

## Quadrados mínimos parciais uni e multivariado aplicados na seleção genômica para características de carcaça em suínos

Uni and multivariate partial least squares applied to genomic selection for carcass traits in pigs

Camila Ferreira Azevedo<sup>I</sup> Fabyano Fonseca e Silva<sup>I</sup> Marcos Deon Vilela de Rezende<sup>II</sup>  
Luiz Alexandre Peternelli<sup>I</sup> Simone Eliza Facione Guimarães<sup>III</sup> Paulo Sávio Lopes<sup>III</sup>

### RESUMO

A principal contribuição da genética molecular é a utilização direta das informações de DNA no processo de identificação de indivíduos geneticamente superiores. Sob esse enfoque, idealizou-se a seleção genômica ampla (Genome Wide Selection - GWS), a qual consiste na análise de marcadores SNPs (Single Nucleotide Polymorphisms) amplamente distribuídos no genoma. Devido a esse grande número de SNPs, geralmente maior que o número de indivíduos genotipados, e à alta colinearidade entre eles, métodos de redução de dimensionalidade são requeridos. Dentre estes, destaca-se o método de regressão via Quadrados Mínimos Parciais (Partial Least Squares - PLS), que além de solucionar tais problemas, permite uma abordagem multivariada, considerando múltiplos fenótipos. Diante do exposto, objetivou-se aplicar e comparar a regressão PLS univariada (UPLS) e multivariada (MPLS) na GWS para características de carcaça em uma população  $F_2$  de suínos Piau×Comercial. Os resultados evidenciaram a superioridade do método MPLS, uma vez que este proporcionou maiores valores de acurácia em relação à abordagem univariada.

**Palavras-chave:** seleção genômica, redução de dimensionalidade, análise multivariada.

### ABSTRACT

The main contribution of molecular genetics is the direct use of DNA information to identify genetically superior individuals. Under this approach, genome-wide selection (GWS) can be used with this purpose. GWS consists in analyzing of a large number of SNP markers widely distributed in the genome, and due to the fact that the number of markers is much larger than the number of genotyped individuals and also to the fact that such markers are highly correlated special statistical methods, like

Partial Least Squares (PLS), are widely required. Thus, the aim of this paper was to propose an application of Uni (UPLS) and Multivariate (MPLS) Partial Least Squares regression to GWS of carcass traits in an  $F_2$  (Piau × commercial) pig population. The results showed that MPLS method provided most accurate genomic breeding values estimates than univariate method.

**Key words:** genomic selection, dimensionality reduction, multivariate analysis.

### INTRODUÇÃO

O principal atrativo da genética molecular em prol do melhoramento animal é a utilização direta das informações provenientes do DNA no processo de identificação de indivíduos geneticamente superiores. Diante de novas classes de marcadores moleculares, como o SNP (Single Nucleotide Polymorphisms), MEUWISSEN et al. (2001) idealizaram a seleção genômica ampla (Genome Wide Selection - GWS), a qual visa o aumento da acurácia de seleção por meio da utilização de um grande número de marcadores amplamente distribuídos no genoma.

Devido à alta densidade de marcadores, é possível assumir que alguns deles estejam em desequilíbrio de ligação com QTLs (Quantitative Trait Loci), possibilitando sua utilização direta na estimação do valor genético dos indivíduos, inclusive os não fenotipados. No âmbito da seleção genômica,

<sup>I</sup>Departamento de Estatística, Centro de Ciências Exatas, Universidade Federal de Viçosa (UFV), Avenida P.H. Rolfs, s/nº, 36570-000, UFV, Viçosa, Minas Gerais, Brasil. Email: camila.azevedo@ufv.br. Autor para correspondência.

<sup>II</sup>Embrapa Florestas, Departamento de Engenharia Florestal, UFV, Viçosa, Minas Gerais, Brasil.

<sup>III</sup>Departamento de Zootecnia, Centro de Ciências Exatas, UFV, Viçosa, Minas Gerais, Brasil.

a aplicação prática dessas informações é um desafio, pois geralmente não é possível a utilização adequada de métodos tradicionais baseados em quadrados mínimos (*Least Squares*) para estimar o efeito de cada SNP no fenótipo, uma vez que geralmente o número de marcadores é maior que o número de animais genotipados (alta dimensionalidade) e alguns marcadores são altamente correlacionados (multicolinearidade) devido ao desequilíbrio de ligação. Dessa forma, métodos estatísticos baseados na redução de dimensionalidade são requeridos. Dentre estes, destaca-se o método de regressão via Quadrados Mínimos Parciais (*Partial Least Squares* - PLS), que além de solucionar esses problemas, também permite uma abordagem multivariada, considerando múltiplos fenótipos, a qual não pode ser implementada diretamente em métodos tradicionais de seleção genômica, tais como Bayes A, Bayes B e LASSO Bayesiano, devido à complexidade desses métodos.

