

Regressão via componentes independentes aplicada à seleção genômica para características de carcaça em suínos

Camila Ferreira Azevedo⁽¹⁾, Marcos Deon Vilela de Resende⁽²⁾, Fabyano Fonseca e Silva⁽¹⁾, Paulo Sávio Lopes⁽³⁾ e Simone Eliza Facioni Guimarães⁽³⁾

⁽¹⁾Universidade Federal de Viçosa (UFV), Departamento de Estatística, Avenida P.H. Rolfs, CEP 36570-000 Viçosa, MG. E-mail: camila.azevedo@ufv.br, fabyanofonseca@ufv.br ⁽²⁾Embrapa Florestas, Estrada da Ribeira, Km 111, Caixa Postal 319, CEP 83411-000 Colombo, PR. E-mail: marcos.deon@gmail.com ⁽³⁾UFV, Departamento de Zootecnia. E-mail: plopes@ufv.br, sfacioniguima@gmail.com

Resumo – O objetivo deste trabalho foi avaliar a eficiência do método de regressão via componentes independentes (ICR) na estimação de valores genéticos genômicos e dos efeitos de marcadores SNP para características de carcaça de uma população F₂ de suínos (Piau x linhagem comercial). Os métodos foram avaliados por meio da concordância entre os valores genéticos preditos e os fenótipos corrigidos, observados por validação cruzada, e também foram comparados com outros métodos geralmente utilizados para os mesmos propósitos, tais como RR-BLUP, PCR e PLS. Os métodos ICR e PCR apresentam resultados similares, mas o método ICR apresenta maiores valores de acurácia.

Termos para indexação: *Sus scrofa*, BLUP, marcadores SNP, regressão aleatória, redução dimensional.

Independent component regression applied to genomic selection for carcass traits in pigs

Abstract – The objective of this work was to evaluate the efficiency of the independent component regression (ICR) method for the estimation of genomic values and of SNP marker effects for carcass traits in a F₂ pig population (Piau x commercial line). The methods were evaluated by the agreement between the genetic predicted values and the corrected phenotypes observed by cross-validation, and they were also compared with other methods generally used for the same purposes, such as RR-BLUP, PCR, and PLS. The ICR and PCR methods show similar results, but ICR has the highest accuracy prediction values.

Index terms: *Sus scrofa*, BLUP, SNP markers, random regression, dimensional reduction.

Introdução

Efeitos genéticos de marcadores do tipo polimorfismo de nucleotídeo único (SNP), em desequilíbrio de ligação com locos de características quantitativas (QTL), podem ser utilizados como critério para identificação de indivíduos candidatos à seleção, o que aumentaria a acurácia na avaliação genética (Meuwissen et al., 2001). Porém, na prática, a utilização deste conjunto de marcadores geralmente apresenta dificuldade, pois o número de marcadores é muito maior que o número de indivíduos genotipados. De acordo com Gianola et al. (2003), a alta colinearidade entre os marcadores, que são variáveis de natureza discreta, requer métodos estatísticos diferenciados para a análise de dados na área de seleção genômica.

No geral, os métodos mais utilizados são os de regressão penalizada sob os enfoques frequentistas (RR-BLUP e G-BLUP) e bayesianos (Bayes A e B

e Lasso Bayesiano). Porém, métodos de redução de dimensionalidade, como regressão via quadrados mínimos parciais (PLS) e via componentes principais (PCR) também são de grande aplicabilidade (Resende et al., 2011). Na seleção genômica, os métodos de redução de dimensionalidade são úteis para remover a redundância entre as variáveis explicativas, neste caso, os marcadores SNP, os quais são naturalmente redundantes em razão do conceito de desequilíbrio de ligação entre eles.

Segundo Solberg et al. (2009), os métodos PCR e PLS fornecem uma análise rápida de grandes quantidades de dados usados na obtenção dos valores genéticos genômicos dos indivíduos. Para isso, necessita-se apenas do pressuposto de aditividade dos efeitos de marcadores. Além destes métodos, a regressão via componentes independentes (ICR) também pode ser enquadrada nesse contexto, pois é definida como um método de redução de dimensionalidade. Contudo,

até o momento, essa técnica ainda não foi aplicada à seleção genômica, de forma que sua utilização e comparação com métodos tradicionais podem caracterizar um ganho tecnológico na área de genética e melhoramento.

As características de carcaça são muito importantes para o desenvolvimento da suinocultura, sobretudo aquelas relacionadas à menor deposição de gordura, para que se possa atender ao crescente e cada vez mais exigente mercado consumidor (Bertol et al., 2010). Assim, várias pesquisas foram desenvolvidas com a finalidade de utilizar marcadores moleculares no melhoramento de suínos para essas características (Silva et al., 2011), com resultados relevantes diante do alto número de QTLs identificados.

