_{k2}, ..., c_{km})$ is the vector of random SNP effects with known incidence matrix $\mathbf{M}$ with (345x3) rows and

(237x3) columns of SNP genotypes (coded as AA, AB, or BB), assumed as $\mathbf{c} \mid \boldsymbol{\Sigma}_c \sim N(\mathbf{0}, \boldsymbol{\Sigma}_c \otimes \mathbf{I})$, where $\mathbf{I}$ and $\boldsymbol{\Sigma}_c$ are, respectively, an identity and SNP markers (co)variance matrices. The $\mathbf{X}$ and $\mathbf{Z}$ are the incidence matrices corresponding to systematic and additive polygenic effects, respectively; and $\mathbf{e} = (e_{a1}, e_{a2}, ..., e_{an}, e_{b1}, e_{b2}, ..., e_{bn}, e_{k1}, e_{k2}, ..., e_{kn})$ is the residuals vector, assumed as $\mathbf{e} \mid \boldsymbol{\Sigma}_e \sim N(\mathbf{0}, \boldsymbol{\Sigma}_e \otimes \mathbf{I})$, where $\mathbf{I}$ and $\boldsymbol{\Sigma}_e$ are, respectively, an identity and residual (co)variance matrices.

Based on the results of M2 (Eq. [4]), we defined two new models (M4 and M5) that included an additive polygenic effect and reduced sets of SNP markers. In M4 (Eq. [6]), we selected the markers that explain at least 0.5% on the SNP variance for the three parameters simultaneously. Further, in M5 (Eq. [7]), only markers explaining 0.5% of the variance for at least one of the considered parameters were selected. The incidence matrix for these two different set of SNPs were given by M4 and M5 respectively.

$$\boldsymbol{\theta} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{M}_4\mathbf{c} + \mathbf{e}, \quad (6)$$

$$\boldsymbol{\theta} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{M}_5\mathbf{c} + \mathbf{e}, \quad (7)$$

The inference

The joint posterior distribution for individual growth curve parameters, their variance components, and systematics, additive polygenic and SNP effects was accessed under a hierarchical framework following the Bayes theorem:

$$f(\boldsymbol{\theta}, \sigma_{\varepsilon j}, \boldsymbol{\beta}, \mathbf{u}, \mathbf{c}, \boldsymbol{\Sigma}_c, \boldsymbol{\Sigma}_g, \boldsymbol{\Sigma}_e \mid \mathbf{y}) \propto f(\mathbf{y} \mid \boldsymbol{\theta}, \sigma_{\varepsilon j})$$
$$f(\boldsymbol{\theta} \mid \boldsymbol{\beta}, \mathbf{u}, \mathbf{c}, \boldsymbol{\Sigma}_c, \boldsymbol{\Sigma}_g, \boldsymbol{\Sigma}_e)\ f(\sigma_{\varepsilon j})\, f(\boldsymbol{\beta})\ f(\mathbf{u} \mid \boldsymbol{\Sigma}_g)$$
$$f(\boldsymbol{\Sigma}_g) f(\mathbf{c} \mid \boldsymbol{\Sigma}_c)\, f(\boldsymbol{\Sigma}_c)\ f(\boldsymbol{\Sigma}_e)$$

Assuming independence among individuals, the conditional distribution of data $\mathbf{y}$, given the growth curve parameters, was a product of normal distributions:

$$f(\mathbf{y} \mid \boldsymbol{\theta}, \sigma_{\varepsilon j}) = \prod_{i=1}^{N}\prod_{j=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2_{\varepsilon j}}} \exp\left\{ - \frac{[y_{ij} - (a_i \exp^{(-b_i)\exp^{(-k_i t_{ij})}})]^2}{2\sigma^2_{\varepsilon j}} \right\}, \quad (8)$$

where N is the total number of individuals; n the number of records within individual; a_i, b_i and k_i are the parameters of the Gompertz growth function for the animal i; and t_{ij} the age (days) at time j.

The prior distribution for the growth curve parameters given the additive polygenic, systematic and SNP effects, as well as the (co)variance components, was assumed as a multivariate Gaussian distribution given by:

$$f(\boldsymbol{\theta} \mid \boldsymbol{\beta}, \mathbf{u}, \boldsymbol{\Sigma}_g, \mathbf{c}, \boldsymbol{\Sigma}_c, \boldsymbol{\Sigma}_e) = \mid \boldsymbol{\Sigma}_e \mid^{-N/2} \exp\left[-\frac{1}{2}(\boldsymbol{\theta} - \mathbf{X}\beta - \mathbf{Z}\mathbf{u} - \mathbf{M}\mathbf{c})'(\boldsymbol{\Sigma}_e \otimes \mathbf{I})^{-1}(\boldsymbol{\theta} - \mathbf{X}\beta - \mathbf{Z}\mathbf{u} - \mathbf{M}\mathbf{c}) \right] \quad (9)$$

where $\boldsymbol{\theta}$ is the vector containing the parameters "a", "b" and "k"; $\boldsymbol{\Sigma}_g$ is the additive polygenic genetic (co)variance matrix; $\boldsymbol{\Sigma}_c$ is the SNP markers (co)variance matrix; $\boldsymbol{\Sigma}_e$ the residual (co)variance matrix between the parameters "a", "b", and "k"; and $\mathbf{I}$ is an identity matrix.

Following the proposed hierarchical approach, Gaussian prior distributions were assumed for the systematics, additive polygenic and SNP effects:

$$f(\boldsymbol{\beta} \mid \mathbf{s}, \mathbf{V}) \propto \mid \mathbf{V} \mid^{-1/2} \exp\left[-\frac{1}{2}(\boldsymbol{\beta} - \mathbf{s})'\mathbf{V}^{-1}(\boldsymbol{\beta} - \mathbf{s}) \right],$$

$$f(\mathbf{u} \mid \boldsymbol{\Sigma}_g, \mathbf{A}) \propto \mid \boldsymbol{\Sigma}_g \mid^{-N_A/2} \exp\left[-\frac{1}{2}\mathbf{u}'(\boldsymbol{\Sigma}_g \otimes \mathbf{A})^{-1}\mathbf{u} \right], \text{ and}$$

$$f(\mathbf{c} \mid \boldsymbol{\Sigma}_c) \propto \mid \boldsymbol{\Sigma}_c \mid^{-m/2} \exp\left[-\frac{1}{2}\mathbf{c}'(\boldsymbol{\Sigma}_c \otimes \mathbf{I})^{-1}\mathbf{c} \right],$$

where $\mathbf{s}$ and $\mathbf{V}$ are subjective means and (co)variances for the prior beliefs about the systematic effects, N_A is the number of animals in the genealogy, $\mathbf{I}$ is an identity matrix

of order m, and **A** is the numerator relationship matrix. Bounded uniform distributions were assumed for $\sigma_{\varepsilon j}$ and (co)variance matrices ($\boldsymbol{\Sigma}_g$, $\boldsymbol{\Sigma}_c$ and $\boldsymbol{\Sigma}_e$) (Varona et al., 1998; Forni et al., 2007).

The sampling methods require random independent draws from the conditional posterior distribution for each parameter. Thus, if θ_{ik} is the k^{th} growth curve parameter for the ith animal, θ_{-ik} are the other parameters for the i^{th} animal and all parameters for all other animals. Thus we have:

$$f(\theta_{ik} \mid \theta_{-ik}, \boldsymbol{\beta}, \mathbf{u}, \mathbf{c}, \boldsymbol{\Sigma}_c, \boldsymbol{\Sigma}_g, \boldsymbol{\Sigma}_e, \sigma_{\varepsilon j}, \mathbf{y}) \propto$$
$$f(\theta_{ik} \mid \theta_{-ik}, \sigma_{\varepsilon j}, \mathbf{y})\, f(\theta_{ik} \mid \theta_{-ik}, \boldsymbol{\beta}, \mathbf{u}, \mathbf{c}, \boldsymbol{\Sigma}_c, \boldsymbol{\Sigma}_g, \boldsymbol{\Sigma}_e)$$

The fully conditional distributions for all parameters of the hierarchical multistage models were derived according to Varona et al. (1999). In the present study, these distributions for growth curve parameters are the products of the conditional distribution of data (Eq.[8]) and the prior distributions of the growth curve parameters (Eq. [9]).

The fully conditional distribution of parameter "a" can be written as:

$$f(a_i \mid b_i, k_i, \boldsymbol{\beta}, \mathbf{u}, \boldsymbol{\Sigma}_g, \mathbf{c}, \boldsymbol{\Sigma}_c, \boldsymbol{\Sigma}_e, \sigma_{\varepsilon j}, \mathbf{y}) \propto$$
$$f(a_i \mid b_i, k_i, \sigma_{\varepsilon j}, y_i)\; f(a_i \mid b_i, k_i, \boldsymbol{\beta}, \mathbf{u}, \boldsymbol{\Sigma}_g, \mathbf{c}, \boldsymbol{\Sigma}_c, \boldsymbol{\Sigma}_e),$$

where,

$$f(a_i \mid b_i, k_i, \sigma_{\varepsilon j}, y_i) \sim N\left[\frac{\sum_{j=1}^{n} y_{ij}(\exp^{(-b_i)}\exp^{(-k_i t_{ij})})}{\sum_{j=1}^{n}[(\exp^{(-b_i)}\exp^{(-k_i t_{ij})})]^2}, \frac{\sigma_{\varepsilon j}}{\sum_{j=1}^{n}[(\exp^{(-b_i)}\exp^{(-k_i t_{ij})})]^2}\right].$$

The fully conditional distribution of parameter "b" can be written as:

$$f(b_i \mid a_i, k_i, \boldsymbol{\beta}, \mathbf{u}, \boldsymbol{\Sigma}_g, \mathbf{c}, \boldsymbol{\Sigma}_c, \boldsymbol{\Sigma}_e, \sigma_{\varepsilon j}, \mathbf{y}) \propto$$
$$f(b_i \mid a_i, k_i, \sigma_{\varepsilon j}, y_i)\; f(b_i \mid a_i, k_i, \boldsymbol{\beta}, \mathbf{u}, \boldsymbol{\Sigma}_g, \mathbf{c}, \boldsymbol{\Sigma}_c, \boldsymbol{\Sigma}_e),$$

where,

$$f(b_i \mid a_i, k_i, \sigma_{\varepsilon j}, r_b, y_i) \propto \prod_{j=1}^{n} \exp\left[- \frac{[y_{ij} - (a_i \exp^{(-b_i)} \exp^{(-k_i t_{ij})})]^2}{2\sigma_{\varepsilon j}} \right].$$

The fully conditional distribution of parameter "k" can be written as:

$$f(k_i \mid a_i, b_i, \boldsymbol{\beta}, \mathbf{u}, \boldsymbol{\Sigma}_\mathbf{g}, \mathbf{c}, \boldsymbol{\Sigma}_\mathbf{c}, \boldsymbol{\Sigma}_\mathbf{e}, \sigma_{\varepsilon j}, \mathbf{y}) \propto$$
$$f(k_i \mid a_i, b_i, \sigma_{\varepsilon j}, y_i)\ f(k_i \mid a_i, b_i, \boldsymbol{\beta}, \mathbf{u}, \boldsymbol{\Sigma}_\mathbf{g}, \mathbf{c}, \boldsymbol{\Sigma}_\mathbf{c}, \boldsymbol{\Sigma}_\mathbf{e}),$$

where,

$$f(k_i \mid a_i, b_i, \sigma_{\varepsilon j}, y_i) \propto \prod_{j=1}^{n} \exp\left[- \frac{[y_{ij} - (a_i \exp^{(-b_i)} \exp^{(-k_i t_{ij})})]^2}{2\sigma_{\varepsilon j}} \right].$$

The parameter "a" could be sampled from a normal distribution by using Gibbs sampling algorithm, but the conditional posterior distribution for the parameters "b" and "k" did not have a closed form. In these cases, the Metropolis-Hastings algorithm with normal proposal distribution centered on the values of b_i and k_i sampled in the immediately previous iteration was used (Forni et al., 2007). The mixed model equations were constructed assuming as observed traits the growth curve parameters (θ) obtained from earlier steps:

$$\begin{bmatrix} \mathbf{X'R^{-1}X} + \mathbf{V^{-1}} & \mathbf{X'R^{-1}Z} & \mathbf{X'R^{-1}M} \\ \mathbf{Z'R^{-1}X} & \mathbf{Z'R^{-1}Z} + (\boldsymbol{\Sigma}_\mathbf{g} \otimes \mathbf{A})^{-1} & \mathbf{Z'R^{-1}M} \\ \mathbf{M'R^{-1}X} & \mathbf{M'R^{-1}Z} & \mathbf{M'R^{-1}M} + (\boldsymbol{\Sigma}_\mathbf{c} \otimes \mathbf{I})^{-1} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \\ \hat{\mathbf{c}} \end{bmatrix} = \begin{bmatrix} \mathbf{X'R^{-1}\theta} \\ \mathbf{Z'R^{-1}\theta} \\ \mathbf{M'R^{-1}\theta} \end{bmatrix} \quad (10)$$

where, $\mathbf{R} = \boldsymbol{\Sigma}_\mathbf{e} \otimes \mathbf{I}$.

The conditional posterior distributions for each location parameter β_l, u_i, and c_h were given by normal distributions defined by the coefficients and the right-hand side (RHS) of the mixed model equations (Eq. [10]):

$$f(\beta_l \mid \beta_{-l}, \boldsymbol{\theta}, \mathbf{u}, \boldsymbol{\Sigma_g}, \mathbf{c}, \boldsymbol{\Sigma_c}, \boldsymbol{\Sigma_e}, \sigma_{\varepsilon j}, \mathbf{y}) \sim N\left[\frac{RHS_i - \sum_{j \neq i} \lambda_{ij} t_j}{\lambda_{ii}}, \frac{1}{\lambda_{ii}}\right],$$

$$f(u_i \mid u_{-i}, \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\Sigma_g}, \mathbf{c}, \boldsymbol{\Sigma_c}, \boldsymbol{\Sigma_e}, \sigma_{\varepsilon j}, \mathbf{y}) \sim N\left[\frac{RHS_i - \sum_{j \neq i} \lambda_{ij} t_j}{\lambda_{ii}}, \frac{1}{\lambda_{ii}}\right], \text{ and}$$

$$f(c_h \mid c_{-h}, \boldsymbol{\theta}, \boldsymbol{\beta}, \mathbf{u}, \boldsymbol{\Sigma_g}, \boldsymbol{\Sigma_c}, \boldsymbol{\Sigma_e}, \sigma_{\varepsilon j}, \mathbf{y}) \sim N\left[\frac{RHS_i - \sum_{j \neq i} \lambda_{ij} t_j}{\lambda_{ii}}, \frac{1}{\lambda_{ii}}\right],$$

where β_{-l}, u_{-i}, and c_{-h}, are the vectors including the current values of these effects after discarding the i^{th} one, h represent the SNP markers (h = 1, 2, …, 237) and λ is the corresponding element from the coefficient matrix of the mixed models equations.

The conditional posterior distributions for the (co)variance matrices were the following inverted Wishart distributions:

$$f(\boldsymbol{\Sigma_g} \mid \boldsymbol{\theta}, \boldsymbol{\beta}, \mathbf{u}, \mathbf{c}, \boldsymbol{\Sigma_c}, \boldsymbol{\Sigma_e}, \sigma_{\varepsilon j}, \mathbf{y}) \sim IW[(\mathbf{u}'\mathbf{A}^{-1}\mathbf{u}), N_a - (n_p + 1)],$$
$$f(\boldsymbol{\Sigma_c} \mid \boldsymbol{\theta}, \boldsymbol{\beta}, \mathbf{u}, \boldsymbol{\Sigma_g}, \mathbf{c}, \boldsymbol{\Sigma_e}, \sigma_{\varepsilon j}, \mathbf{y}) \sim IW[(\mathbf{c}'\mathbf{c}), N_h - (n_p + 1)], \text{ and}$$
$$f(\boldsymbol{\Sigma_e} \mid \boldsymbol{\theta}, \boldsymbol{\beta}, \mathbf{u}, \boldsymbol{\Sigma_g}, \mathbf{c}, \boldsymbol{\Sigma_c}, \sigma_{\varepsilon j}, \mathbf{y}) \sim IW[(\mathbf{e}'\mathbf{e}), N - (n_p + 1)].$$

where n_p is the number of parameters assumed in the growth curve and N_h is the total number of SNP markers.