Na suinocultura moderna, as características de carcaça são importantes por estarem diretamente relacionadas com as exigências do mercado consumidor (ROSA et al., 2008; ZANGERONIMO et al., 2009), o qual tende a optar pela baixa deposição de gordura. Porém, em programas de melhoramento, essa tendência deve ser contemplada simultaneamente com outras variáveis, tais como rendimento e comprimento de carcaça. Assim, estimativas do mérito genético dos indivíduos, considerando ou não informações genômicas, geralmente, são estimadas com maior acurácia, quando provenientes de análises que consideram simultaneamente todas essas características, ou seja, sob um enfoque multivariado (SILVA et al., 2011). Embora alguns trabalhos envolvendo seleção genômica para características de carcaça em suínos já tenham sido publicados (FONTANESI et al., 2010; UIMARI et al., 2011), até o momento não há relatos na literatura a respeito de abordagens multivariadas.

Diante do exposto, o presente trabalho teve por objetivo utilizar e comparar o método PLS uni e multivariado na seleção genômica para características de carcaça (espessuras de toucinho, bacon e rendimentos de carcaça) em uma população  $F_2$  de suínos (Piau × Comercial), utilizando um painel de marcadores SNPs.

## MATERIAL E MÉTODOS

Os dados fenotípicos e de marcadores utilizados no presente estudo são provenientes da Granja de Melhoramento de Suínos do Departamento de Zootecnia da Universidade Federal de Viçosa, em

Viçosa, Minas Gerais, Brasil, no período de novembro de 1998 a julho de 2001. Os dados fenotípicos foram corrigidos para os efeitos fixos de sexo, lote e presença do gene halotano e desregressados (RESENDE et al., 2010). Neste experimento, a população  $F_2$  é composta de 345 suínos provenientes do cruzamento de dois varrões da raça naturalizada brasileira Piau, com 18 fêmeas de linhagem desenvolvida na UFV, pelo acasalamento de animais de raça comercial (Landrace × Large White × Pietrain).

Os detalhes dos procedimentos utilizados, cuja extração do DNA foi realizada no Laboratório de Biotecnologia Animal do DZO da UFV, podem ser encontrados em PEIXOTO et al. (2006). A genotipagem foi realizada via tecnologia Golden Gate/Vera Code R, no Laboratório de Genética Animal (LGA), Embrapa Recursos Genéticos e Biotecnologia (CENARGEN), Brasília, DF, conforme descrito por HIDALGO et al. (2011).

Os marcadores SNPs estão distribuídos da seguinte forma nos cromossomos de *Sus scrofa*: SSC1 (56), SSC4 (54), SSC7 (59), SSC8 (31), SSC17 (25), totalizando assim 237 marcadores identificados para os animais  $F_2$ . Esses marcadores foram obtidos apenas em regiões nas quais se observaram a presença QTL em estudos prévios (HIDALGO et al., 2011) nessa mesma população, caracterizando assim um mapeamento fino apenas em regiões cromossômicas de interesse, o que explica o número reduzido de marcadores utilizados.

As características analisadas foram espessuras de toucinho medidas imediatamente após a última costela na linha dorso-lombar (ETUC); a 6,5 cm da linha dorso-lombar (ETO); entre a última e a penúltima vértebra lombar (ETUL); menor espessura na região acima da última vértebra lombar, na linha dorso-lombar (ETL); espessura de bacon (EBACON); comprimento de carcaça pelo método de classificação americano (MLC) e rendimento de carcaça (RCARC).

Quadrados Mínimos Parciais Multivariados (*Multivariate Partial Least Squares* – MPLS) consistem na obtenção de estimadores para as variáveis dependentes  $Y_1, \dots, Y_n$  a partir de componentes  $T_i$  ( $i = 1, \dots, p$ ), que são combinações lineares das variáveis explicativas  $X_1, \dots, X_m$  (marcadores). Sob esse enfoque, a equação de predição é expressa por:

$$Y_k = \beta_{k0} + \beta_{k1}T_1 + \beta_{k2}T_2 + \dots + \beta_{kp}T_p + e_k \quad (1)$$

em que  $Y_k$  é o vetor de dados fenotípicos da

característica  $k$  ( $k = 1, \dots, n$ ),  $\beta_{ki}$  são os coeficientes provenientes da regressão entre  $Y_k$  e os componentes  $T_i$  ( $i = 1, \dots, p$ ) e  $e$  é o vetor de resíduos. A principal característica do MPLS é que a correlação de qualquer par de componentes igual a zero, solucionando os problemas de multicolinearidade.

