

Método para mapeamento de locos controladores de características oligogênicas

Genetic mapping of oligogenic traits using maximum likelihood functions

Rodrigo Barros Rocha^I William Silva Barros^{II} Júpiter Israel Muro-Abad^{III}
Rafael Simões Tomaz^{IV} Cosme Damião Cruz^{IV} Everaldo Gonçalves de Barros^{IV}
Elza Fernandes de Araújo^{IV}

RESUMO

Características oligogênicas de distribuição discreta e expressão governada por poucos genes de maior efeito têm se mostrado importantes na condução dos programas de melhoramento, com destaque para a resposta de resistência das plantas às doenças. Métodos tradicionais de detecção de QTL's, que pressupõem normalidade e herança governada por múltiplos fatores, não deveriam ser utilizados para mapeamento dessas características de distribuição discreta e interação epistática predominante. O objetivo deste trabalho é avaliar os resultados de um método para mapeamento e detecção de locos controladores da expressão de características oligogênicas, OTL's (Oligogenic Trait Loci). Esse método, definido como MMCO (Método de Mapeamento de Características Oligogênicas), utiliza funções de verossimilhança para obtenção de estimativas de ligação fatorial entre locos marcadores e locos controladores de características oligogênicas. Os resultados indicam que o método foi adequado para detecção de OTL's em populações F2 relativamente pequenas, compostas por 200 indivíduos, e que a determinação a priori do padrão de herança é condição necessária para a utilização dessa estratégia, que se diferencia por atender as pressuposições de análise, não necessitar de informação prévia de ordenamento entre as marcas e por permitir a obtenção de estimativas a partir da informação contida em todas as classes genotípicas.

Palavras-chave: mapeamento genético, locos controladores de características oligogênicas (OTL's)

ABSTRACT

Oligogenic traits are distinguished by their heritage ruled by higher effect genes and by their importance for

cultivated plants, with emphasis to the plant disease resistance inheritance. The qualitative nature and epistatic interaction of these traits results in a heritage pattern that should not be interpreted through traditional QTL detection strategies. The objective of this work was to propose a method for Oligogenic Traits Loci (OTL) detection. This method, defined as Oligogenic Trait Mapping Method (OTMM) uses maximum likelihood probability functions to obtain adjusted "r" estimates that express the distance among the molecular markers and the OTL loci. The results show that the method was adequate for OTL detection even in relatively small F2 populations. The prior definition of the oligogenic heritage pattern is one the main requirements of this method that focus in the attainment of the analysis presumptions without previous markers order information.

Key words: genetic mapping, oligogenic trait loci (QTL's).

INTRODUÇÃO

Características oligogênicas, de distribuição discreta e expressão governada por poucos genes de maior efeito, têm se mostrado importantes na condução dos programas de melhoramento de várias culturas (RAO & LI, 2000). Para o melhoramento vegetal, a resistência das plantas à ação de patógenos é provavelmente a característica mais importante de natureza oligogênica (ALZATE-MARIN et al., 2005; BARBOSA-PRESTES et al., 2008; GAGLIARDI &

^IEmpresa Brasileira de Pesquisa Agropecuária (Embrapa), Centro de Pesquisa Agroflorestal de Rondônia, 76815-800, Porto Velho, RO, Brasil. E-mail: rodrigo@cpafro.embrapa.br. Autor para correspondência.

^{II}Departamento de Matemática e Estatística, Universidade Federal de Pelotas (UFPel), Pelotas, RS, Brasil.

^{III}Aracruz Celulose, Vitória, ES, Brasil

^{IV}Departamento de Biologia Geral, Universidade Federal de Viçosa (UFV), Viçosa, MG, Brasil.

CAMARGO, 2009). Comparado com as características de distribuição contínua, menor número de métodos estão disponíveis para a análise de locos controladores de características oligogênicas - OTL (*Oligogenic Trait Loci*) (VARSHNEY et al., 2005).

A natureza discreta da resposta associada à interação epistática de múltiplos fatores resulta em uma característica que não deveria ser mapeada pelos métodos tradicionais, que pressupõem normalidade e herança governada por muitos genes de pequeno efeito (LANDER & BOTSTEIN, 1989; ZENG, 1994; LI, 2004). Essa limitação foi avaliada em outros trabalhos: VISSCHER et al. (1996) compararam métodos utilizando regressão linear e regressão linear generalizada a partir de dados binários e concluíram não haver diferença na acurácia da predição e no poder do teste das metodologias. XU & ATCHLEY (1996) estenderam a metodologia de mapeamento por intervalo composto para dados binários utilizando modelos de regressão logística. RAO & LI (2000) apresentaram estratégias para análise de variáveis discretas utilizando diferentes modelos alélicos. E ainda existem autores que apresentam aproximações bayesianas (YI et al., 2004) e equações de modelos mistos para detecção do OTL (CHEN & LIU, 2009). Embora possam ser encontrados na literatura modelos para análise de características oligogênicas em herodogramas familiares (TANG & SIEGMUND, 2001), é necessário considerar mais apropriadamente as condições da experimentação vegetal, tendo em vista a estrutura das populações, o número de indivíduos avaliados e o padrão de herança das características de maior relevância.