The conditional posterior distribution for the residual standard deviation ($\sigma_{\varepsilon j}$) have not closed form, thus the Metropolis-Hasting algorithm was used:

$$f(\sigma_{\varepsilon j} \mid \boldsymbol{\theta}, \boldsymbol{\beta}, \mathbf{u}, \boldsymbol{\Sigma_g}, \mathbf{c}, \boldsymbol{\Sigma_c}, \boldsymbol{\Sigma_e}, \mathbf{y})$$
$$\propto \prod_{j=1}^{n} \exp\left[-\frac{[y_{ij} - (a_i \exp^{(-b_i)} \exp^{(-k_i t_{ij})})]^2}{2\sigma_{\varepsilon j}}\right]$$

We applied a Metropolis-Hastings algorithm with a uniform proposal distribution centered at the current values b_i and k_i (as mentioned earlier). The choice of the limits for this distribution determines the acceptance rate. If the width of such an interval is too small, the proposed values will be closed to the current ones, the rejection rate will be low

but the process will move slowly throughout the parameter space. On the other hand, if it is too large, the proposed values are far away from the current ones and these results in a high rejection rate (Blasco et al., 2003). The above choice led to acceptance rates ranging between 50.74% and 52.45% (M1), 45.00% and 48.76% (M2), 48.80% and 50.67% (M3), 49.50% and 52.64% (M4), 46.66% and 51.28% (M5).

A total of 400,000 samples were generated, assuming a burn-in period and sampling interval (thin) of 100,000 and 10 iterations, respectively. The convergence of the MCMC chains was verified by graphical inspection and BOA (Smith, 2007) R software. Convergence was assessed using the Heidelberg e Welch (1983), Geweke (1992) and Raftery and Lewis (1992) methods.

Model testing

The goodness of fit analyzes for the considered models was based on the deviance information criterion (DIC) developed by Spiegelhalter et al. (2002): $DIC = D(\bar{\theta}) + 2p_D$, where $D(\bar{\theta})$ is a point estimate of the deviance obtained by replacing the parameters by their posterior means estimates in the likelihood function and *pD* is the effective number of parameters in the model, where $p_D = \bar{D}(\theta) - D(\bar{\theta})$. Models with smaller DIC should be preferred to models with larger DIC.

In addition to the goodness of fitting, we also calculated the predictive ability by cross-validation, which involved training one subset of the population (300 animals), and validating on the remaining individuals (45 animals). Here, we randomly split the data sets into two groups from the original data set (345 animals), these two subset were redefined 10 times, D1, D2, ..., D10. Finally, the average of the 10 correlation coefficients between the predicted and observed phenotypes was obtained.

The predicted weight ($\hat{y}_{ij}$) for animal i in time j based on Gompertz model was calculated as follows: $\hat{y}_i = \hat{a}_i \exp(-\hat{b}_i \exp(-\hat{k}_i t_{ij}))$, where $\hat{a}_i$, $\hat{b}_i$ and $\hat{k}_i$ are elements of the estimated vector $\hat{\theta}$ given by: $\hat{\boldsymbol{\theta}} = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{Z}\hat{\mathbf{u}} + \mathbf{M}\hat{\mathbf{c}}$. Thus, the solutions for these animals of the validation population were obtained based on the solutions of the training population animals.

The five tested models were applied to the 10 cross-validation replicates. In each replicate, systematic, genetic effects and SNPs markers effects were estimated and provided a phenotypic prediction for the masked animals. Finally, the predictive ability used to measure the efficiency of the models was given by the correlation between observed and predicted phenotypes from the validation population.

QTL identification

Based on SNP markers that were considered as relevant based on M2 (Eq. [4]) we verified the existence of QTL already described for growth traits by using the PigQTLdb tool (National Animal Genome Research Program, 2016). The traits which have been used in the PigQTLdb were body weights 34 weeks and at slaughter (related to parameter "a") and average daily gain (related to parameter "k").

RESULTS AND DISCUSSION

Model comparison

The M2 (Eq. [4]) was used only to estimate SNP variance and thus to fit reduced models M4 (Eq. [6]) and M5 (Eq. [7]) based on the results from M2, in this way their results are not shown or discussed.

Models were compared using the Deviance Information Criteria (DIC). The following results were obtained: model M3 (DIC=10572.73), model M1 (DIC=10745.97), model M4 (DIC=11083.42) and model M5 (DIC=10823.73) (Table 1).

The M3 that considered the pedigree of animals associated with the information of all SNP markers presented the best fit based on the lower DIC value. Thus, we can note that the selection of markers explaining the higher percentage of variance did not improve the goodness of fit, and that SNP markers explained a lower percentage of variance could be relevant to explain the observed covariance between relatives.

Table 1. Deviance Information Criterion (DIC) for models.

Models	DIC
Pedigree and markers (M3)[1]	10,572.73
Pedigree and markers (M4)[2]	11,083.42
Pedigree and markers (M5)[3]	10,823.73
Pedigree (M1)[4]	10,745.97

[1]complete model (pedigree and SNP markers information); [2]pedigree and SNP markers information (considers the SNP markers higher effect for commons parameters – a, b and k); [3]pedigree and SNP markers information (considers the SNP markers higher effect for different parameters – a, b and k); [4]Only pedigree information.

Correlation coefficient between all predicted and observed phenotypic values were also used to access the goodness of fit (Table 2). The same model indicated by DIC (Pedigree and markers) was considered as the best one since presented higher correlation coefficients at all ages, except at birth. The superiority of this model was remarkable at last age (150 days), which has the greater economic relevance because correspond to weight at slaughter.

This result is in agreement with de los Campos et al. (2009) that, analyzing a mice population, concluded that the model that considered the pedigree with SNP's markers effects showed the best goodness of fit.

Table 2. Correlation coefficient between predicted and observed values from models including different sources of genetic information (pedigree, and pedigree and markers) and their standard errors (SE) and below means of the correlations of the 10 groups of the cross-validation and their standard deviation (SD).

Dataset	Age	Pedigree and markers (M3)[1]	Pedigree and markers (M4)[2]	Pedigree and markers (M5)[3]	Pedigree (M1)[4]
Full data	1	0.450 [0.048]	0.450 [0.048]	0.454 [0.048]	0.440 [0.049]
	21	0.795 [0.033]	0.772 [0.034]	0.770 [0.034]	0.770 [0.035]
	42	0.864 [0.027]	0.844 [0.029]	0.845 [0.029]	0.844 [0.029]
	63	0.887 [0.025]	0.871 [0.026]	0.879 [0.025]	0.868 [0.027]
	77	0.932 [0.019]	0.919 [0.021]	0.927 [0.020]	0.916 [0.021]
	105	0.922 [0.021]	0.916 [0.021]	0.917 [0.021]	0.916 [0.021]
	150	0.814 [0.031]	0.760 [0.035]	0.797 [0.032]	0.737 [0.036]
Cross-validation	1	0.198 [0.031]	0.204 [0.030]	0.208 [0.032]	0.202 [0.030]
	21	0.224 [0.052]	0.238 [0.057]	0.280 [0.058]	0.262 [0.053]
	42	0.311 [0.032]	0.326 [0.030]	0.361 [0.033]	0.360 [0.030]
	63	0.376 [0.025]	0.390 [0.030]	0.386 [0.026]	0.393 [0.031]
	77	0.408 [0.042]	0.426 [0.042]	0.423 [0.041]	0.423 [0.044]
	105	0.393 [0.036]	0.410 [0.037]	0.407 [0.036]	0.459 [0.022]
	150	0.245 [0.028]	0.258 [0.027]	0.257 [0.022]	0.247 [0.023]

[1]complete model (pedigree and SNP markers information); [2]pedigree and SNP markers information (considers the SNP markers with higher effect for commons parameters – a, b and k); [3]pedigree and SNP markers information (considers the SNP markers with higher effect for different parameters – a, b and k); [4]Only pedigree information.

Predictive abilities were also calculated for all tested models in each evaluated ages (Table 2). All models presented lower predictive ability for initial phase of growth curve. Nevertheless, they were able to predict with higher predictive ability the weights at ages above 21 days. This lower predictive ability may be related to the fact that growth

models do not fit well to the initial age, once prenatal growth of animals is not measured. This period is known for the maximum rate of tissues and organs, so will determine traits such as weight birth of piglets and the consequences established during prenatal life will be continuous throughout the life of the animal (Fall et al., 2003; Foxcroft and Town, 2004).

A slight decrease in the correlations at later ages (105 to 150 days) was also observed. This decay can be explained because the last age considered in this study is not the age of maturity itself, but the age at slaughter (150 days - 65 kg), i.e., the animals continue to growing after this period, as can be seen in Peloso et al. (2010) that evaluated carcass traits and meat quality in five distinct genetic groups of pigs with animals up to 202 days old.

Correlations between the predicted and observed phenotypes at different ages in growth functions are seldom used in the literature. However, the importance of these results is remarkable because they are useful in identifying factors in animal production that may be modified in order to change growth trajectories.

Variances components and heritability

The marginal posterior densities of the variance components showed that a large part of adult weight variation is due to additive genetic effects (Table 3). Higher influence of additive genetic factors on these growth curves parameters was also reported by Koivula et al. (2008) and Cai et al. (2012) in pigs, and by Forni et al. (2007) in beef cattle. The "a" parameter of the growth curve can be used as a selection criterion to control adult body weight that increases when selecting for growth rate, especially in situations in which the slaughter weight is reached before the maturity, as occurred in this study. Also "k" parameter can be used as a selection criterion indicating the rate that animals approach the adult weight (Table 3).

Table 3. Features of the marginal posterior distributions of additive genetic and residual variance components and heritability and highest probability density (HPD) of growth curve parameters from models including different sources of genetic information (pedigree, markers, and pedigree and markers) for each parameter.

Traits	Pedigree and markers (M3)[1]	Pedigree and markers (M4)[2]	Pedigree and markers (M5)[3]	Pedigree (M1)[4]
Additive Genetic Variance and HPD				
a	30.42	24.55	19.23	23.95
	[0.28, 68.10]	[0.61, 53.00]	[0.47, 41.46]	[0.40, 51.43]
b	0.10	0.10	0.08	0.10
	[$5x10^{-2}$, $2x10^{-1}$]	[$6x10^{-2}$, $2x10^{-1}$]	[$4x10^{-2}$, $3x10^{-1}$]	[$6x10^{-6}$, $1.5x10^{-1}$]
k	$2x10^{-7}$	$1x10^{-7}$	$1x10^{-7}$	$1x10^{-7}$
	[$2x10^{-8}$, $3x10^{-7}$]	[$3x10^{-8}$, $2x10^{-7}$]	[$3x10^{-8}$, $2x10^{-7}$]	[$4x10^{-8}$, $2x10^{-7}$]
Residual Variance and HPD				
a	22.35	20.26	16.69	20.18
	[0.90, 48.16]	[1.05, 43.08]	[0.70, 36.86]	[0.78, 43.06]
b	0.03	0.03	0.03	0.02
	[$4x10^{-3}$, $5x10^{-2}$]	[$4x10^{-3}$, $5x10^{-2}$]	[$6x10^{-3}$, $6x10^{-2}$]	[$5x10^{-3}$, $5x10^{-2}$]
k	$1.3x10^{-7}$	$9x10^{-8}$	$9x10^{-8}$	$9x10^{-8}$
	[$2x10^{-8}$, $2x10^{-7}$]	[$3x10^{-8}$, $2x10^{-7}$]	[$3x10^{-8}$, $2x10^{-7}$]	[$3x10^{-8}$, $2x10^{-7}$]
Heritability and HPD				
a	0.56	0.54	0.53	0.53
	[0.15, 0.94]	[0.12, 0.93]	[0.12, 0.94]	[0.11, 0.95]
b	0.77	0.80	0.73	0.79
	[0.54, 0.97]	[0.60, 0.97]	[0.47, 0.95]	[0.60, 0.97]

k	0.57 [0.27, 0.87]	0.55 [0.25, 0.84]	0.55 [0.24, 0.84]	0.55 [0.24, 0.84]

[1]complete model (pedigree and SNP markers information); [2]pedigree and SNP markers information (considers the SNP markers with higher effect for commons parameters – a, b and k); [3]pedigree and SNP markers information (considers the SNP markers with higher effect for different parameters – a, b and k); [4]Only pedigree information.

We opted to show only the results obtained for parameters "a" and "k", since the parameter "b" has no biological interpretation. The heritability estimates (Table 3) indicate that "a" and "k" parameters can be an alternative for pig breeding programs that aim to produce animals with higher growth rate. Estimated heritability values of the present study were higher than those found by Koivula et al. (2008) (a=0.44, b=0.55 and k=0.31), working on Finnish Yorkshire pigs also using the Gompertz model. This may be due to the effect of the variance of performance given the production function parameters, which was not considered in the analyses of those authors, what causes the estimation noise to be absorbed by the residual variances (Varona et al., 1999).

Considering the model that showed the best goodness of fit to the data (Pedigree and markers - M3), genetic correlation between the growth curve parameters was obtained in order to assess whether the traits ("a" and "k") are relevant for a breeding program. Direct selection for a high value of "a" parameter will also imply in selection for higher value of "k" parameter (as indicated by the high and positive genetic correlation, 0.78, between the two parameters), and therefore the selection will result in animals more precocious (high maturation rate) and heavier animals. This high and positive correlation between the parameters "a" and "k" was also reported in others growth curve studies, e.g., Cai et al. (2012) in pigs, which have obtained the same value reported here, and Forni et al. (2007) in beef cattle.

The use of pedigree associated with SNP markers may capture extra sources of genetic variance compared with models based only on pedigree (de los Campos et al., 2009). Similarly, Calus and Veerkamp (2007) working with simulated data, concluded that the inclusion of polygenic effects associated with marker information improved the variance components estimation. Results similar to those reported by these authors, we could see in this study (Table 3), in which the genetic variance was higher in model M3 (Pedigree and markers) compared with model M1 (Pedigree).

Small number and sparse distribution of SNP markers in the whole genome could be a limitation of the approach used at the present work. However, these markers were located in regions where QTLs have been found in previous studies in this same population (Silva et al., 2010; Hidalgo et al., 2013), thus generating a SNP marker panel that was able to capture the genetic variation on the considered traits (a, b and k parameters). Despite the relatively small number of animals evaluated, the population was structured with a F2 design, which results in large linkage disequilibrium blocks that improve the capture of genetic variance, even in low-density marker panels (Costa et al., 2015).

QTLs identification

The list of relevant SNPs based on the joint analysis, that affect the adult weight (a) and the maturity rate (k) in pigs, as well as their genome positions and the related QTLs (PigQTLdb - National Animal Genomes Research Program, 2016) are shown in a Supplementary Material. We considered only the markers that explained at least 0.5% of the total genetic variance (Figure 1). A total of 22 SNPs for the "a" parameter, 17 SNPs for the “b” parameter and 26 SNPs for the "k" parameter, distributed in chromosomes (SSC) 1, 4, 7, 8, 17 and X were selected. We opted to show only relevant markers that have influenced "a", "k" and both parameters simultaneously.

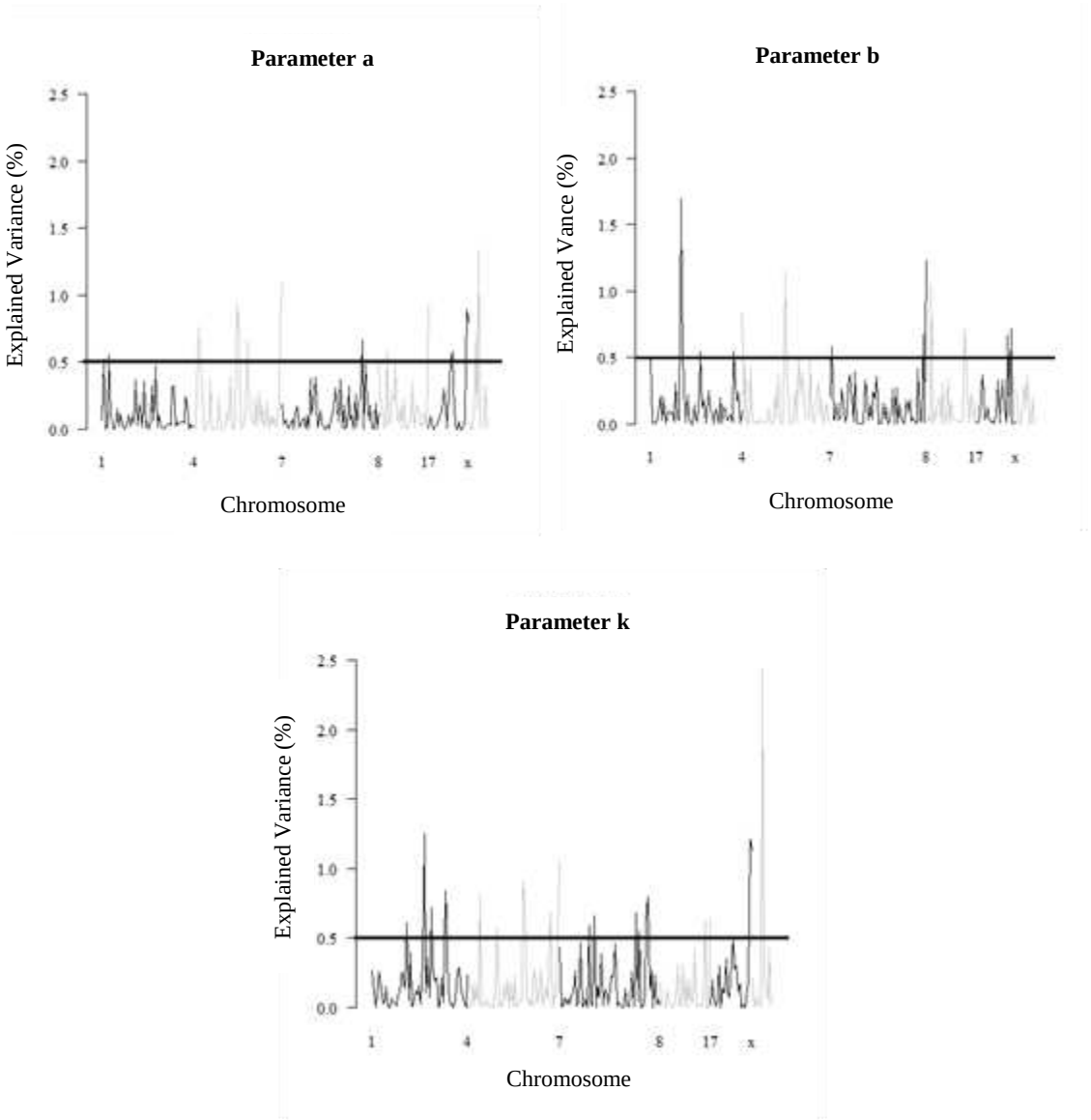

Figure 1. Percentage of total variance explained by each SNP for: (a) parameter “a”, (b) parameter “b”, (c) parameter “k”.