Generalizando, suponha que já foram determinados  $T_i$  ( $i \geq 1$ ),  $V_{ij}$  para  $j = 1, \dots, m$  e  $R_{ik}$  para  $k = 1, \dots, n$ . Por definição,  $V_{(i+1)j}$  é o resíduo proveniente da regressão entre  $T_i$  e  $V_{ij}$  e  $R_{(i+1)k}$  é o resíduo proveniente da regressão entre  $R_{ik}$  e  $T_i$  expressos, respectivamente, por:

$$V_{(i+1)j} = V_{ij} - \{t'_i v_{ij} / (t'_i t_i)\} T_i, \quad (2)$$

$$R_{(i+1)k} = R_{ik} - \{t'_i r_{ik} / (t'_i t_i)\} T_i, \quad (3)$$

em que  $t_i$  é o vetor de valores de  $T_i$ ,  $v_{ij}$  é o vetor de valores de  $V_{ij}$ ,  $r_{ik}$  é o vetor de valores de  $R_{ik}$ ,  $t'_i v_{ij} / (t'_i t_i)$  e  $t'_i r_{ik} / (t'_i t_i)$  os coeficientes da regressão. Se  $i=1$ ,  $V_{1j}$  é a variável centrada de  $X_j$  ( $V_{1j} = X_j - \bar{X}_j$ , para  $j = 1, \dots, m$ ) e  $R_{1k}$  é a variável centrada de  $Y_k$  ( $R_{1k} = Y_k - \bar{Y}_k$ , para  $k = 1, \dots, n$ ).

O processo é repetido sucessivamente a fim de obter as matrizes  $R_{(i+1)} = (r_{(i+1)1}, \dots, r_{(i+1)n})$  e  $V_{(i+1)} = (v_{(i+1)1}, \dots, v_{(i+1)m})$ . Assim, utilizando as matrizes  $R_{(i+1)}$  e  $V_{(i+1)}$ , um novo vetor,  $u_{i+1}$ , deve ser determinado por meio da multiplicação  $R_{i+1} c_{i+1}$ , sendo  $c_{i+1}$  o autovetor correspondente ao autovalor de  $R'_{i+1} V_{i+1} V'_{i+1} R_{i+1}$ . Segundo GARTHWAITE (1994), geralmente, define-se o componente  $T_{i+1}$  como sendo uma média ponderada dada por:

$$T_{(i+1)} = \sum_{j=1}^m w_{(i+1)j} \hat{b}_{(i+1)j} V_{(i+1)j}, \quad (4)$$

em que  $w_{(i+1)j} \propto \text{var}(V_{(i+1)j}) = (v'_{(i+1)j} v_{(i+1)j}) / (n-1)$  é o peso, e  $\hat{b}_{(i+1)j} = v'_{(i+1)j} u_{(i+1)} / (v'_{(i+1)j} v_{(i+1)j})$  é a estimativa do coeficiente de regressão entre  $U_{(i+1)}$  e  $V_{(i+1)j}$ .

O método é repetido para obter os componentes  $T_{(i+2)}$ ,  $T_{(i+3)}$ , ...,  $T_p$ , os quais são regredidos para cada  $Y_k$  e esses componentes, assim por meio do método dos quadrados mínimos ordinários (*Ordinary Least Squares* – OLS), são determinados os coeficientes do modelo (1), obtendo a seguinte equação de predição:

$$\hat{Y}_k = \hat{\beta}_{k0} + \hat{\beta}_{k1} \hat{T}_1 + \hat{\beta}_{k2} \hat{T}_2 + \dots + \hat{\beta}_{kp} \hat{T}_p, \quad (5)$$

sendo  $\hat{Y}_k$  o vetor de estimativas do valor genômico dos indivíduos para a característica  $k$ .

Os coeficientes associados aos componentes na equação (5) não possuem interpretação biológica, porém é possível estimar os coeficientes associados às variáveis originais (efeitos dos marcadores) combinando os modelos (5) e (4). Dessa forma, tem-se:

$$\hat{m}_{pls} = \hat{B} W \hat{\beta}, \quad (6)$$

em que  $W$  é a matriz de pesos,  $\hat{B}$  é matriz cujos elementos são os coeficientes provenientes da regressão de  $U_i$  em relação a  $V_{ij}$  e  $\hat{\beta}$  é o vetor de coeficientes  $\hat{\beta}_{ki}$ .