O método avaliado neste trabalho baseia-se na utilização de funções de verossimilhança para obtenção de estimativas de recombinação de melhor ajuste aos dados, considerando as pressuposições de herança e limitações da experimentação vegetal no desenvolvimento dos modelos. Além da expectativa teórica de que o atendimento mais apropriado das pressuposições da herança proporciona resultados de maior acurácia, estimativas de ligação obtidas utilizando funções de verossimilhança apresentaram maior poder de teste em comparação com as metodologias de marca simples e intervalo simples (ROCHA et al., 2008). A natureza mais restritiva das funções de verossimilhança em associação com o fato de não serem afetadas pela ocorrência de dois QTL's em um mesmo grupo de ligação são diferenciais desse método previamente caracterizados.

O objetivo deste trabalho é apresentar os resultados de método de detecção e mapeamento de locos controladores da expressão de características oligogênicas, baseado em funções de verossimilhança

com ênfase em segregações e interações epistáticas governadas por dois e três genes em populações F2.

MATERIALE MÉTODOS

Método de mapeamento de características oligogênicas (MMCO)

Estimativas de máxima verossimilhança são obtidas a partir de um vetor de parâmetros desconhecidos, Θ , e a função $L(\Theta; x)$ representa o valor de Θ que corresponde à maior probabilidade de produzir os dados observados, x . De acordo com SCHUSTER & CRUZ (2004), a notação que expressa a verossimilhança de Θ , dado o valor de dados observados x , pode ser representada por:

$$L(\Theta; x) = \prod_{i=1}^n f(\Theta; x_i) \quad (1), \text{ em que } \Theta: \text{ vetor de parâmetros}$$

desconhecidos, x : dados observados, n : número de observações.

Nessa aplicação, a função de máxima verossimilhança representa o produto das funções de densidade e probabilidade individuais da segregação da característica oligogênica em n observações associados à ocorrência das classes genotípicas nas populações segregantes. SCHUSTER & CRUZ (2004), demonstraram que a segregação de marcadores moleculares binomiais ou multicategóricos segue distribuição multinomial: $f(x) = \lambda p_1^{n_1} p_2^{n_2} \dots p_n^{n_n}$ (2);

$$\lambda = \frac{N!}{n_1! n_2! \dots n_n!} \quad (3); \quad N = \sum_{i=1}^n n_i \quad (4), \text{ em que } x: \text{ variável}$$

aleatória, n_i : i -ésimo evento, p_i : probabilidade de ocorrência do i -ésimo evento.

O Método de Mapeamento de Características Oligogênicas (MMCO) utiliza funções de verossimilhança para a obtenção de estimativas de distância entre os marcadores e o loco OTL, por meio dos estimadores de probabilidade condicional que expressam a razão entre a frequência conjunta de ocorrência dos genótipos do OTL e dos marcadores e a frequência marginal de ocorrência dos marcadores:

$$(5), \text{ em que } P(Q/A_i B_j C_k) = \frac{P(Q \cap A_i B_j C_k)}{P(A_i B_j C_k)}: \text{ probabilidade}$$

condicional de manifestar o fenótipo binário dado que possui o i -ésimo genótipo do marcador no loco A, o j -ésimo genótipo do marcador no loco B e o k -ésimo genótipo do marcador no loco C; $P(Q \cap A_i B_j C_k)$: probabilidade conjunta de manifestar o fenótipo binário dado que possui o i -ésimo genótipo do marcador no loco A, o j -ésimo genótipo do marcador no loco B e o k -ésimo genótipo do marcador no loco C

e $P(A_i B_j C_k)$: probabilidade marginal do genótipo $A_i B_j C_k$ na população.

Desta forma, a verossimilhança para N indivíduos é obtida pelo produto das funções de densidade individuais (SCHUSTER & CRUZ, 2004). $L(y) = \prod_{i=1}^N f(y_i; M_i)$ (6), em que y : vetor de valores fenotípicos; y_i : valor observado da característica fenotípica no i -ésimo indivíduo e M_i : genótipo do marcador no i -ésimo indivíduo.

$f(y_i; A_i B_j C_k) = \lambda P(Q/A_i B_j C_k)^{n_1} P(q/A_i B_j C_k)^{n_2}$ (7), sendo $P(Q/A_i B_j C_k)$: probabilidade condicional de manifestar o fenótipo binário considerando a ocorrência do i -ésimo genótipo no loco A, do j -ésimo genótipo no loco B e do k -ésimo genótipo no loco C; $P(q/A_i B_j C_k)$: probabilidade condicional de não manifestar o fenótipo binário considerando a ocorrência do i -ésimo genótipo no loco A, do j -ésimo genótipo no loco B e do k -ésimo genótipo no loco C; n_1 : número de indivíduos que expressaram o fenótipo binário na população e n_2 : número de indivíduos que não expressaram o fenótipo binário na população.

Para testar a ligação entre um marcador e a característica oligogênica, foi utilizado o teste estatístico da razão de verossimilhança, que por facilidade de interpretação é avaliado utilizando o logaritmo na base 10 da razão de verossimilhança:

$$Z = \text{Log}_{10} \left[L \frac{(r/x)}{(r=0,5/x)} \right] \quad (8), \text{ em que: } r = \acute{e} \text{ o valor de}$$

ligação fatorial entre o loco marcador e o OTL.

