

Detecção de QTL em dados de famílias estruturadas como as de um núcleo MOET por meio do método da regressão

[*QTL detection in MOET nucleus family structure-like data using the regression method*]

M.G.C.D. Peixoto¹, M.L. Martinez¹, R.L. Teodoro¹, M.A. Machado¹, M.R.S. Carvalho²,
F.C. Gomes¹, R.S. Verneque¹

¹Embrapa Gado de Leite
Rua Eugênio do Nascimento, 610
36038-330 – Juiz de Fora, MG

²Instituto de Ciências Biológicas - UFMG – Belo Horizonte, MG

Recebido em 11 de setembro de 2008

Aceito em 15 de maio de 2009

E-mail: gaby@cnpgl.embrapa.br

Apoio: FAPEMIG

RESUMO

O objetivo deste trabalho foi avaliar a eficiência do método de regressão em detectar QTL com base na utilização de dados da estrutura de família (irmãos completos e meios-irmãos), como aqueles gerados em um núcleo MOET. Foram simulados dados fenotípicos e genotípicos em uma estrutura de núcleo MOET fechado de seleção. Três arquivos foram analisados, contendo: a) informações conjuntas de irmãos completos e meios-irmãos; b) apenas informações de irmãos completos e c) apenas informações de meios-irmãos. Verificou-se que o método da regressão, para dados discretos ou contínuos, foi capaz de detectar associações entre marcador e QTL em níveis bastante expressivos de significância ($P < 0.001$ e $P < 0.0001$), quando se utilizou o arquivo que continha informações conjuntas de irmãos completos e meios-irmãos. Os resultados indicaram a possibilidade de utilização dessa metodologia para estudos de detecção/validação de QTL em rebanhos ou núcleos de seleção que utilizam MOET.

Palavras-chave: QTL, poder de detecção, marcador molecular, estrutura de família

ABSTRACT

The objective of this study was to evaluate the efficiency of the regression method to detect QTL using data from full and half-sib families, like those generated in a MOET nucleus. For this study, genotypic and phenotypic data were simulated in a structure of a closed selection MOET nucleus. Three files were analyzed containing: a) the joint information of full and half sibs; b) only full sibs data; and c) only half sibs data. The method of regression, for continuous or discrete data, was able to detect associations of marker and QTL in very expressive levels of significance ($P < 0.001$ $P < 0.0001$), when the file containing the joint information of full and half sisters was used. The results indicated the possibility of using this methodology for studies of QTL detection / validation in MOET nucleus or herds under selection.

Keywords: QTL, power of detection, molecular marker, family structure

INTRODUÇÃO

A utilização de marcadores tem sido indicada como importante ferramenta para o monitoramento da diversidade genética, o planejamento de acasalamentos e a seleção assistida (Brascamp et al., 1993). Apesar de a eficiência da seleção assistida por marcadores não estar bem clara e documentada (Dekkers, 2004), tal seleção constitui uma das muitas estratégias propostas para aumentar o progresso genético em núcleos MOET de seleção, minimizando a endogamia.

A detecção de marcadores associados a QTL depende 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 (Martinez e Vukasinovic, 2000; 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 de mapeamento

por meio desses 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 para investigar a ligação marcador-QTL. O delineamento mais eficiente baseia-se na obtenção de linhagens endogâmicas divergentes, pois o desequilíbrio máximo de ligação é produzido na geração F1 do cruzamento entre as linhagens.

Uma série de metodologias foi desenvolvida, permitindo utilizar toda a informação disponível e maximizar o poder de detecção de QTL (Nicholas, 2006). Dentre estas: análise de informações de segregação incerta por meio do método da 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, por meio da amostragem de Gibbs e do método da Cadeia de Markov/Monte Carlo (Bink e van Arendonk, 1999; Bink et al., 2000).

Apesar da perda de poder de detecção, considerando que nem todos os indivíduos são informativos e a existência de diferenças individuais quanto à fase de ligação, a análise dentro de uma população exogâmica pode tornar evidentes as diferenças na segregação do QTL (Lynch e Walsh, 1998). Em gado de leite, usualmente, o desequilíbrio de ligação entre marcador e QTL pode ser encontrado dentro de famílias de meias-irmãs paternas, sendo que, quanto maior o número de filhas por reprodutor, maior o poder de detecção do QTL. Com vistas a aumentar ainda mais o poder de detecção, um algoritmo utilizando as informações de famílias de irmãs-completas e meias-irmãs paternas foi proposto (Le Roy et al., 1998).