A principal diferença entre o MPLS e o método Quadrados Mínimos Parciais Univariado (*Univariate Partial Least Squares* - UPLS) está no modo como  $u_{i+1}$  é construído. No método UNPLS, é o vetor de resíduos provenientes da regressão entre  $Y$  (única variável dependente) e os componentes  $T_i$  ( $i = 1, \dots, p$ ). Enquanto no método MPLS aplica-se uma regressão entre cada variável  $Y_k$  (várias variáveis dependentes) e os componentes  $T_i$  ( $i = 1, \dots, p$ ) e  $u_{i+1}$  é visto como uma combinação linear entre os vetores de resíduos.

Como qualquer análise de regressão múltipla, um passo importante na execução do MPLS e UPLS é a determinação do número ótimo de componentes (novas variáveis regressoras,  $T_i$ ,  $i = 1, \dots, p$ , caracterizadas como combinações

lineares dos marcadores SNPs) que devem ser incluídos na equação (5). Neste estudo, o número ótimo de componentes, utilizados para a predição dos valores genômicos, foi escolhido como sendo o número mínimo que estabilizasse a curva do coeficiente de determinação ( $R^2$ ) em relação aos componentes.

A população  $F_2$  de suínos foi dividida em três populações distintas, cada qual com 115 indivíduos. Assim, duas dessas populações foram assumidas como população de treinamento e utilizadas para obter os efeitos dos marcadores SNPs. A outra população, denominada de validação, foi utilizada para avaliar a concordância entre valores genéticos preditos via estimativas provenientes da população de treinamento e fenótipos corrigidos observados.

A acurácia obtida para cada método foi calculada por meio da seguinte expressão (RESENDE et al., 2010):  $r_{g,\hat{g}} = r_{y,s} / \sqrt{h_{sm}^2}$ , sendo  $h_{sm}^2$  a herdabilidade da segregação mendeliana dada por  $h_{sm}^2 = 0,5h^2 / [0,5h^2 + (1-h^2)]$  e  $h^2$  a herdabilidade do caráter, estimada pelo método REML. Esses valores de acurácia também foram comparados com acurácias esperadas, determinadas conforme equação proposta por RESENDE (2008), na qual esses valores esperados são funções do tamanho efetivo da população, do tamanho dos cromossomos (em Morgans), do número de locos e da herdabilidade de cada característica.

Em posse do melhor método para cada característica, os efeitos dos marcadores foram estimados e padronizados, considerando toda a população  $F_2$  de suínos fazendo uso da equação (6). A partir dessas informações, o Manhattan plot foi construído, sendo que cada ponto representa um SNP, com o eixo X mostrando sua localização no cromossomo e o eixo Y a magnitude de seu efeito. Nesses gráficos, a opção pelo módulo dos efeitos padronizados no eixo Y diz respeito ao fato de não ter sido utilizadas regressões simples para detectar efeito significativo de cada marcador, o que resultaria em p-valores a serem usados nesse eixo.

Todas as análises realizadas foram implementadas no software livre R (R DEVELOPMENT CORE TEAM, 2010), utilizando a função *pls* do pacote *pls*.