The SNPs explaining higher percentage of variance for the "k" parameter were associated with average daily gain. Approximately 23% of these SNPs are located in the SSC7. These results are in agreement with Ai et al. (2012) who found QTL for growth traits in this same chromosome, in a F2 pig population (White Duroc vs Erhualian); and Ruckertz and Bennewitz (2010), who reported QTLs in SSC7 for daily weight gain in crossbred pigs (European wild boars vs Meishan females).

The "a" parameter was associated with the weight at slaughter and the weight at 34 weeks, with the most relevant SNPs located on chromosomes 1, 4 and 8. These findings are in agreement with Koning et al. (2001) who found QTLs associated with final weight traits in these chromosomes in a F2 population (Meishan vs Large White, Dutch Landrace); Ai et al. (2012), who reported QTLs for weight at slaughter on chromosomes 4 and 8; and Liu et al. (2007) who detected QTLs for carcass composition and average daily gain in the SSC1, suggesting multiple QTLs in this chromosome in crossbred pigs (Pietrain vs Duroc).

We identified three relevant SNPs (55840514 bp at SSC17, 55814469 at SSC17 and 76475804 at SSC X) for "a" and "k" simultaneously, and three SNPs affecting only "a" (292758 bp at SSC1, 67319 bp at SSC8 and 50290193 bp at SSC17), that are located in genome regions not previously described in the literature (see Supplementary Material).

In summary, whereas the genome association analysis is an impartial scan of the entire genome without any assumption about the role of a certain gene, the QTL approach allows researchers to investigate the region where a specific marker of the gene underlying a complex trait is located. When combining these two approaches in the same study, we have the advantage of identifying QTLs from the same population in which relevant markers for the traits of interest were identified. In this context, a joint genomic association analysis of multiple potentially correlated traits, like growth curve parameters, may be advantageous. This approach has increased the power of QTL detection as reported by Galesloot et al. (2014), when comparing several multitrait and single trait GWAS methods. In addition, these authors suggested that the multitrait method may be able to identify genetic variants that are currently not identifiable by standard single trait analysis.

CONCLUSIONS

Markers may allow capturing fractions of additive variance that would be lost if pedigrees are the only source of genetic information used. The model including marker and pedigree information had better goodness-of-fit than pedigree-based or marker-based models.

The heritability estimates for mature weight ("a") and maturity rate ("k") indicated that these traits is a feasible alternative for breeding programs aiming to change the shape of growth curves in pig breeding programs.

The multitrait GWAS was efficient to report QTLs associated with functions related to biological processes of growth in pigs. Relevant SNPs are located in genome regions not previously described in the literature. Future studies targeting these areas could provide further knowledge to uncover the genetic architecture underlying growth curves in pigs.

LITERATURE CITED

AI, H., REN, J., ZHANG, Z., MA, J., GUO, Y., YANG, B., HUANG, L., 2012. Detection of quantitative trait loci for growth- and fatness-related traits in a large-scale White Duroc x Erhualian intercross pig population. **Anim. Genet.**, 43, 383–391.

BAND, G. de O., GUIMARÃES, S.E.F., LOPES, P.S., PEIXOTO, J. de O., FARIA, D.A., Pires, A.V., Figueiredo, F.de C., Nascimento, C.S.do, Gomide, L.A.M., 2005. Relationship between the Porcine Stress Syndrome gene and carcass and performance traits in F2 pigs resulting from divergent crosses. **Genet. Mol. Biol.**, 28 (1), 92-96.

BLASCO, A., PILES, M., VARONA, L., 2003. A Bayesian analysis of the effect of selection for growth rate on growth curves in rabbits. **Genet. Sel. Evol.**, 35, 21-41.

CAI, W., KAISER, M.S., DEKKERS, J.C.M., 2012. Bayesian analysis of the effect of selection for residual feed intake on growth and feed intake curves in Yorkshire swine. **J. Anim. Sci.**, 90, 127–141.

CALUS, M.P.L., VEERKAMP, R.F., 2007. Accuracy of breeding values when using and ignoring the polygenic effect in genomic breeding value estimation with a marker density of one SNP per cM. **J. Anim. Breed. Genet.**, 124, 362–368.

COSTA, E.V., DINIZ, D.B., VERONEZE, R., RESENDE, M.D.V., AZEVEDO, C.F., GUIMARÃES, S.E.F., SILVA, F.F., LOPES, P.S., 2015. Estimating additive and dominance variances for complex traits in pigs combining genomic and pedigree information. **Genet. Mol. Res.**, 14 (2), 6303-6311.

De los CAMPOS, G., NAYA, H., GIANOLA, D., CROSSA, J., LEGARRA, A., MANFREDI, E., WEIGEL, K., COTES, J.M., 2009. Predicting quantitative traits with regression models for dense molecular markers and pedigree. **Genet.**, 182, 375–385.

FALL, C.H.D., YAJNIK, C.S., RAO, S., DAVIES, A.A., BROWN, N., FARRANT, H.J.W., 2003. Micronutrients and fetal growth. **J. Nutr.**, 133, 1747–1756.

FORNI, S., PILES, M., BLASCO, A., VARONA, L., OLIVEIRA, H.N., LÔBO, R.B., ALBUQUERQUE, L.G., 2009. Comparison of different nonlinear functions to describe Nelore cattle growth. **J. Anim. Sci.**, 87, 496–506.

FORNI, S., PILES, M., BLASCO, A., VARONA, L., OLIVEIRA, H.N., LÔBO, R.B., ALBUQUERQUE, L.G., 2007. Analysis of beef cattle longitudinal data applying a nonlinear model. **J. Anim. Sci.**, 85, 3189–3197.

FOXCROFT, G.R., TOWN, S.C., 2004. Prenatal programming of postnatal performance – the unseen cause of variance. **Adv. Pork. Produc.**, 15, 269.

GALESLOOT, T.E., VAN STEEN, K., KIEMENEY, L.A.L.M., JANSS, L.L., VERMEULEN, S.H., 2014. A comparison of multivariate genome-wide association methods. **PLoS One**, 9 (4), e95923.

GEWEKE, J. Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments. In: **Bayesian Statistics 4**, Oxford: Oxford University Press, 1992. p. 169- 193.

HEIDELBERGER, P. AND WELCH, P.D. Simulation Run Length Control in the Presence of an Initial Transient. **Operations Research** 31:1109-1144. 1983.

HIDALGO, A.M., LOPES, P.S., PAIXÃO, D.M., SILVA, F.F., BASTIAANSEN, J.W.M., PAIVA, S.R., FARIA, D.A., GUIMARÃES, S.E.F., 2013. Fine mapping and single nucleotide polymorphism effects estimation on pig chromosomes 1, 4, 7, 8, 17 and X. **Genet. Mol. Biol.**, 36, 511-519.

IBÁÑEZ-ESCRICHE, N., BLASCO, A., 2011. Modifying growth curve parameters by multitrait genomic selection. **J. Anim. Sci.**, 89, 661–668.

KOIVULA, M., SEVÓN-AIMONEN, M.-L., STRANDÉN, I., MATILAINEN, K., SERENIUS, T., STALDER, K.J., MÄNTYSAARI, E.A., 2008. Genetic (co)variances and breeding value estimation of Gompertz growth curve parameters in Finnish Yorkshire boars, gilts and barrows**. J. Anim. Breed. Genet.**, 125, 168–175.

KONING, D.J. de, RATTINK, A.P., HARLIZIUS, B., GROENEN, M.A.M., BRASCAMP, E.W., VAN ARENDONK, J.A.M., 2001. Detection and characterization of quantitative trait loci for growth and reproduction traits in pigs. **Livest. Prod. Sci.**, 72, 185–198.

LIU, G., JENNEN, D.G.J., THOLEN, E., JUENGST, H., KLEINWÄCHTER, T., HÖLKER, M., TESFAYE, D., ÜN, G., SCHREINEMACHERS, H.-J., MURANI, E., PONSUKSILI, S., KIM, J.-J., SCHELLANDER, K., WIMMERS, K., 2007. A genome scan reveals QTL for growth, fatness, leanness and meat quality in a Duroc-Pietrain resource population. **Anim. Genet.**, 38, 241–252.

National Animal Genome Research Program, 2016. PigQTLdb. Available at: (http://www.animalgenome.org/QTLdb/pig.html).

PELOSO, J.V., LOPES, P.S., GOMIDE, L.A.M., GUIMARÃES, S.E.F., CARNEIRO, P.L.S., 2010. Carcass and ham quality characteristics of heavy pigs from

different genetic groups intended for the production of dry-cured hams. **Meat Sci.**, 86, 371–376.

RAFTERY, A.E. AND LEWIS, S. How Many Iterations in the Gibbs Sampler? In: **Bayesian Statistics 4**, Oxford: Oxford University Press, 1992. p. 763-773.

RAMOS, A.M., CROOIJMANS, R.P.M.A., AFFARA, N.A., AMARAL, A.J., ARCHIBALD, A.L., BEEVER, J.E., BENDIXEN, C., CHURCHER, C., CLARK, R., DEHAIS, P., et al., 2009. Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. **PLoS One**, 4, e6524.

RÜCKERT, C., BENNEWITZ, J., 2010. Joint QTL analysis of three connected F2-crosses in pigs. **Genet. Sel. Evol.**, 42, 40.

SILVA, F.F., RESENDE, M.D.V., ROCHA, G.S., DUARTE, D.A.S., LOPES, P.S., BRUSTOLINI, O.J.B., THUS, S., VIANA, J.M.S., GUIMARÃES, S.E.F., 2013. Genomic growth curves of an outbred pig population. **Genet. Mol. Biol.**, 36, 4, 520-527.

SILVA, F.F., ROSA, G.J.M., GUIMARÃES, S.E.F., LOPES, P.S., de los CAMPOS, G., 2011. Three-step Bayesian factor analysis applied to QTL detection in crosses between outbred pig populations. **Livest. Sci.**, 142, 210–215.

SILVA, K.M., BASTIAANSEN, J.W.M., KNOL, E.F., MERKS, J.W.M., et al., 2010. Meta-analysis of results from quantitative trait loci mapping studies on pig chromosome 4. **Anim. Genet.**, 42, 280-292.

SMITH, B.J., 2007. BOA: An R Package for MCMC Output Convergence Assessment and Posterior Inference. **J. Stat. Softw.**, 21(11).

SPIEGELHALTER, D.J., BEST, N.G., CARLIN, B.P., VAN DER LINDE, A., 2002. Bayesian measures of model complexity and fit. **J. R. Statist. Soc.** B, 64, 583–639.

VARONA, L., MORENO, C., GARCÍA CORTÉS, L.A., YAGÜE, G., ALTARRIBA, J., 1999. Two-step versus joint analysis of Von Bertalanffy function. **J. Anim. Breed. Genet.**, 116, 331–338.

VARONA, L., MORENO, C., GARCIA CORTES, L.A., ALTARRIBA, J., 1998. Bayesian analysis of wood's lactation curve for Spanish dairy cows. **J. Dairy Sci.**, 81, 1469–1478.

VARONA, L., MORENO, C., GARCÍA CORTÉS, L.A., ALTARRIBA, J., 1997. Multiple trait genetic analysis of underlying biological variables of production functions. **Livest. Prod. Sci.**, 47, 201-209.

VERARDO, L.L., SILVA, F.F., VARONA, L., RESENDE, M.D.V., BASTIAANSEN, J.W.M., LOPES, P.S., GUIMARÃES, S.E.F., 2015. Bayesian GWAS and network analysis revealed new candidate genes for number of teats in pigs. **J. Appl. Genet.**, 56, 123–132.

SUPPLEMENTAR MATERIAL

Relevant SNPs for the parameters (a, and k) in pigs, their positions in base pairs (bp) at pig chromosome (chr) with their references, and the marker selected that explain at least 0.5% on the SNP variance for the three parameters.

SNP	Param.	chr	Position (bp)	Variance (%)	Evidenced by
ALGA0007908	k	1	190856045	0.7802	Kim et al. (2000); Harmegnies et al. (2006); Liu et al. (2007); Ruckert and Bennewitz (2010)
ALGA0007015	a	1	150993073	0.4760	Koning et al. (2001); Liu et al. (2007); Liu et al. (2008); Ruckert and Bennewitz (2010)
ALGA0001557	a	1	16104793	0.5500	Koning et al. (2001); Liu et al. (2007); Ruckert and Bennewitz (2010)
ALGA0000022	a	1	292758	0.5241	
ALGA0006721	a, k	1	142016866	1.2464	Koning et al. (2001); Liu et al. (2007); Liu et al. (2008); Ruckert and Bennewitz (2010)
ALGA0007897	k	1	190593688	0.8416	Kim et al. (2000); Harmegnies et al. (2006); Liu et al. (2007); Ruckert and Bennewitz (2010)
ALGA0007023	k	1	151750737	0.7175	Koning et al. (2001); Liu et al. (2007); Liu et al. (2008); Ruckert and Bennewitz (2010)
ALGA0005071	a	1	80441657	1.7000	Koning et al. (2001); Evans et al. (2003); Liu et al. (2007); Ruckert and Bennewitz (2010)

ALGA0005714	k	1	105010422	0.6111	Evans et al. (2003); Liu et al. (2007); Ruckert and Bennewitz (2010)
ALGA0029783	a, k	4	127966743	1.0901	Knott et al. (1998); Nagamine et al. (2003)
ALGA0027472	k	4	100262783	0.7273	Andersson et al. (1994); Knott et al. (1998); Wang et al. (1998); Marklund et al. (1999); Walling et al. (2000); Bidanel et al. (2001); Koning et al. (2001); Knott et al. (2002); Cepica et al. (2003); Nagamine et al. (2003); Mercade et al. (2005); Murani et al. (2006); Fontanesi et al. (2010); Ruckert and Bennewitz (2010); Tortereau et al. (2010)
ALGA0022414	a	4	3097092	0.7633	Nagamine et al. (2003); Edwards et al. (2008); Liu et al. (2008)
ALGA0026242	a	4	80196806	1.1400	Andersson et al. (1994); Knott et al. (1998); Wang et al. (1998); Marklund et al. (1999); Walling et al.(2000); Bidanel et al. (2001); Knott et al. (2002); Cepica et al. (2003); Nagamine et al. (2003); Mercade et al. (2005); Murani, et al. (2006); Sanchez et al. (2006); Fontanesi et al. (2010); Tortereau et al. (2010)
ALGA0029781	a, k	4	127915978	0.7701	Knott et al. (1998); Nagamine et al. (2003)
ALGA0027463	a, k	4	100209536	0.9127	Andersson et al. (1994); Knott et al. (1998); Walling et al. (1998); Wang et al. (1998); Marklund et al.

