

Avaliação de metodologias para validação de QTL utilizando a estrutura de família de um núcleo MOET

Evaluation of methodologies for QTL validation using a MOET nucleus family structure

Peixoto, MGCD^{1*}; Martinez, ML¹; Teodoro, RL¹; Carvalho, MRS²; Machado, MA¹; Euclides, RF³; Gomes, FC; Verneque, RS¹

¹ Embrapa Gado de Leite, Brasil. *gaby@cnpgl.embrapa.br

² Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Brasil.

³ Departamento de Zootecnia, Universidade Federal de Viçosa, Brasil.

Suporte financeiro da Fapemig

Resumo

A detecção de marcadores associados a QTL pode auxiliar os programas de melhoramento. O poder de detecção depende de vários aspectos, tais como o desequilíbrio de ligação. Vários métodos já foram desenvolvidos para detecção de QTL. Nosso objetivo foi avaliar a utilização de dados da estrutura de família do núcleo MOET (irmãos completos e meio-irmãos) na validação de QTL mediante regressão. Verificou-se que estas análises foram capazes de detectar associações marcador-QTL em níveis expressivos de significância ($p < 0,0001$). Os resultados indicam a possibilidade de utilização destes dados para validação de QTL.

Palavras-chave: QTL, Poder de detecção, Marcadores moleculares, Estrutura de família.

Abstract

Detection of markers associated to QTL can help breeding programs. The power of detection depends on several aspects, such as linkage disequilibrium. Many methods have been developed to QTL detection. Our objective was to evaluate the MOET nucleus family structure (full and half-sibs) data for QTL validation using regression. It was verified that these analyses were able to detect marker-QTL associations in an expressive significance level ($p < 0.0001$). Results pointed out the possibility of using these data to QTL validation.

Key words: QTL, Power of detection, Molecular markers, Family structure.

Introdução

A utilização de marcadores tem sido indicada como importante ferramenta para o monitoramento da diversidade genética, planejamento de acasalamentos e seleção assistida (Brascamp et al., 1993), apesar da eficiência da seleção assistida por marcadores não estar bem clara e documentada (Dekkers, 2004).

A detecção/mapeamento de um QTL dependem da estrutura de pedigree, da aferição sistemática das características, do número de indivíduos medidos, do grau de desequilíbrio de ligação marcador-QTL, do número de marcadores genotipados e do efeito do QTL sobre a característica (Israel e Weller, 2002). Em animais de produção, o mapeamento de QTL por meio de marcador depende também da fonte de desequilíbrio e da história prévia da população.

Os primeiros métodos de detecção basearam-se na utilização de marcadores únicos, considerados efeitos fixos na análise de variância. Existem, entretanto, questionamentos sobre o poder de detecção e resolução do mapeamento por meio destes métodos. Outros delineamentos baseados no cruzamento entre linhagens divergentes (Simpson, 1989), cruzamentos entre populações segregantes (Beckmann e Soller, 1989) e estudos de família (Kashi et al., 1990) foram propostos. O mais eficiente baseia-se no cruzamento entre linhagens divergentes, pois o máximo de desequilíbrio é produzido na geração F1.

Várias metodologias foram desenvolvidas, permitindo-se utilizar toda a informação disponível e maximizar o poder de detecção. Dentre estas: análise de informações de segregação incerta pela máxima verossimilhança (Fernando e Grossman, 1989), mapeamento de grupo de ligação em populações exogâmicas (Haley e Knott, 1992), mapeamento por intervalo para cruzamentos F2 e retrocruzamentos (Darvasi et al., 1993), delineamento de filhas, netas e bisnetas (Coppieters et al., 1999) e análise de informações de populações exogâmicas pela amostragem de Gibbs ou Cadeia de Markov/Monte Carlo (Bink et al., 2000).

Apesar da perda de poder de detecção, das diferenças individuais quanto à fase de ligação e de todos os indivíduos não serem informativos, a análise de população exogâmica pode evidenciar diferenças na

segregação do QTL (Lynch e Walsh, 1997). Em gado de leite, usualmente, o desequilíbrio de ligação marcador-QTL pode ser encontrado em famílias de meio-irmãs paternas, sendo que quanto maior o número de filhas por reprodutor maior o poder de detecção. Para aumentar o poder de detecção, um algoritmo utilizando informações de famílias de irmãs-completas e meio-irmãs paternas foi proposto (Le Roy et al., 1998).

Em qualquer delineamento, a detecção de QTL baseia-se na probabilidade condicional de presença de um genótipo de um QTL quando um genótipo do marcador é encontrado, apesar desta probabilidade ser função do delineamento experimental usado. Seleção em núcleos MOET tem sido implementada baseada na avaliação de dados de família de irmãos completos e meio-irmãos (Schrooten e van Arendonk, 1992). Nosso objetivo foi avaliar o poder de algumas metodologias para validar QTL utilizando esta estrutura de família.