No mapeamento de QTL por meio de marcadores, com quaisquer dos métodos ou delineamentos empregados, as informações fenotípicas sobre a característica quantitativa de interesse referem-se a irmãs completas ou a meias-irmãs (van der Beek et al., 1995). Esquemas de seleção em núcleos MOET têm sido implantados em alguns países (Strathie e McGuirk, 1995; Penna et al., 1998; Jokinen et al., 2003). Nesses esquemas, são produzidas famílias de irmãos completos e meios-irmãos, cujos dados produtivos são utilizados na avaliação genética dos animais. O objetivo deste estudo de simulação foi avaliar a eficiência de metodologias baseadas no método de regressão em detectar QTL com base na utilização de dados da estrutura de família (irmãos completos e meios-irmãos), como aqueles gerados em um núcleo MOET.

MATERIAL E MÉTODOS

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

Foram gerados 50 marcadores e 100 QTL, ambos dialélicos, distribuídos em cromossomos autossômicos, segundo regiões de consenso detectadas em estudos de mapeamento de QTL para 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 (cinco touros e 25 vacas) foi selecionada de 25 touros e 500 vacas. A cada geração, foram obtidos cinco irmãos completos por família e 125 filhos (irmãos completos e meios-irmãos) 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. Baseando-se nesse arquivo, dois outros foram aleatoriamente 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 qualquer estrutura de família, a covariância entre os irmãos j e j' dentro da família i , assumindo-se endogamia nula, é igual a:

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

Detecção de QTL em dados...

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

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

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

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

Para o estudo de associação, optou-se pelo método da regressão, para dados discretos (R) ou contínuos (AV), utilizando-se os seguintes arquivos, contendo diferentes fontes de informação: irmãos completos + meios-irmãos (IC+MI), apenas irmãos completos (IC) ou apenas meios-irmãos (MI). Foram realizadas análises de marcadores únicos, usando-se os procedimentos PROC REG, PROC GLM e PROC CORR do pacote computacional SAS/2003. O modelo fixo utilizado foi:

$y = Xb + e$, em que:

y= vetor de valores genéticos para a característica;

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

b= vetor de efeitos fixos dos marcadores;

e= resíduos.

A significância da variação entre as classes genótípicas 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 se verificar o desvio de dominância, os contrastes compararam a média dos heterozigotos à média das classes de homozigotos. Para ambos, utilizou-se o teste t. Os efeitos de substituição gênica foram estimados por meio de contrastes específicos.

RESULTADOS E DISCUSSÃO

Nas Tab. 1, 2, 3 e 4 são apresentadas as frequências genótípicas e alélicas dos marcadores nos diferentes conjuntos de dados utilizados nas análises de associação. Pode-se verificar que, com qualquer dos conjuntos de dados (IC+MI, IC ou MI), as frequências genótípicas e alélicas geradas para cada um dos marcadores foram diversificadas, porém, quando consideradas as frequências para um mesmo marcador em cada arquivo, verificou-se que estas se mantiveram semelhantes, o que permitiu efetuar comparações dos resultados obtidos nas análises com cada um dos arquivos.

Tabela 1. Frequências genótípicas observadas nos locos marcadores gerados para todos os indivíduos (irmãos completos e meios-irmãos)

Genótipo	Marcador									
	M01	M02	M03	M04	M05	M06	M07	M08	M09	M10
0	41,25	24,68	6,08	35,96	44,44	24,00	9,08	24,52	16,56	48,44
1	48,00	51,20	38,88	47,68	43,76	49,68	40,08	50,88	48,08	42,92
2	10,48	24,12	55,04	16,36	11,80	26,32	50,84	24,60	35,36	8,64
	M11	M12	M13	M14	M15	M16	M17	M18	M19	M20
0	11,04	32,56	53,28	13,16	6,28	6,04	79,84	60,88	24,88	9,12
1	47,08	49,28	38,68	45,48	35,56	36,40	18,84	34,88	50,96	42,44
2	41,88	18,16	8,04	41,36	58,16	57,56	1,32	4,24	24,16	48,44
	M21	M22	M23	M24	M25	M26	M27	M28	M29	M30
0	2,56	2,00	11,32	16,24	16,12	17,16	31,96	3,76	13,04	20,92
1	27,08	26,40	43,04	45,16	45,36	49,88	49,92	30,60	48,72	48,64
2	70,36	71,60	45,84	38,60	38,52	32,96	18,12	65,64	38,24	30,44
	M31	M32	M33	M34	M35	M36	M37	M38	M39	M40
0	28,84	38,68	42,24	42,48	29,44	36,36	5,12	77,84	11,20	26,24
1	48,80	46,72	47,96	47,68	48,40	48,64	38,24	20,64	88,80	50,88
2	22,36	14,60	9,80	9,84	22,16	15,00	56,64	1,52	-	22,88
	M41	M42	M43	M44	M45	M46	M47	M48	M49	M50
0	15,04	8,48	43,08	18,48	66,16	12,72	23,16	54,04	31,32	28,40
1	48,80	40,76	45,28	50,08	30,84	44,08	49,56	40,72	48,28	48,20
2	36,16	50,76	11,64	31,44	3,00	43,20	27,28	5,24	20,40	23,40