O objetivo deste trabalho foi avaliar a eficiência do método de regressão via componentes independentes na estimação de valores genéticos genômicos e dos efeitos de marcadores SNP para características de carcaça de uma população F_2 de suínos (Piau x linhagem comercial).

Material e Métodos

Os dados foram obtidos na Granja de Melhoramento de Suínos, do Departamento de Zootecnia, da Universidade Federal de Viçosa (UFV), MG, no período de novembro de 1998 a julho de 2001. Foi utilizada uma população constituída de 345 indivíduos F_2 de suínos provenientes do cruzamento de dois varrões da raça naturalizada Piau, com 18 fêmeas de linhagem desenvolvida na UFV – pelo acasalamento de animais de linha comercial (Landrace x Large White x Pietrain).

A extração do DNA foi realizada no Laboratório de Biotecnologia Animal, do Departamento de Zootecnia, da UFV, segundo Peixoto et al. (2006). A genotipagem foi realizada via tecnologia Golden Gate/Vera Code R, no Laboratório de Genética Animal, da Embrapa Recursos Genéticos e Biotecnologia, Brasília, DF, conforme descrito por Paixão et al. (2011). Os marcadores SNP estão distribuídos da seguinte forma nos cromossomos de *Sus scrofa*: SSC1 (56), SSC4 (54), SSC7 (59), SSC8 (31) e SSC17 (25), o que totalizou 237 marcadores.

As características analisadas foram: espessura de toucinho medida imediatamente após a última costela na linha dorso-lombar (ETUC), a 6,5 cm da linha

dorso-lombar (ETO), e entre a última e a penúltima vértebra lombar (ETUL); menor espessura de toucinho na região acima da última vértebra lombar, na linha dorso-lombar (ETL); maior espessura de toucinho na região da copa, na linha dorso-lombar (ETSH); espessura de bacon (EBACON); comprimento de carcaça pelo método de classificação americano (MLC); e rendimento de carcaça (RCARC).

Os dados fenotípicos foram corrigidos para efeitos fixos dos genitores e desregressados (“deregressing”) (Resende et al., 2010), para trabalhar com uma melhor aproximação do mérito genético verdadeiro de indivíduos não aparentados. Esses procedimentos foram realizados nos programas ASREML (Gilmour et al., 2000) e R (R Development Core Team, Vienna, AT).

O método “random regression best linear unbiased predictor” (RR-BLUP), também denominado SNP-BLUP, faz uso de preditores do tipo BLUP no qual os efeitos das marcas são covariáveis de efeitos aleatórios. A predição desses efeitos é baseada no modelo linear misto (Meuwissen et al., 2001; Resende, 2007): $y = Wb + X m_{rr-blup} + e$, em que y é o vetor de fenótipos corrigidos com dimensão $I \times 1$, na qual I é o número de indivíduos genotipados e fenotipados; b é o vetor de efeitos fixos com matriz de incidência W (no caso, b contém apenas o efeito da média geral μ); $m_{rr-blup}$ é o vetor de efeitos de marcas com matriz de incidência X com dimensão $J \times 1$, na qual J é o número de marcadores. Dessa forma, a predição do valor genético genômico (GBV) do indivíduo i via RR-BLUP é obtida pela expressão:

$$\hat{y} = \hat{\mu} + \sum_{j=1}^J X_{ij} \hat{m}_{rr-blup} \quad (1),$$

em que $\sum_{j=1}^J X_{ij}$ é a linha da matriz X , correspondente ao indivíduo i .

De forma análoga ao método RR-BLUP, no desenvolvimento dos métodos de redução de dimensionalidade, a matriz X é definida como a matriz de incidência dos marcadores SNP, e y é o vetor de dados fenotípicos corrigidos e desregressados.

Segundo Garthwaite (1994), o método PLS apresenta similaridades com o PCR. A principal diferença é que o PCR leva em consideração apenas as variáveis explicativas na construção dos componentes, enquanto o PLS leva em consideração, também, a variável dependente. Ambos os métodos

são apropriados para situações nas quais se têm mais covariáveis do que observações, conforme constatado na seleção genômica. Além disso, com esses métodos, garante-se que a correlação entre qualquer par de componentes (combinações lineares de todos os SNP) seja igual a zero, o que é de suma importância para a seleção genômica ampla (GWS), em razão da abundância de marcadores SNP presente no genoma e da multicolinearidade existente entre eles.

O método PCR é definido de acordo com a seguinte combinação de variáveis:

$$\hat{Z} = X\hat{P} \quad (2),$$

em que P é a matriz de autovetores da matriz de covariância entre as covariáveis (X), e Z é a matriz dos componentes principais ($Z_v, v = 1, \dots, n_{pc}$), os quais representam combinações lineares das covariáveis originais.