Material e Métodos

O programa GENESYS foi utilizado para simulação dos dados genealógicos, fenotípicos, genotípicos e moleculares, em uma estrutura populacional de núcleo MOET (famílias de irmãos completos/meio-irmãos), usando um modelo gênico (Euclides, 1996). Os métodos numéricos implementados incluem procedimentos para mapear QTL em populações exogâmicas, com meio-irmãos e irmãos completos, usando múltiplos marcadores, para característica única.

Foram gerados 50 marcadores e 100 QTL, ambos dialélicos, distribuídos em cromossomos autossômicos, segundo regiões de consenso levantadas em estudos de mapeamento de QTL de características de produção de leite. Os efeitos genéticos aditivos e de meio foram normalmente distribuídos, de modo que a variância resultasse na herdabilidade de 0,29. Interações dos efeitos aditivo e não-aditivo foram ignoradas. Assumiram-se, inicialmente, frequências alélicas intermediárias (0,5) e uniformes em ambos os sexos.

A população-base do núcleo (5 touros e 25 vacas) foi selecionada de 25 touros e 500 vacas. A cada geração, foram obtidos 5 irmãos completos por família e 125 filhos por touro, dos quais se selecionaram os pais da geração seguinte, evitando-se acasalamento entre indivíduos aparentados. Após 20 gerações, acumularam-se 2500 registros dos indivíduos, de seus respectivos pai e mãe, sexo, genótipos para 50 marcadores e de seu valor genético para a característica. A partir deste arquivo, foram constituídos um, com dados de uma família por touro e apenas irmãos completos (IC), em um total de 100 indivíduos, e outro com apenas um filho por família e apenas meios-irmãos (MI) paternos, em um total de 500 indivíduos.

Para as análises de associação, optou-se, em função da simplicidade e velocidade computacional, pelo método da regressão para dados discretos ou contínuos, para análise de marcadores únicos, utilizando-se os procedimentos do SAS (SAS, 2003). O modelo fixo utilizado foi:

$$y = Xb + e$$

onde:

y= vetor de valores genotípicos;

X= matriz de incidência relacionando os efeitos fixos dos marcadores aos valores genotípicos;

b= vetor de efeitos fixos dos marcadores;

e= resíduos.

Para qualquer estrutura de família, a covariância entre os irmãos j e j' dentro da família i, assumindo-se endogamia nula, é igual a:

$$V_i = Var \begin{bmatrix} Y_{ij} \\ Y_{ij'} \end{bmatrix} = \sigma^2 C_i$$

onde

$$C_i = \begin{bmatrix} 1 & Cov[Y_{ij}, Y_{ij'}] \\ Cov[Y_{ij}, Y_{ij'}] & 1 \end{bmatrix}$$

$$Cov[Y_{ij}, Y_{ij'}] = \pi_q \sigma^2 + \frac{1}{2} \sigma^2 a \quad (\text{irmãos completos})$$

$$Cov[Y_{ij}, Y_{ij'}] = \pi_q \sigma^2 + \frac{1}{4} \sigma^2 a \quad (\text{meio-irmãos})$$

sendo π_q a proporção de alelos idênticos por descendência no loco candidato compartilhado pelos irmãos.

A significância da variação entre as classes genotípicas foi acessada pelo teste de 'F' e a dos efeitos aditivos associados a cada marcador, por meio de contrastes que compararam as médias das duas classes de homozigotos. Para o efeito de dominância, os contrastes compararam a média dos heterozigotos à média das classes de homozigotos. Para ambos, utilizou-se o teste de 't'. Os efeitos de substituição gênica foram estimados por meio de contrastes específicos.

Resultados e Discussão

Verificaram-se frequências genotípicas e alélicas diversificadas para cada um marcadores gerados, porém, as frequências de um mesmo marcador em cada arquivo (IC+MI, IC, MI) foram semelhantes, o que permitiu comparar os resultados das análises destes arquivos. Na Tabela 1 são apresentados os resultados das análises de associação para os marcadores significativos. A magnitude do efeito de cada marcador sobre a variação genotípica, em função dos níveis de probabilidade das análises do arquivo IC+MI, mostraram que os níveis mínimos de significância para se aceitar a associação marcador-QTL não foram atingidos quando os efeitos do QTL eram pequenos (Israel e Weller, 2002). A magnitude mínima dos efeitos estimados, para valor de $p < 0,01$, foi 0,0027 e, para valor de $p < 0,0001$, 0,0073.

Em magnitudes elevadas, mesmo sob frequências alélicas extremas, a possibilidade de detecção foi maior. Isto pode ser constatado para M01, M18 e M28 ($p < 0,0001$). Para M02, M14 e M24, encontrou-se significância mesmo para os efeitos de magnitude pequena, provavelmente porque as frequências alélicas estavam próximas às intermediárias. Estes são aspectos importantes na definição de marcadores para o mapeamento fino de regiões candidatas (Lynch e Walsh, 1997).

