



MAPEAMENTO DE QTL EM FAMÍLIAS DE MEIOS-IRMÃOS POR MEIO DE GENES IDÊNTICOS POR DESCENDÊNCIA

MARCOS VINICIUS GUALBERTO BARBOSA DA SILVA¹, GUSTAVO HENRIQUE DE SOUZA², MÁRIO LUIZ MARTINEZ¹, ROBLEDO DE ALMEIDA TORRES³, PAULO SÁVIO LOPES³, MARCO ANTÔNIO MACHADO², ANA LUISA SOUZA AZEVEDO⁴

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

² Estudante de Doutorado do DZO-UFV

³ Professor da Universidade Federal de Viçosa – DZO, CEP 36570-000, Viçosa-MG

⁴ Bolsista da FAPEMIG

RESUMO - O procedimento de mapeamento por intervalo baseado em modelo aleatório foi aplicado para estudar a robustez e propriedades no mapeamento de QTL em populações simuladas, constituídas de famílias de meios-irmãos. As famílias de meios-irmãos foram simuladas considerando um segmento cromossômico de 100 cM com seis marcadores igualmente distribuídos no cromossomo a intervalos de 20 cM, com seis alelos de mesma frequência. O número de alelos do suposto QTL foi de 2, 4, 6 ou 10 alelos com frequências iguais. O número de famílias foi de 10 ou 100 e os tamanhos de famílias foram de 25 ou 50 com 10, 50 ou 100% da variância genética devido ao QTL. Na estimação dos parâmetros do QTL observou-se pouca influência do número de alelos do QTL, sendo a proporção da variância genética e o número de indivíduos analisados os fatores que mais influenciaram as estimativas.

PALAVRAS-CHAVE: genes idênticos por descendência, mapeamento por intervalo, QTL

MAPPING QTL IN HALF-SIBS FAMILIES USING IDENTICAL BY DESCENT GENES

ABSTRACT – An interval mapping procedure based on the random model approach was applied to investigate its robustness and properties for QTL mapping in simulated populations with half-sib structures. Populations were simulated taking into account a chromosomal segment of 100 cM with six markers equally distributed in 20 cM intervals with six alleles with same frequencies. The number of alleles of the putative QTL was set to 2, 4, 6 and 10 with the same frequencies. The number of families was set to 10 and 100, family sizes of 25 and 50 with 10, 50 and 100% of the genetic variation proportion due to QTL. In the estimation of QTL parameters little effect was observed for the number of QTL alleles. The proportion of genetic variance and the number of analyzed individuals were the factors that most influenced the estimates.

KEYWORDS: identical by descent genes, interval mapping, QTL

INTRODUÇÃO

O principal objetivo dos estudos de mapeamento de QTLs é identificar marcadores que possam ser utilizados em programas de melhoramento por meio de seleção assistida por marcadores (MAS). Como um exemplo da aplicação da MAS em gado de leite pode ser citada a pré-seleção de touros jovens para entrada em testes de progênie. Nesse caso, segundo Khatkar et al. (2004), as vantagens seriam os aumentos no diferencial de seleção, no ganho genético e também na redução do intervalo de gerações.

Grande número de delineamentos experimentais, envolvendo populações segregantes ou provenientes de cruzamentos entre linhagens endogâmicas, podem ser usados para o mapeamento de QTLs. Em espécies domésticas, como gado de leite, o custo elevado e o tempo necessário para formação de populações cruzadas podem ser proibitivos, sendo mais interessante o uso de rebanhos comerciais. Nesse caso, assume-se equilíbrio de ligação, passando o número de famílias e o tamanho delas a ter importância fundamental, dado que a fase de ligação pode ser diferente para cada família. Adicionalmente, o número de alelos do QTL está relacionado à quantidade de genótipos possíveis, tornando a análise mais complicada e demorada em razão do número de distribuições normais para construir a distribuição geral dos valores genotípicos.

Desta forma, objetivou-se, com este trabalho, verificar a robustez dos modelos aleatórios no mapeamento de QTL, a partir de informações provenientes de populações de meios-irmãos, com

diferentes números e tamanhos de famílias, de alelos do QTL e de percentagens da variação genética explicada pelo QTL.

MATERIAL E MÉTODOS

A técnica de simulação de Monte Carlo foi usada para gerar informações genotípicas e fenotípicas. Para o mapeamento de um QTL foram simuladas populações considerando um cromossomo de 100 cM, com a presença de um QTL a 50 cM e seis marcadores espaçados a cada 20 cM. Considerou-se a herdabilidade da característica como 0,25, o número de famílias igual a 10 e 100 repetições por tratamento, variando-se o número de filhas (25 e 50), o número de alelos por QTL (2, 4, 6 e 10) e a variação genética devido ao QTL (10, 50 e 100%).