					(1999); Walling et al. (2000); Bidanel et al. (2001); Koning et al. (2001); Knott et al. (2002); Cepica et al. (2003); Nagamine et al. (2003); Mercade et al. (2005); Murani et al. (2006); Fontanesi et al. (2010); Ruckert and Bennewitz (2010); Tortereau et al. (2010)
ALGA0026446	a	4	85013698	0.6947	Andersson et al. (1994); Knott et al. (1998); Wang et al. (1998); Marklund et al. (1999); Walling et al.(2000); Bidanel et al. (2001); Knott et al. (2002); Nagamine et al. (2003); Nagamine et al. (2004); Mercade et al. (2005); Murani et al. (2006); Sanchez et al. (2006); Fontanesi et al. (2010); Tortereau et al. (2010)
ALGA0022429	a	4	3183518	0.8508	Nagamine et al. (2003); Edwards et al. (2008); Liu et al. (2008)
ALGA0024036	k	4	20554086	0.8087	Knott et al. (1998); Walling et al. (1998); Walling et al. (2000); Nagamine et al. (2003); Liu et al. (2008)
ALGA0029474	k	4	122989609	0.6906	Knott et al. (1998); Walling et al. (2000); Malek et al. (2001); Bidanel et al. (2001); Knott et al. (2002); Nagamine et al. (2003)
ALGA0025382	k	4	60311651	0.5680	Andersson et al. (1994); Knott et al. (1998); Walling et al. (1998); Wang et al. (1998);

					Marklund et al. (1999); Walling et al. (2000); Bidanel et al. (2001); Knott et al. (2002); Koning et al. (2003); Nagamine et al. (2003); Evans et al. (2003); Mercade et al. (2006); Murani et al. (2006); Fontanesi et al. (2010); Tortereau et al. (2010); Ai et al. (2012)
ALGA0040318	k	7	35289714	0.4684	Koning et al. (2003); Evans et al. (2003); Nagamine et al. (2004); Sanchez et al., (2006); Liu et al. (2008); Ruckert and Bennewitz (2010); Ai et al. (2012)
ALGA0045009	a, k	7	120884969	0.6200	Nagamine et al. (2003)
ALGA0040948	k	7	45579337	0.5843	Bidanel et al. (2001); Quintanilla et al. (2002); Kim et al. (2006); Liu et al. (2008); Ruckert and Bennewitz (2010); Ai et al. (2012)
ALGA0044983	a	7	120615826	0.6660	Nagamine et al. (2003)
ALGA0045338	k	7	125022538	0.7978	Onteru et al. (2013);
ALGA0044302	k	7	110744562	0.6815	Nagamine et al. (2003); Edwards et al. (2008); Wang et al. (2015);
ALGA0041266	k	7	50275200	0.6610	Nezer et al. (2002); Quintanilla et al. (2002); Nagamine et al. (2003); Kim et al. (2006); Liu et al. (2008); Ruckert and Bennewitz (2010)
ALGA0044524	k	7	115268795	0.5526	Nagamine et al. (2003); Wang et al. (2015)
ALGA0049235	k	8	55138365	1.7988	Koning et al.,(2001); Ai et al. (2012)
ALGA0047819	a	8	20455978	0.4827	Koning et al.,(2001); Beeckmann et al. (2003); Evans et al. (2003);

					Ruckert and Bennewitz (2010); Ai et al. (2012)
ALGA0050287	a, k	8	66561233	0.9245	Koning et al. (2001); Ai et al. (2012)
ALGA0049233	k	8	55126130	0.6233	Koning et al. (2001); Ai et al. (2012)
ALGA0047008	a	8	10356235	0.5783	Koning et al. (2001); Quintanilla et al. (2002); Ai et al. (2012)
ALGA0046028	a	8	67319	0.4970	
ALGA0096707	a, k	17	55840514	1.1371	
ALGA0094911	a, k	17	35020233	0.4920	Pierzchala et al. (2003)
ALGA0096701	a, k	17	55814469	1.2122	
ALGA0094915	a	17	35099305	0.5832	Pierzchala et al. (2003)
ALGA0096093	a	17	50290193	0.6675	
MARC0099472	a, k	x	76475804	2.4284	
ALGA0099785	a	x	35172136	0.6496	Cepica et al. (2003); Geldermann et al. (2003)

LITERATURE CITED

AI, H., REN, J., ZHANG, Z., MA, J., GUO, Y., YANG, B., HUANG, L., 2012. Detection of quantitative trait loci for growth- and fatness-related traits in a large-scale White Duroc x Erhualian intercross pig population. **Anim. Genet.**, 43, 383–391.

ANDERSSON L., HALEY C.S., ELLEGREN H., KNOTT S.A., JOHANSSON M., ANDERSSON K., ANDERSSON-EKLUND L., EDFORS-LILJA I., FREDHOLM M., HANSSON I., HAKANSSON J., LUNDSTROM K., 1994. Genetic Mapping of Quantitative Trait Loci for Growth and Fatness in Pigs. **Science**, 263, 1771-1774.

BEECKMANN P., SCHRÖEFFEL J.JR., MOSER G., BARTENSCHLAGER H., REINER G., GELDERMANN H., 2003. Linkage and QTL mapping for Sus scrofa chromosome 1. **J. Anim. Breed. Genet.**, 120(1), 1–10.

BIDANEL, J.P., MILAN, D., IANNUCCELLI, N., AMIGUES, Y., BOSCHER, M.Y., BOURGEOIS, F., JEAN-CARITEZ, C., GRUAND, J., LE ROY, P., LAGANT, H., QUINTANILLA, R., RENARD, C., GELLIN, J., OLLIVIER, L., CHEVALET, C., 2001. Detection of quantitative trait loci for growth and fatness in pigs. **Genet. Sel. Evol.**, 33, 289–309.

CEPICA, S., STRATIL, A., KOPECNY, M., BLAZKOVA, P., SCHRÖFFEL JR, J., DAVOLI, R., FONTANESI, L., REINER, G., BARTENSCHLAGER, H., MOSER, G., GELDERMANN, H., 2003. Linkage and QTL mapping for Sus scrofa chromosome 4. **J. Anim. Breed. Genet.**, 120 (1), 28–37.

EDWARDS, D.B., ERNST, C.W., TEMPELMAN, R.J., ROSA, G.J.M., RANEY, N.E., HOGE, M.D., BATES, R.O., 2008. Quantitative trait loci mapping in an F2 Duroc × Pietrain resource population: I. Growth traits. **J. Anim. Sci.**, 86, 241–253.

EVANS, G.J., GIUFFRA, E., SANCHEZ, A., KERJE, S., DAVALOS, G., VIDAL, O., ILLÁN, S., NOGUERA, J.L., VARONA, L., VELANDER, I., SOUTHWOOD, O.I., KONING, D.-J. de, HALEY, C.S., PLASTOW, G.S., ANDERSSON, L., 2003. Identification of Quantitative Trait Loci for production traits in commercial pig populations. **Genet.**, 164, 621–627.

FONTANESI, L., SCOTTI, E., BUTTAZZONI, L., DALL'OLIO, S., DAVOLI, R., RUSSO, V., 2010. A single nucleotide polymorphism in the porcine cathepsin K (CTSK) gene is associated with back fat thickness and production traits in Italian Duroc pigs. **Mol. Biol. Rep.**, 37, 491–495.

GELDERMANN H., MÜLLER E., MOSER G., REINER G., BARTENSCHLAGER H., CEPICA S., STRATIL A., KURYL J., MORAN C., DAVOLI R., BRUNSCH C., 2003. Genome-wide linkage and QTL mapping in porcine F2 families generated from Pietrain, Meishan and Wild Boar crosses. **J. Anim. Breed. Genet.**, 120, 363–393.

HARMEGNIES, N., DAVIN, F., de SMET, S., BUYS, N., GEORGES, M., COPPIETERS, W., 2006. Results of a whole-genome quantitative trait locus scan for growth, carcass composition and meat quality in a porcine four-way cross. **Anim. Genet.**, 37, 543–553.

KIM, C.W., HONG, Y.H., YUN, S.-I., LEE, S.-R., KIM, Y.H., KIM, M.-S., CHUNG, K.H., JUNG, W.Y., KWON, E.J., HWANG, S.S., PARK, D.H., CHO, K.K., YEO, J.S., CHANG, K.-T., 2006. Use of microsatellite markers to detect quantitative trait loci in Yorkshire pigs. **J. Reprod. Dev.**, 52(2), 229-237.

KIM, K.S., LARSEN, N., SHORT, T., PLASTOW, G., ROTHSCHILD, M.F., 2000. A missense variant of the porcine melanocortin-4 receptor (MC4R) gene is associated with fatness, growth, and feed intake traits. **Mamm. Genome**, 11, 131–135.

KNOTT, S.A., MARKLUND, L., HALEY, C.S., ANDERSSON, K., DAVIES, W., ELLEGREN, H., FREDHOLM, M., HANSSON, I., HOYHEIM, B., LUNDSTRÖM,

K., MOLLER, M., ANDERSSON, L., 1998. Multiple marker mapping of quantitative trait loci in a cross between outbred wild boar and Large White pigs. **Genet.**, 149, 1069–1080.

KNOTT, S.A., NYSTRÖM, P.E., ANDERSSON-EKLUND, L., STERN, S., MARKLUND, L., ANDERSSON L., HALEY, C.S. 2002. Approaches to interval mapping of QTL in a multigeneration pedigree: the example of porcine chromosome 4. **Anim. Genet.**, 33, 26-32.

KONING, D.J.DE, PONG-WONG, R., VARONA, L., EVANS, G. J, GIUFFRA, E., SANCHEZ, A., PLASTOW, G., NOGUERA, J.L., ANDERSSON, L., HALEY, C.S., 2003. Full pedigree quantitative trait locus analysis in commercial pigs using variance components. **J. Anim. Sci.**, 81, 2155–2163.

KONING, D.J. DE, RATTINK, A.P., HARLIZIUS, B., GROENEN, M.A.M., BRASCAMP, E.W., VAN ARENDONK, J.A.M., 2001. Detection and characterization of quantitative trait loci for growth and reproduction traits in pigs. **Livest. Prod. Sci.**, 72, 185–198.

LIU, G., KIM, J.J., JONAS, E., WIMMERS, K., PONSUKSILI, S., MURANI, E., PHATSARA, C., THOLEN, E., JUENGST, H., TESFAYE, D., CHEN, J.L., SCHELLANDER, K., 2008. Combined line-cross and half-sib QTL analysis in Duroc–Pietrain population. **Mamm. Genome**, 19, 429–438.

LIU, G., JENNEN, D.G.J., THOLEN, E., JUENGST, H., KLEINWÄCHTER, T., HÖLKER, M., TESFAYE, D., ÜN, G., SCHREINEMACHERS, H.-J., MURANI, E., PONSUKSILI, S., KIM, J.-J., SCHELLANDER, K., WIMMERS, K., 2007. A genome scan reveals QTL for growth, fatness, leanness and meat quality in a Duroc-Pietrain resource population. Anim. Genet., 38, 241–252.

MALEK, M., DEKKERS, J.C.M., LEE, H.K., BAAS, T.J., ROTHSCHILD, M.F., 2001. A molecular genome scan analysis to identify chromosomal regions influencing economic traits in the pig. I. Growth and body composition. **Mamm. Genome**, 12, 630–636.

MARKLUND L., NYSTRÖM P., STERN S., ANDERSSON-EKLUND L., ANDERSSON L., 1999. Confirmed quantitative trait loci for fatness and growth on pig chromosome 4. **Heredity**, 82, 134-141.

MERCADE, A., ESTELLE, J., NOGUERA, J. L., FOLCH, J.M., VARONA, L., SILIO, L., SANCHEZ, A., PEREZ-ENCISO, M., 2005. On growth, fatness, and form: A

further look at porcine Chromosome 4 in an Iberian x Landrace cross. **Mamm. Genome**, 16, 374–382.

MERCADE, A., PEREZ-ENCISO, M., VARONA, L., ALVES, E., NOGUERA, J.L., SANCHEZ, A., FOLCH, J.M., 2006. Adipocyte fatty-acid binding protein is closely associated to the porcine FAT1 locus on chromosome 4. **J. Anim. Sci.**, 84, 2907–2913.

MURANI, E., MURÁNIOVA, M., PONSUKSILI, S., SCHELLANDER, K., WIMMERS, K., 2006. Molecular characterization and evidencing of the porcine CRH gene as a functional-positional candidate for growth and body composition. **Biochem. Biophys. Res. Commun.**, 342, 394–405.

NAGAMINE, Y., HALEY, C.S., SEWALEM, A., VISSCHER, P.M., 2003. Quantitative trait loci variation for growth and obesity between and within lines of pigs (Sus scrofa). **Genet.**, 164, 629–635.

NAGAMINE, Y., VISSCHER, P.M., HALEY, C.S., 2004. QTL detection and allelic effects for growth and fat traits in outbred pig populations. **Genet. Sel. Evol.**, 36, 83–96.

NEZER, C., MOREAU, L., WAGENAAR, D., GEORGES, M., 2002. Results of a whole genome scan targeting QTL for growth and carcass traits in a Piétrain × Large White intercross. **Genet. Sel. Evol.**, 34, 371–387.

ONTERU, S.K., GORBACH, D.M., YOUNG, J.M., GARRICK, D.J., DEKKERS, J.C.M., ROTHSCHILD, M.F., 2013. Whole Genome Association Studies of Residual Feed Intake and Related Traits in the Pig. **PLoS One**, 8, e61756.

PIERZCHALA, M., CIESLAK, D., REINER, G., BARTENSCHLAGER, H., MOSER G., 2003. Linkage and QTL mapping for Sus scrofa chromosome 17. **J. Anim. Breed. Genet.**, 120(1), 132–137.

QUINTANILLA, R., MILAN, D., BIDANEL, J.P., 2002. A further look at quantitative trait loci affecting growth and fatness in a cross between Meishan and Large White pig populations. **Genet. Sel. Evol.**, 34, 193–210.

RÜCKERT, C., BENNEWITZ, J., 2010. Joint QTL analysis of three connected F2-crosses in pigs. Genet. Sel. Evol., 42, 40.

SANCHEZ, M.P., RIQUET, J., IANNUCCELLI, N., GOGUE, J., BILLON, Y., DEMEURE, O., CARITEZ, J.C., BURGAUD, G., FEVE, K., BONNET, M., PERY, C., LAGANT, H., LE ROY, P., BIDANEL, J.P., MILAN, D., 2006. Effects of quantitative trait loci on chromosomes 1, 2, 4, and 7 on growth, carcass, and meat quality traits in backcross Meishan x Large White pigs. **J. Anim. Sci.**, 84, 526–537.

TORTEREAU, F., GILBERT, H., HEUVEN, H.C.M., BIDANEL, J.P., GROENEN, M.A.M., RIQUET, J., 2010. Combining two Meishan F2 crosses improves the detection of QTL on pig chromosomes 2, 4 and 6. **Genet. Select. Evol.**, 42, 42.

WALLING, G.A., VISSCHER, P.M., ANDERSSON, L., ROTHSCHILD, M.F., WANG, L., MOSER, G., GROENEN, M.A.M., BIDANEL, J.P., CEPICA, S., ARCHIBALD, A.L., GELDERMANN, H., KONING, D.J. de, MILAN, D., HALEY, C.S., 2000. Combined Analyses of Data From Quantitative Trait Loci Mapping Studies: Chromosome 4 Effects on Porcine Growth and Fatness. **Genet.**, 155, 1369–1378.

WALLING G.A., ARCHIBALD A.L., CATTERMOLE J.A., DOWNING A.C., FINLAYSON H.A., NICHOLSON D., VISSCHER P.M., WALKER C.A., HALEY C.S., 1998. Mapping of quantitative trait loci on porcine chromosome 4. **Anim. Genet.**, 29, 415-424.

WANG, K., LIU, D., HERNANDEZ-SANCHEZ, J., CHEN, J., LIU, C., WU, Z., FANG, M., LI, N., 2015. Genome Wide Association Analysis Reveals New Production Trait Genes in a Male Duroc Population. **PLoS One**, 10(9), e0139207.

WANG, L., YU, T.-P., TUGGLE, C.K., LIU, H.-C., ROTHSCHILD, M.F., 1998. A Directed Search for Quantitative Trait Loci on Chromosomes 4 and 7 in Pigs. **J. Anim. Sci.**, 76, 2560–2567.