A consistência, normalidade, eficiência assintótica e possibilidade de obtenção das estimativas de variância, pelo inverso do índice de informatividade, são propriedades desejadas, características desse método.

Simulação dos dados

Para a simulação das progênies, foi utilizado o aplicativo computacional GQMOL (CRUZ, 2008), de acordo com as seguintes etapas: Simulação do genoma: foi considerado um genoma hipotético com quatro grupos de ligação constituídos por 30 marcas equidistantes codominantes e nível de saturação médio de 5cM. Simulação dos genitores: foram gerados de maneira a serem heterozigotos e divergentes em relação aos locos avaliados. Simulação das populações: foram geradas 500 populações F2 constituídas por 200 indivíduos para a interpretação de 100 repetições para cada característica. Simulação da característica oligogênica: cinco características governadas por dois e três genes de interação epistática dominante com proporções esperadas entre resistentes (1) e suscetíveis (0) de 9:7, 13:3, 15:1, 63:1 e 27:37, considerando cinco

condições de herança oligogênica (Tabela 1). Locos que alteram a expressão da característica estão presentes nos grupos de ligação 1 e 2 (OTL1 – localizado entre os marcadores 4 e 5 do grupo de ligação 1 e OTL2 – localizado entre os marcadores 14 e 15 do grupo de ligação 2); o grupo de ligação 3 apresenta loco que influencia a expressão do OTL somente quando a característica apresenta interações epistáticas entre três genes (OTL3- entre os marcadores 25 e 26). O grupo de ligação 4 não apresenta em nenhuma das situações consideradas locos que alteram a expressão da característica.

Análises genômicas

Foram aplicados testes de qui-quadrado (χ^2) para verificar a segregação dos marcadores. As estimativas de recombinação foram obtidas pelo método gráfico (SCHUSTER & CRUZ, 2004). Os grupos de ligação obtidos considerando ($r_{\max} = 30\%$) e o LOD mínimo ($\text{LOD}_{\min} = 3$) tiveram sua melhor ordem estimada pelo método SARF (“Sum of Adjacent Recombination Fractions”).

RESULTADOS E DISCUSSÃO

Parâmetros para simulação

A definição dos parâmetros de simulação é fundamental para a interpretação mais parcimoniosa dos dados. Trabalhos que objetivam o desenvolvimento de mapas-referência e detecção de QTL's têm se baseado na avaliação de pelo menos 200 indivíduos (COSTA & SILVA et al., 2007; ROCHA et al., 2007). Espécies diploides representam diversas espécies vegetais cultivadas (BORÉM, 2001). Quatro grupos de ligação são suficientes para explorar os fenômenos de epistasia que serão estudados, e a saturação avaliada (5cM) é um valor considerado conservador no desenvolvimento de mapas de ligação em plantas (MACKAY, 2001).

Estimativas de probabilidade condicional e funções de verossimilhança

Os estimadores de probabilidade condicional consideram as seguintes pressuposições em relação ao padrão de herança: i) expressão binária do tipo presença ou ausência, ii) herança governada por poucos genes, iii) segregação correlacionada com uma marca molecular previamente mapeada, iv) distribuição multinomial. As funções de verossimilhança desenvolvidas para mapeamento das OTL's estão mostradas a seguir, para as seguintes pressuposições:

1- Dois genes, marcadores de natureza dominante,

Tabela 1 - Valores paramétricos e estimativas MMCO das distâncias entre marcadores moleculares e OTL obtidas

Proporção	Marcador	LOD	\hat{r}	r	Condição	Marcador	LOD	\hat{r}	r
9p7	M3-GL1	8,79	9,72	10	63p1	M12-GL2	3,92	9,72	10
	M4-GL1	11,27	6,23	5		M13-GL2	6,85	0	5
	M5-GL1	18,8	0	0		M14-GL2	6,22	0	0
	M6-GL1	11,87	6,13	5		M2-GL1	6,1	9,62	15
	M11-GL2	13,47	7,56	15		M5-GL1	3,2	0	0
	M12-GL2	15,3	5,62	10		M6-GL1	3,05	0	5
	M13-GL2	21,02	1,4	5		M14-GL2	3,01	0	0
	M14-GL2	26,62	0	0		M15-GL2	4,28	6,54	5
	M15-GL2	16,63	6,64	5		M25-GL3	3,8	0	0
13p3	M3-GL1	16,96	7,45	10	27p37	M26-GL3	3,01	0	5
	M4-GL1	20,63	4,41	5		M3-GL1	7,16	7,15	10
	M5-GL1	26,32	0	0		M4-GL1	8,39	4,71	5
	M6-GL1	20,62	4,41	5		M5-GL1	13,04	0	0
	M7-GL1	15,33	9	10		M6-GL1	11,96	0	5
	M8-GL1	15,74	8,58	15		M7-GL1	9,87	2,4	10
	M13-GL2	6,63	0	5		M8-GL1	9,53	2,5	15
	M14-GL2	6,38	0	0		M13-GL2	9,67	7,35	5
	M15-GL2	6,51	0	5		M14-GL2	16,99	0	0
	M16-GL2	4,3	3,51	10		M15-GL2	10,54	5,93	5
	15p1	M4-GL1	5,82	3,3		5	M24-GL3	7,04	4,9
M5-GL1		7,85	0	0	M25-GL3	9,1	0	0	
M6-GL1		5,69	3,41	5	M26-GL3	8,6	0	5	
M7-GL1		4,12	7,45	10	M27-GL3	7,4	7,8	10	