Tabela 2. Frequências alélicas observadas nos locos marcadores gerados para todos os indivíduos (irmãos completos e meios-irmãos)

Alelo	Marcador									
	M01	M02	M03	M04	M05	M06	M07	M08	M09	M10
A ₁	0,65	0,50	0,26	0,60	0,66	0,49	0,29	0,50	0,41	0,70
A ₂	0,35	0,50	0,74	0,40	0,34	0,51	0,71	0,50	0,59	0,30
	M11	M12	M13	M14	M15	M16	M17	M18	M19	M20
A ₁	0,35	0,57	0,73	0,36	0,24	0,24	0,89	0,78	0,50	0,30
A ₂	0,65	0,43	0,27	0,64	0,76	0,76	0,11	0,22	0,50	0,70
	M21	M22	M23	M24	M25	M26	M27	M28	M29	M30
A ₁	0,16	0,15	0,33	0,39	0,39	0,42	0,57	0,19	0,37	0,45
A ₂	0,84	0,85	0,67	0,61	0,61	0,58	0,43	0,81	0,63	0,55
	M31	M32	M33	M34	M35	M36	M37	M38	M39	M40
A ₁	0,53	0,62	0,66	0,66	0,54	0,61	0,24	0,88	0,56	0,52
A ₂	0,47	0,38	0,34	0,34	0,46	0,39	0,76	0,12	0,44	0,48
	M41	M42	M43	M44	M45	M46	M47	M48	M49	M50
A ₁	0,39	0,34	0,66	0,44	0,82	0,35	0,48	0,74	0,55	0,52
A ₂	0,61	0,66	0,34	0,56	0,18	0,65	0,52	0,26	0,45	0,48

Tabela 3. Frequências genotípicas observadas nos locos marcadores gerados considerando apenas famílias de meios-irmãos

Genótipo	Marcador									
	M01	M02	M03	M04	M05	M06	M07	M08	M09	M10
0	44,20	23,60	5,20	36,60	44,60	24,40	8,40	26,40	13,80	46,00
1	46,80	52,80	41,20	46,20	42,80	49,00	42,60	47,20	52,40	45,20
2	9,00	23,60	53,60	17,20	12,60	26,60	49,00	26,40	33,80	8,80
	M11	M12	M13	M14	M15	M16	M17	M18	M19	M20
0	9,40	33,80	52,00	12,60	5,80	5,20	78,40	62,40	27,60	8,20
1	49,40	47,00	39,40	46,20	36,80	34,40	20,40	32,60	49,00	43,00
2	41,20	19,20	8,60	41,20	57,40	60,40	1,20	5,00	23,40	48,80
	M21	M22	M23	M24	M25	M26	M27	M28	M29	M30
0	3,00	2,00	11,32	17,80	17,80	15,80	30,40	3,40	12,60	21,80
1	28,60	27,00	45,60	45,00	45,20	50,60	50,20	30,00	48,80	47,60
2	68,40	71,00	43,20	37,20	37,00	33,60	19,40	66,60	38,60	30,60
	M31	M32	M33	M34	M35	M36	M37	M38	M39	M40
0	28,20	38,80	42,20	41,80	30,80	37,60	4,00	77,00	-	26,20
1	49,80	48,80	49,40	49,80	47,00	45,80	40,60	21,20	12,80	49,40
2	22,00	12,40	8,40	8,40	22,20	16,60	55,40	1,80	87,20	24,40
	M41	M42	M43	M44	M45	M46	M47	M48	M49	M50
0	14,00	6,60	42,00	19,20	67,60	13,20	23,40	56,40	31,40	28,40
1	48,00	42,40	47,20	49,80	29,40	43,40	48,80	37,60	47,00	49,00
2	38,00	51,00	10,80	31,00	3,00	43,40	27,80	6,00	21,60	22,60