Chapter 2

Genetic evaluation of age at first calving for Brazilian Brahman cattle using censored Bayesian models

ABSTRACT: The purpose of this study was to estimate genetic parameters and compare models for handling censored data of age at first calving (AFC) in Brahman Brazilian beef cattle. Censored records were defined as AFC records out of range of 731 and 1824 days. Data including information of 53,703 Brahman cows were analyzed using 4 different methods: conventional linear method (LM), considering only uncensored records; simulation method (SM), which data are augmented by drawing random samples from positive truncated normal distributions; penalty method (PM), in which a constant of 21 days was added to censored records and a bitrait threshold-linear method model (TLcens) considering any prior information about censored records. The additive genetic variance components estimated from LM and PM were similar. Heritability estimates for AFC ranged from 0.09 (TLcens) to 0.20 (LM). In general, genetic breeding values correlations from different methods and the percentage of in common selected animals indicated moderate reranking, ranging from 0.82 (LM x SM) to 0.97 (LM x PM) and 32.70% (SM x TLcens) to 89.12% (LM x PM), respectively. Comparisons based on cross-validation analyses, indicated LM as a suitable alternative for predicting breeding values for AFC in this Brahman population.

Key words: age at first calving, censored data, Brahman cattle, threshold analysis

INTRODUCTION

One of the most relevant selection criteria for the genetic improvement of reproductive efficiency in beef cattle is the age at first calving (AFC). Despite easiness of recording, AFC may show recording mistakes due to non-occurrence and/or delay in communication of the calving until a given pre-fixed time period. Thus, this trait is widely referred as censored (Tarrés et al., 2006).

The simplest option to handle this problem is omitting these observations, which leads to loss of a number of information and may decreases the prediction accuracy. Furthermore, it may distort the real variability of the trait and to mask genetic differences between animals (Guo et al., 2001; Dias et al., 2004). Another option is to use suitable statistical models in studies with Zebu real data which the censored observations are more effectively exploited in the genetic evaluation.

Some methods have been proposed to deal with censured traits in genetic evaluations. One is based on simulation of censored records from positive truncated normal distributions taking into account the estimated effects of the model (Donoghue et al., 2004a; Korsgaard et al., 2003). Another one is the penalty methodology proposed by Johnston and Bunter (1996), which consists to impute information by adding a constant (number of days) to real data. For AFC, 21 days are often included based on the assumption that the heifer should be fertile in the subsequent estrous cycle. The linear-threshold bivariate analysis considers the censoring status (threshold binary trait) as an additional trait to improve the accuracy of genetic parameter estimates. In the last method, it assumes that the correlation between fertility traits (e.g., AFC) and the censoring status might improve the prediction accuracy (Urioste et al., 2007a).

In this context, we aimed to compare the mentioned methods under a Bayesian framework for genetic evaluation of AFC in Brazilian Brahman cattle by accessing predictive performance via cross-validation.

MATERIALS AND METHODS

Data

Brahman fertility data were provided by Brazilian Association of Zebu Cattle Breeders (ABCZ). Age at first calving (AFC) was defined as the interval between birth and first parity of the cows. The AFC records were obtained during the period of 1960 and 2014. Before editing the data, there were 59,929 trait records available in the database. Data from females with AFC above 1825 days of age were assumed to have failed to calve, *i.e.* censored data. Data editing was performed by removing 1) animals with incomplete records; 2) single record by contemporary groups (CG); 3) animals belonging to contemporary groups consisting of only noncalvers; 4) outliers based on 3 standard deviation within CG.

The CG were formed as the combination of herd, year and birth season. CG with eight or more animals with phenotypic information were kept in the database for analysis. The values of censored records were generated using two different strategies: 1) Females without phenotypic information received as censored record the biggest CG value of AFC; 2) Adding the highest AFC value a 21-day penalty, within contemporary group which correspond to an estrous cycle (Johnston and Bunter, 1996). A complete description of the databases is presented in the table 1:

Table 1. Description of the database structure to trait to the age at first calving (AFC) as the number of records, descriptive statistics and information groups contemporary

Records[1]	N	(%)	Mean (days)	SD (days)	Minimum	Maximum
UR	50,630	94.28	1,189.97	239.89	731	1,824
CR	3,073	5.72	1,744.57	112.90	859	1,825

[1]UR – uncensored data; CR – censored data

Methodologies

AFC data were analyzed using four different Bayesian methods to deal with censored phenotypic records.

Linear Method (LM) was based on uncensored AFC data (DS1). The LM was used to evaluate the scenario when censored records were not used (DS1 dataset). For this, the following standard animal model was fitted:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Wcg} + \boldsymbol{Za} + \boldsymbol{e} \qquad [1]$$

where $\boldsymbol{y}$ is the vector of AFC records; $\boldsymbol{\beta}$ is the vector of systematic effects (mean, registration class and mating type); $\boldsymbol{cg}$ is the vector of contemporary group (herd-year-season) effects; $\boldsymbol{a}$ is the vector of additive genetic effects; $\boldsymbol{e}$ is the residual vector; and $\boldsymbol{X}$, $\boldsymbol{W}$ and $\boldsymbol{Z}$ are the incidence matrices associated with $\boldsymbol{\beta}$, $\boldsymbol{cg}$ and $\boldsymbol{a}$, respectively. It was assumed that $\boldsymbol{\beta} \sim N(\boldsymbol{0}, \boldsymbol{I}\sigma^2_{\beta})$, being σ^2_{β} a known variance with value 1e+10 (large variance) to represent vague prior knowledge, $\boldsymbol{cg} \sim N(\boldsymbol{0}, \boldsymbol{I}\sigma^2_{cg})$, $\boldsymbol{a} \sim N(\boldsymbol{0}, \boldsymbol{A}\sigma^2_{a})$ and $\boldsymbol{e} \sim N(\boldsymbol{0}, \boldsymbol{I}\sigma^2_{e})$, being $\boldsymbol{A}$ the numerator relationship matrix, σ^2_{cg} the contemporary group variance, σ^2_{a} the additive genetic variance, $\boldsymbol{I}$ the identity matrix, and σ^2_{e} the residual variance. The vector $\boldsymbol{\beta}$ included registration class (animals registered as pure by origin or in open book) and mating type (artificial insemination, embryo transfer, fertilization *in vitro*, natural mating and controlled mating).

Simulation Method (SM) is based in the same model presented in Eq. [1], however the dataset DS2 (censored data + uncensored data) was used instead DS1. For heifers with censored record, the methodology considers data simulation. Thus, $\boldsymbol{y}' = [\boldsymbol{y}_{ur} \ \boldsymbol{y}_{cr}]'$ is a vector in which $\boldsymbol{y}_{ur}$ is the vector of uncensored records of AFC, and $\boldsymbol{y}_{cr}$ the vector of simulated values for censored records. Using Gibbs sampling approach (Sorensen et al., 1998; Guo et al., 2001), $\boldsymbol{y}_{cr}$ were sampled from their respective predictive distributions. It was assumed that $\boldsymbol{y}_{cr}$ values followed a truncated Gaussian distribution whose lower limit is defined by the maximum values of AFC within the corresponding contemporary group. Thus, the augmented data $\boldsymbol{y}_{cr}$ were considered within each iteration of the Gibbs sampler as an observation for each censored record (Donoghue et al., 2004a; Korsgaard et al., 2003).

Penalty Method (PM) is equivalent to SM, however the censored records were replaced by a set of augmented records by adding a constant of 21 days over the highest AFC value within each contemporary group (DS3). The penalty suggested that the cows failing to become pregnant would conceive if they had another opportunity, as an extra estrous cycle (Donoghue et al., 2004b; Hou et al., 2009).

Threshold – Linear censored (TLcens) method is based on bitrait analysis where one trait is continuous and the another one is a threshold binary trait (calving success), which indicates the censored status. Females that calved were coded as 2, and cows without a recorded calving were assigned a 1 (failure). The binary records were associated to liability values given by a latent continuous variable (Sorensen and Gianola, 2002). At each MCMC iteration, the binary records generate a liability value below or over a given threshold. This model considers $\boldsymbol{y}' = [\boldsymbol{y}_{ur} \ \boldsymbol{y}_{cr}]'$, where $\boldsymbol{y}_{cr}$ are the higher day of AFC records within contemporary group (DS4).

$$\begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{l} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}_y & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{X}_l \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_y \\ \boldsymbol{\beta}_l \end{bmatrix} + \begin{bmatrix} \boldsymbol{W}_y & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{W}_l \end{bmatrix} \begin{bmatrix} \boldsymbol{cg}_y \\ \boldsymbol{cg}_l \end{bmatrix} + \begin{bmatrix} \boldsymbol{Z}_y & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Z}_l \end{bmatrix} \begin{bmatrix} \boldsymbol{a}_y \\ \boldsymbol{a}_l \end{bmatrix} + \begin{bmatrix} \boldsymbol{e}_y \\ \boldsymbol{e}_l \end{bmatrix} \qquad [2]$$

where $\boldsymbol{y}$ is the vector of AFC records; $\boldsymbol{l}$ is the vector of liability generated from censored status; $\boldsymbol{X}$, $\boldsymbol{\beta}$, $\boldsymbol{W}$, $\boldsymbol{Z}$, $\boldsymbol{cg}$, $\boldsymbol{a}$ and $\boldsymbol{e}$ is the same of model [1]. The following priori distributions were assumed:

$$\begin{bmatrix} \boldsymbol{a}_y \\ \boldsymbol{a}_l \end{bmatrix} \sim \boldsymbol{N}(\mathbf{0}, \mathbf{G}_0 \otimes \mathbf{A}), \begin{bmatrix} \boldsymbol{\beta}_y \\ \boldsymbol{\beta}_l \end{bmatrix} \sim \boldsymbol{N}(\mathbf{0}, \boldsymbol{\Sigma}_\beta \otimes \mathbf{I}), \begin{bmatrix} \boldsymbol{cg}_y \\ \boldsymbol{cg}_l \end{bmatrix} \sim \boldsymbol{N}(\mathbf{0}, \mathbf{CG}_0 \otimes \mathbf{I})$$

$$\text{and} \begin{bmatrix} \boldsymbol{e}_y \\ \boldsymbol{e}_l \end{bmatrix} \sim \boldsymbol{N}(\mathbf{0}, \mathbf{R}_0 \otimes \mathbf{I})$$

where $\boldsymbol{G_0}$ and $\boldsymbol{R_0}$ are the additive genetic and residual (co)variance matrices, respectively (as proposed by Varona et al., 1999); $\boldsymbol{CG_0}$ is the (co)variance matrix for CG effects; and $\boldsymbol{\Sigma}_\beta$ is a diagonal known matrix with values 1e+10 (large variances) to represent vague prior knowledge for systematic effects.

Markov Chain Monte Carlo (MCMC) Sampling

Parameters were drawn from the posterior distributions using Gibbs sampling, as implemented in the programs TM, kindly provided by Andres Legarra, INRACastanet Tolosan, France (Legarra et al., 2008). A total of 400,000 samples were generated, assuming a burn-in period and sampling interval (thin) of 100,000 and 10 iterations, respectively. The convergence of the MCMC chains was verified by graphical inspection and R package BOA (Smith, 2007). For all analyses, convergence was assessed using methodology presented by Heidelberg e Welch (1983), Geweke (1992) and Raftery and Lewis (1992).

Methods Comparisons

The predictive ability was accessed by cross-validation, which was implemented considering in the training dataset all censored records, and 70% of uncensored records obtained randomly within each contemporary group and validating on the remaining individuals. Here, we randomly split the data sets into two groups from the original data set, as specified above, these two subset were redefined 10 times, D1, D2, ..., D10. Finally, the average of the 10 correlation coefficients between the predicted and observed phenotypes was obtained.

The predicted phenotypes vector was calculated as $\hat{\boldsymbol{y}} = \boldsymbol{X}\hat{\boldsymbol{\beta}} + \boldsymbol{W}\hat{\boldsymbol{d}} + \boldsymbol{Z}\hat{\boldsymbol{a}}$. Thus, the solutions for the animals in the validation population were obtained based on the solutions of the training population animals. Finally, the predictive ability used to measure the efficiency of each method was given by the correlation between observed and predicted phenotypes from the validation population and mean square errors.

The individual accuracy of breeding value for each animal *i* was also calculated and used to compare the considered methods. The accuracy (r) was calculated as showed in Eq. [3]:

$$r_i = \sqrt{1 - \frac{(SD_i)^2}{\sigma_a^2}}, \qquad [3]$$

where SD_i is the posterior standard deviation of the breeding value of each animal *i* and σ^2_a is the additive genetic variance.

Spearman's correlation coefficients between predicted breeding values from different methods were computed to infer on differences in the ranking of animals. In addition, the percentage of in common animals selected at different percentiles (TOP1% and TOP10%) based on the compared methods were also calculated.

RESULTS

Variance Components

Posterior means with respective standard deviation and HPD95% region for variance components and heritability from different methods are presented in table 2. The LM only considers AFC uncensored values and its results should be used as reference (simplest model). The additive genetic variance obtained from LM and PM were similar and showed overlapping of HPD intervals. The corresponding estimates from SM and TLcens methods presented the higher and lower values, respectively. Since the SM method was based on simulated random numbers based on a truncated normal distribution, thus inserting some source of variation in the data, this method produces higher estimates for additive genetic variance. The lowest posterior means for the residual variance was observed for LM and PM, whereas the highest value was reported for SM. This higher value for SM suggests that obtaining random numbers based on a truncated normal distribution for all censored records can overestimate this variance component.

Table 2. Posterior means, standard deviation and highest posterior density region (95%) for heritability and variance components and genetic parameters for AFC in Brahman cattle.

Method[1]	$\mathbf{h^2}$	$\boldsymbol{\sigma^2_a}$	$\boldsymbol{\sigma^2_{gc}}$	$\boldsymbol{\sigma^2_e}$
LM	0.20 (0.01)	11,887.01 (598.41)	5,780.80 (766.84)	43,205.65 (516.72)
	[0.18; 0.21]	[10,765.13; 13,070.92]	[4,331.42; 7,309.51]	[42,198.82; 44,205.45]
SM	0.19 (0.01)	31,481.96 (1,378.36)	54,378.23 (6,188.77)	80,271.21 (1,110.02)
	[0.17; 0.21]	[28,748.36; 34,210.96]	[42,632.60; 66,547.76]	[78,032.36; 82,409.43]
PM	0.18 (0.01)	14,354.36 (712.52)	17,560.54 (2,063.04)	46,658.04 (598.47)
	[0.16; 0.20]	[13,014.01; 15,783.45]	[13,715.58; 21,650.96]	[45,494.81; 47,833.18]
TLcens	0.09 (0.006)	8,221.12 (518.38)	20,960.41 (2,476.74)	64,038.80 (503.09)
	[0.08; 0.10]	[7,281.80; 9,261.24]	[16,191.94; 25,910.80]	[63,064.86; 64,972.09]

h^2= heritability, σ^2_a, σ^2_{gc} and σ^2_e = additive genetic, contemporary group and residual variances, respectively;

[1]LM, SM, PM, TLcens: linear, simulation, penalty, threshold-linear censored methodologies.

The estimates of heritability were similar and ranged from 0.09 (TLcens data set) to 0.20 (LM data set). Heritability estimate using LM, SM and PM were similar and greater than that from TLcens (Table 2). As result of the lower estimate of the additive variance under TLcens, its posterior mean of heritability was smaller than the other methods.

Methods Comparisons

The predictive ability was performed by correlation coefficients (between observed and predicted phenotype) using cross-validation approach and mean square error (Table 3). Higher correlation and lower MSE were found for LM, indicating that this method is recommended to be used in AFC genetic evaluation with censored records in Brazilian Brahman cattle.

Table 3. Average mean square error (MSE) and correlations between observed and predicted phenotypes from 10 fold of the cross-validation with respective standard deviation (SD)

Comparison method		
Method[1]	**MSE (SD)**	**Correlation (SD)**
LM	52,541.85 (643.26)	0.30 (0.006)
SM	81,104.72 (8,137.71)	0.19 (0.078)
PM	55,267.17 (1,320.13)	0.25 (0.008)
TLcens	56,660.74 (541.39)	0.22 (0.005)

[1]LM, SM, PM, TLcens: linear, simulation, penalty, threshold-linear censored

Spearman correlation coefficients, accuracy of breeding values and percentage of in common selected animals (considering different percentiles, TOP1% and TOP10%) between breeding values predicted from different methodologies are shown in Table 4. Spearman correlations between the LM method and all other were higher among predictions obtained from linear model.

Concordance between the selected top 1% animals ranged from 32.70% (SM and TLcens) to 82.96% (LM and PM). For the top 10% of animals it was slightly higher and

ranged from 59.48% (SM and TLcens) to 89.12% (LM and PM). These results show that although most animals in the top 1% or 10% were not ranked similarly under methods, there were exceptions.