O método PLS é definido de acordo com as seguintes decomposições das matrizes X e Y, as quais são efetuadas de forma simultânea:

$$X = TL' + E_1 \quad (3),$$

$$Y = Uq' + e_2 \quad (4),$$

em que T e U são matrizes de componentes; L e q são matrizes de carregamento; e E_1 e E_2 são vetores de resíduos. A decomposição não é independente, o que possibilita estabelecer uma relação entre componentes de X e Y de forma que, para cada fator, a seguinte relação é obtida:

$$\hat{u}_l = \hat{b}_l t_l \quad (5),$$

em que $\hat{b}_l = (u'_l t_l) / (t'_l t_l)$ ($l = 1, \dots, n_{pls}$) são coeficientes estimados via quadrados mínimos ordinários (OLS) e agrupados em matriz diagonal B.

Para estabelecer a relação entre Y e os componentes, utilizou-se a regressão linear múltipla para obter as equações de predição do PCR e do PLS, respectivamente:

$$\hat{y} = \hat{\alpha}_0 + \hat{\alpha}_1 \hat{z}_1 + \hat{\alpha}_2 \hat{z}_2 + \dots + \hat{\alpha}_{n_{pc}} \hat{z}_{n_{pc}} \quad (6),$$

$$\hat{y} = \hat{TB}\hat{q}' \quad (7),$$

em que α_v é o coeficiente da regressão entre Y e Z, obtido por meio do método OLS.

Os coeficientes Bq' e $\hat{\alpha}$ não apresentam interpretação biológica; porém, é possível estimar os coeficientes associados às variáveis originais (efeitos dos marcadores) ao se combinar as equações (2) e (6), (4) e (7) dos métodos PCR e PLS, respectivamente. Dessa forma, tem-se:

$$\hat{m}_{per} = \hat{P}\hat{\alpha} \quad (8),$$

$$\hat{m}_{pls} = \hat{L}\hat{B}\hat{q}' \quad (9).$$

O método ICR, proposto por Comon (1994), consiste em decompor a matriz de covariáveis X em combinações de componentes independentes, o que garante a retirada da multicolinearidade dos dados, além de reduzir a dimensionalidade. Por esse método, existe o pressuposto de que os dados sejam provenientes de uma distribuição não normal. Portanto, ele pode ser aplicado de forma eficiente à GWS, em que a matriz de marcas X está parametrizada com os valores 0, 1 e 2 (distribuição não normal). Assim, tem-se a decomposição dada por:

$$X' = A'S' \quad (10),$$

em que A é definida como uma f(KR), na qual K é uma matriz de ortogonalização de S obtida por meio da decomposição espectral e R é uma matriz ortogonal que maximiza a independência estatística das colunas de S, na qual S é a matriz dos componentes independentes $S_k, k = 1, \dots, n_{icr}$.

O algoritmo desenvolvido por Hyvärinen (1998) é utilizado no ICR para encontrar a matriz R, com base no princípio da máxima entropia (J(r)). Desse modo, obtém-se a seguinte aproximação:

$$J(r) \propto [E\{G_1(r)\} - E\{G_1(v)\}]^2 \quad (11),$$

em que r e v são as variáveis padronizadas, e $G_1(u) = -\exp(-u^2/2)$, em que u é uma variável normal padronizada. Após o processo iterativo, tem-se a matriz de componentes dada por:

$$\hat{S} = XKR \quad (12),$$

em que KR é uma aproximação de A'. Dessa forma, obtém-se a equação de predição baseada no método ICR, expressa por:

$$\hat{y} = \hat{\gamma}_0 + \hat{\gamma}_1 \hat{s}_1 + \hat{\gamma}_2 \hat{s}_2 + \dots + \hat{\gamma}_{n_{icr}} \hat{s}_{n_{icr}} \quad (13),$$

em que S_k é o componente independente e γ_k é o coeficiente da regressão determinado por meio do método OLS, $k = 1, \dots, n_{icr}$. Semelhantemente aos outros métodos, pode-se obter os efeitos de marcadores a partir das equações (12) e (13), por meio da equação:

$$\hat{m}_{icr} = KR\hat{\gamma} \quad (14),$$

em que $\hat{\gamma}$ é o vetor de estimativas dos coeficientes provenientes da regressão entre Y e S.

Um passo importante dos métodos de redução dimensional é a escolha do número ótimo de componentes a serem inseridos no modelo. Um critério de decisão para o PLS e o PCR é adotar

uma porcentagem da variação total explicada pelos componentes, a qual, no presente trabalho, foi de 70%. Essa porcentagem também foi considerada para o método ICR, uma vez que o número de componentes no ICR pode ser o mesmo que o no PCR (Cadavid et al., 2008).