LOD= Valores das estimativas de LODscore, \hat{r} = estimativas de ligação entre os marcadores moleculares e o OTL, r = valores paramétricos de distância entre os marcadores moleculares e o OTL.

configuração de aproximação entre loco marcador e loco OTL:

$$9:7 - L(p_i/n_i) = \lambda \left(\frac{3}{8} + \frac{3}{16} s_1^2 \right)^{n_1} \left(\frac{6}{16} - \frac{3}{16} s_1^2 \right)^{n_2} \left(\frac{3}{16} - \frac{3}{16} s_1^2 \right)^{n_3} \left(\frac{1}{16} + \frac{3}{16} s_1^2 \right)^{n_4} \quad (9);$$

$$13:3 \text{ (aprox. e repulsão)} - L(p_i/n_i) = \lambda \left(\frac{10}{16} - \frac{1}{16} s_1^2 \right)^{n_1} \left(\frac{2}{16} + \frac{1}{16} s_1^2 \right)^{n_2} \left(\frac{3}{16} + \frac{1}{16} s_1^2 \right)^{n_3} \left(\frac{1}{16} - \frac{1}{16} s_1^2 \right)^{n_4} \quad (10);$$

$$15:1 - L(p_i/n_i) = \lambda \left(\frac{11}{16} + \frac{1}{16} s_1^2 \right)^{n_1} \left(\frac{1}{16} - \frac{1}{16} s_1^2 \right)^{n_2} \left(\frac{4}{16} - \frac{1}{16} s_1^2 \right)^{n_3} \left(\frac{1}{16} s_1^2 \right)^{n_4} \quad (11).$$

2- Dois genes, técnica de marcador molecular de natureza dominante, configuração de repulsão entre loco marcador e loco OTL.

$$9:7 - L(p_i/n_i) = \lambda \left(\frac{6}{16} + \frac{3}{4} s_2 \right)^{n_1} \left(\frac{6}{16} - \frac{3}{4} s_2 \right)^{n_2} \left(\frac{3}{16} - \frac{3}{4} s_2 \right)^{n_3} \left(\frac{1}{16} + \frac{3}{4} s_2 \right)^{n_4} \quad (12);$$

$$13:3 \text{ (aproximação)} - L(p_i/n_i) = \lambda \left(\frac{10}{16} - \frac{1}{4} s_2 \right)^{n_1} \left(\frac{2}{16} + \frac{1}{4} s_2 \right)^{n_2} \left(\frac{3}{16} + \frac{1}{4} s_2 \right)^{n_3} \left(\frac{1}{16} - \frac{1}{4} s_2 \right)^{n_4} \quad (13);$$

$$13:3 \text{ (repulsão)} - L(p_i/n_i) = \lambda \left(\frac{9}{16} + \frac{3}{4} s_2 \right)^{n_1} \left(\frac{3}{16} - \frac{3}{4} s_2 \right)^{n_2} \left(\frac{4}{16} - \frac{3}{4} s_2 \right)^{n_3} \left(\frac{3}{4} s_2 \right)^{n_4} \quad (14);$$

$$15:1 - L(p_i/n_i) = \lambda \left(\frac{11}{16} + \frac{1}{4} s_2 \right)^{n_1} \left(\frac{1}{16} - \frac{1}{4} s_2 \right)^{n_2} \left(\frac{4}{16} - \frac{1}{4} s_2 \right)^{n_3} \left(\frac{1}{4} s_2 \right)^{n_4} \quad (15)$$

3- Dois genes, marcadores de natureza codominante, configuração de aproximação ou repulsão entre loco marcador e loco OTL.

$$9:7 - L(p_i/n_i) = \lambda \left(\frac{3}{16} - \frac{3}{16} r^2 \right)^{n_1} \left(\frac{1}{16} + \frac{3}{16} r^2 \right)^{n_2} \left(\frac{6}{16} + \frac{6}{16} r^2 - \frac{6}{16} r \right)^{n_3} \left(\frac{2}{16} - \frac{6}{16} r^2 + \frac{6}{16} r \right)^{n_4} \left(-\frac{3}{16} r^2 + \frac{6}{16} r^2 \right)^{n_5} \left(\frac{4}{16} + \frac{3}{16} r^2 - \frac{6}{16} r \right)^{n_6} \quad (16)$$

$$13:3 \text{ (aprox.)} - L(p_i/n_i) = \lambda \left(\frac{3}{16} - \frac{3}{16} r^2 \right)^{n_1} \left(\frac{1}{16} - \frac{1}{16} r^2 \right)^{n_2} \left(\frac{6}{16} - \frac{2}{16} r^2 + \frac{2}{16} r \right)^{n_3} \left(\frac{2}{16} + \frac{2}{16} r^2 - \frac{2}{16} r \right)^{n_4} \left(\frac{4}{16} + \frac{1}{16} r^2 - \frac{2}{16} r \right)^{n_5} \left(-\frac{1}{16} r^2 + \frac{2}{16} r \right)^{n_6} \quad (17)$$