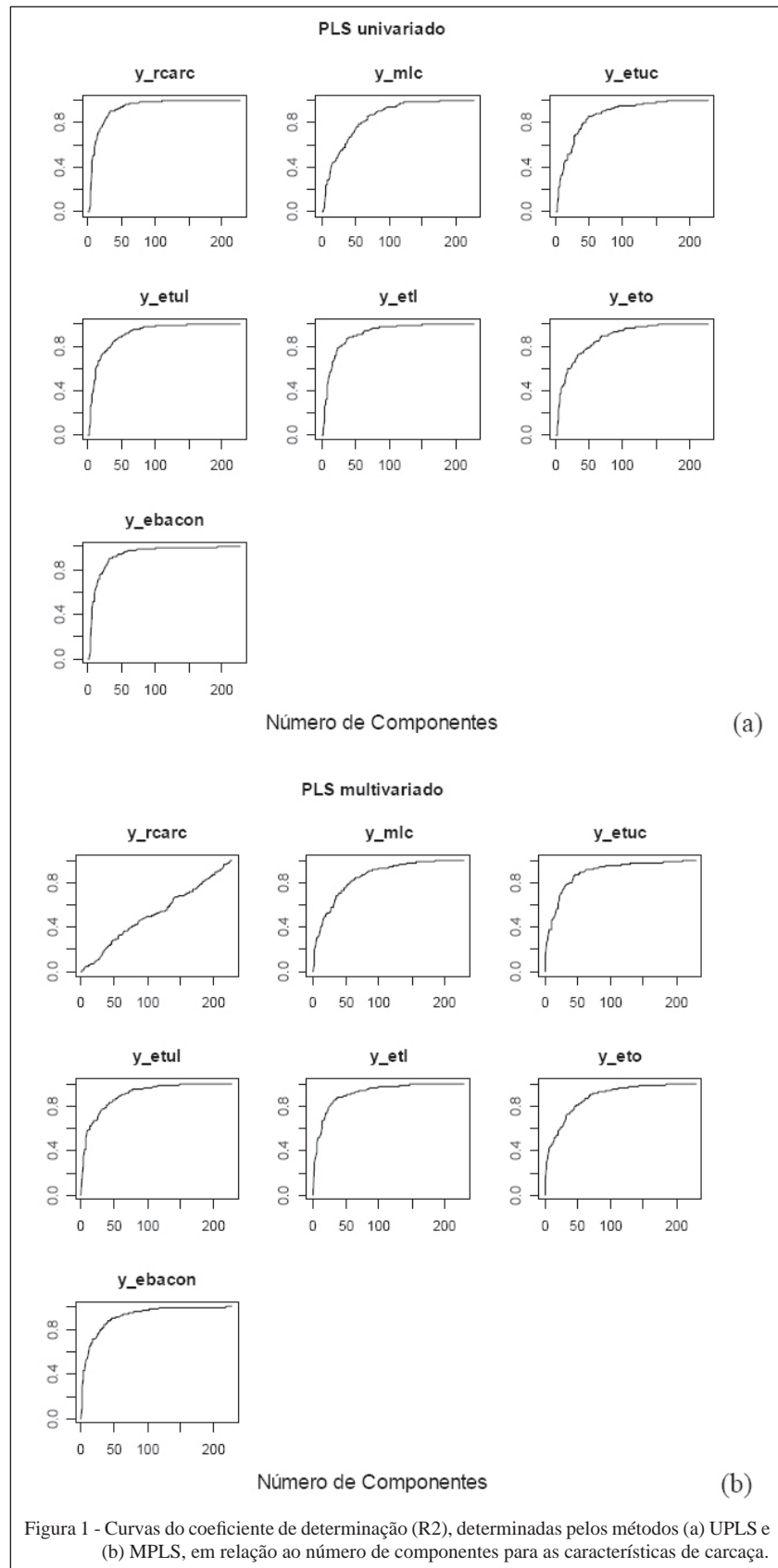
## RESULTADOS E DISCUSSÃO

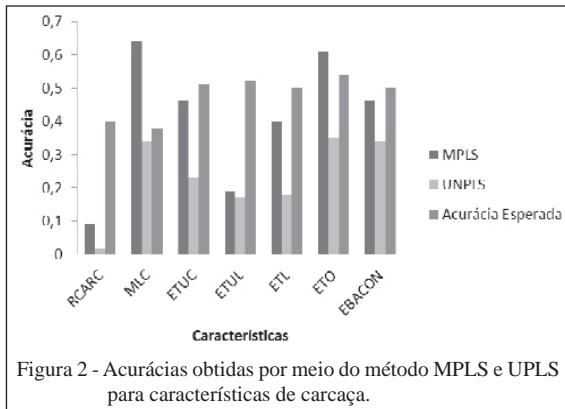
Conforme apresentado, optou-se pelo número mínimo de variáveis latentes que estabilizasse a curva do coeficiente de determinação ( $R^2$ ) do modelo, que, neste trabalho, foi de 50 componentes (Figura 1), portanto, uma redução de 79% na dimensionalidade em relação às variáveis regressoras originais ( $\mathbf{X}_j$ ,  $j=1,\dots,m$ ), uma vez que 50 componentes correspondem a 21% do número total de variáveis originais (237 SNPs). Esse número de componentes foi ideal para todas as características, uma vez que a variação explicada pela regressão foi em torno de 80% ao se especificar tal número de componentes, porém, para a característica RCARC no MPLS, talvez um número superior de componentes seja requerido devido ao comportamento linear crescente (Figura 1b).

As acurácias obtidas por meio do método MPLS bem como pelo UPLS, para cada característica de carcaça, são apresentadas na figura 2. A superioridade do MPLS quanto à predição do valor genético genômico é evidente, tendo em vista que a eficiência matemática média do MPLS foi de mais de cem por cento sobre o método UPLS. Esses resultados são intuitivamente esperados, isso porque o MPLS captura de forma realística a natureza das características, considerando as correlações entre elas.