Os métodos apresentados foram comparados por meio de estudo de validação cruzada (Resende et al., 2011), realizado separadamente para cada característica de carcaça. Para tanto, a população F_2 de suínos foi dividida em três populações distintas, cada uma com 115 indivíduos. Assim, a cada repetição da análise, duas dessas populações foram consideradas como população de estimação e utilizadas para obter os efeitos dos marcadores SNP. A outra população, denominada população de validação, foi utilizada para avaliar a concordância (via correlação $r_{y,\hat{y}}$) entre valores genéticos preditos (via estimativas provenientes da população de estimação) e fenótipos corrigidos observados. O processo foi repetido de forma que a cada etapa uma das três populações constituísse a população de validação. Dessa forma, a acurácia proveniente de cada método foi dada pela expressão: $r_{q,\hat{q}} = r_{y,\hat{y}} / (h^2_{sm})^{0.5}$ em que h^2_{sm} é a herdabilidade da segregação mendeliana dada por $h^2_{sm} = 0,5h^2 / [0,5h^2 + (1 - h^2)]$, e h^2 é a herdabilidade do caráter estimada pelo método REML (Resende Júnior et al., 2012).

De posse do melhor método para cada característica, eleito de acordo com os valores de acurácia, o efeito dos marcadores em valor absoluto foi estimado e padronizado por meio das equações (1), (8), (9) e (14), respectivamente, para os métodos RR-BLUP, PLS, PCR e ICR. A partir dessas informações, foram elaborados os gráficos “Manhattan plot”, nos quais cada ponto representa o efeito de cada SNP com sua respectiva localização no cromossomo. As rotinas computacionais dos métodos RR-BLUP, PLS, PCR e ICR foram implementadas no programa R (R Development Core Team, Vienna, AT) com os pacotes rrBLUP, pls, caret e as funções mixed.solve, pls, pcr e icr, respectivamente. Essas rotinas se encontram em Resende et al. (2012).

Resultados e Discussão

A opção pelo número de variáveis latentes que explicassem 70% da variável X resultou em 47, 34 e 34 componentes, respectivamente, para os métodos PLS, PCR e ICR.

Os valores de acurácia, obtidos pelos métodos de redução dimensional, para cada característica de carcaça, e pelo método RR-BLUP, apresentaram baixa magnitude (Tabela 1). Entretanto, esses valores eram esperados em razão do número reduzido de 230 indivíduos na população de estimação.

O método RR-BLUP mostrou-se mais eficiente para as características MLC e ETL, enquanto o método ICR mostrou-se adequado para ETSH, ETUC, ETUL e EBACON. O método PCR se destacou para o caráter ETO, porém apresentou resultado similar ao RR-BLUP para a característica restante, RCARC.

Desse modo, o método que mais se destacou nessas análises foi o ICR. A principal vantagem observada em relação aos demais métodos de redução dimensional é o fato de este método considerar a independência entre os componentes, que, além de garantir a ausência de relação linear entre as variáveis latentes, também garante a ausência de qualquer dependência funcional não linear. Contudo, o PLS apresentou os mais baixos valores de acurácia para as características de carcaça, possivelmente porque o método, em seu desenvolvimento, não contempla a dependência existente entre as variáveis explicativas (marcadores SNP). Além disso, este método é sensível ao grau de simetria da distribuição dos dados e pode ser afetado pela multicolinearidade existente entre variáveis observadas e latentes (Cassel et al., 2009).

O PLS é o método que mais se diferencia dos demais, uma vez que as correlações dos resultados de acurácia

Tabela 1. Acurácias na predição de valores genéticos genômicos obtidas pelos métodos PLS, PCR, ICR e RR-BLUP, para características de carcaça em suínos.

Características ⁽¹⁾	Acurácia			
	PLS	PCR	ICR	RR-BLUP
RCARC (%)	0,20	0,29	0,24	0,30
MLC (cm)	0,05	0,21	0,15	0,25
ETSH (mm)	0,11	0,35	0,45	0,34
ETUC (mm)	0,01	0,47	0,51	0,29
ETUL (mm)	0,03	0,27	0,39	0,27
ETL (mm)	0,03	0,42	0,45	0,47
ETO (mm)	0,11	0,37	0,29	0,24
EBACON (mm)	0,08	0,40	0,48	0,24

⁽¹⁾RCARC, rendimento de carcaça; MLC, comprimento de carcaça pelo método americano; ETSH, maior espessura de toucinho na região da copa, na linha dorso-lombar; ETUC, espessura de toucinho imediatamente após a última costela, na linha dorso-lombar; ETUL, espessura de toucinho medida entre a última e a penúltima vértebra lombar; ETL, menor espessura de toucinho na região acima da última vértebra lombar, na linha dorso-lombar; ETO, espessura de toucinho medida imediatamente após a última costela, a 6,5 cm da linha dorso-lombar; EBACON, espessura de bacon.

obtidos entre ele e os métodos PCR, ICR e RR-BLUP, respectivamente, foram -0,38, -0,57 e -0,11. Já os métodos mais concordantes foram o PCR e o ICR, com correlação de 0,86, pois o ICR pode ser considerado uma extensão do método PCR.

