

Algoritmo para Cálculo da Proporção de Genes Idênticos por Descendência, para Mapear QTL em Famílias de Meio-irmãos

Mario Luiz Martinez¹, Natascha Vukasinovic²

RESUMO - O desenvolvimento de mapas de ligação tem estimulado a procura por métodos que permitam mapear genes responsáveis por variação em *loci* de características quantitativas (QTL). O método de pares-de-irmãos com base na relação entre indivíduos idênticos por descendência (IBD) tem sido utilizado para mapear QTLs. O objetivo deste trabalho foi estender o método de HASEMAN e ELSTON (1972) para se estimar a proporção (π) de genes IBD em famílias de meio-irmãos. Os resultados obtidos sugerem que os procedimentos usados foram eficientes para se estimar π , mesmo quando os genótipos dos pais foram parcial ou totalmente desconhecidos.

Palavras-chave: genes idênticos por descendência, meio-irmãos, QTL

Algorithm to Calculate Proportion of Genes Identity by Descent, to Map QTL in Half-sibs Families

ABSTRACT - The development of linkage maps has stimulated the search for methods to map genes involved with quantitative trait *loci* (QTL). A sib par method based on the relation between individuals identic by descent (IBD) has been used to map QTLs. The objective of this paper was to extend the HASEMAN and ELSTON (1972) method to estimate the proportion (π) of genes IBD in half-sib families. The obtained results suggested that the used procedures were efficient to estimate π even when parents genotypes were partial or totally unknown.

Key Words: genes identity by descent, half-sibs, QTL

Introdução

O desenvolvimento de mapas de ligação com grande número de marcadores moleculares tem estimulado a procura por métodos que permitam mapear genes responsáveis por variação em características quantitativas. HASEMAN e ELSTON (1972) propuseram o método de pares-de-irmãos com base na relação entre indivíduos idênticos por descendência. O método assume que a diferença no valor fenotípico da característica entre dois irmãos é menor quando eles têm grande proporção de genes idênticos por descendência (IBD) que influenciam a característica. Assim, se houver ligação entre um *locus* de característica quantitativa (QTL) que influencie a característica e o *locus* do marcador, a proporção esperada de IBD do QTL pode ser estimada por intermédio da proporção de IBD do *locus* marcador. A detecção da ligação entre QTL e marcador pode ser realizada por meio de simples regressão linear. Este método de pares-de-irmãos tem mostrado ser robusto quanto a diversos tipos de distribuição dos

dados e independe do modelo genético do QTL (AMOS, 1989). Todavia, o método em sua forma original é limitado, pois há confundimento entre o efeito genético do QTL e a taxa de recombinação entre QTL e o *locus* marcador. Isto significa que se pode apenas detectar a ligação entre o QTL e o marcador, mas não é possível estimar se é devido a um QTL de grande efeito localizado longe do marcador, ou devido a um QTL de pequeno efeito, mas muito próximo do marcador.

FULKER e CARDON (1994) desenvolveram o procedimento para mapeamento de QTL pelo par-de-irmãos associado ao método do intervalo, que utiliza dois marcadores simultaneamente para separar o efeito da localização do QTL de seu efeito propriamente dito. GOLDGAR (1990) desenvolveu um método de pontos múltiplos de IBD com base na máxima verossimilhança, para estimar a variância genética devido à determinada região do cromossomo. Este método foi expandido por SCHORK (1993) para estimar simultaneamente a variância de várias regiões do cromossomo e o efeito de ambiente comum

¹ Pesquisador da Embrapa Gado de Leite - CNPGL - 36038-330 - Juiz de Fora - MG. E. mail: martinez@cnpgl.embrapa.br

² Swiss Federal Institute of Technology. Clausiusstr. 50, 8092 Zurich, Switzerland. E.mail: vukasinovic@inw.agrl.ethz.ch

partilhado por todos os irmãos em um mesmo ambiente. XU e ATCHLEY (1995) desenvolveram um procedimento geral para o mapeamento de QTL, assumindo uma única distribuição normal dos valores fenotípicos do QTL.