Tabela 4. Frequências genotípicas observadas nos locos marcadores gerados considerando apenas famílias de irmãos completos

Genótipo	Marcador									
	M01	M02	M03	M04	M05	M06	M07	M08	M09	M10
0	45,00	16,00	5,00	30,00	48,00	31,00	4,00	36,00	17,00	47,00
1	39,00	43,00	51,00	45,00	42,00	48,00	35,00	50,00	55,00	41,00
2	16,00	41,00	44,00	25,00	10,00	21,00	61,00	14,00	28,00	12,00
	M11	M12	M13	M14	M15	M16	M17	M18	M19	M20
0	16,00	28,00	68,00	18,00	2,00	4,00	85,00	57,00	26,00	7,00
1	47,00	51,00	24,00	40,00	30,00	42,00	15,00	34,00	45,00	37,00
2	37,00	21,00	8,00	42,00	68,00	54,00	-	9,00	29,00	56,00
	M21	M22	M23	M24	M25	M26	M27	M28	M29	M30
0	3,00	5,00	5,00	14,00	13,00	15,00	43,00	-	16,00	24,00
1	16,00	43,00	37,00	41,00	42,00	46,00	49,00	20,00	53,00	45,00
2	81,00	52,00	58,00	45,00	45,00	39,00	8,00	80,00	31,00	31,00
	M31	M32	M33	M34	M35	M36	M37	M38	M39	M40
0	17,00	46,00	32,00	35,00	23,00	35,00	2,00	78,00	-	25,00
1	44,00	45,00	59,00	56,00	46,00	43,00	40,00	17,00	3,00	62,00
2	39,00	9,00	9,00	9,00	31,00	22,00	58,00	5,00	97,00	13,00
	M41	M42	M43	M44	M45	M46	M47	M48	M49	M50
0	11,00	3,00	41,00	4,00	64,00	22,00	37,00	67,00	32,00	32,00
1	57,00	36,00	46,00	54,00	36,00	47,00	36,00	31,00	44,00	48,00
2	32,00	61,00	13,00	42,00	-	31,00	27,00	2,00	24,00	20,00

Na Tab. 5 são apresentados os resultados das análises de associação para todos os marcadores, ressaltando-se aqueles cujos efeitos foram significativos. Os resultados para M39 foram ignorados em função da existência de apenas dois dos três genótipos possíveis (Tab. 1). Considerando-se o arquivo IC+MI, como dado discreto (R)

Detecção de QTL em dados...

ou contínuo (AV), foi verificado que a análise de regressão de cada marcador sobre a variação genética não foi capaz de detectar associações entre o marcador e o QTL quando a magnitude dos efeitos do QTL foi pequena (<0,0015, para R, e <0,0020, para AV), nem mesmo assumindo-se valores menos conservadores de probabilidade ($P<0,05$) para a aceitação de associação. A magnitude mínima dos efeitos estimados, que permitiu a detecção de associação para cada valor de probabilidade, foi de 0,0015 ($P<0,05$), 0,0027 ($P<0,01$) e 0,0073 ($P<0,0001$).

Tabela 5. Efeitos do marcador, valor de probabilidade e significância obtidos com os diferentes métodos e arquivos utilizados