$$13:3 \text{ (repulsão)} - L(p_i/n_i) = \lambda \left(\frac{4}{16} - \frac{3}{16} r^2 \right)^{n_1} \left(\frac{3}{16} r^2 \right)^{n_2} \left(\frac{8}{16} + \frac{6}{16} r^2 - \frac{6}{16} r \right)^{n_3} \left(-\frac{6}{16} r^2 + \frac{6}{16} r \right)^{n_4} \left(\frac{1}{16} - \frac{3}{16} r^2 + \frac{6}{16} r \right)^{n_5} \left(\frac{3}{16} + \frac{3}{16} r^2 - \frac{6}{16} r \right)^{n_6} \quad (18)$$

$$15:1- L(p_i / n_i) = \lambda \left(\frac{4}{16} - \frac{1}{16} r^2 \right)^{n1} \left(\frac{1}{16} r^2 \right)^{n2} \left(\frac{8}{16} + \frac{2}{16} r^2 - \frac{2}{16} r \right)^{n3} \cdot \left(-\frac{2}{16} r^2 + \frac{2}{16} r \right)^{n4} \left(\frac{3}{16} - \frac{1}{16} r^2 + \frac{2}{16} r \right)^{n5} \left(\frac{1}{16} + \frac{1}{16} r^2 - \frac{2}{16} r \right)^{n6} \quad (19)$$

4 - Três genes, marcadores de natureza dominante, configuração de aproximação ou repulsão entre loco marcador e loco OTL.

$$63:1- L(p_i / n_i) = \lambda \left(\frac{47}{64} - \frac{1}{16} s_3 \right)^{n1} \left(\frac{1}{64} - \frac{1}{16} s_3 \right)^{n2} \cdot \left(\frac{16}{64} - \frac{1}{16} s_3 \right)^{n3} \left(\frac{1}{16} s_3 \right)^{n4} \quad (20)$$

$$27:37- L(p_i / n_i) = \lambda \left(\frac{18}{64} + \frac{9}{16} s_3 \right)^{n1} \left(\frac{30}{64} - \frac{9}{16} s_3 \right)^{n2} \cdot \left(\frac{9}{64} - \frac{9}{16} s_3 \right)^{n3} \left(\frac{7}{64} + \frac{9}{16} s_3 \right)^{n4} \quad (21)$$

5 - Três genes, marcadores de natureza codominante, configuração de aproximação ou repulsão entre loco marcador e loco OTL.

$$63:1- L(p_i / n_i) = \lambda \left(\frac{16}{64} - \frac{1}{64} r^2 \right)^{n1} \left(\frac{1}{64} r^2 \right)^{n2} \left(\frac{32}{64} + \frac{2}{64} r^2 - \frac{2}{64} r \right)^{n3} \cdot \left(-\frac{2}{64} r^2 + \frac{2}{64} r \right)^{n4} \left(\frac{15}{64} - \frac{1}{64} r^2 + \frac{2}{64} r \right)^{n5} \left(\frac{1}{64} + \frac{1}{64} r^2 - \frac{2}{64} r \right)^{n6} \quad (22)$$

$$27:37- L(p_i / n_i) = \lambda \left(\frac{9}{64} - \frac{9}{64} r^2 \right)^{n1} \left(\frac{7}{64} + \frac{9}{64} r^2 \right)^{n2} \left(\frac{18}{64} + \frac{18}{64} r^2 - \frac{18}{64} r \right)^{n3} \cdot \left(\frac{14}{64} - \frac{18}{64} r^2 + \frac{18}{64} r \right)^{n4} \left(\frac{9}{64} r^2 + \frac{18}{64} r \right)^{n5} \left(\frac{16}{64} + \frac{9}{64} r^2 - \frac{18}{64} r \right)^{n6} \quad (23)$$

Em que: r = é o valor de ligação fatorial entre o loco marcador e o OTL, $s_1 = (1-r)$, $s_2 = 1/4(1-r)^2$.

A inclusão de outros valores de probabilidade condicional permite adequar as funções de verossimilhança para quaisquer outros tipos de populações e características para a obtenção de estimativas de variância mínima, estimadas a partir da informação contida em todas as classes genotípicas que compõem a população (SCHUSTER & CRUZ, 2004). O maior número de cálculos necessários é a principal limitação para o uso dessa ferramenta.

Mapeamento genético

A interpretação dos dados de mapeamento teve o objetivo de inferir a qualidade do mapa estimado a partir das condições analíticas avaliadas. Previamente à análise de pares de marcas, os marcadores tiveram sua segregação individual testada pelo teste de qui-quadrado com correção de Bonferroni para testes múltiplos. Do conjunto total de marcas avaliadas, dois marcadores apresentaram segregação distorcida. No caso específico, em que a obtenção de eventos de recombinação deve-se ao acaso, foi feita a opção por não descartar os marcadores que apresentaram segregação distorcida.