No enfoque da seleção genômica, não há conhecimento de relatos na literatura a respeito da utilização dos métodos de redução dimensional em melhoramento de suínos. No entanto, em estudos com outras espécies, os métodos PLS e PCR já foram utilizados. Moser et al. (2009), por exemplo, compararam cinco métodos para dados de gado leiteiro, entre eles o PLS e o RR-BLUP, e obtiveram acurácias próximas para os dois métodos. Solberg et al. (2009), ao avaliarem os métodos PLS e PCR, observaram acurácias similares (0,47 e 0,45) para as metodologias. Macciota et al. (2010), em trabalho com gado leiteiro, obtiveram acurácias de 0,28–0,46 para o PCR.

Os métodos PCR e ICR apresentaram, respectivamente, 19 e 26% a mais de eficiência relativa do que o RR-BLUP (Tabela 2). Quanto à característica EBACON, os métodos PCR e ICR foram, respectivamente, 67 e 100% mais eficientes que o método RR-BLUP. Entretanto, o PLS novamente apresentou resultados inferiores aos demais, tendo sido 75% menos eficiente, em média, do que o RR-BLUP.

De acordo com as estimativas das correlações dos valores genéticos genômicos dos indivíduos, obtidos pelos métodos, a maior correlação foi verificada entre o ICR e o PCR (Tabela 3). O método PLS apresentou menores

Tabela 2. Eficiência relativa na seleção genômica obtida pelos métodos de redução de dimensionalidade em relação ao RR-BLUP, para cada característica de carcaça em suínos.

Características ⁽¹⁾	Eficiência relativa		
	PLS	PCR	ICR
RCARC (%)	0,67	0,97	0,80
MLC (cm)	0,20	0,84	0,60
ETSH (mm)	0,32	1,03	1,32
ETUC (mm)	0,03	1,62	1,76
ETUL (mm)	0,11	1,00	1,44
ETL (mm)	0,06	0,89	0,96
ETO (mm)	0,30	1,54	1,21
EBACON (mm)	0,33	1,67	2,00

⁽¹⁾RCARC, rendimento de carcaça; MLC, comprimento de carcaça pelo método americano; ETSH, maior espessura de toucinho na região da copa, na linha dorso-lombar; ETUC, espessura de toucinho imediatamente após a última costela, na linha dorso-lombar; ETUL, espessura de toucinho medida entre a última e a penúltima vértebra lombar; ETL, menor espessura de toucinho na região acima da última vértebra lombar, na linha dorso-lombar; ETO, espessura de toucinho medida imediatamente após a última costela, a 6,5 cm da linha dorso-lombar; EBACON, espessura de bacon.

correlações com os demais métodos. Vale ressaltar que as correlações entre RR-BLUP e ICR e PCR foram altas, tendo variado de 0,87 a 0,91.

A identificação de marcadores de grandes efeitos é de suma importância para a GWS, pois possibilita verificar a existência de QTL que afetam o caráter quantitativo nas regiões desses marcadores. Assim, dos efeitos de SNP estimados pelas equações (1), (8), (9) e (14) (Figura 1), a característica que apresentou comportamento poligênico mais acentuado foi a ETL, uma vez que os efeitos se distribuíram uniformemente ao longo dos cromossomos, e a característica que apresentou este comportamento menos acentuadamente foi o RCARC. Quanto aos métodos ICR e PCR, não foi possível discriminar de forma clara o comportamento dos efeitos de marcadores.

Tabela 3. Estimativas das correlações dos valores genéticos genômicos (GBVs) provenientes dos métodos PLS, PCR, ICR e RR-BLUP, para as características de carcaça em suínos.

Características ⁽¹⁾	Métodos	PLS	PCR	ICR	RR-BLUP
RCARC (%)	PLS	1	0,35	0,35	0,50
	PCR		1	0,97	0,87
	ICR			1	0,87
MLC (cm)	PLS	1	0,34	0,33	0,57
	PCR		1	0,95	0,87
	ICR			1	0,87
ETSH (mm)	PLS	1	0,35	0,35	0,50
	PCR		1	0,97	0,87
	ICR			1	0,87
ETUC (mm)	PLS	1	0,47	0,46	0,63
	PCR		1	0,98	0,90
	ICR			1	0,90
ETUL (mm)	PLS	1	0,48	0,47	0,64
	PCR		1	0,98	0,89
	ICR			1	0,89
ETL (mm)	PLS	1	0,44	0,43	0,61
	PCR		1	0,98	0,91
	ICR			1	0,91
ETO (mm)	PLS	1	0,43	0,42	0,63
	PCR		1	0,96	0,88
	ICR			1	0,88
EBACON (mm)	PLS	1	0,46	0,46	0,61
	PCR		1	0,95	0,89
	ICR			1	0,89

⁽¹⁾RCARC, rendimento de carcaça; MLC, comprimento de carcaça pelo método americano; ETSH, maior espessura de toucinho na região da copa, na linha dorso-lombar; ETUC, espessura de toucinho imediatamente após a última costela, na linha dorso-lombar; ETUL, espessura de toucinho medida entre a última e a penúltima vértebra lombar; ETL, menor espessura de toucinho na região acima da última vértebra lombar, na linha dorso-lombar; ETO, espessura de toucinho medida imediatamente após a última costela, a 6,5 cm da linha dorso-lombar; EBACON, espessura de bacon.