Os dados fenotípicos foram simulados, com distribuição normal, média zero e variância igual a um, a partir de um modelo que incluiu os efeitos do genótipo do QTL, da contribuição do pai e da mãe para o valor poligênico da progênie, da amostragem mendeliana e do erro. As progênies foram geradas sob suposição de ausência de interferência e as frequências de recombinação foram calculadas usando-se a função de Haldane.

Nas análises foi usado o modelo proposto por Goldgar (1990), sob a pressuposição de ausência de interferência e de dominância, e que os *crossing-overs* seguiam a distribuição de Poisson e eram uniforme e independentemente distribuídos na região analisada. A variância fenotípica era igual a soma dos componentes do efeito genético do QTL (σ_q^2), do poligênico (σ_a^2) e do erro (σ_e^2). A covariância entre dois meios-irmãos não-endogâmicos da mesma família foi considerada como $Cov(y_{ij}, y_{ij'}) = \pi_q \sigma_q^2 + \frac{1}{4} \sigma_a^2$, onde π_q é a proporção de alelos IBD do suposto QTL (Martinez e Vukasinovic, 2000).

Não foram considerados efeitos fixos na simulação. Os tamanhos dessas famílias (25 e 50) foram escolhidos de modo a contemplar diferentes situações possíveis em populações comerciais de gado de leite. Ainda, foi escolhido o valor igual a 0,25 para a herdabilidade da característica e três diferentes proporções de variância explicada pelo QTL (10, 50 e 100%). O procedimento de máxima verossimilhança (ML), no mapeamento por intervalo, foi aplicado aos dados simulados, em que o cromossomo de 100 cM foi percorrido a passos de 2 cM da esquerda para a direita. A função de verossimilhança foi maximizada usando o algoritmo simplex. A posição com a maior razão de verossimilhança (LR) foi aceita como a mais provável localização do QTL. Para cada combinação de parâmetros, a simulação e as análises foram repetidas cem vezes. As acurácias das estimativas das posições dos QTL foram determinadas a partir da variância entre repetições e calculadas como quatro vezes o erro-padrão empírico. O valor empírico de limiar foi definido como o 95º percentil da distribuição empírica do teste LR sob a hipótese nula de que não havia nenhum QTL no segmento cromossômico.

RESULTADOS E DISCUSSÃO

As estimativas para as localizações dos QTL, com o correspondente intervalo de confiança, para os diferentes números de famílias, tamanhos de família, número de alelos do QTL e proporções de variância genética devido aos QTL são apresentados na Tabela 1.

Ao se considerar as famílias com 25 progênies, o aumento no número de alelos do QTL teve pouca influência sobre a localização do QTL. Com 10% da variância genética devido ao QTL, as estimativas de localização foram subestimadas e com 50% da variância genética, as estimativas de localização (49,92 à 53,88) foram as mais próximas do valor real (50) principalmente para 4, 6 e 10 alelos do QTL. Para 100% da variância genética devido ao QTL, as estimativas de localização foram superestimadas (53,24 a 56,16), havendo o indício da existência de um QTL, embora o valor de LR nesta situação tenha ficado abaixo do valor empírico de limiar (9,2776).

De modo geral, o número de alelos do QTL não influenciou nas estimativas dos intervalos de confiança. Ocorreram pequenas variações dentro de cada combinação de tamanho de família e variância genética. O aumento da proporção da variância genética devido ao QTL de 10 para 100% resultou na diminuição do intervalo de confiança das estimativas.

Quando foram consideradas dez famílias e tamanho de família igual a 50, o número de alelos do QTL também não influenciou na localização deste, sendo o melhor resultado obtido com 2 alelos do QTL e 100% da variância genética devido ao QTL. Para todas as três variâncias estudadas, os valores máximos de LR obtidos ficaram abaixo do valor de limiar.

Ao se considerar baixa proporção da variância genética devido ao QTL (10%), houve a tendência de indicar erroneamente a presença de QTLs, independentemente do tamanho de família e número de alelos do QTL. Uma possível explicação para este fato, segundo Martinez et al. (1999), pode ser a baixa herdabilidade da característica, aliada à reduzida variação genética explicada pelo QTL e ao número de famílias utilizado, tornando as estimativas viesadas.