Para a obtenção dos grupos de ligação, foram considerados valores tradicionalmente utilizados em trabalhos de mapeamento (LOD=3,0 e $r=0,30$). Forte evidência de ligação fatorial foi evidenciada nos valores de LOD score observados (LOD_{mínimo}=30,26), e não foram observadas marcas não ligadas. Foram recuperados os quatro grupos de ligação originais com saturação média de 5,24cM. Os desvios entre os

valores estimados e os valores paramétricos são comparáveis com os caracterizados por COSTA & SILVA et al., 2007.

Outros trabalhos sugerem a suficiência dos parâmetros considerados na simulação para a obtenção das populações de retrocruzamento. A avaliação do conteúdo médio de informação e dos intervalos de confiança associados às estimativas de recombinação indica que a população de F2 em aproximação mapeada com marcadores codominantes é uma das que conferem as estimativas de recombinação mais precisas, juntamente com as populações de retrocruzamento e a população de duplos haploides (ROCHA et al., 2004).

Deteção de OTL's

Os dados mostraram a identificação dos locos controladores das características oligogênicas nos intervalos paramétricos simulados (marcadores 4 e 5 no GL1, marcadores 14 e 15 no GL2, e marcadores 25 e 26 no GL3) (Tabela 1), e em média foram identificados 4,3 marcadores candidatos por loco controlador. O mapa de ligação acrescido ao loco OTL mostrou uma alteração na distância entre os marcadores 4 e 5 (GL1) de 12,1 para 11,5 e marcadores 14 e 15 (GL2) de 13,2 para 12,97 (Figura 1).

O método permite a obtenção de valores da distância em unidades de mapa entre os locos marcadores e o OTL associados a valores de *LOD score*, que associam a probabilidade de ocorrência da hipótese de ligação em relação à hipótese nula de segregação independente, sem a necessidade de conhecimento prévio de ordenamento entre as marcas, por estimar individualmente a distância entre os marcadores e o loco OTL. Em relação a outras propriedades, o maior poder de teste comparado aos métodos de detecção por marca simples e ao mapeamento por intervalo e o não confundimento de efeitos de QTLs localizados no mesmo grupo de ligação são características diferenciadas dessa estratégia (ROCHA et al., 2008). A determinação *a priori* do padrão de herança da característica é um dos principais requisitos para o uso dessa metodologia.

Poder do teste e determinação do padrão de herança

A hipótese testada neste trabalho é de que a metodologia que considera as pressuposições de herança das características oligogênicas é adequada para detecção de OTL's, tendo em vista as condições analíticas dos trabalhos de mapeamento em plantas. Com esse objetivo, o poder do teste foi interpretado como a frequência em que a posição do OTL foi identificada corretamente em comparação com a hipótese alternativa de segregação independente entre os locos controladores da característica oligogênica e as marcas moleculares.

Os resultados da análise com a determinação apropriada do padrão de herança encontram-se na diagonal da tabela 2. Os padrões epistáticos de detecção