Marcador	Efeito R	Prob.	Efeito AV	Prob.	Efeito AV-MI	Prob.	Efeito AV-IC	Prob.	Efeito R-MI	Prob.	Efeito R-IC	Prob.
1	0,0220	0,0001*	0,0220	0,0001*	0,0290	0,0007*	0,0420	0,1246ns	0,0377	0,0528*	0,0242	0,0005*
2	0,0073	0,0001*	0,0073	0,0001*	0,0024	0,5553ns	0,0422	0,1236ns	0,0046	0,5028ns	0,0001	0,8097ns
3	0,0154	0,0001*	0,0164	0,0001*	0,0375	0,0001*	0,0228	0,3275ns	0,0080	0,3768ns	0,0374	0,0001*
5	0,0002	0,5185ns	0,0002	0,7834ns	0,0004	0,9074ns	0,0498	0,0839ns	0,0498	0,0256*	0,0000	0,8971ns
13	0,0018	0,0354*	0,0027	0,0347*	0,0030	0,4705ns	0,0050	0,7828ns	0,0050	0,4856ns	0,0012	0,4438ns
14	0,0089	0,0001*	0,0094	0,0001*	0,0079	0,1378ns	0,0560	0,0609ns	0,0556	0,0182*	0,0068	0,0657ns
15	0,0096	0,0001*	0,0097	0,0001*	0,0136	0,0334*	0,0673	0,0341*	0,0633	0,0116*	0,0135	0,0093*
16	0,0014	0,0625ns	0,0015	0,1544ns	0,012	0,0502*	0,0181	0,4128ns	0,0176	0,1877ns	0,0095	0,0292*
17	0,0162	0,0001*	0,0172	0,0001*	0,0213	0,0047*	0,0082	0,3709ns	0,0082	0,3709ns	0,0207	0,0012*
18	0,0230	0,0001*	0,0230	0,0001*	0,0321	0,0003*	0,0114	0,5728ns	0,0019	0,6704ns	0,0289	0,0001*
20	0,0121	0,0001*	0,0128	0,0001*	0,018	0,0110*	0,0306	0,2219ns	0,0291	0,0895ns	0,0150	0,0061*
21	0,0043	0,0010*	0,0051	0,0017*	0,0026	0,5245ns	0,0313	0,2137ns	0,002	0,6590ns	0,0010	0,4749ns
24	0,0056	0,0002*	0,0056	0,0009*	0,01	0,0828ns	0,0025	0,8873ns	0,0003	0,8621ns	0,0098	0,0271*
25	0,0055	0,0002*	0,0055	0,0011*	0,0097	0,0878ns	0,0036	0,8397ns	0,0005	0,8269ns	0,0095	0,0295*
26	0,0185	0,0001*	0,0186	0,0001*	0,008	0,1362ns	0,0095	0,6280ns	0,003	0,5869ns	0,0077	0,0497*
27	0,0005	0,2827ns	0,0024	0,0473*	0,0003	0,9250ns	0,0326	0,2002ns	0,0008	0,7857ns	0,0003	0,7040ns
28	0,0760	0,0001*	0,0760	0,0001*	0,1013	0,0001*	0,0175	0,1893ns	0,0175	0,1893ns	0,1004	0,0001*
29	0,0121	0,0001*	0,0122	0,0001*	0,0091	0,1030ns	0,0396	0,1407ns	0,0192	0,1690ns	0,0085	0,0391*
30	0,0040	0,0016*	0,004	0,0063*	0,0067	0,1899ns	0,0216	0,3460ns	0,0011	0,7444ns	0,0044	0,1382ns
32	0,0019	0,0275*	0,0042	0,0053*	0,0064	0,2016ns	0,0567	0,0590ns	0,0113	0,2921ns	0,0001	0,8419ns
33	0,0007	0,1744ns	0,0008	0,3639ns	0,0051	0,2814ns	0,0598	0,0503*	0,0467	0,0308*	0,0036	0,1790ns
34	0,0009	0,1268ns	0,0010	0,2860ns	0,0050	0,2888ns	0,0629	0,0429*	0,0521	0,0223*	0,0034	0,1942ns
35	0,0027	0,0098*	0,0029	0,0267*	0,0021	0,5950ns	0,0120	0,5567ns	0,0015	0,7035ns	0,0006	0,5812ns
36	0,0015	0,0492*	0,0016	0,1389ns	0,0004	0,9137ns	0,0571	0,0578ns	0,0092	0,3435ns	0,0001	0,8419ns
37	0,0019	0,0282*	0,0020	0,0822ns	0,0015	0,6825ns	0,0227	0,3283ns	0,0014	0,7113ns	0,0003	0,7154ns
40	0,0014	0,0594ns	0,0029	0,0250*	0,0029	0,4816ns	0,0046	0,7981ns	0,0027	0,6065ns	0,0011	0,4589ns
42	0,0100	0,0001*	0,0104	0,0001*	0,0191	0,0083*	0,0603	0,0489*	0,0241	0,1230ns	0,0151	0,0059*
43	0,0023	0,0169*	0,0028	0,0318	0,0059	0,2305ns	0,0168	0,4401ns	0,0160	0,2092ns	0,0001	0,8531ns
44	0,0029	0,0076*	0,0037	0,0101	0,0046	0,3165ns	0,0186	0,4015ns	0,0040	0,5312ns	0,0046	0,1292ns
46	0,0021	0,0205*	0,0024	0,0495	0,0061	0,2176ns	0,0116	0,5673ns	0,0042	0,5213ns	0,0056	0,0945ns
49	0,0152	0,0001*	0,0158	0,0001	0,0072	0,1675ns	0,0053	0,7722ns	0,0041	0,5248ns	0,0070	0,0610ns
50	0,0030	0,0059*	0,0034	0,0142	0,0015	0,6929ns	0,0670	0,0345	0,0084	0,3657ns	0,0014	0,3982ns

*significativo e ns:não-significativo.