Sob o enfoque da seleção genômica, embora não existam relatos na literatura a respeito da utilização dos métodos UPLS e MPLS em melhoramento de suínos, em outras espécies, o método UPLS já foi utilizado, e, de alguma forma, os valores de acurácia reportados para ele podem fornecer uma ideia a respeito da eficiência desse método. MOSER et al. (2009) realizou um estudo de comparação entre cinco métodos para dados de gado leiteiro, sendo que o UPLS mostrou acurácia de 0,42, não diferindo dos outros métodos utilizados. SOLBERG et al. (2009) realizaram um estudo fazendo uso também do método UPLS e observou 0,46 de acurácia, também bastante semelhantes àqueles proporcionados pelos outros métodos utilizados.

As acurácias esperadas e estimadas por cada um dos métodos utilizados são apresentadas na figura 2. Nota-se que o método MPLS mostrou uma diferença média (0,07) menor que aquela apresentada pelo método UPLS (0,24), evidenciando que gerou estimativas de acurácia mais próximas daquelas calculadas por meio de parâmetros específicos dessa população  $F_2$ . Nota-se também que, para as variáveis MLC e ETO, o método MPLS mostrou valores de acurácia maiores que os esperados, situação esta que não foi observada para nenhuma variável ao se





considerar o método UPLS. Portanto, diante de tais resultados, é possível inferir que o método MPLS é mais apropriado para seleção genômica que o método UPLS em situações envolvendo vários fenótipos, naturalmente correlacionados (Tabela 1), tais como esses referentes a avaliações de carcaça em suínos.

As estimativas das correlações genéticas (Tabela 1) entre espessura de toucinho e a característica comprimento de carcaça obtiveram valores negativos (-0,38 a -0,57), o que garante a eficiência na seleção. As espessuras de toucinho apresentaram correlações genéticas positivas com o rendimento de carcaça (0,27 a 0,33). Além disso, essas características são altamente correlacionadas entre si e com a espessura de bacon (0,46 a 0,74), conforme encontrado por MENDONÇA et al. (2012), utilizando esse mesmo conjunto de dados, porém sob o enfoque de modelos mistos com inclusão da matriz de parentesco.

A identificação de marcadores de grandes efeitos é de suma importância para a GWS, pois possibilita determinar a posição desses SNPs, a fim de verificar a existência de QTLs que afetam o

caráter quantitativo nessas regiões. Assim, uma vez que o método MPLS apresentou melhores resultados, a identificação dos marcadores de maiores efeitos foi realizada por meio das estimativas dos parâmetros do modelo de regressão (efeitos de SNPs), contemplado pelo mesmo método. Em posse do melhor método para cada característica e a fim de facilitar tal identificação, os gráficos *Manhattan* foram confeccionados e estão dispostos na figura 3.

De forma geral, os marcadores de maior relevância na característica RCARC se encontram no primeiro terço do cromossomo 7, tal como EVANS et al. (2004) encontrou um QTL significativo para suínos de linhagem comercial. Enquanto que para as outras características, esses SNPs se encontram no terço intermediário (MLC, ETUL e ETO) e no terço final (ETUC e ETL) do cromossomo 1, concordando com os resultados reportados por BEECKMANN et al. (2003) e ROHRER et al. (1998) para suínos das raças Meishan, Pietrain e European Wild Boar; e provenientes do cruzamento Meishan × White, respectivamente.

**CONCLUSÃO**

O método mais eficiente para a estimação do valor genético genômico é o MPLS, o qual, por contemplar todos os fenótipos simultaneamente, resulta em maior acurácia.

O marcadores SNPs de maior efeito identificados por este método estão localizados em regiões cromossômicas previamente relatadas em outros estudos, como regiões relacionadas à presença de QTLs que afetam características de carcaça em suínos.

Tabela 1 - Estimativas das correlações genéticas, entre as características, dos valores genéticos genômicos dos indivíduos.

Caract.	RCARC	MLC	ETUC	ETUL	ETL	ETO	EBACON
RCARC	1	-0,61±0,01	0,31±0,02	0,31±0,03	0,27±0,01	0,33±0,03	0,14±0,02
MLC		1	-0,57±0,02	-0,38±0,03	-0,57±0,01	-0,56±0,02	-0,55±0,02
ETUC			1	0,53±0,04	0,65±0,02	0,73±0,04	0,57±0,03
ETUL				1	0,51±0,02	0,58±0,05	0,46±0,04
ETL					1	0,76±0,02	0,64±0,01
ETO						1	0,74±0,04
EBACON							1

RCARC = rendimento de carcaça (%); MLC = comprimento de carcaça pelo Método Americano (cm); ETUC = espessura de toucinho imediatamente após a última costela, na linha dorso-lombar (mm); ETUL = espessura de toucinho medida entre a última e a penúltima vértebra lombar (mm); ETL= menor espessura de toucinho na região acima da última vértebra lombar, na linha dorso-lombar (mm); ETO = espessura de toucinho medida imediatamente após a última costela, a 6,5cm da linha dorso-lombar (mm); EBACON= espessura de bacon (mm).