A aplicação de qualquer um dos procedimentos acima mencionados requer o conhecimento da proporção de IBD para cada par-de-irmão de uma família. Em geral, o valor real desta proporção não é conhecida, mas pode ser estimada. HASEMAN e ELSTON (1972) usaram o método Bayesiano para estimar a proporção de IBD (π) entre irmãos completos, quando os genótipos dos irmãos e dos pais eram conhecidos. Esses autores propuseram também um algoritmo para calcular π em famílias de irmãos completos, quando apenas os genótipos dos irmãos eram conhecidos.

O objetivo deste trabalho foi estender o algoritmo de HASEMAN e ELSTON (1972), para estimar a proporção de IBD entre irmãos em famílias de meio-irmãos, em situações nas quais se conhece o genótipo de todos os indivíduos da família, apenas dos meio-irmãos e do pai ou apenas dos meio-irmãos.

Material e Métodos

Segundo HASEMAN e ELSTON (1972), para o caso geral de sistemas multialélicos com quatro ou mais alelos, haverá sete diferentes tipos de acasalamentos e semelhantemente sete diferentes tipos de pares-de-irmãos, conforme definido por KEMPTHORNE (1957). Assim, considerando-se um *locus* A com os alelos i, j, k, l, ter-se-ão os seguintes tipos:

- I. $A_iA_i - A_iA_i$
- II. $A_iA_i - A_jA_j$
- III. $A_iA_i - A_iA_j$
- IV. $A_iA_i - A_jA_k$
- V. $A_iA_j - A_iA_j$
- VI. $A_iA_j - A_iA_k$
- VII. $A_iA_j - A_kA_l$

Observa-se que o tipo I envolve apenas um alelo, os tipos II, III e V dois alelos, os tipos IV e VI três alelos e o tipo VII quatro alelos, e, independente do número total de alelos em um *locus*, ter-se-ão no máximo quatro diferentes alelos em determinado acasalamento.

É importante observar ainda que $A_iA_i \times A_iA_i$ e $A_jA_j \times A_jA_j$ são dois acasalamentos genotipicamente diferentes, mas ambos de um mesmo tipo I, pois são acasalamentos de indivíduos homocigotos idênticos

(A_iA_i ou A_jA_j). Semelhantemente, $A_iA_j \times A_iA_j$, $A_iA_k \times A_iA_k$, $A_iA_l \times A_iA_l$, $A_jA_k \times A_jA_k$, $A_jA_l \times A_jA_l$ e $A_kA_l \times A_kA_l$ são todos de um mesmo tipo V de acasalamento.

Assumindo que **a** represente qualquer par de meio-irmãos, a estimativa da proporção de IBD ($\hat{\pi}_a$) entre dois meio-irmãos será dada por:

$$\hat{\pi}_a = \frac{1}{2} f_a \quad (1)$$

em que f_a é a probabilidade que o par de meio-irmãos **a** tenha um IBD em um *locus* ao acaso, condicional ao genótipo do par-de-irmãos e dos pais.

Famílias de meio-irmãos têm apenas um pai em comum e, portanto, o número possível de combinações de acasalamento é igual a $[n(n+1)/2]^3$, em que **n** é o número de alelos em um determinado *locus*, e o número de genótipos possíveis é igual a $(n+1)/2$, pois existem três indivíduos (um pai e duas mães) envolvidos para a produção de um determinado par de meio-irmãos. A princípio pode parecer uma tarefa quase que impossível estimar a proporção de IBD para todas as combinações possíveis, visto que um *locus* com quatro alelos produz 1000 combinações. Entretanto, como mencionado anteriormente, muitos acasalamentos genotipicamente diferentes pertencem a um mesmo tipo de acasalamento, e isto reduz substancialmente o número de combinações de acasalamentos. No caso de quatro alelos, o número de combinações de acasalamentos de tipos diferentes se reduz a 25 (10 para o caso em que o pai comum é homocigoto e 15 quando o pai comum é heterocigoto). *Algoritmo para o cálculo de π em famílias de meio-irmãos*