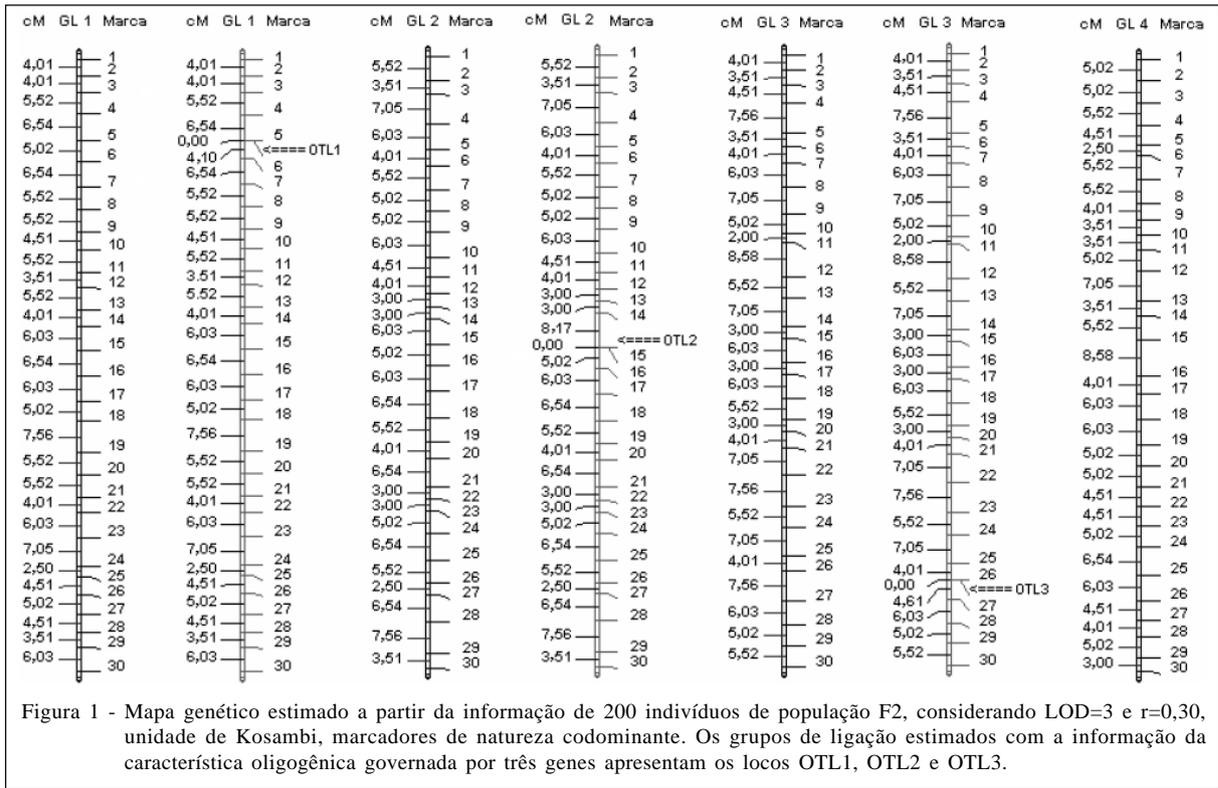


Figura 1 - Mapa genético estimado a partir da informação de 200 indivíduos de população F2, considerando LOD=3 e r=0,30, unidade de Kosambi, marcadores de natureza codominante. Os grupos de ligação estimados com a informação da característica oligogênica governada por três genes apresentam os locos OTL1, OTL2 e OTL3.

Tabela 2 - Estimativas da ocorrência de erros do tipo I e do tipo II na rejeição da hipótese H_0 de segregação independente entre os locos responsáveis pela expressão e a característica oligogênica em diferentes padrões de segregação. As estimativas do poder do teste do método estão marcadas em negrito na diagonal da tabela.

		Interações epistáticas														
Loco		-----9p7-----			-----13p3-----			-----15p1-----			-----63p1-----			-----27p37-----		
		Ho (OTL1)	Tipo I	Tipo II	Ho (OTL1)	Tipo I	Tipo II	Ho (OTL1)	Tipo I	Tipo II	Ho (OTL1)	Tipo I	Tipo II	Ho (OTL1)	Tipo I	Tipo II
OTL1	9p7	100	0	0	1	94	5	2	94	4	0	100	0	99	1	0
	13p3	64	36	0	94	6	0	96	3	1	9	91	0	91	9	0
	15p1	0	100	0	0	100	0	90	10	0	68	14	18	2	98	0
	63p1	0	100	0	9	91	0	68	14	18	73	27	0	0	100	0
	27p37	99	1	0	91	9	0	2	98	0	0	100	0	98	2	0
OTL2	9p7	100	0	0	64	36	0	2	94	40	0	100	0	99	1	0
	13p3	69	31	0	94	6	0	0	99	1	0	100	0	48	52	0
	15p1	0	100	0	0	99	1	90	10	0	68	14	18	2	98	0
	63p1	0	100	0	0	100	0	68	14	18	73	27	0	0	100	0
	27p37	99	1	0	48	52	0	2	98	0	0	100	0	98	2	0
OTL3	9p7	**	**	**	**	**	**	**	**	**	**	**	**	**	**	**
	13p3	**	**	**	**	**	**	**	**	**	**	**	**	**	**	**
	15p1	**	**	**	**	**	**	**	**	**	**	**	**	**	**	**
	63p1	0	100	0	0	100	0	0	100	0	73	27	0	0	100	0
	27p37	0	100	0	0	100	0	0	100	0	0	100	0	98	2	0

** As proporções 9:7, 13:3, 15:1 correspondem às proporções esperadas de epistasia entre dois genes e não se aplicam ao loco C326, relacionado às proporções de 63:1 e 27:37 de epistasia entre três genes. $H_{0(OTL1)}$: frequência da rejeição apropriada da hipótese de segregação independente no intervalo em que se encontra o OTL, Tipo I: frequência de erros tipo I, Tipo II: frequência de erros tipo II.

mais acurada foram respectivamente 9:7, 27:37, 13:3 (Tabela 2). *A priori* sabe-se que padrões de herança com menor frequência de uma das classes (15:1 e 63:1) necessitam de maior número de eventos meióticos para representação da classe mais rara. Populações de tamanho superior a 200 indivíduos devem ser utilizadas para obtenção de melhores estimativas do poder do teste (superior a 90% para a proporção de 15:1 e a 73% para a proporção de 63:1) (Tabela 2).

A ocorrência de erros do tipo I (não detecção do loco OTL no intervalo apropriado) e do tipo II (detecção do loco OTL em posição equivocada), devido à declaração incorreta do padrão de herança oligogênico, indica que a maior acurácia das estimativas desse método depende da correta declaração do padrão de herança. Por outro lado, essa limitação também pode ser explorada como um diferencial para a avaliação mais apropriada de uma característica que sabidamente se distribui em classes.

CONCLUSÕES

Os resultados indicam que o método foi adequado para detecção de OTL's em populações F2 compostas por 200 indivíduos e que a declaração correta do padrão de herança é condição necessária para a utilização dessa estratégia, tendo em vista as reduções do poder de teste da metodologia causadas por declarações equivocadas do padrão de herança.