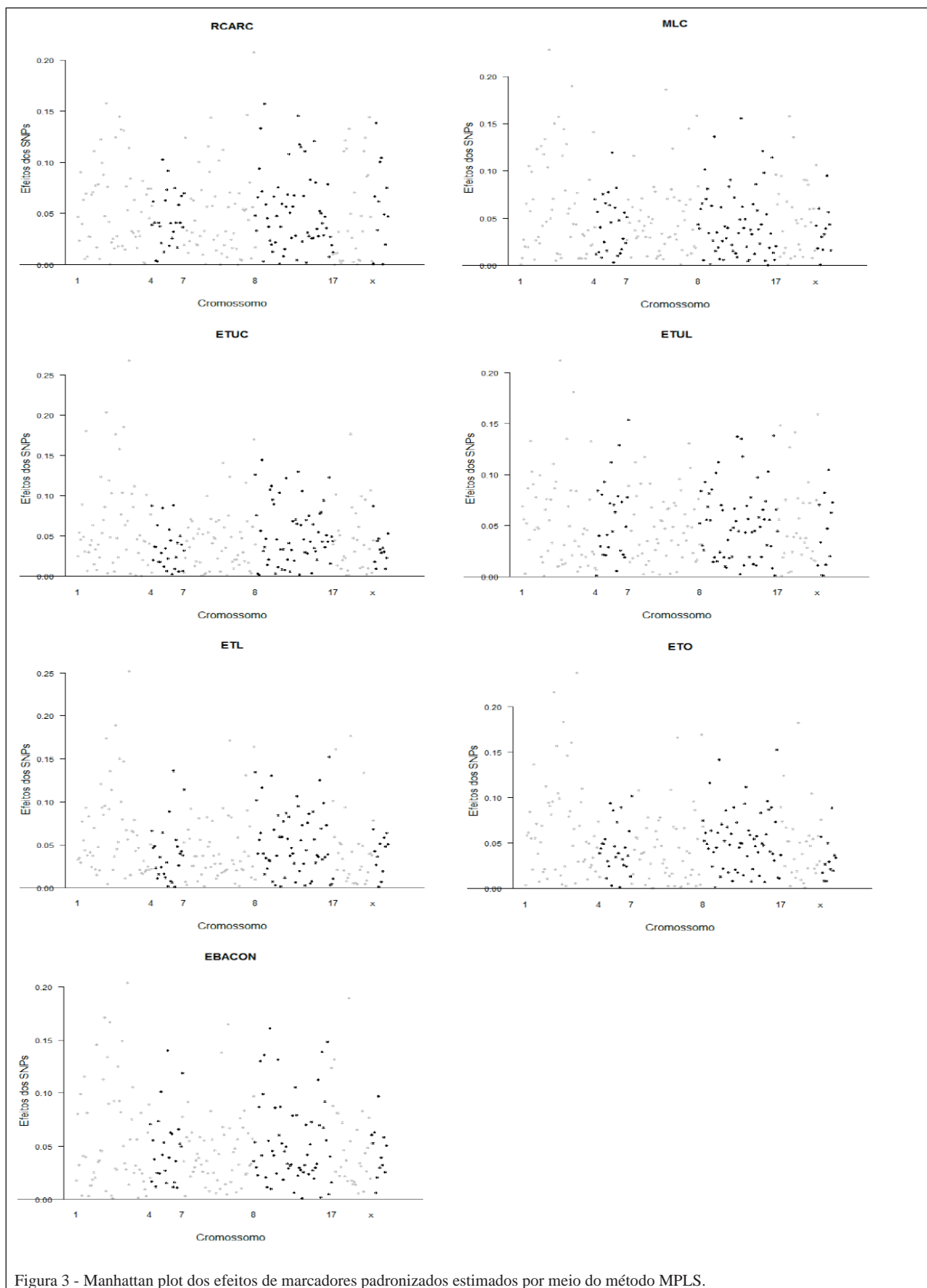


Figura 3 - Manhattan plot dos efeitos de marcadores padronizados estimados por meio do método MPLS.