TABELA 1. Efeito do número de alelos do QTL, número e tamanho da família e proporção da variância genética devido ao QTL sobre a estimativa da posição (cM) e intervalos de confiança (IC)

NF	TF	NA	Proporção da variância genética devido ao QTL (%)					
			10		50		100	
			cM	IC	cM	IC	cM	IC
10	25	2	46,38	12,55	53,88	9,81	56,16	7,90
		4	47,28	12,30	50,36	10,45	53,24	8,99
		6	48,48	12,28	49,92	10,33	53,50	8,22
		10	48,28	12,25	50,76	9,98	53,40	8,62
10	50	2	59,36	12,21	56,30	12,14	52,90	11,37
		4	61,20	12,44	55,80	13,24	54,26	11,43
		6	59,60	12,91	55,40	13,19	54,60	11,27
		10	62,36	12,49	57,54	12,50	54,82	11,57
100	50	4	-	-	-	-	50,30	2,40

NF = número de famílias

TF = Tamanho da família; NA = número de alelos do QTL; cM = local estimado do QTL; IC = intervalo de confiança.

Aumentando-se o número de famílias para 100 e o número de indivíduos dentro de cada família para 50, foi possível a localização quase exata do QTL (50,3), com reduzido intervalo de confiança (2,4) (Figura 1). O aumento no número de famílias melhorou as estimativas de localização e contribuiu para a redução dos valores dos intervalos de confiança, melhorando a acurácia das estimativas, indicando que como uma das possíveis causas para o viés observado pode ser o reduzido número de famílias utilizadas.

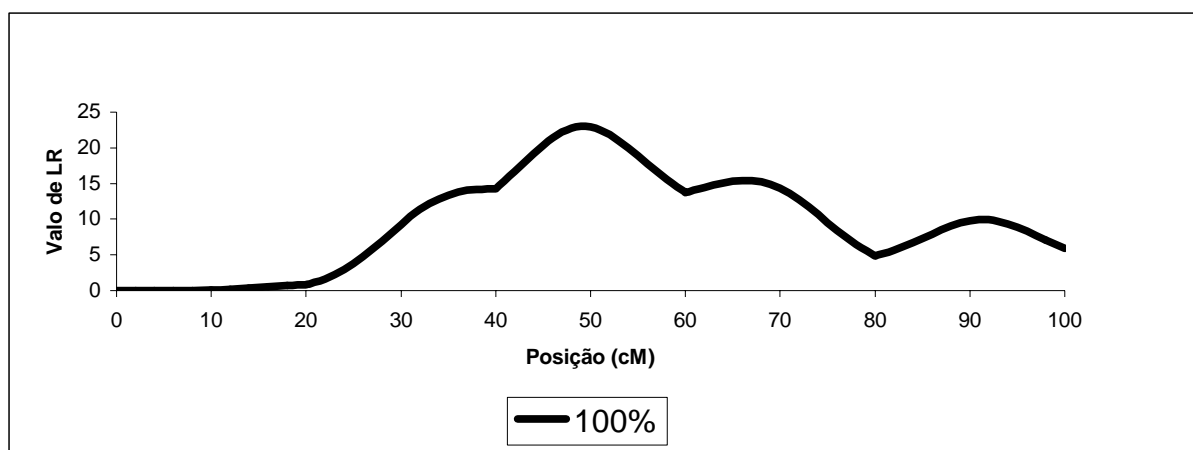


FIGURA 1. Valores de LR para 100% da proporção da variância genética devido ao QTL, com número de família = 100 e tamanho de família = 50, com número de alelos do QTL = 4

CONCLUSÕES

Na estimação dos parâmetros do QTL, observou-se pouca influência do número de alelos do QTL, sendo a proporção da variância genética e o número de indivíduos analisados dentro de cada família os fatores que mais influenciaram nas estimativas.



REFERÊNCIAS BIBLIOGRÁFICAS

- GOLDGARD, D. E. Multipoint analysis of human quantitative genetic variation. **Am. J. Hum. Genet.**, v. 47, p. 957-967. 1990.
- KHATKAR, M. S.; THOMSON, P. C.; TAMMEN, I. et al. Quantitative trait loci mapping in dairy cattle: review and meta-analysis. **Genet. Sel. Evol.**, v. 36, p. 163-190, 2004.
- MARTINEZ, M. L.; VUKASINOVIC, N.; FREEMAN, A. E. Random model approach for QTL mapping in half-sib families. **Genet. Sel. Evol.**, v. 31, p. 319-340, 1999.
- MARTINEZ, M. L.; VUKASINOVIC, N. Algoritmo para cálculo da proporção de gens idênticos por descendência, para mapear QTL em famílias de meio-irmãos. **Rev. Bras. Zoot.**, v. 29, p. 443-451, 2000.