REFERÊNCIAS

- ALZATE-MARIN A.L. et al. Seleção assistida por marcadores moleculares visando ao desenvolvimento de plantas resistentes a doenças, com ênfase em feijoeiro e soja. **Fitopatologia Brasileira**, v.30, n.4, p.333-342, 2005. Disponível em: <<http://biostatistics.oxfordjournals.org>>. Acesso em 29 jun. 2009. doi: 10.1590/S0100-41582005000400001.
- BARBOSA-PRESTES M.M. et al. Controle genético da resistência parcial à ferrugem da folha em aveia (*Avena sativa* L.). **Ciência Rural**, v.38, n.2, p.308-314, 2008. Disponível em: <<http://www.ufsm.br/ccr/revista>>. Acesso em 29 jun. 2009. doi: 10.1590/S0103-84782008000200002.
- BORÉM, A. **Melhoramento de plantas**. 3.ed. Viçosa: UFV, 2001. 500p.
- CHEN Z.; LIU J. Mixture generalized linear models for multiple interval mapping of quantitative trait loci in experimental crosses. **Biometrics**, v.65, n.2, p.470-477, 2009. Disponível em: <<http://www.biometrics.tibs.org>>. Acesso em 29 jun. 2009. doi: 10.1111/j.1541-0420.2008.01100.x.
- COSTA e SILVA, L. et al. Simulation of population size and genome saturation level for genetic mapping of recombinant inbred lines (RILs). **Genetics and Molecular Biology**, v.30, n.4, p.1101-1108, 2007. Disponível em: <<http://www.gmb.org.br>>. Acesso em 29 jun. 2009. doi: 10.1590/S1415-47572007000600013.
- CRUZ, C.D. **Programa para análises de dados moleculares e quantitativos – GQMOL**. Viçosa: UFV, 2008. (Software em desenvolvimento). Disponível em: <<http://www.ufv.br/dbg/genes/gdown.htm>>. Acesso em 29 jun. 2009.
- GAGLIARDI P.R.; CAMARGO L.E.A. Resistência de variedades comerciais de cana-de-açúcar ao agente causal do raquitismo-da-soqueira. **Ciência Rural**, v.39, n.4, p.1211-1214, 2009. Disponível em: <<http://www.ufsm.br/ccr/revista>>. Acesso em 29 jun. 2009. doi: 10.1590/S0103-84782009005000015.
- LANDER E.S.; BOTSTEIN D. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. **Genetics**, v.121, n.1, p.185-199, 1989.
- LI J. **Development of multiple interval mapping for mapping QTL in ordinal traits**. Raleigh, North Carolina: North Carolina State University, 2004. 124p.
- MACKAY, T.F.C. Genetic architecture of quantitative traits. **Annual Review of Genetics**, v.35, p.303-339; 2001. Disponível em: <<http://arjournals.annualreviews.org/loi/genet>>. Acesso em 29 jun. 2009. doi:10.1146/annurev.genet.35.102401.090633.
- RAO, S.; LI X. Strategies for genetic mapping of quantitative traits. **Genetica**, v.109, p.183-197, 2000.
- ROCHA, R.B. et al. Comparisons of segregating populations for genetic mapping. **Crop Breeding and Applied Biotechnology**, v.4, p.408-415, 2004.
- ROCHA, R.B. et al. Mapping of QTLs related with wood quality and developmental characteristics in hybrids (*Eucalyptus grandis* x *Eucalyptus urophylla*). **Revista Árvore**, v.31, n.1, p.13-24, 2007. Disponível em: <<http://www.revistaarvore.ufv.br>>. Acesso em 29 jun. 2009. doi: 10.1590/S0100-67622007000100003.
- ROCHA, R.B. et al. Genetic mapping of oligogenic traits using maximum likelihood functions - backcrossing populations. **Revista Brasileira de Biometria**, v.26, p.7-21, 2008.
- SCHUSTER I.; CRUZ, C.D. **Estatística genômica aplicada a populações derivadas de cruzamentos**. Viçosa: UFV, 2004. 585p.
- TANG H.K.; SIEGMUND D. Mapping quantitative trait loci in oligogenic models. **Biostatistics**, v.2, n.2, p.147-162, 2001. Disponível em: <<http://biostatistics.oxfordjournals.org>>. Acesso em 29 jun. 2009. doi: 10.1093/biostatistics/2.2.147.
- VARSHNEY, R.K. et al. Genic microsatellite markers in plants: features and applications. **Trends in Biotechnology**, v.23, n.1, p.48-55, 2005. Disponível em: <<http://www.trends.com/tibtech/default.htm>>. Acesso em 29 jun. 2009. doi: 10.1016/j.tibtech.2004.11.005.
- VISSCHER, P.M., et al. Mapping QTLs for binary traits in backcross and F2 populations. **Genetical Research**, v.68, p.55-63, 1996. Disponível em: <<http://journals.cambridge.org>>. Acesso em 29 jun. 2009. doi: 10.1017/S0016672300033887.
- ZENG, Z.B. Precision mapping of quantitative trait loci. **Genetics**, v.136, p.1457-1468, 1994.
- XU, S.; ATCHLEY, W.R. Mapping quantitative trait loci for complex binary diseases using line crosses. **Genetics**, v.143, p.1417-1424, 1996.
- YI, N. et al. Mapping multiple quantitative trait loci for ordinal trait. **Behavior Genetics**, v.34, n.1, p.3-15, 2004.