R: regressão; AV: análise de variância; MI: apenas dados de meios-irmãos; IC: apenas irmãos completos.

Quando a magnitude dos efeitos foi elevada, mesmo sob frequências alélicas extremas, a possibilidade de detecção foi maior. Isso pode ser constatado para M01, M18 e M28 ($P<0,0001$). Para M02, M14 e M24, encontrou-se significância até mesmo para os efeitos de pequena magnitude, provavelmente porque as frequências alélicas estavam próximas às intermediárias. O efeito do marcador e a frequência de seus alelos são aspectos importantes na definição dos marcadores a serem utilizados para o mapeamento fino de regiões candidatas (Lynch e Walsh, 1998). Israel e Weller (2002) encontraram pequenas diferenças na estimativa dos efeitos do QTL ao utilizar diferentes métodos de detecção de QTL em dados em que pequena fração da população foi genotipada, assumindo-se diferentes frequências alélicas para os marcadores. Os autores, porém,

não verificaram diferenças expressivas no vício das estimativas ou nas estimativas das médias dos erros-padrão a frequências intermediárias ou extremas, apesar de os melhores resultados terem sido obtidos quando a frequência era intermediária.

Estes resultados são de grande interesse à validação de QTL identificados em algumas populações, com o objetivo de auxílio à seleção em outras populações. Uma vez comprovada a eficiência desta metodologia, a detecção/validação de associação poderá se dar em populações como a de um núcleo MOET; pequenas, mas com estrutura de família complexa o suficiente para esses estudos. Os resultados obtidos, embora promissores, devem, porém, ser avaliados com cautela, visto que as covariâncias entre irmãos completos ($\frac{1}{2}$) e meios-irmãos ($\frac{1}{4}$) não foram consideradas no modelo e podem, portanto, viciar os resultados. Métodos que consideram a covariância genética entre parentes, neste caso entre irmãos completos e meios-irmãos, deverão ser implementados para confirmação dos resultados encontrados neste estudo.

Ao comparar os resultados das regressões, utilizando-se o arquivo IC+MI, quando a variável foi considerada classificatória (R) àqueles em que a variável foi considerada 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 R, enquanto M27 e M40 ($P < 0,05$ e $P < 0,02$, respectivamente), apenas para AV (Tab. 5). Nas análises comparativas dos resultados de associação obtidos por meio dos diferentes métodos (R e AV) e apresentados nas Tab. 6 e 7, observou-se alta correlação entre os resultados para o efeito do marcador ($r = 0,99$) e sua significância ($r = 0,88$), considerando-se todos os resultados. Se apenas valores mais conservadores de probabilidade ($P < 0,01$) fossem considerados nas análises de correlação, esta se tornaria igual a um, mostrando que tais métodos teriam a mesma eficiência para detectar a associação entre marcador e o suposto QTL com dados de IC+MI. Segundo Haley e Knott (1992), as análises de regressão simples para detecção de QTL poderão ser mais eficientes se utilizarem marcadores flanqueadores.

Tabela 6. Correlações entre os efeitos estimados nas análises pelas diferentes metodologias

Método	R	AV	R - MI	AV - MI	R - IC	AV - IC
REG	1	0,99	0,08	0,94	0,95	-0,02
AV		1	0,08	0,94	0,95	-0,02
REG - MI			1	0,11	0,10	0,77
AV - MI				1	0,97	0,02
REG - IC					1	-0,04

R: regressão; AV: análise de variância; MI: apenas dados de meios-irmãos; IC: apenas irmãos completos.

Tabela 7. Correlações entre os valores de probabilidades alcançados nas análises para cada um dos marcadores gerados

Método	REG	AV	REG - MI	AV - MI	REG - IC	AV - IC
REG	1	0,88	0,05	0,66	0,52	0,21
AV		1	-0,11	0,65	0,50	0,11
REG - MI			1	0,19	0,08	0,70
AV - MI				1	0,75	0,14
REG - IC					1	0,04

R: regressão; AV: análise de variância; MI: apenas dados de meios-irmãos; IC: apenas irmãos completos.