Ao se comparar os resultados das regressões quando a variável foi considerada classificatória (REG), àqueles considerando-a contínua (AV), verificaram-se pequenas diferenças na significância dos efeitos, sendo que M36 e M37 foram significativos ($P < 0,05$ e $P < 0,03$, respectivamente) apenas para REG, enquanto M27 e M40 ($P < 0,05$ e $P < 0,02$, respectivamente), apenas para AV. Nas análises de correlação entre as estimativas do efeito do marcador e entre as significâncias dos mesmos, resultantes das regressões utilizando o arquivo IC+MI para $p < 0,05$, observaram-se valores altos ($r = 0,99$ e $0,62$, respectivamente). Se apenas valores mais conservadores ($p < 0,01$) fossem considerados nas análises de correlação, esta se tornaria 100%, o que evidencia eficiência semelhante destas metodologias na detecção de associação entre marcador e suposto QTL em populações exogâmicas (Haley e Knott, 1992).

As análises de correlação mostraram também que as AV, utilizando apenas informações de MI, e as REG, utilizando apenas informações de IC ($r = 0,94$ e $0,95$, respectivamente), detectaram valores semelhantes para os efeitos dos marcadores. Porém, considerando-se os níveis de probabilidade, estas análises não detectaram significância na associação do marcador ao suposto QTL. Estes resultados podem ser atribuídos ao tamanho amostral destes arquivos, pois, neste caso, a acurácia das estimativas da variância genética será menor, diminuindo o poder de detecção de alguns QTL. Apesar da diminuição significativa do poder de detecção de associação, uma vez que não se conhecem a fase de ligação e o grau de desequilíbrio de ligação, e que o número de indivíduos genotipados é pequeno, este estudo verificou que a utilização conjunta das informações IC+MI, por meio de REG e AV, em populações exogâmicas, permitiram detectar a associação marcador-QTL (Le Roy et al., 1998).

Na Tabela 2, são apresentados o nível de probabilidade e a magnitude dos efeitos dos marcadores. M01, M03, M17, M18 e M28 mostraram maior efeito aditivo ($p < 0,001$). M27 e M40 apresentaram efeito de dominância ($p < 0,05$).