Table 4. Spearman correlation between all animals (above diagonal) and accuracy (below diagonal) of predicted breeding values of AFC trait and percentage of animals in common between methods at 1% (above diagonal) and 10% (below diagonal) of selected individuals

Method[1]	LM	SM	PM	TLcens
	Spearman correlations and accuracy			
LM	_	0.82	0.97	0.95
SM	0.95	_	0.88	0.83
PM	0.98	0.96	_	0.95
TLcens	0.87	0.88	0.89	_
	Percentage of animals in common			
LM	_	46.96	82.96	55.65
SM	61.23	_	52.17	32.70
PM	89.12	66.48	_	51.48
TLcens	81.70	59.48	78.33	_

[1]LM, SM, PM, TLcens: linear, simulation, penalty, threshold-linear censored methodologies.

Threshold-Linear Analysis with Calving success (CS)

Summary of the posterior distributions of (co)variance components, heritabilities, and genetic correlation from the CS-AFC bitrait analysis are presented in Table 5. For the AFC trait analyzed joint with CS in the TLcens method, the contemporary group and residual variances were more variable than the additive genetic variance, illustrating the importance of environmental effects in this trait (AFC). For the CS trait, the posterior means for additive genetic variance and heritability in this study were low in magnitude (Table 5).

Table 5. Estimates of (co)variance components and genetic parameters and their standard deviation (SD) from bitrait analysis of age at first calving in days (AFC, trait 1) and calving success (CS, trait 2) – TLcens method

	Trait		
	AFC		CS
σ^2_a	8,221.12 (518.38) [7,281.80; 9,261.24]		0.01 (0.001) [0.006; 0.015]
σ^2_{cg}	20,960.41 (2,476.74) [16,191.94; 25,910.80]		0.61 (0.071) [0.48; 0.75]
σ^2_e	64,038.80 (503.09) [63,064.86; 64,972.09]		1.00
h^2	0.09 (0.006) [0.08; 0.10]		0.01 (0.001) [0.004; 0.009]
σ_{a12}	—	-8.72 (1.08) [-10.71; -6.60]	—
σ_{cg12}	—	-82.58 (11.55) [-106.20; -60.77]	—
σ_{e12}	—	-249.60 (1.09) [-251.65; -247.44]	—
r_{a12}	—	-0.95 (0.031) [-0.98; -0.90]	—
r_{p12}	—	-0.87 (0.009) [-0.89; -0.85]	—

σ^2_a - additive genetic variance; σ^2_{cg} – contemporary group variance; σ^2_e – residual variance; h^2 – heritability; σ_{a12} - additive genetic covariance; σ_{cg12} – contemporary group covariance; σ_{e12} – residual covariance; r_{a12} – genetic correlation; r_{p12} phenotypic correlation

DISCUSSION

LM and PM reported similar estimates for all variance components, indicating the correspondence among the simplest and penalty methods. Using either LM or PM to handle censored fertility records, a small impact on estimation of its variance components was expected. On the other hands, Urioste et al. (2007b) found similar additive genetic and residual variances estimates for SM and PM considering days to calving in Angus cattle. Probably, the correlation generated by the tritrait analysis (three calving intervals) used for these authors could have affected the variance components estimates. However, Forni and Albuquerque (2003) reported that imputation of censored data did not improve the identification of genetic differences between animals.

All methods showed medium heritabilities for AFC trait (0.18 to 0.20, Table 2) indicating a good scope for selection, except the TLcens method that, showed low heritability (0.09). Garcia et al. (2016) reported low heritability estimate (0.14) for AFC in Nellore cattle using TLcens method. In general, heritability estimates for AFC from field data reported in the literature also oscillated as observed in the present study, ranging from 0.10 to 0.37 in Brazilian Zebu cattle (Boligon and Albuquerque, 2011; Barrozo et al., 2012; Moreira et al., 2015). Differences on heritability estimates observed in the literature, compared to our study, may reflect differences in populations, in trait definitions, management practices that eventually confounding genetic and environmental effects estimates, or the influence of the data structure.

Donoghue et al. (2004a) and Donoghue et al. (2004b) using simulation and real data of fertility for days to calving, respectively, reported similar results among heritability estimates in relation to our study using penalty (PM) and simulation (SM) methods. On the other hand, the heritability results under threshold analysis conflicts with some authors (Johnston and Bunter, 1996; Morris et al., 2000; Phocas and Sapa, 2004). These authors have

reported higher heritability estimates using TLcens models than linear models, in disagreement with the results presented here. TLcens estimates were lower than the other methods, however, according to Boligon et al. (2008), different populations and models may affect genetic parameter estimates for AFC.

Heritability estimate for CS (calving success) (Table 5) in a bitrait analysis was 0.01. Estimates for CS or related traits, using threshold models, have been reported varying between 0.03 (Donoghue et al., 2004c) and 0.25 to 0.27 (Rust and Groeneveld, 2002). In general, unlike observed in our study, the use of threshold models have estimated higher values for heritabilities in comparison with linear models (Johnston and Bunter, 1996; Morris et al., 2000; Phocas and Sapa, 2004).

Although the genetic correlation between AFC and CS (-0.95, negative and genetically favorable) has been little exploited in the literature, especially in Brahman cattle, it is a relevant finding. Johnston and Bunter (1996) reported a very high negative genetic correlation (−0.97) for days to calving and CS in Angus females, suggesting that they can be considered genetically the same trait. Donoghue et al. (2004c), working with field data from first-calf Angus females, reported genetic correlation equal to -0.73. The higher negative genetic correlation indicate that AFC could act as an indicator trait of CS and could be implemented as a selection criterion for fertility traits, since females with a higher probability of calving success will also present a lower AFC. Furtheremore, selection to increase probability of CS would result as a correlated response, since CS has lower heritability estimate (0.01) than AFC (0.09), as can be seen in Table 5. The high negative genetic correlation (-0.95) between the two traits takes the AFC to be measured earlier in life of the animal. Nevertheless, disadvantage may be its implementation in large data sets.

Urioste et al. (2007a) using predictive ability on fertility traits (days to calving and calving success) of Uruguayan Aberdeen Angus cattle found similar correlations for PM and

SM, and PM and TLcens. However, Urioste et al. (2007b) found the main disadvantage implementation difficulties for large data sets for threshold analysis compared to linear models. In our study, the LM method besides presenting less MSE and higher predictive ability (Table 3) demanded less computational effort compared to other methods.

Spearman correlations and accuracy (table 4) were very similar for all methods suggesting there are small differences in the ranking of animals. Similar to results reported by Donoghue et al. (2004a), we observed that correlations and accuracies based on the uncensored data were slightly higher than censored.

The Spearman correlation between LM and SM was lower than other methods (0.82). This result is similar to Donoghue et al. (2004b) that worked with days to calving in Australian Angus cattle (0.81). According to the authors, these results indicated some reranking of animals when censored records were ignored when compared with methods that included noncalving females in the analysis. Assuming that the former approach is inferior, as it is ignoring an important source of genetic variation in fertility, these results highlight the need to include records from noncalving cows in order to accurately estimate differences in fertility for animals. On the other hand, in our study the Spearman correlation between LM and PM was higher than for other methods (0.97). It may reflect differences in populations, since these authors used only sires for correlation analysis and different trait definition.

Despite the differences observed for predictive abilities and genetic parameter estimates, the Spearman correlations and accuracies among methods were similar and closed to unity, indicating that no major reranking would be expected across these methodologies. These similarities and another previously results suggests that either approaches (LM and PM) could be used for genetic evaluations of AFC trait.

Percentage of animals in common (Table 4) showed the same pattern as observed for Spearman correlations of breeding value predictions and accuracy estimate, which were higher among linear model. The percentages of animals in common were higher between LM and PM, indicating no great reranking of animals when censored records are used. These percentages were, even, smaller when only TOP1% of the animals was considered compared to the TOP10%. These results indicate high changes in the ranks, although most animals in the top 1% or 10% were similarly ranked under different methods, there were exceptions.

Similar to the present study, Garcia et al. (2016) working with Nellore cattle found coincidence in sire ranking approaching, higher values among linear models. On the other hand, Hou et al. (2009) found high coincidence of TOP10 bulls in Danish Holstein when breeding values were predicted using PM, SM and TLcens. We can infer that the choice of methodology in our study could have large effect in the identification of the best animals in this population.

CONCLUSIONS

A linear model using censored data was the most accurate method for genetic evaluation of AFC in Brazilian Brahman cattle. Penalty method is also an alternative method to genetic evaluations of AFC data.

The larger estimate of the residual variance under the simulation method (SM) suggests that this approach does not provide a good method for handling censored records in beef fertility data.

The genetic correlation reported between AFC and CS indicates a strong negative correlation. Selecting animals with shorter age first to calving genetically will lead to correlated increases in calving success.

LITERATURE CITED

BARROZO, D.; BUZANSKAS, M.E.; OLIVEIRA, J.A. et al. Genetic parameters and environmental effects on temperament score and reproductive traits of Nellore cattle. **Animal**, v.6, n.1, p.36-40, 2012.

BOLIGON, A.A.; VOZZI, P.A.; NOMELINI, J.; et al. Genetic parameters for age at first calving estimated by different models for Nellore breed herds. **Ciência Rural**, v.38, n.2, p.432-436, 2008.

BOLIGON, A.A.; ALBUQUERQUE, L.G. Genetic parameters and relationships of heifer pregnancy and age at first calving with weight gain, yearling and mature weight in Nelore cattle. **Livestock Science**, v.141, n.1, p.12-16, 2011.

DIAS, L.T.; EL FARO, L.; ALBUQUERQUE, L.G. Heritability Estimates for Age at First Calving in Nelore Cattle. **Revista Brasileira de Zootecnia**, v.33, n.1, p.97-102, 2004.

DONOGHUE, K.A.; REKAYA, R.; BERTRAND, J.K. Comparison of methods for handling censored records in beef fertility data: Simulation study. **Journal Animal Science**, v.82, p.351-356, 2004a.

DONOGHUE, K. A.; REKAYA, R.; BERTRAND, J. K. Comparison methods for handling censored records in beef fertility data: Field study. **Journal of Animal Science**, v.82, n.2, p.357-361, 2004b.

DONOGHUE, K. A.; REKAYA, R.; BERTRAND, J. K.; MISZTAL, I.Threshold-linear analysis of measures of fertility in artificial insemination data and days to calving in beef cattle. **Journal of Animal Science**, v.82, p.987-993, 2004c.

FORNI, S.; ALBUQUERQUE, L.G. 2003. Environmental effects and heritability estimates for days to calving in Nelore cattle. **Proc**. 40th Reunião Anual da Sociedade Brasileira de Zootecnia, Santa Maria, Rio Grande do Sul, Brazil.

GARCIA, D.A.; ROSA, G.J.; VALENTE B.D.; CARVALHEIRO, R.; ALBUQUERQUE, L.G. Comparison of models for the genetic evaluation of reproductive traits with censored data in Nellore cattle. **Journal of Animal Science**, v.94, n.6, p.2297-306, 2016.

GEWEKE, J. Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments. In: **Bayesian Statistics 4**, Oxford: Oxford University Press, 1992. p. 169- 193.

GUO, S.F.; GIANOLA, D.; SHORT, R.R.T. Bayesian analysis of lifetime performance and prolificacy in Landrace sows using a linear mixed model with censoring. **Livestock Production Science**, v.72, p.243-252, 2001.

HEIDELBERGER, P. AND WELCH, P.D. Simulation Run Length Control in the Presence of an Initial Transient. **Operations Research** 31:1109-1144. 1983.

HOU, Y.; MADSEN, P.; LABOURIAU, R.; et al. Genetic analysis of days from calving to first insemination and days open in Danish Holsteins using different models and censoring scenarios. **Journal of Dairy Science**, v.92, n.3, p.1229-1239, 2009.

JOHNSTON, D. J.; BUNTER, K. L. Days to calving in Angus cattle: Genetic and environmental effects, and covariances with other traits. **Livestock Production Science**, v.45, n.1, p.13-22, 1996.

KORSGAARD, I.R.; LUND, M.S.; SORENSEN, D.; et al. Multivariate Bayesian analysis of Gaussian, right censored Gaussian, ordered categorical and binary traits using Gibbs sampling. **Genetics Selection Evolution**, v.35, p.159-183, 2003.

LEGARRA, A.; VARONA, L.; LÓPEZ DE MATURANA, E. **TM: Threshold Model.** (http://snp.toulouse.inra.fr/~alegarra), 2008.

MOREIRA, H.L.; BUZANSKAS, M.E.; MUNARI, D.P.; et al. Reproductive traits selection in Nelore beef cattle. **Ciência e Agrotecnologia**, v.39, n.4, p.355-362, 2015.

MORRIS, C.A.; WILSON, J.A.; BENNETT, G.L.; et al. Genetic parameters for growth, puberty, and beef cow reproductive traits in a puberty selection experiment. **New Zealand Journal of Agricultural Research**, v.43, p.83–91, 2000.

PHOCAS, F.; SAPA, J. Genetic parameters for growth, reproductive performance, calving ease and suckling performance in beef cattle heifers. **Journal of Animal Science**, v.79, p.41-48, 2004.

RAFTERY, A.E. AND LEWIS, S. How Many Iterations in the Gibbs Sampler? In: **Bayesian Statistics 4**, Oxford: Oxford University Press, 1992. p. 763-773.

RUST, T.; GROENEVELD, E. Variance component estimation of female fertility traits in two indigenous and two European beef cattle breeds of South Africa. **South African Journal of Animal Science**, v.32 (1), p.23-29, 2002.

SMITH, B.J. BOA: An R Package for MCMC Output Convergence Assessment and Posterior Inference. **Journal of Statistical Software**, 21(11), 2007.

SORENSEN, D.A.; GIANOLA, D.; KORSGAARD, I.R. Bayesian mixed-effects model analysis of a censored normal distribution with animal breeding applications. **Acta Agriculturae Scandinavica A—Animal Sciences**, v.48, n.4, p.222-229, 1998.

SORENSEN, D.; GIANOLA, D. **Likelihood, bayesian and MCMC methods in quantitative genetics**. Springer, 2002.

TARRÉS, J.; PIEDRAFITA, J.; DUCROCQ, V. Validation of an approximate approach to compute genetic correlations between longevity and linear traits. **Genetics Selection Evolution**, v.38, p.65-83, 2006.

URIOSTE, J.I.; MISZTAL, I.; BERTRAND, J.K. Fertility traits in spring-calving Aberdeen Angus cattle. 2. Model comparison. **Journal of Animal Science**, v.85, n.11, p.2861-2865, 2007a.

URIOSTE, J.I.; MISZTAL, I.; BERTRAND, J.K. Fertility traits in spring-calving Aberdeen Angus cattle. 1. Model development and genetic parameters. **Journal of Animal Science**, v.85, n.11, p.2854-2860, 2007b.

Chapter 3

Genetic parameter estimates for age at first calving and visual scores in Brazilian Brahman cattle through Bayesian multitrait models

ABSTRACT: We aimed to evaluate the genetic association between visual scores (body structure, precocity and muscularity) and age at first calving (AFC) using Bayesian multitrait models in Brazilian Brahman cattle. A total of 7,539 records of body structure, precocity and muscularity and 50,630 records of AFC were used to estimate genetic parameters. Heritability estimates were 0.59, 0.44, 0.38 and 0.20 for body structure, precocity, muscularity and AFC using bitrait models; and 0.60, 0.44, 0.40 and 0.20, respectively, using full multitrait model. The genetic correlations were 0.57 between body structure and precocity, 0.56 between body structure and muscularity, and 0.82 between precocity and muscularity. The genetic correlations between visual scores and AFC were negative and moderate (-0.29, -0.24 and -0.31 for body structure, precocity and muscularity using bitrait model and -0.29, -0.22 and -0.29 for body structure, precocity and muscularity using full multitrait model). The results indicate that visual scores can be used as extra selection criteria in Brahman breeding programs since favorable correlated responses with age at first calving were observed.

Keywords: body structure, precocity, muscularity, bitrait, multitrait

INTRODUCTION

The advantage of including visual scores in breeding programs is that a large number of animals can be evaluated without being subjected to the stress of measurements, a fact that makes the process faster and more economically feasible (Jorge Júnior et al., 2001, 2004). According to Koury Filho et al. (2009) studies on visual scores are relevant to understand genetic correlations with other traits of interest, such as age at first calving (AFC).

Since visual scoring of body structure, precocity and muscularity is a recent visual assessment method, there are few studies correlating these traits with other economically important traits (for example, age at first calving - AFC). Studies estimating genetic parameters for these visual scores are therefore needed. In Brazil, there are few scientific studies approaching Brahman cattle, a fact that makes investigations on genetic breeding relevant for the development of this breed in the country (Bertipaglia et al., 2012).

The inclusion of reproductive traits as selection criteria is fundamental due to predominant poor fertility in Zebu breeds, characterized by a long postpartum anestrous period (Nava-Trujillo et al., 2010). Thus, the attempt is to select for sexual precocity in one of the most important fertility trait, the age at first calving (AFC), since lower AFC values are associated with heifer precocity, high lifetime productivity, increase in the number of calves and allows higher genetic progress rate (Bazzoli et al., 2014).