Os resultados das análises dos diferentes conjuntos de dados e métodos mostraram também que as AV, utilizando apenas informações de MI, e as R, utilizando apenas informações de IC, apresentaram alta correlação com AV (IC+MI) e R(IC+MI), e detectaram valores semelhantes para os efeitos dos marcadores. Porém, considerando-se os valores de probabilidade, essas análises não detectaram significância na associação do marcador ao suposto QTL. Tais resultados podem ser atribuídos ao número de dados 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 R ou de AV, em populações com a estrutura de núcleos MOET, permitiu detectar a associação marcador-QTL (Le Roy et al., 1998).

Na Tab. 8 são apresentados o valor de probabilidade e a magnitude dos efeitos associados aos marcadores. Os marcadores M01, M03, M17, M18 e M28 foram os de maior efeito médio de substituição gênica ($P < 0,0001$). Assumindo-se níveis menos conservadores de significância, apenas os marcadores M27 e M40 apresentaram

Detecção de QTL em dados...

desvio significativo ($P < 0,03$ e $P < 0,05$, respectivamente) de dominância. Para M32, os efeitos aditivo e de dominância foram significativos ($P < 0,02$ e $P < 0,05$, respectivamente), o que pode ser reflexo da interferência de um QTL próximo, fato que, em dados reais, poderia ser explicado também pela ação epistática. Porém, níveis menos conservadores de significância foram assumidos e, portanto, estes resultados devem ser avaliados com reserva.

Tabela 8. Valores e significância dos efeitos genéticos aditivos (A) e de dominância (D) dos marcadores de efeitos significativos, obtidos das análises de variância, considerando-se todos os indivíduos (irmãos completos e meios-irmãos)

Marcador	A	Prob.	D	Prob.	Marcador	A	Prob.	D	Prob.
M01	4,87	0,0001	9,87	0,9873	M27	-0,94	0,1345	-1,88	0,0261
M02	-2,60	0,0001	-5,21	0,8418	M28	10,66	0,0001	21,33	0,8970
M03	5,30	0,0001	10,59	0,1012	M29	3,38	0,0001	6,76	0,6291
M13	-2,09	0,0096	-4,19	0,1299	M30	-1,93	0,0015	-3,86	0,6878
M14	-3,23	0,0001	-6,46	0,2617	M32	-1,88	0,0042	-3,75	0,0177
M15	3,64	0,0001	7,28	0,6916	M35	-1,50	0,0126	-2,997	0,4474
M17	8,61	0,0001	17,22	0,1246	M40	-1,09	0,0737	-2,18	0,0506
M18	5,25	0,0001	10,49	0,6159	M42	-3,78	0,0001	-7,55	0,2824
M20	4,13	0,0001	8,25	0,2104	M43	-1,85	0,0087	-3,71	0,2753
M21	-4,23	0,0018	-8,46	0,1550	M44	-1,48	0,0184	-2,95	0,1517
M24	-2,24	0,0004	-4,49	0,9386	M46	-1,67	0,0145	-3,33	0,4229
M25	-2,23	0,0004	-4,47	0,9563	M49	-3,59	0,0001	-7,18	0,2167
M26	-4,16	0,0001	-8,32	0,7236	M50	-1,68	0,0050	-3,35	0,3328

CONCLUSÕES

Apesar da pouca sensibilidade dos delineamentos baseados em populações exogâmicas e de tamanho pequeno, a metodologia de análise conjunta de dados de famílias de irmãos completos e de meios-irmãos por meio de regressão permitiu detectar a associação entre o marcador e o suposto QTL, até mesmo em níveis de significância conservadores. A frequência alélica e o efeito do marcador foram aspectos importantes para a eficiência dessa metodologia.

AGRADECIMENTOS

Ao Prof. Ricardo Frederico Euclides, do Departamento de Zootecnia da UFV, pela gentileza de gerar os dados utilizados neste estudo.