Ambos os pais e as progênes têm o genótipo conhecido

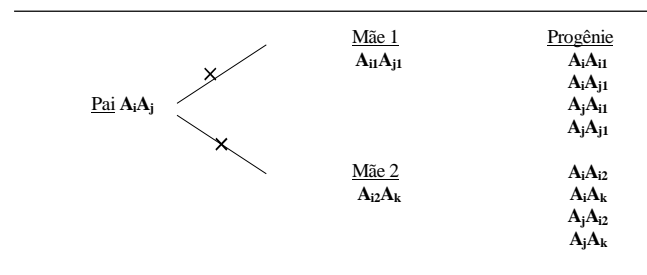
Considere o acasalamento de um pai com duas mães em que não há nenhum parentesco entre eles (entre pai com as mães e entre as mães). Considere um determinado *locus* A que tenha seis alelos com a mesma frequência. Assuma que o pai é heterocigoto para este *locus* e tenha o genótipo A_iA_j . Assuma que a primeira mãe tenha o mesmo genótipo $A_{i1}A_{j1}$, em que o índice 1 indica que os alelos (i e j) do pai e dessa mãe não são idênticos por descendência, ou seja, os dois animais não são parentes. Assuma que o genótipo da segunda mãe é $A_{i2}A_{k2}$, em que o índice 2 indica que o alelo i dessa mãe não tem a mesma origem que os alelos do pai e da primeira mãe. Segundo a definição dada por HASEMAN e ELSTON (1972), o acasalamento do pai com a primeira mãe é do tipo V,

e o acasalamento do pai com a segunda mãe, do tipo VI. Assim, para obter-se $\hat{\pi}_a$, deve-se proceder da seguinte maneira:

1º.) Obter todos os possíveis genótipos das progênes resultantes dos acasalamentos deste pai com as duas mães, como apresentado na Figura 1.

Figura 1 - Genótipos^a resultantes do acasalamento de um pai e duas mães

Figure 1 - Genotypes from mating one sire with two dams



^aNesta notação, o primeiro alelo é originário do macho.

^aIn this notation, the first allele is from the sire.

2º.) Determinar os tipos para todas as possíveis combinações de pares-de-irmãos. Isso pode ser feito estabelecendo-se uma tabela de dupla entrada, conforme apresentado na Tabela 1. Observa-se que, neste exemplo, têm-se um par-de-irmãos do tipo I, um do tipo II, seis do tipo III, dois do tipo IV, dois do tipo V e quatro do tipo VI. O número máximo de combinações será sempre 16, independente do número de alelos ser maior do que quatro.

3º.) Determinar o número de alelos IBD entre os pares-de-irmãos. Para o caso de meio-irmãos, quando os pais não são consanguíneos, o número de alelos IBD será 1 (se ambos os irmãos receberam o mesmo alelo do pai comum) ou 0 (se os irmãos receberam alelos diferentes do pai comum). O número de alelos IBD para o exemplo apresentado encontra-se na Tabela 1. Observa-se que o par-de-irmão tipo I (A_iA_{i1} - A_iA_{i2}), embora tenha apenas o alelo A_i , tem apenas um alelo IBD (no caso representado por i), os demais A_{i1} e A_{i2} são provenientes das mães que não são parentes e, portanto, não são IBD. O mesmo raciocínio se aplica aos demais pares-de-irmãos.

4º.) Calcular a probabilidade (f_a) de que determinado par-de-irmão **a** tenha um alelo IBD no locus **A**. Para se calcular f_a , usa-se:

$$f_a = \frac{N_{IBD}}{N_t} \quad (2)$$

em que N_{IBD} é o número de pares-de-irmãos do tipo **a**, que têm um alelo IBD, e N_t é o número total de pares-de-irmãos do tipo **a**.

Neste exemplo, no caso do par-de-irmãos do tipo III, têm-se $N_t = 6$ e $N_{IBD} = 4$ e, portanto,

$$f_{III} = \frac{4}{6} = \frac{2}{3} = 0,6667$$

5º.) Calcular a proporção de alelos IBD ($\hat{\pi}_a$) para determinado par-de-irmão do tipo **a**.

A proporção ($\hat{\pi}_a$) é calculada por intermédio da expressão (1):

$$\hat{\pi}_a = \frac{1}{2} f_a$$

Neste exemplo, para o par-de-irmãos do tipo III, têm-se:

$$\begin{aligned} \hat{\pi}_{III} &= \frac{1}{2} \times f_{III} \\ &= \frac{1}{2} \times \frac{2}{3} = \frac{1}{3} = 0,3333 \end{aligned}$$

Assim, utilizando-se das expressões f_a , π_a e das informações da Tabela 1, calculam-se as demais proporções de alelos IBD para cada um dos tipos de pares-de-irmãos. O resultado para a combinação de acasalamento deste exemplo encontra-se na Tabela 2.