The use of appropriate methodologies to estimate genetic correlations between categorical morphological and continuous reproduction traits through multitrait framework has great interest in animal breeding (Faria et al., 2009a). However, studies correlating visual scores with reproductive performance of cows are scarce in the literature (Faria et al. 2009b; Boligon and Albuquerque 2010). Thus, we aimed to estimate genetic parameters between age at first calving and visual scores (body structure,

precocity and muscularity) using linear bitrait and full multitrait Bayesian models in Brazilian Brahman cattle by accessing predictive performance via cross-validation.

MATERIALS AND METHODS

Data

Brahman fertility and visual scores data were provided by Brazilian Association of Zebu Cattle Breeders (ABCZ). The categorical traits of visual scores of body structure (S), precocity (P) and muscularity (M), and the reproductive trait age at first calving (AFC) were studied.

Age at first calving was defined as the interval between birth and first parity of the cows. These records were obtained during the period from 1995 to 2012. Before editing the data, a total of 59,929 traits records were available in the database. The total number of animals in the pedigree was 61,616. Data from females AFC ranged from 731 to 1824 days of age. Data editing was performed by removing: animals with incomplete records; single record by contemporary groups (CG); outliers based on three standard deviation within CG. The animals were visually evaluated and received scores varying from one to six for the traits: body structure, precocity and muscularity. The animal that was considered to be intermediate to the traits (body structure, precocity and muscularity) received the score three or four, the animals that were considered inferior received the scores one or two and the best animals for the traits (S, P and M) received highest scores.

Records of visual scores of body structure, precocity and muscularity were collected according with the method of Morphological Evaluation System (MES, Sistema de Avaliação Morfológica - SAM) developed by company Brasilcomz, which applies modern procedures to collecting data on visual scores. S is evaluated by the quantity of meat in the carcass, using measurement of body length and height of the animal, with

larger animals receiving higher scores; P is evaluated by the measurement of the ratio of rib depth to limb height, with higher scores corresponding to animals that will deposit fat earlier; and M is evaluated by the determination of muscle distribution, volume and length, with animals with more convex musculature receiving higher scores. These scores were assigned to each animal. On the figure 1 are shown the distribution of body structure, precocity and musculature scores.

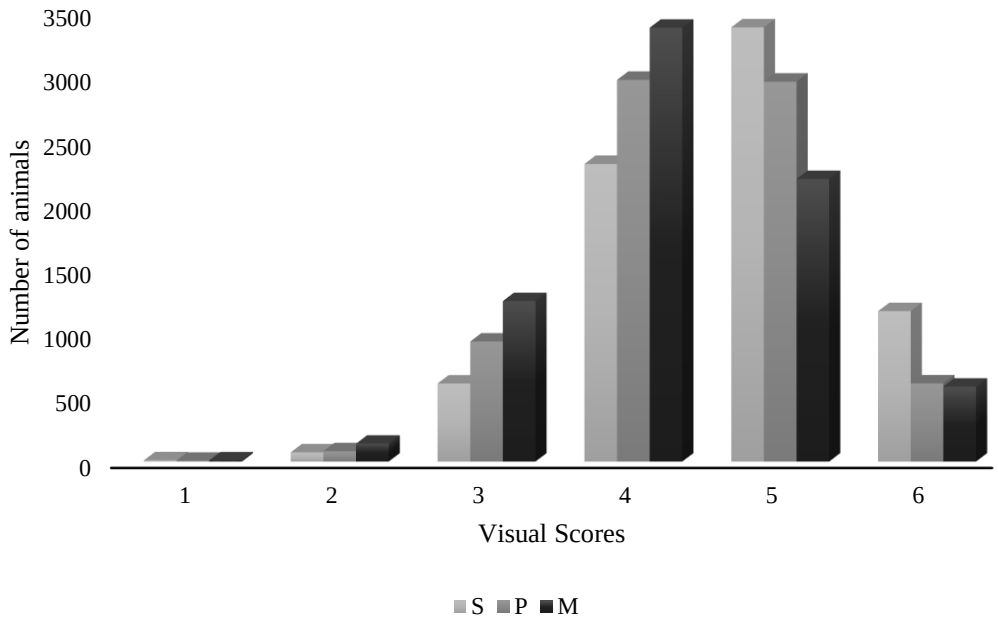

Figure 1. Distribution of visual body structure (S), precocity (P) and muscularity (M) scores in 18,530 Brahman females (worse = 1; best = 6). For the visual score one, the number of animals were 10, 6 and 9 respectively.

The visual scores were obtained at the sobreano (550 days) and varied from 490 days to 610 days, thus the definition of contemporary groups for visual scores traits (CG_{VS}) were defined, taking into account the farm, year, season of birth, management lot and the diet. And the contemporary groups for age at first calving (CG_{AFC}) were formed as the combination of herd, year and birth season. CG with eight or more animals with

phenotypic information were kept in the database for analysis. The descriptive statistics for the studied traits are shown in Table 1.

Table 1. Descriptive statistics of body structure (S), precocity (P), muscularity (M) and age at first calving (AFC) in Brahman cattle

Traits	N	GC	Mean	SD	CV (%)	Minimum	Maximum
AFC (days)	50,630	186	1,189.97	239.89	20.16	731	1824
S (points)	7,539	96	4.65	0.88	18.86	1	6
P (points)	7,539	96	4.40	0.85	19.30	1	6
M (points)	7,539	96	4.24	0.89	21.00	1	6

N - number of observations; SD - standard deviation; CV - coefficient of variation (%).

Methodologies

The (co)variance components and genetic parameters were estimated using bitrait and full multitrait Bayesian models. A total of four models were fitted, being three bitrait (AFC with S, P and M) and one multitrait (AFC, S, P and M):

The following general animal multitrait model was considered:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Z}_1\boldsymbol{a} + \boldsymbol{Z}_2\boldsymbol{cg}_{AFC} + \boldsymbol{Z}_3\boldsymbol{cg}_{VS} + \boldsymbol{e}, \quad [1]$$

where $\boldsymbol{y}$ is the vector containing records for the traits AFC body structure (S), precocity (P) and muscularity (M) scores; $\boldsymbol{\beta}$ is the vector of systematic effects (mean, registration class and mating type); $\boldsymbol{cg}_{AFC}$ is the vector of contemporary group for age at first calving (herd-year-season) effects; $\boldsymbol{cg}_{VS}$ is the vector of contemporary group for visual scores (herd-year-season-diet) effects; $\boldsymbol{a}$ is the vector of additive genetic effects; $\boldsymbol{e}$ is the residual vector; $\boldsymbol{X}$, $\boldsymbol{Z}_1$, $\boldsymbol{Z}_2$, and $\boldsymbol{Z}_3$ are the incidence matrices associated with $\boldsymbol{\beta}$, $\boldsymbol{a}$, $\boldsymbol{cg}_{AFC}$, and $\boldsymbol{gc}_{VS}$, respectively. It was assumed that $\boldsymbol{\beta} \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}_\beta \otimes \boldsymbol{I})$, being $\boldsymbol{\Sigma}_\beta$ a known diagonal matrix

with values 1e+10 (large variances) to represent vague prior knowledge, $\boldsymbol{cg}_{\text{AFC}} \mid \boldsymbol{\Sigma}_{cg_{\text{AFC}}} \sim \text{N}(\mathbf{0}, \boldsymbol{\Sigma}_{cg_{\text{AFC}}} \otimes \text{I})$, $\boldsymbol{cg}_{\text{VS}} \mid \boldsymbol{\Sigma}_{cg_{\text{VS}}} \sim \text{N}(\mathbf{0}, \boldsymbol{\Sigma}_{cg_{\text{VS}}} \otimes \text{I})$, $\boldsymbol{a} \mid \boldsymbol{\Sigma}_a, \mathbf{A} \sim \text{N}(\mathbf{0}, \boldsymbol{\Sigma}_a \otimes \mathbf{A})$, and $\mathbf{e} \mid \boldsymbol{\Sigma}_{\text{e}} \sim \text{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\text{e}} \otimes \mathbf{I})$, being $\mathbf{A}$ the numerator relationship matrix, $\boldsymbol{\Sigma}_{\text{cg}_{\text{AFC}}}$ is the contemporary group for age at first calving (co)variance matrix, $\boldsymbol{\Sigma}_{\text{cg}_{\text{VS}}}$ is the contemporary group fo visual scores (co)variance matrix, $\boldsymbol{\Sigma}_a$ is the additive genetic (co)variance matrix, $\boldsymbol{I}$ the identity matrix, and $\boldsymbol{\Sigma}_e$ is the residual (co)variance matrix. The vector $\boldsymbol{\beta}$ included registration class (registered animals as pure by origin or in open book) and mating type (artificial insemination, embryo transfer, fertilization *in vitro*, natural mating and controlled mating). For (co)variances matrices of random effects the inverted Wishart was defined as *prior* distribution. Thus, $\boldsymbol{y}' = [\boldsymbol{y}_{\text{AFC},} \ \boldsymbol{y}_{\text{VS}}]'$ is a vector in which $\boldsymbol{y}_{\text{AFC}}$ represents the vector of AFC records and $\boldsymbol{y}_{\text{VS}}$ the vector of visual scores records (S, P or M and the three scores together).

Markov Chain Monte Carlo (MCMC) Sampling

Parameters were drawn from the posterior distributions using Gibbs sampling using the TM (Legarra et al., 2008) software. A total of 400,000 samples were generated, assuming a burn-in period and sampling interval (thin) of 100,000 and 10 iterations, respectively. The convergence of the MCMC chains was verified by graphical inspection and BOA (Smith, 2007) R software. Convergence was assessed using the Heidelberg e Welch (1983), Geweke (1992) and Raftery and Lewis (1992) methods.

Models Comparisons

The predictive ability was accessed by cross-validation, which involved training one subset of the population (about 70% of the animals), and validating on remaining

individuals. Here, we randomly split the data sets into two groups from the original data set, as specified above, these two subset were redefined 10 times, D1, D2, ..., D10. Finally, the average of the 10 correlation coefficients between the predicted and observed phenotypes was obtained.

The predicted phenotype vector was calculated as $\hat{\boldsymbol{y}} = \boldsymbol{X}\hat{\boldsymbol{\beta}} + \boldsymbol{Z}_1\hat{\boldsymbol{a}} + \boldsymbol{Z}_2\boldsymbol{c}\hat{\boldsymbol{g}}_{AFC} + \boldsymbol{Z}_3\boldsymbol{c}\hat{\boldsymbol{g}}_{VS}$. Thus, the solutions for the animals in the validation population were obtained based on the solutions of the training population animals. Finally, the predictive ability used to measure the efficiency obtained by full multitrait and bitrait models was given by the correlations between observed and predicted phenotypes from the validation population.

RESULTS

Variance Components

Posterior means with respective standard deviation and HPD95% region for variance components obtained by full multitrait and bitrait models are presented in table 2. In general, the estimates provided by these two different models were very similar and showed overlapping through HPD intervals.

Heritabilities, Genetic and Phenotypic correlations

The heritability estimates with respective standard deviation and HPD95% region, genetic and phenotypic correlations among AFC, body structure, precocity and muscularity obtained by multitrait and bitrait analyses are presented in Table 3. The heritabilities reported in the present study for the visual scores traits presented high magnitude, being higher for body structure (~ 0.60) when compared to precocity (~ 0.44) and muscularity (~ 0.40). The heritability estimate for AFC was moderate (around 0.20) for all analysis.

The genetic and phenotypic correlations between visual scores traits were medium to high, ranged from 0.57 to 0.82 and 0.48 to 0.67, respectively. Both the genetic and phenotypic correlations among AFC, body structure, precocity and muscularity were medium magnitude and negative in both bitrait and full multitrait models (Table 3).

Table 2. Posterior means, standard deviation and highest posterior density region (95%) for variance components for age at first calving (AFC), body structure (S), precocity (P) and muscularity (M) obtained by full multitrait and bitrait Bayesian models in Brahman cattle

Additive genetic variance (σ^2_a)				
Models	**AFC**	**S**	**P**	**M**
Full Multitrait	11,994.53 (597.89) [10,812.34; 13,153.76]	0.48 (0.04) [0.40; 0.55]	0.32 (0.03) [0.26; 0.37]	0.30 (0.03) [0.23; 0.35]
Bitrait AFC-S	11,958.91 (595.70) [10,816.70; 13,137.72]	0.48 (0.03) [0.40; 0.54]	—	—
Bitrait AFC-P	12,041.68 (594.40) [10,911.03; 13,225.75]	—	0.31 (0.03) [0.25; 0.37]	—
Bitrait AFC-M	11,974.18 (596.13) [10,857.16; 13,169.28]	—	—	0.28 (0.03) [0.22; 0.35]
Contemporary group variance (σ^2_{cg})				
Full Multitrait	5,453.07 (737.26) [4,083.47; 6,947.96]	0.03 (0.008) [0.01; 0.04]	0.03 (0.008) [0.01; 0.04]	0.03 (0.01) [0.01; 0.05]
Bitrait AFC-S	5,449.61 (734.82) [4,055.81; 6,933.64]	0.02 (0.006) [0.009; 0.03]	—	—
Bitrait AFC-P	5,451.43 (735.03) [4,093.95; 6,974.59]	—	0.02 (0.007) [0.01; 0.03]	—
Bitrait AFC-M	5,437.82 (732.90) [4,072.32; 6,939.07]	—	—	0.02 (0.008) [0.01; 0.04]
Residual variance (σ^2_e)				
Full Multitrait	43,268.98 (516.59) [42,254.55; 44,279.37]	0.31 (0.02) [0.26; 0.36]	0.40 (0.02) [0.35; 0.43]	0.45 (0.02) [0.40; 0.49]
Bitrait AFC-S	43,280.62 (516.10) [42,233.85; 44,258.35]	0.31 (0.02) [0.27; 0.36]	—	—

Bitrait AFC-P	43,219.68 (515.33) [42,186.79; 44,216.91]	—	0.40 (0.02) [0.35; 0.44]	—
Bitrait AFC-M	43,273.27 (516.19) [42,217.40; 44,243.40]	—	—	0.45 (0.02) [0.40; 0.49]

SPM - body structure, precocity and muscularity together

Table 3. Heritability (diagonal), standard deviation and highest posterior density region (95%), genetic (above the diagonal) and phenotypic correlations (below the diagonal) for age at first calving (AFC), body structure (S), precocity (P) and muscularity (M) obtained by full multitrait and bitrait Bayesian models in Brahman cattle

Full Multitrait model				
Traits	**AFC**	**S**	**P**	**M**
AFC	0.20 (0.01) [0.17, 0.21]	-0.29 (0.06) [-0.40, -0.18]	-0.22 (0.05) [-0.32, -0.13]	-0.29 (0.04) [-0.41, -0.15]
S	-0.19 (0.02) [-0.22, -0.15]	0.60 (0.04) [0.53, 0.67]	0.57 (0.05) [0.48, 0.66]	0.56 (0.05) [0.46, 0.66]
P	-0.17 (0.02) [-0.20, -0.13]	0.48 (0.01) [0.45, 0.50]	0.44 (0.04) [0.37, 0.51]	0.82 (0.03) [0.76, 0.87]
M	-0.22 (0.02) [-0.25, -0.18]	0.49 (0.01) [0.46, 0.51]	0.67 (0.008) [0.66, 0.69]	0.40 (0.03) [0.33, 0.46]
Bitrait model				
Traits	**$h_{1,2}$**	**h_{22}**	**$rG_{1,2}$**	**$rP_{1,2}$**
AFC-S	0.20 (0.01) [0.18, 0.21]	0.59 (0.03) [0.52, 0.66]	-0.29 (0.05) [-0.40, -0.18]	-0.18 (0.01) [-0.21, -0.15]
AFC-P	0.20 (0.01) [0.18, 0.21]	0.44 (0.03) [0.36, 0.50]	-0.24 (0.04) [-0.32, -0.14]	-0.17 (0.02) [-0.19, -0.13]
AFC-M	0.20 (0.01) [0.18, 0.21]	0.38 (0.03) [0.31, 0.46]	-0.31 (0.06) [-0.44, -0.18]	-0.22 (0.02) [-0.25, -0.18]

$h_{1,2}$ – heritability of AFC in the presence of S, P or M; $h_{2,2}$ – heritability of S, P or M in the presence of AFC; $rG_{1,2}$ – genetic correlation between AFC and S, P or M; $rP_{1,2}$ – phenotypic correlation between AFC and S, P or M

Figure 2 shows the posterior distributions of the genetic correlations between AFC and body structure, precocity and muscularity scores obtained by bitrait analyzes.