REFERÊNCIAS BIBLIOGRÁFICAS

- BECKMAN, J.S.; SOLLER, M. Detection of linkage between marker loci and loci affecting quantitative traits in crosses between segregating populations. *Theor. Appl. Genet.*, v.76, p.228-236, 1989.
- BINK M.C., JANSS L.L., QUAAS R.L. Markov chain Monte Carlo for mapping a quantitative trait locus in outbred populations. *Genet Res.*, 75:231-241, 2000.
- BINK, M.C.; VAN ARENDONK, J.A. Detection of quantitative trait loci in outbred populations with incomplete marker data. *Genetics*, v.151, p.409-420, 1999.
- BRASCAMP, E.W.; VAN ARENDONK, J.A.M.; GROEN, A.F. Economic appraisal of the utilization of genetic markers in dairy cattle breeding. *J. Dairy Sci.*, v.76, p.1204-1213, 1993.
- COPPIETERS, W.; KVASZ, A.; ARRANZ J.J. et al. The great-grand-daughter design: a simple strategy to increase the power of a grand-daughter design for QTL mapping. *Genet. Res.*, v.74, p.189-199, 1999.
- DARVASI A., WEINREB A., MINKE V. et al. Detecting marker-QTL linkage and estimating QTL gene effect and map location using a saturated genetic map. *Genetics*, v.134, p.943-951, 1993.
- DEKKERS, J.C.M. Commercial application of marker- and gene-assisted selection in livestock: Strategies and lessons. *J. Anim. Sci.*, v.82, p.313-328, 2004.

- EUCLYDES, R.F. Uso do sistema para simulação Genesys na avaliação de métodos de seleção clássicos e associados a marcadores moleculares. 1996. 149f. Tese (Doutorado) - Universidade Federal de Viçosa, Viçosa, MG.
- FERNANDO, R.L.; GROSSMAN, M. Marker assisted selection using best linear unbiased prediction. *Genet. Sel. Evol.*, v.21, p.467-477, 1989.
- HALEY, C.S.; KNOTT, S.A. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity*, v.69, p.315-324, 1992.
- ISRAEL, C.; WELLER, J.I. Estimation of quantitative trait loci effects in dairy cattle populations. *J. Dairy Sci.*, v.85, p.1285-1297, 2002.
- JOKINEN, J.; VANHAMÄKI, T.; TAMMIRANTA, N. et al. Five years of ASMO breeding scheme. Jokioinen: MTT Agrifood Res. Finland, NJF, 2003. Disponível em: <http://portal.mtt.fi/portal/page/portal/AGRONET/YHTEISET_HANKEET/NJF/NJF2003/3.PDF>, Acessado em: 17 set. 2007.
- KASHI, Y.; HALLERMAN, E.; SOLLER, M. Marker-assisted selection of candidate bulls for progeny testing programmes. *Anim. Prod.*, v.51, p.63-74, 1990.
- LE ROY, P.; ELSEN, J.M.; BOICHARD, D. et al. An algorithm for QTL detection in mixture of full and half sib families. In: WORLD CONGRESS ON GENETIC APPLIED TO LIVESTOCK PRODUCTION, 6., 1998, Armidale. *Proceedings...* Armidale: WCGALP, 1998.
- LYNCH, M.; WALSH, B. *Genetic and analysis of quantitative traits*. Sunderland: Sinauer Association, 1998. 980p.
- MARTINEZ, M.L.; VUKASINOVIC, N. Algoritmo para cálculo da proporção de genes idênticos por descendência para mapear QTL em famílias de meio-irmãs. *Rev. Bras. Zootec.*, v.29, p.443-451, 2000.
- NICHOLAS, F.W. Discovery, validation and delivery of DNA markers. *Aust. J. Exp. Agric.*, v.46, p.155-158, 2006.
- PENNA, V.M.; MADALENA, F.E.; ALVIM, M.T.T. Open MOET nucleus of selection in Guzerá. In: WORLD CONGRESS ON GENETIC APPLIED TO LIVESTOCK PRODUCTION, 6., 1998, Armidale. *Proceedings...* Armidale: WCGALP, 1998. (CD-ROM).
- SIMPSON, S.P. Detection of linkage between quantitative trait loci and restriction fragment length polymorphisms using inbred lines. *Theor. Appl. Genet.*, v.77, p.815-819, 1989.
- STRATHIE, R.J.; MCGUIRK, B.J. Developments with MOET dairy breeding scheme. *Br. Cattle Breed. Club Dig.*, v.50, p.9-15, 1995.
- van der BEEK, S.; VAN ARENDONK, J.A.M.; GRÖEN, A.F. Power of two and three-generation QTL mapping experiments in an outbred population containing full-sib or half-sib families. *Theor. Appl. Genet.*, v.91, p.1115-1124, 1995.