Utilizando-se o algoritmo acima descrito, obteve-se a Tabela 3, que contém a proporção de alelos IBD para todas as possíveis combinações de acasalamentos, quando ambos, pais e progênes, têm genótipos conhecidos.

Progênie e pai em comum têm os genótipos conhecidos

Quando apenas o genótipo do pai em comum e das progênes é conhecido, π_a pode ser aproximado por intermédio da seguinte expressão:

$$\hat{\pi}_a = \frac{\sum P_a \times \hat{\pi}_a}{\sum_{m_1 m_2} P_a} \quad (3)$$

em que P_a = prob (par-de-irmão tipo **a** | tipos de acasalamento m_1 e m_2), ou seja, P_a é a probabilidade da ocorrência de determinado tipo **a** de par-de-irmãos, dado os tipos de acasalamentos m_1 e m_2 , em que m_1 e m_2 são tipos de combinações de acasalamentos que possam gerar determinado tipo **a** de par-de-irmãos.

Como exemplo, considere que o pai é heterozigoto com genótipo A_iA_j ; os genótipos das mães são desconhecidos; e o par-de-irmãos é do tipo VI. Neste caso, para calcular π_{VI} , pode-se usar a expressão (3), substituindo os valores apresentados na Tabela 3 para os pares-de-irmãos do tipo VI, obtidos quando o pai comum tem o genótipo A_iA_j . Como pode ser visto

Tabela 1 - Tipos de pares-de-irmãos e número de alelos IBD entre um par-de-irmãos (entre parênteses)

Table 1 - Sib pair type and number of alleles IBD between a sib pair (in parenthesis)

Genótipos das progênes do acasalamento do tipo V <i>Progenies genotypes from mating type V</i>	Genótipos das progênes do acasalamento do tipo VI <i>Progenies genotypes from mating type VI</i>			
	$A_i A_{i2}$	$A_i A_k$	$A_j A_{i2}$	$A_j A_k$
$A_i A_{i1}$	I(1)	III(1)	III(0)	IV(0)
$A_i A_{j1}$	III(1)	VI(1)	V(0)	VI(0)
$A_j A_{i1}$	III(0)	VI(0)	V(1)	VI(1)
$A_j A_{j1}$	II(0)	IV(0)	III(1)	III(1)

Tabela 2 - Probabilidades de determinado tipo a de par-de-irmãos (g_a), de um alelo IBD (f_a) e proporção ($\hat{\pi}_a$) de alelos IBD, conforme os acasalamentos de indivíduos dos tipos V e VITable 2 - Probabilities of a sib pair type a (g_a), one allele IBD (f_a) and proportion ($\hat{\pi}_a$) of alleles IBD, according to the matings of individuals of types V e VI

Tipo de par-de-irmãos <i>Sib pair type</i>	g_a	f_a	$\hat{\pi}_a$
I	1/16	1	1/2
II	1/16	0	0
III	3/8	2/3	1/3
IV	1/8	0	0
V	1/8	1/2	1/4
VI	1/4	1/2	1/4
VII	0	0	0

na Tabela 3, existem várias combinações de tipos de acasalamentos que geram pares-de-irmãos do tipo VI. A primeira combinação que interessa é a do acasalamento do tipo III ($A_i A_j \times A_i A_i$) com a do acasalamento do tipo III ($A_i A_j \times A_k A_k$). Essa combinação pode produzir três diferentes tipos de pares-de-irmãos: III, IV e VI. Neste caso, o par-de-irmãos do tipo VI tem uma probabilidade de ocorrência igual a 1/2 e $\hat{\pi} = 1/4$. A segunda combinação de tipos de acasalamento é a do tipo III ($A_i A_j \times A_i A_i$) com a do tipo VI ($A_i A_j \times A_i A_k$), que pode produzir pares-de-irmãos do tipo I, III, IV, V e VI. Nessa situação, o par-de-irmãos do tipo VI tem 1/4 de probabilidade de ocorrer e $\hat{\pi} = 1/4$. No total, existem 22 combinações de tipos de acasalamentos que podem gerar o par-de-irmãos do tipo VI. Assim, o numerador da equação 3 é dado pela soma do produto da probabilidade de ocorrência do tipo VI com $\hat{\pi}_{VI}$, ou seja:

$$\sum_{m_1 m_2} P_a \times \hat{\pi}_a =$$

$$1/2 \times 1/4 + 1/4 \times 1/4 + 1/4 \times 1/4 + 1/2 \times 1/4 + 1/2 \times 0 + 1/2 \times 1/2$$

$$+ 1/2 \times 1/4 + 1/2 \times 1/8 + 1/2 \times 3/8 + 1/2 \times 1/4 + 1/4 \times 1/2 + 1/4 \times 1/4 + 1/2 \times 1/4 + 3/8 \times 1/6 + 3/8 \times 1/3 + 3/8 \times 1/6 + 3/8 \times 1/3 + 1/2 \times 1/4 + 1/2 \times 3/8 + 1/2 \times 1/4 + 1/2 \times 3/8 + 1/2 \times 1/2 = 135/48$$

e o denominador é:

$$\sum_{m_1 m_2} P_a =$$

$$1/2 + 1/4 + 1/4 + 1/2 + 1/2 + 1/2 + 1/2 + 1/2 + 1/2 + 1/2 + 1/2 + 1/4 + 1/2 + 3/8 + 3/8 + 3/8 + 3/8 + 1/2 + 1/2 + 1/2 + 1/2 = 39/4$$

Finalmente,

$$\hat{\pi}_{VI} = (135/48)/(39/4) = 0,288$$

Da mesma forma, $\hat{\pi}$ para todos os demais tipos de pares-de-irmãos pode ser calculado como no exemplo apresentado, utilizando-se das informações da Tabela 3 e considerando-se todas as combinações possíveis de acasalamentos.

Apenas os genótipos das progênes são conhecidos

Quando os genótipos dos pais são desconhecidos, mas os genótipos das progênes são conhecidos, a proporção de alelos IBD entre pares-de-irmãos pode ser estimada utilizando-se da frequência dos alelos estimada na população. Assim, se um pai é acasalado a duas mães não-aparentadas, escolhidas ao acaso na população, a proporção de IBD em determinado *locus* será:

$$\hat{\pi}_a = P_{(hom)} \times \hat{\pi}_{a(hom)} + P_{(het)} \times \hat{\pi}_{a(het)} \quad (4)$$

$$= \left(\sum_{j=1}^n p^2 j \right) \times \hat{\pi}_{a(hom)} + \left(1 - \sum_{j=1}^n p^2 j \right) \times \hat{\pi}_{a(het)}$$

em que $\hat{\pi}_a$ é a média estimada da proporção de

Tabela 3 - Proporção de alelos idênticos por descendência ($\hat{\pi}$) quando ambos os pais e os genótipos dos irmãos são conhecidos
 Table 3 - Proportion of alleles identical by descent ($\hat{\pi}$) when both parental and sib genotypes are known