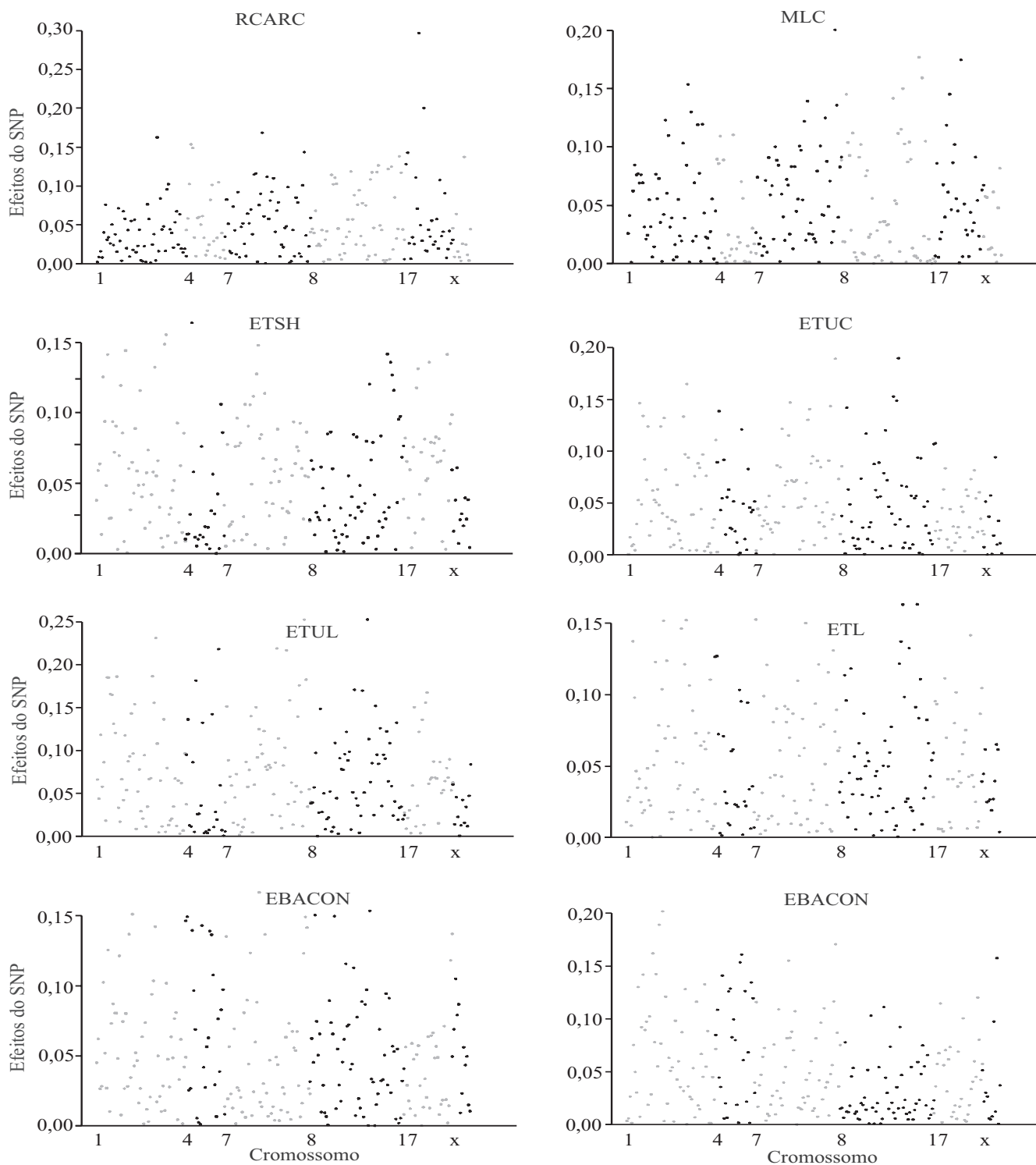


Figura 1. “Manhattan plot” dos módulos dos efeitos de marcadores para as características de carcaça de suínos F_2 (comercial x Piau) avaliadas. RCARC, rendimento de carcaça; MLC, comprimento de carcaça pelo método americano; ETSH, maior espessura de toucinho na região da copa, na linha dorso-lombar; ETUC, espessura de toucinho imediatamente após a última costela, na linha dorso lombar; ETUL, espessura de toucinho medida entre a última e a penúltima vértebra lombar; ETL, menor espessura de toucinho na região acima da última vértebra lombar, na linha dorso-lombar; ETO, espessura de toucinho medida imediatamente após a última costela, a 6,5 cm da linha dorso-lombar; EBACON, espessura de bacon.