## REFERÊNCIAS

- BEECKMANN, P. et al. Linkage and QTL mapping for Sus scrofa chromosome 1. **Journal of Animal Breeding and Genetics**, v.120, p.1-10, 2003. Disponível em: <<http://www.animalgenome.org/QTLdb/references/ISU0007.pdf>>. Acesso em: 10 mai. 2012. doi: 0.1046/j.0931-2668.2003.00418.x.
- EVANS, G.J. et al. Identification of quantitative trait loci for production traits in commercial pig populations. **Genetics**, v.164, p.621-627, 2003. Disponível em: <<http://www.genetics.org/content/164/2/621.full.pdf>> Acesso em: 20 mai. 2012.
- FONTANESI, L. et al. The IGF2 intron3-g.3072G>A polymorphism is not the only SSC2p mutation affecting meat production and carcass traits in pigs: evidences from the effects of a cathepsin D (CTSD) gene polymorphism. **Journal of Animal Science**, v.88, n.7, p.:2235-2245,2010. Disponível em: <<http://www.animal-science.org/content/early/2010/04/09/jas.2009-2560>>. Acesso em: 10 mai. 2012. doi: 10.2527/jas.2009-2560.
- GARTHWAITE, P.H. An interpretation of partial least squares. **Journal of the American Statistical Association**, v.89, p.122-127, 1994.
- HIDALGO, A.M. et al. **Fine mapping and single nucleotide polymorphism effects estimation on pig chromosomes 1,4,7,8,17 and x. p.31, 2011.** Dissertation of Genetics and Breeding - University Federal of Viçosa, Viçosa, MG.
- MENDONÇA, P.T. et al. Estimação de parâmetros genéticos de uma população F2 de suínos. **Revista Brasileira de Saúde e Produção Animal**, v.13, p.330-343, 2012. Disponível em: <<http://revistas.ufba.br/index.php/rbspa/article/viewArticle/2199>>. Acesso em: 20 mai. 2012.
- MEUWISSEN, T.H.E. et al. Prediction of total genetic value using genome wide dense marker maps. **Genetics**, v.157, p.1819-1829, 2001. Disponível em: <<http://www.genetics.org/content/157/4/1819.full>>. Acesso em: 08 ago. 2011.
- MOSER, G. et al. A comparison of five methods to predict genomic breeding values of dairy bulls from genome-wide SNP markers. **Genetics Selection Evolution**, v.41, p.41-53, 2009. Disponível em: <<http://www.gsejournal.org/content/41/1/56>>. Acesso em: 05 abr. 2012. doi:10.1186/1297-9686-41-56.
- PEIXOTO, J.O. et al. Associations of leptin gene polymorphisms with production traits in pigs. **Journal of Animal Breeding and Genetics**, v.123, p.378-383, 2006.
- R DEVELOPMENT CORE TEAM. R: a language and environment for statistical computing. Viena: R Foundation for Statistical Computing, 2010. p. 409.
- RESENDE, M.D.V. **Genômica quantitativa e seleção no melhoramento de plantas perenes e animais.** Colombo: Embrapa Florestas, 2008. 330p.
- RESENDE, M.D.V. et al. **Computação da seleção genômica ampla (GWS).** Colombo: Embrapa Florestas, 2010. 79p.
- ROHRER, G.A. et al. Identification of quantitative trait loci affecting carcass composition inswine: I. Fat deposition traits. **Journal of Animal Science**, v.76, p.47-54, 1998. Disponível em: <<http://naldc.nal.usda.gov/download/12904/PDF>>. Acesso em: 13 de jun 2012.
- ROSA, A.F. et al. Características de carcaça de suínos de três linhagens genéticas em diferentes idades ao abate. **Ciência Rural**, v.38, p.1718-1724, 2008.
- SILVA, F.F. et al. Three-step Bayesian factor analysis applied to QTL detection in crosses between outbred pig populations. **Livestock Science**, v.142, p.210-215, 2011.
- SOLGERG, T.R. et al. Reducing dimensionality for prediction of genome-wide breeding values. **Genetics Selection Evolution**, v.41, p.29, 2009. Disponível em: <<http://www.gsejournal.org/content/41/1/29>>. Acesso em: 01 de jun 2012. doi:10.1186/1297-9686-41-29.
- UIMARI, P. et al. Whole-genome SNP association analysis of reproduction traits in the Finnish Landrace pig breed. **Genetics Selection Evolution**, v.43, p.42, 2011.
- ZANGERONIMO, M.G. et al. Desempenho e características de carcaça de suínos dos 20 aos 50kg recebendo rações com reduzido teor de proteína bruta e diferentes níveis de lisina digestível verdadeira. **Ciência Rural**, v.39, p.1507-1513, 2009.