Pai em comum <i>Common parent</i>	Acasalamentos <i>Matings</i>		Tipo de acasalamento <i>Mating type</i>		Frequência do Acasalamento <i>Mating frequency</i>	Tipos de pares-de-irmãos <i>Sib-pairs type</i>						
	1º	2º	1º	2º		I	II	III	IV	V	VI	VII
A _i A _i	A _i A _i	A _i A _i	I	I	p _i ⁶	1/4 (1) ¹						
A _i A _i	A _i A _i	A _j A _j	I	II	p _i ⁴ p _j ²			1/4 (1)				
A _i A _i	A _i A _i	A _j A _j	I	III	p _i ⁵ p _j	1/4 (1/2)		1/4 (1/2)				
A _i A _i	A _i A _i	A _j A _k	I	IV	p _i ⁴ p _j p _k			1/4 (1)				
A _i A _i	A _j A _j	A _j A _j	II	II	p _i ² p _j ⁴					1/4 (1)		
A _i A _i	A _j A _j	A _k A _k	II	II	p _i ² p _j ² p _k ²						1/4 (1)	
A _i A _i	A _j A _j	A _i A _j	II	III	p _i ³ p _j ³			1/4 (1/2)		1/4 (1/2)		
A _i A _i	A _k A _k	A _i A _j	II	III	p _i ³ p _j p _k ²			1/4 (1/2)				1/4 (1/2)
A _i A _i	A _j A _j	A _k A _l	II	IV	p _i ² p _j ² p _k p _l							1/4 (1)
A _i A _i	A _j A _j	A _j A _k	II	IV	p _i ² p _j ³ p _k					1/4 (1/2)		1/4 (1/2)
A _i A _i	A _i A _j	A _i A _j	III	III	p _i ⁴ p _j ²	1/4 (1/4)		1/4 (1/2)		1/4 (1/4)		
A _i A _i	A _i A _j	A _i A _k	III	III	p _i ⁴ p _j p _k	1/4 (1/4)		1/4 (1/2)				1/4 (1/4)
A _i A _i	A _i A _j	A _j A _k	III	IV	p _i ³ p _j ² p _k			1/4 (1/2)		1/4 (1/4)		1/4 (1/4)
A _i A _i	A _i A _j	A _k A _l	III	IV	p _i ³ p _j p _k p _l			1/4 (1/2)				1/4 (1/2)
A _i A _i	A _j A _k	A _j A _k	IV	IV	p _i ² p _j ² p _k ²					1/4 (1/2)		1/4 (1/2)
A _i A _i	A _j A _k	A _j A _l	IV	IV	p _i ² p _j ² p _k p _l					1/4 (1/4)		1/4 (3/4)
A _i A _i	A _j A _k	A _l A _m	IV	IV	p _i ² p _j p _k p _l p _m							1/4 (1)
A _i A _j	A _i A _i	A _i A _i	III	III	p _i ⁵ p _j	1/2 (1/4)		0 (1/2)		1/2 (1/4)		
A _i A _j	A _i A _i	A _j A _j	III	III	p _i ³ p _j ³		0 (1/4)	1/2 (1/2)		0 (1/4)		
A _i A _j	A _j A _j	A _k A _k	III	III	p _i p _j ³ p _k ²			1/2 (1/4)	0 (1/4)			1/4 (1/2)
A _i A _j	A _j A _j	A _i A _j	III	V	p _i ² p _j ⁴	1/2 (1/8)	0 (1/8)	1/4 (1/2)		1/4 (1/4)		
A _i A _j	A _i A _i	A _i A _k	III	VI	p _i ⁴ p _j p _k	1/2 (1/8)		1/6 (3/8)	0 (1/8)	1/2 (1/8)		1/4 (1/4)
A _i A _j	A _j A _j	A _i A _k	III	VI	p _i ² p _j ³ p _k		0 (1/8)	1/2 (3/8)	0 (1/8)	0 (1/8)		1/4 (1/4)
A _i A _j	A _i A _i	A _k A _l	III	VII	p _i ³ p _j p _k p _l			1/2 (1/4)	0 (1/4)			1/4 (1/2)
A _i A _j	A _k A _k	A _k A _k	IV	IV	p _i p _j p _k ⁴					1/2 (1/2)		0 (1/2)
A _i A _j	A _k A _k	A _l A _l	IV	IV	p _i p _j p _k ² p _l ²						1/2 (1/2)	0 (1/2)
A _i A _j	A _k A _k	A _i A _j	IV	V	p _i ² p _j ² p _k ²			1/2 (1/4)	0 (1/4)			1/4 (1/2)
A _i A _j	A _k A _k	A _i A _k	IV	VI	p _i ² p _j p _k ³			1/2 (1/8)	0 (1/8)	1/2 (1/4)		1/8 (1/2)
A _i A _j	A _k A _k	A _j A _l	IV	VI	p _i p _j ² p _k ² p _l			1/2 (1/8)	0 (1/8)			3/8 (1/2)
A _i A _j	A _k A _k	A _k A _l	IV	VII	p _i p _j p _k ³ p _l					1/2 (1/4)		1/4 (1/2)
A _i A _j	A _k A _k	A _l A _m	IV	VII	p _i p _j p _k ² p _l p _m							1/2 (1/2)

Continua na próxima página...

Continuação da Tabela 3...