O SNP de maior efeito para a característica de espessura de toucinho ETSH encontra-se no terço inicial do cromossomo 4. Nessa região, Gonçalves et al. (2005) detectaram dois QTL significativos para espessura de toucinho, em suínos resultantes do cruzamento entre Meishan, Large White e Landrace. Nas características comprimento de carcaça e espessura de toucinho, os SNP de maior efeito foram encontrados no terço final do cromossomo 7, como observado por Nezer et al. (2002) e Sousa et al. (2011). Também no cromossomo 4, Silva et al. (2011) detectaram QTL significativo na mesma população F₂, para a característica EBACON, o que está de acordo com o obtido no presente trabalho, no qual observaram-se marcadores SNP relevantes nesta região. Para a característica ETO, foram encontrados SNP relevantes no terço inicial do cromossomo 8. Bidanel et al. (2001) analisaram uma população F₂ de Meishan e Large White e encontraram QTL significativo nessa mesma região, para essa característica, além de SNP relevantes para a característica ETUL, no cromossomo 8.

Conclusões

1. O método de regressão via componentes independentes (ICR) apresenta maiores valores de acurácia e maior eficiência relativa para a predição de valores genéticos genômicos de características de carcaça em suínos.

2. Os métodos ICR e de regressão via componentes principais (PCR) apresentam maiores similaridades, enquanto o método de regressão via quadrados mínimos parciais (PLS) apresenta valores mais discrepantes.

Referências

- BERTOL, T.M.; CAMPOS, R.M.L. de; COLDEBELLA, A.; SANTOS FILHO, J.I. dos; FIGUEIREDO, E.A.P. de; TERRA, N.N.; AGNES, I.B.L. Qualidade da carne e desempenho de genótipos de suínos alimentados com dois níveis de aminoácidos. **Pesquisa Agropecuária Brasileira**, v.45, p.621-629, 2010. DOI: 10.1590/S0100-204X2010000600012.
- BIDANEL, J.P.; MILAN, D.; IANNUCELLI, N.; AMIGUES, Y.; BOSCHER, M.Y.; BOURGEOIS, F.; CARITEZ, J.C.; GRUAND, J.; LE ROY, P.; LAGANT, H.; QUINTANILLA, R.; RENARD, C.; GELLIN, J.; OLLIVIER, L.; CHEVALET, C. Detection of quantitative trait loci for growth and fatness in pigs. **Genetics Selection Evolution**, v.33, p.289-309, 2001. DOI: 10.1186/1297-9686-33-3-289.
- CADAVID, A.C.; LAWRENCE, J.K.; RUZMAIKIN, A. Principal components and independent component analysis of solar and space data. **Solar Physics**, v.248, p.247-261, 2008. DOI: 10.1007/s11207-007-9026-2.
- CASSEL, C.; HACKL, P.; WESTLUND, A.H. Robustness of partial least-squares method for estimating latent variable quality structures. **Journal of Applied Statistics**, v.26, p.435-446, 2009. DOI: 10.1080/02664769922322.
- COMON, P. Independent component analysis, a new concept. **Signal Processing**, v.36, p.287-314, 1994. DOI: 10.1016/0165-1684(94)90029-9.
- GARTHWAITE, P.H. An interpretation of partial least squares. **Journal of the American Statistical Association**, v.89, p.122-127, 1994. DOI: 10.1080/01621459.1994.10476452.
- GIANOLA, D.; PEREZ-ENCISO, M.; TORO, M.A. On marker-assisted prediction of genetic value: beyond the ridge. **Genetics**, v.163, p.347-365, 2003.
- GILMOUR, A.R.; CULLIS, B.R.; WELHAM, S.J. THOMPSON, R. **ASReml-R reference manual**. Orange: NSW Agriculture, 2000. 218p.
- GONÇALVES, T. de M.; OLIVEIRA, H.N. de; BOVENHUIS, H.; BINK, M.; ARENDONK, J.V. Modelos alternativos para detecção de locos de características quantitativas (QTL) de carcaça e crescimento nos cromossomos 4, 5 e 7 de suínos. **Revista Brasileira de Zootecnia**, v.34, p.1540-1552, 2005. DOI: 10.1590/S1516-35982005000500014.
- HVYÄRINEN, A. New approximations of differential entropy for independent component analysis and projection pursuit. **Advances in Neural Information Processing Systems**, v.10, p.273-279, 1998.
- MACCIOTA, N.P.P.; PINTUS, M.A.; STERI, R.; PIERAMATI, C.; NICOLAZZI, E.L.; SANTUS, E.; VICARIO, D.; VAN KAAM, J.T.; NARDONE, A.; VALENTINI, A.; AJMONE-MARSAN, P. Accuracies of direct genomic breeding values estimated in dairy cattle with a principal component approach. **Journal of Dairy Science**, v.93, p.532-533, 2010.
- MEUWISSEN, T.H.E.; HAYES, B.J.; GODDARD, M.E. Prediction of total genetic value using genome wide dense marker maps. **Genetics**, v.157, p.1819-1829, 2001.
- MOSER, G.; TIER, B.; CRUMP, R.E.; KHATKAR, M.S.; RAADSMA, H.W. 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. DOI: 10.1186/1297-9686-41-56.
- NEZER, C.; MOREAU, L.; WAGENAAR, D.; GEORGES, M. Results of a whole genome scan targeting QTL for growth and carcass traits in a Pietrain × Large White intercross. **Genetics Selection Evolution**, v.34, p.371-387, 2002. DOI: 10.1186/1297-9686-34-3-371.
- PAIXÃO, D.M.; CARNEIRO, P.L.S.; PAIVA, S.R.; SOUSA, K.R.S.; VERARDO, L.L.; BRACCINI NETO, J.; PINTO, A.P.G.; HIDALGO, A.M.; NASCIMENTO, C.S. do; PÉRISSÉ, I.V.; LOPES, P.S.; GUIMARÃES, S.E.F. Mapeamento de QTL nos cromossomos 1, 2, 3, 12, 14, 15 e X em suínos: características de carcaça e qualidade de carne. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v.64, p.974-982, 2012. DOI: 10.1590/S0102-09352012000400026.