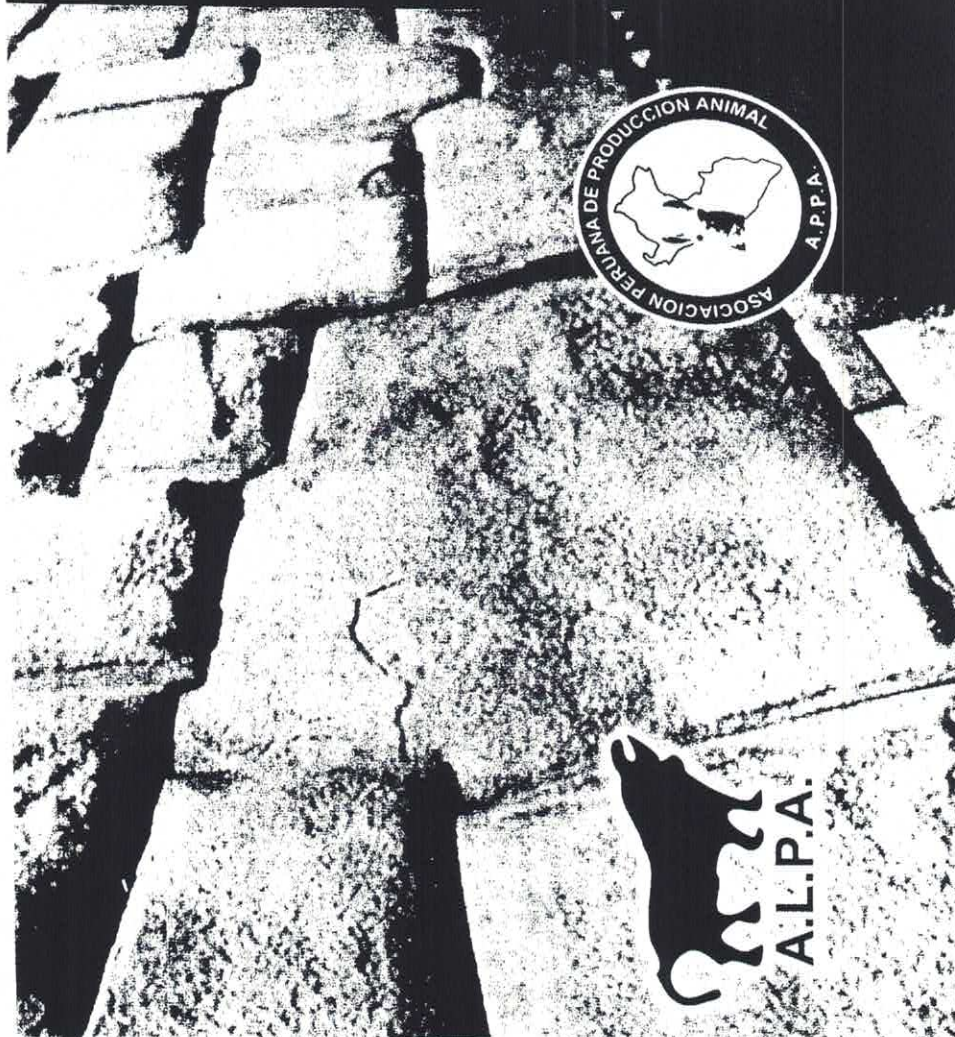
Conclusões

Apesar da pouca sensibilidade dos delineamentos baseados em populações exogâmicas e de tamanho pequeno, verificou-se que a mistura de dados de famílias de irmãos completos e meio-irmãos, analisados como variável discreta ou contínua por meio de regressão, permitiu validar a associação entre o marcador e o suposto QTL, até mesmo em níveis de significância conservadores.

Referências Bibliográficas

- Beckman, J.S.; Soller, M. 1989. Detection of linkage between marker loci and loci affecting quantitative traits in crosses between segregating populations. *Theor. Appl. Genet.* 76: 228-236.
- Bink, M.; Janss, LLG; Quaas, RL. 2000. Markov chain Monte Carlo for mapping a quantitative trait locus in outbred populations. *Genetical Research*, 75: 231-241.

- Brascamp, E.W.; van Arendonk, J.A.M.; Groen, A.F. 1993. Economic appraisal of the utilization of genetic markers in dairy cattle breeding. *J. Dairy Sci.* 76: 1204-1213.
- Coppieters, W.; Kvasz, A.; Arranz J.J. et al. 1999. The great-grand-daughter design: a simple strategy to increase the power of a grand-daughter design for QTL mapping. *Genet. Res. Camb.* 74: 189-199.
- Darvasi, A.; Weinreb, A.; Minke, V. et al. 1993. Detecting marker-QTL linkage and estimating QTL gene effect and map location using a calculated genetic map. *Genetics*, 134: 943-951.
- Dekkers, J.C.M. 2004. Commercial application of marker- and gene-assisted selection in livestock: Strategies and lessons. *J. Anim. Sci.* 82:313-328.
- Euclides, R.F. 1996. Sistema para Simulação Genesys.
- Fernando, R.L.; Grossman, M. 1989. Marker assisted selection using best linear unbiased prediction. *Genet. Sel. Evol.* 21: 467-477.
- Haley, C.S.; Knott, S.A. 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* 69: 315-324.
- Israel, C.; Weller, J.I. 2002. Estimation of quantitative trait loci effects in dairy cattle populations. *J. Dairy Sci.* 85: 1285-1297.
- Kashi, Y.; Hallerman, E.; Soller, M. 1990. Marker-assisted selection of candidate bulls for progeny testing programmes. *Anim. Prod.* 51: 63-74.
- Le Roy, P.; Elsen, J.M.; Boichard, D. et al. 1998. An algorithm for QTL detection in mixture of full and half sib families. *Proceedings of the 6th World Congress on Genetic Applied to Livestock Production*, 26: 257-260.
- Lynch, M.; Walsh, B. 1997. Genetic and analysis of quantitative traits. Sinauer Association, Sunderland. 980 p.
- SAS Institute, Inc.. 2003. SAS / SAS User's guide, Version 9.1th. 9th ed. SAS Ins., Inc., Cary.
- Schrooten, C.; van Arendonk, J.A.M. 1992. Stochastic simulation of dairy cattle breeding schemes: genetic evaluations of nucleus size type. *J. Anim. Breed. Genet.* 109: 1-15.
- Simpson, S.P. 1989. Detection of linkage between quantitative trait loci and restriction fragment length polymorphisms using inbred lines. *Theor. Appl. Genet.* 77: 815-819.



V CONGRESO INTERNACIONAL DE GANADERIA DE DOBLE PROPOSITO



XX REUNION ASOCIACION LATINOAMERICANA DE PRODUCCION ANIMAL (ALPA)
XXX REUNION ASOCIACION PERUANA DE PRODUCCION ANIMAL (APPA)
V CONGRESO INTERNACIONAL DE GANADERIA DE DOBLE PROPOSITO

Del 21 al 25 Octubre 2007 - Cuzco - Perú

INFORMES: appa.alpa2007@gmail.com inscripciones.alpa2007@gmail.com

www.alpa.org.pe