Pai em comum	Acasalamentos		Tipo de acasalamento		Frequência do Acasalamento	Tipos de pares-de-irmãos									
	Matings		Mating type			Sib-pairs type									
	1 ^o	2 ^o	1 ^o	2 ^o		I	II	III	IV	V	VI	VII			
A _i A _j	A _i A _j	A _i A _j	V	V	p _i ³ p _j ³	1/2(1/8)	0(1/8)	1/4(1/2)	1/4(1/4)	1/4(1/4)					
A _i A _j	A _i A _k	A _i A _k	V	VI	p _i ³ p _j ² p _k	1/2(1/16)	0(1/16)	1/3(3/8)	0(1/8)	1/4(1/8)	1/4(1/4)				
A _i A _j	A _k A _l	A _k A _l	V	VII	p _i ³ p _j ² p _k p _l			1/2(1/4)	0(1/4)	0(1/8)	1/4(1/2)				
A _i A _j	A _i A _k	A _i A _k	VI	VI	p _i ³ p _j p _k ²	1/2(1/16)		1/4(1/4)	0(1/8)	1/2(3/16)	1/6(3/8)				
A _i A _j	A _i A _l	A _i A _l	VI	VI	p _i ³ p _j p _k p _l	1/2(1/16)		1/4(1/4)	0(1/8)	1/2(1/16)	1/3(3/8)				0(1/8)
A _i A _j	A _i A _k	A _i A _k	VI	VI	p _i ² p _j ² p _k ²		0(1/16)	1/2(1/4)	0(1/8)	1/3(3/16)	1/6(3/8)				
A _i A _j	A _j A _l	A _j A _l	VI	VI	p _i ² p _j ² p _k p _l		0(1/16)	1/2(1/4)	0(1/8)	0(1/16)	1/3(3/8)				0(1/8)
A _i A _j	A _k A _l	A _k A _l	VI	VII	p _i ² p _j p _k ² p _l			1/2(1/4)	0(1/8)	1/2(1/8)	1/4(1/2)				0(1/8)
A _i A _j	A _m A _n	A _m A _n	VI	VII	p _i ² p _j p _k p _l p _m p _n			1/2(1/8)	0(1/8)	1/2(1/8)	3/8(1/2)				0(1/4)
A _i A _j	A _k A _l	A _k A _l	VII	VII	p _i p _j ² p _k ² p _l ²					1/2(1/4)	1/4(1/2)				0(1/4)
A _i A _j	A _l A _m	A _l A _m	VII	VII	p _i p _j p _k p _l ² p _m					1/2(1/8)	3/8(1/2)				0(3/8)
A _i A _j	A _m A _n	A _m A _n	VII	VII	p _i p _j p _k p _l p _m p _n					1/2(1/2)	1/2(1/2)				0(1/2)

1 Probabilidade de determinado tipo de par-de-irmãos, conforme os tipos de acasalamentos. Probability of the sib pair type according to two mating types.

alelos IBD para um par-de-irmãos do tipo **a**; $\hat{\pi}_{a(\text{hom})}$ e $\hat{\pi}_{a(\text{het})}$ são as médias da proporção de IBD para aquele tipo **a** de par de meio-irmãos, obtida por intermédio de todas as combinações de acasalamento que possam gerar aquele tipo **a**, quando o pai em comum é homocigoto e heterocigoto, respectivamente; e $\hat{\pi}_{a(\text{hom})}$ e $\hat{\pi}_{a(\text{het})}$ podem ser estimados por intermédio da equação (3). P_(hom) e P_(het) são as possibilidades que um pai em comum, escolhido ao acaso na população, possa ser homocigoto ou heterocigoto, respectivamente, para determinado locus e p_j (j = 1, ..., n), a frequência do jth alelo na população.

Cálculo da porcentagem de pares de meio-irmãos

Os genótipos dos pais e das progênes de meio-irmãos foram simulados considerando-se seis loci marcadores, cada um com 2, 4, 6 ou 10 alelos separados entre si por uma distância de 20 cM, por intermédio da alocação ao acaso dos alelos em cada um dos locus e assumindo equilíbrio Hardy-Weinberg e fases de ligação desconhecidas. Cada um dos 100 pais não-aparentados entre si foi acasalado com 50 fêmeas diferentes e não-aparentadas entre si e com o pai, para produzirem uma única progênie. Os genótipos foram simulados assumindo-se que a taxa de recombinação entre dois loci não influencia a taxa de recombinação no intervalo adjacente. Assim, foram simuladas 100 famílias, cada uma com 50 progênes de meio-irmãos. Após a constituição das famílias, procedeu-se à formação dos pares de meio-irmãos, conforme o algoritmo apresentado anteriormente. A porcentagem de pares de meio-irmãos foi estimada considerando-se 637.500 pares de loci, para os casos em que se conheciam os genótipos de todos os indivíduos de cada família, apenas o pai comum e as progênes e apenas as progênes.