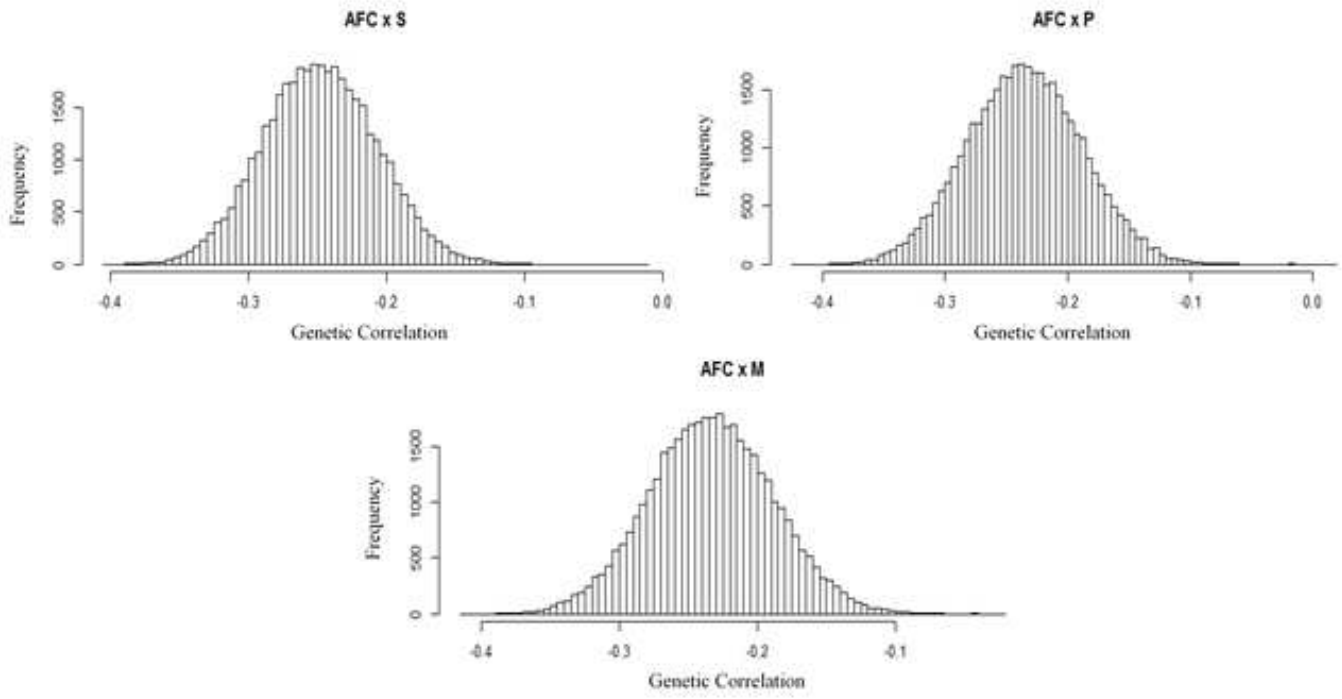

Figure 2. Posterior density of genetic correlations of age at first calving (AFC) with scores of body structure (S), precocity (P) and musculaturity (M) in Brahman cattle by using bitrait models.

Model Comparison

The predictive ability was performed by correlation coefficients (between observed and predicted phenotype) using cross-validation approach (Table 4). The lower standard deviations for this statistic suggest high precision of the cross-validation inference. The correlations were similar between traits, indicating that both models (full multitrait or bitrait) are recommended to be used in AFC genetic evaluation with visual scores body structure, precocity and muscularity in Brazilian Brahman cattle.

Table 4. Correlations (with respective posterior standard deviation) between predicted and observed phenotypes from 10 fold of the cross-validation from full multitrait and bitrait Bayesian models in Brahman cattle.

Models	AFC	S	P	M
Full Multitrait	0.13 (0.08)	0.96 (0.02)	0.97 (0.01)	0.77 (0.20)
Bitrait AFC-S	0.10 (0.09)	0.95 (0.06)	__	__
Bitrait AFC-P	0.10 (0.09)	__	0.96 (0.04)	__
Bitrait AFC-M	0.10 (0.10)	__	__	0.97 (0.01)

AFC – age at first calving; S – body structure; P – precocity; M – muscularity; SPM - body structure, precocity and muscularity together

DISCUSSION

The full multitrait and bitrait models presented similar results in terms of genetic parameter estimates since the regions of credibility have overlapped.

The heritability estimate for AFC (around 0.20) is considered moderate to high for fertility traits. Heritability estimates for AFC from field data reported in the literature ranging from 0.10 to 0.37 in Brazilian Zebu cattle (Boligon and Albuquerque, 2011; Barrozo et al., 2012; Moreira et al., 2015).

Studies estimating heritabilities for body structure, precocity and muscularity in Brahman cattle are scarce. It was noted that heritability estimates found for these traits were high (Table 3), thus, it is expected a high response for direct selection. These results are in agreement with those found by Faria et al. (2009a) that evaluated visual scores in Nellore cattle and also obtained high heritability for body structure (0.68), precocity (0.65) and muscularity (0.62). In order to evaluate the possible use of visual scores as selection criteria to improve carcass quality in Brahman cattle, Bertipaglia et al. (2012) found smaller heritability values for body structure, precocity and muscularity, 0.39, 0.43 and 0.40, respectively. Low heritability estimates have been reported by Shiotsuki et al.

(2009) for conformation (0.15), finishing (0.21) and muscling (0.23) of Nellore heifers exposed to reproduction at 16 months of age.

Cardoso et al. (2004) stated that the differences in heritability estimates for visual scores observed between studies might be due to inconsistencies in the evaluation systems, which vary among examiners and breeding programs. Differences in the estimation models may also affect the magnitude of heritability estimates.

The mean heritabilities for visual scores of body structure, precocity and muscularity were of high magnitude (Table 3) indicating that great part of the variation in these traits are due to genes with additive effects (Falconer and Mackay, 1996). Consequently, their adoption as selection criteria will be effective in this population and expressive genetic gain could be achieved in breeding programs.

The genetic correlations between visual scores of body structure, precocity and muscularity were high (0.57 to 0.82). These findings are in agreement with Koury Filho et al. (2009) that estimated genetic correlations of 0.49, 0.63 and 0.90, S and P, S and M, and P and M, respectively in Nellore cattle. In Brahman cattle, Bertipaglia et al. (2012) found positive association between visual scores, with genetic correlations ranging from 0.79 to 0.91. The high values of genetic correlations estimates indicate that one trait captures a high proportion of the genetic variance of the other two and would be enough in a selection scheme aimed on the improving of the three traits. However, including the three traits in one selection index will have the benefit of avoiding that individuals with extreme scores for one trait will not be seleced even if it would have high values in the other two.

The estimated genetic correlations between visual scores and age at first calving were negative and presented medium magnitude (Table 3) for both models (bitrait and full multitrat). In this sence, the selection of animals with better body composition will

bring as correlated response more eficiente animals in terms of age at first calving. These results corroborate with those obtained by Boligon and Albuquerque (2010), that estimated genetic correlations between visual scores and age at first calving varying from -0.23 to -0.29 in Nellore heifers. Boligon et al. (2012) reported lower genetic correlations between conformation, finishing precocity, muscling and days to calving (-0.11, -0.19 and -0.16 respectively).

The negative genetic correlation, in this case, implies in a favorable association between traits (visual scores and AFC). Although estimates of the genetic correlations obtained have been to vary from -0.22 to 0.31, the selection for animals with the best biotype may lead to favorable responses for AFC. The selection of animals for a desirable biotype, evaluated by visual scores, can to lead animals with higher sexual precocity. The genetic correlations recorded for AFC with the traits of muscularity (M), body structure (S) and precocity (P) in Brahman cattle presenting moderate magnitudes, being favorable to selection. Animals with desirable biotype will present greater fertility and sexual precocity, indicating that selection for visual scores will promote reduction in the age at first calving, a fact that would be beneficial for the Brazilian production systems.

In average, the age at first calving in Brazilian beef cattle is higher than 40 months (Barbosa et al., 2015). Thus, the identification of females that conceive at younger ages should be one of the priorities of the most breeding programs in Brazil. It is relevant to mention that conception is a trait more related to body weight than to the age of the animal, since it is common practice in many herds to adopt minimum weight for the entry of females into reproduction (Mercadante et al., 2000).

The phenotypic correlations of visual scores with AFC were negative and presented medium magnitud (Table 3), nevertheless, may suggest that the improvement in enviromental conditions for visual scores does not almost interfere in the age at first

calving of cows. Since heritability estimates between AFC and visual scores were medium to high magnitude (0.20) for reproductive traits, the phenotypic correlation is mainly determined by genetic correlation. In this context, Boligon et al. (2012) found phenotypic correlations of visual scores with subsequent rebreeding and days to calving close to zero in Nellore heifers. Bertipaglia et al. (2012) reported phenotypic correlations close to zero between visual scores and scrotal circumference (-0.0002 to 0.03) in Brahman cattle. Consequently, joint selection for visual scores and reproductive traits (AFC) induces a favorable genetic correlation among the traits without necessarily expressing a phenotypic association.

A selection index considering visual scores will increase, at long term, the frequency of desired genes for AFC, thus improving cow reproductive performance. Therefore, visual scores are alternative traits to compose new selection indexes, since they present sufficient genetic variability to promote genetic progress. However, studies in this area are scarce, and further investigations are necessary.

Correlation coefficient between predicted and observed phenotypic values were used to access the "predictive ability" (Table 4). The correlations were similar among traits, indicating that both models (full multitrait or bitrait) are recommended to be used in AFC genetic evaluation with visual scores. Finally, based in our study, we conclude that using of the full multitrait model would be the best choice, since the results were very similar to those obtained with the bitrait model. Thus, it would reduce the number of analyzes to be performed.

CONCLUSIONS

The visual scores of body structure, precocity and muscularity can be used as selection criteria, once they show high heritability.

Direct selection for visual scores together with female reproductive trait is recommended to improve carcass composition and increase the fertility of beef cows.

Full multitrait model would be the best option for genetic evaluation of Brazilian Brahman cattle for Bayesian models, since the number of analyzes would be reduced.

LITERATURE CITED

BARBOSA, A.C.B.; OLIVEIRA, I.R.; MALHADO, A.C.M. et al. Genetic improvement of Nellore cattle in the semi-arid region of northeastern Brazil. **Acta Scientiarum. Animal Sciences,** v.37, n.4, p.429-435, 2015.

BARROZO, D.; BUZANSKAS, M.E.; OLIVEIRA, J.A. et al. Genetic parameters and environmental effects on temperament score and reproductive traits of Nellore cattle. **Animal**, v.6, n.1, p.36-40, 2012.

BAZZOLI, I.; De MARCHI, M.; CECCHINATO, A. et al. Factors associated with age at slaughter and carcass weight, price, and value of dairy cull cows. **Journal of Dairy Science**, v. 97, n. 2, p. 1082-1091, 2014.

BERTIPAGLIA, T.S.; CARREÑO, L.O.D.; MACHADO, C.H.C.; et al. Estimates of genetic parameters for visual scores and their correlation with production and reproductive traits in Brahman cattle. **Revista Brasileira de Zootecnia**, v.41, n.6, p.1407-1411, 2012.

BOLIGON, A.A. and ALBUQUERQUE, L.G. Genetic correlations between visual scores and reproductive traits in Nelore cattle using Bayesian inference. **Pesquisa agropecuária brasileira**, v.45, n.12, p.1412-1418, 2010.

BOLIGON, A.A. and ALBUQUERQUE, L.G. Genetic parameters and relationships of heifer pregnancy and age at first calving with weight gain, yearling and mature weight in Nelore cattle. **Livestock Science**, v.141, n.1, p.12-16, 2011.

BOLIGON, A.A.; AYRES, D.R.; PEREIRA, R.J.; et al. Genetic associations of visual scores with subsequente rebreeding and days to first calving in Nellore cattle. **Journal of Animal Breeding and Genetics**, v.129, p.448-456, 2012.

CARDOSO, F.F.; CARDELLINO, R.A.; CAMPOS, L.T. (Co) variance components and genetic parameters of post-weaning traits in Angus cattle. **Revista Brasileira de Zootecnia**, v.33, n.2, p.313-319, 2004.

FALCONER, D.S.; MACKAY, T.F.C. **Introduction to quantitative genetics.** 4.ed. Edinburgh: Longman Group Limited, 1996. 464p.

FARIA, C.U.; KOURY FILHO, W.; MAGNABOSCO, C.U.; et al. Bayesian inference in genetic parameter estimation of visual scores in Nellore beef-cattle. **Genetics and Molecular Biology**, v.32, n.4, p.753-760, 2009a.

FARIA, C.U.; MAGNABOSCO, C.U.; ALBUQUERQUE, L.G.; et al. Bayesian analysis in the estimation of genetic correlations between visual scores and reproductive traits in Nelore cattle using linear-threshold models. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v.61, n.4, p.949-958, 2009b.

GEWEKE, J. Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments. In: **Bayesian Statistics 4**, Oxford: Oxford University Press, 1992. p. 169- 193.

HEIDELBERGER, P. AND WELCH, P.D. Simulation Run Length Control in the Presence of an Initial Transient. **Operations Research** 31:1109-1144. 1983.

JORGE JÚNIOR, J.; DIAS, L.T.; ALBUQUERQUE, L.G. Fatores de correção de escores visuais de conformação, precocidade e musculatura, à desmama, para idade da vaca ao parto, data juliana de nascimento e idade à desmama em bovinos da raça Nelore. **Revista Brasileira de Zootecnia**, v.33, n.6, p.2044-2053, 2004 (supl. 2).

JORGE JÚNIOR, J.; PITA, F.V.C.; FRIES, L.A. et al. Influência de alguns fatores de ambiente sobre os escores de conformação, precocidade e musculatura à desmama em um rebanho da raça Nelore. **Revista Brasileira de Zootecnia**, v.30, n.6, p.1697-1703, 2001.

KOURY FILHO, W.; ALBUQUERQUE, L.G.; ALENCAR, M.M. et al. Estimativas de herdabilidade e correlações para escores visuais, peso e altura ao sobreano em rebanhos da raça Nelore. **Revista Brasileira de Zootecnia**, v.38, n.12, p.2362-2367, 2009.

LEGARRA, A.; VARONA, L.; LÓPEZ DE MATURANA, E. **TM: Threshold Model.** (http://snp.toulouse.inra.fr/~alegarra), 2008.

MERCADANTE, M.E.Z.; LÔBO, R.B.; OLIVEIRA, H.N. Estimates of (Co)Variances among Reproductive and Growth Traits in Female Nellore Cattle. **Revista Brasileira de Zootecnia**, v.30, p.997-1004, 2000.

MOREIRA, H.L.; BUZANSKAS, M.E.; MUNARI, D.P.; et al. Reproductive traits selection in Nelore beef cattle. **Ciência e Agrotecnologia**, v.39, n.4, p.355-362, 2015.

NAVA-TRUJILLO, H.; SOTO-BELLOSO, E.; HOET, A.E. Effects of clinical mastitis from calving to first service on reproductive performance in dual-purpose cows. **Animal Reproduction Science**, v. 121, n. 1, p. 12-16, 2010.

RAFTERY, A.E. AND LEWIS, S. How Many Iterations in the Gibbs Sampler? In: **Bayesian Statistics 4**, Oxford: Oxford University Press, 1992b. p. 763-773.

SHIOTSUKI, L.; SILVA, J.A.V. II; TONHATI, H.; et al. Genetic associations of sexual precocity with growth traits and visual scores of conformation, finishing, and muscling in Nelore cattle. **Journal of Animal Science**, v.87, p.1591-1597, 2009.

SMITH, B.J. BOA: An R Package for MCMC Output Convergence Assessment and Posterior Inference. **Journal of Statistical Software**, 21(11), 2007.

.

GENERAL CONCLUSIONS

In the pig data, the model including marker and pedigree information had better goodness-of-fit than pedigree-based or marker-based models. The heritability estimates for mature weight (“a”) and maturity rate (“k”) indicated that these traits is a feasible alternative for breeding programs aiming to change the shape of growth curves in pig breeding programs.

The multitrait GWAS was efficient to report QTLs associated with functions related to biological processes of growth in pigs. Relevant SNPs are located in genome regions not previously described in the literature.

In Brahman cattle, the age at first calving censored data could be incorporated in genetic evaluations through a linear model. Given the heritability estimates, individual selection should imply in genetic gains for visual scores traits (body structure, precocity and muscularity) and age at first calving. The direct selection for visual scores together with female reproductive trait (AFC) is recommended to improve carcass composition and increase the fertility of beef cows.