- PEIXOTO, J. de O.; GUIMARAES, S.E.F.; LOPES, P.S.; SOARES, M.A.M.; PIRES, A.V.; BARBOSA, M.V.G.; TORRES, R. de A.; SILVA, M. de A.E. e. Associations of leptin gene polymorphisms with production traits in pigs. **Journal of Animal Breeding and Genetics**, v.123, p.378-383, 2006. DOI: 10.1111/j.1439-0388.2006.00611.x.
- RESENDE JÚNIOR, M.F.R.; MUÑOZ, P.; RESENDE, M.D.V.; GARRICK, D.J.; FERNANDO, R.L.; DAVIS, J.M.; JOKELA, E.J.; MARTIN, T.A.; PETER, G.F.; KIRST, M. Accuracy of genomic selection methods in a standard dataset of loblolly pine (*Pinus taeda* L.). **Genetics**, v.190, p.1503-1510, 2012. DOI: 10.1534/genetics.111.137026.
- RESENDE, M.D.V. de. Seleção genômica ampla (GWS) e modelos mistos. In: RESENDE, M.D.V. de. **Matemática e estatística na análise de experimentos e no melhoramento genético**. Colombo: Embrapa Florestas, 2007. p.517-534.
- RESENDE, M.D.V. de; RESENDE JUNIOR, M.F.R.; AGUIAR, A.M.; ABAD, J.I.M.; MISSIAGIA, A.A.; SANSALONI, C.P.; PETROLI, C.D.; GRATTAPALIA, D. **Computação da seleção genômica ampla (GWS)**. Colombo: Embrapa Florestas, 2010. 79p.
- RESENDE, M.D.V. de; SILVA, F.F. e; LOPES, P.S.; AZEVEDO, C.F. **Seleção genômica ampla (GWS) via modelos mistos (REML/BLUP), inferência bayesiana (MCMC), regressão aleatória multivariada e estatística espacial**. Viçosa: Ed. UFV, 2012. 291p.
- RESENDE, M.D.V. de; SILVA, F.F. e; VIANA, J.M.S.; PETERNELLI, L.A.; RESENDE JUNIOR, M.F.R.; DEL VALLE, P.M. **Métodos estatísticos na seleção genômica ampla**. Colombo: Embrapa Florestas, 2011. 105p.
- SILVA, F.F.; ROSA, G.J.M.; GUIMARÃES, S.E.F.; LOPES, P.S.; DE LOS CAMPOS, G. Three-step Bayesian factor analysis applied to QTL detection in crosses between outbred pig populations. **Livestock Science**, v.142, p.210-215, 2011. DOI: 10.1016/j.livsci.2011.07.012.
- SOLBERG, T.R.; SONESSON, A.K.; WOOLLIAMS, J.A.; MEUWISSEN, T.H.E. Reducing dimensionality for prediction of genome-wide breeding values. **Genetics Selection Evolution**, v.41, p.29, 2009. DOI: 10.1186/1297-9686-41-29.
- SOUSA, K.R.S.; GUIMARÃES, S.E.F.; SILVA FILHO, M.I. da; LOPES, M.S.; PINTO, A.P.G.; VERARDO, L.L.; BRACCINI NETO, J.; LOPES, P.S. Mapeamento de locos de características quantitativas nos cromossomos 5, 7 e 8 de suínos. **Revista Brasileira de Zootecnia**, v.40, p.115-123, 2011. DOI: 10.1590/S1516-35982011000100016.

Recebido em 12 de novembro de 2012 e aprovado em 6 de maio de 2013