Resultados e Discussão

As proporções ($\hat{\pi}_a$) estimadas de alelos idênticos por descendência para os diferentes tipos de pares-de-irmãos, assim como a probabilidade da ocorrência de determinado tipo de par-de-irmãos, encontram-se na Tabela 3. Na primeira parte da tabela, estão $\hat{\pi}_a$ para os casos em que o pai em comum é homocigoto. Pode-se observar que neste caso todos os pares-de-irmãos têm $\hat{\pi} = 0,25$, independente das mães serem ou não homocigotas. Segundo ROBERTSON (1973), quando a proporção de genes idênticos por descendência é igual a 0,25 para o caso de pares de meio-irmãos e 0,50 para pares-de-

irmãos completos, estes são ditos "não-informativos", o que significa que estes pares-de-irmãos não contribuem com nenhuma informação para a detecção de ligação entre QTL e marcador. Na segunda parte da Tabela 3, são apresentadas $\hat{\pi}$, quando o pai comum é heterozigoto. Nesta situação, observa-se que $\hat{\pi}$ varia de 0 a 0,5 e um mesmo tipo de par-de-irmão pode ser proveniente de diferentes combinações de acasalamentos dos pais. Observa-se também que os tipos de pares de meio-irmãos II, IV e VII têm sempre $\hat{\pi} = 0$, independente dos genótipos dos pais. Semelhantemente, o par-de-irmão do tipo I terá sempre $\hat{\pi} = 1/4$, quando o pai em comum for homozigoto, ou $\hat{\pi} = 1/2$ quando o pai for heterozigoto.

Na Tabela 4, são apresentadas as proporções ($\hat{\pi}$) de alelos IBD, quando os genótipos das mães são desconhecidos. Pode-se observar que, independente do número de alelos considerados, quatro, seis ou mais de seis, $\hat{\pi}$ é praticamente a mesma e as mesmas combinações que geram pares-de-irmãos não-informativos ($\hat{\pi} = 0,25$) são as mesmas, exceto para um caso em particular em que o tipo de par-de-irmão é $A_iA_j - A_jA_x$ (Tipo VI) proveniente de um pai heterozigoto A_iA_j . Todavia, quando ambos os pais são desconhecidos, a proporção ($\hat{\pi}$) é aproximada (Tabela 5) e os pares-de-irmãos não-informativos deixam de existir. Observa-se nesta tabela que $\hat{\pi}$ é praticamente o mesmo, independente de se considerarem quatro ou dez alelos de igual frequência em um mesmo *locus* A.

A porcentagem para 637.500 pares de meio-irmãos com um ou dois *locus* de marcadores ligados entre si com $\hat{\pi} = 0,25$ estão na Tabela 6. É importante observar que a porcentagem de pares com $\hat{\pi}_1$ e $\hat{\pi}_2$ iguais a 0,25, o que indica que estes pares não contribuirão com nenhuma informação para a detecção e o mapeamento de QTLs, diminui consideravelmente com o aumento do número de alelos dos *locus* marcadores, independente dos genótipos das mães serem conhecidos ou não. Pode-se observar também que há aumento da porcentagem de pares-de-irmãos com $\hat{\pi}_1$ e $\hat{\pi}_2 = 0,25$, quando se aumenta

a porcentagem de mães com genótipos desconhecidos até o máximo de quatro alelos por *locus*. É importante notar que, se o *locus* tiver seis ou mais alelos, a porcentagem de pares de meio-irmãos não-informativos será menor, mesmo quando se desconhece o genótipo das mães. Embora não apresentado, quando se conhece apenas o genótipo dos pares de meio-irmãos, a porcentagem de $\hat{\pi}_1$ e $\hat{\pi}_2 = 0,25$ foi zero, ou seja, não houve nenhum par-de-irmão não-informativo, independentes do número de alelos.

A porcentagem estimada para os diferentes tipos de pares de meio-irmãos resultante de simulação de 100 famílias de 50 meio-irmãos cada uma encontra-se na Tabela 7. Observa-se que os tipos IV, VI e VII de pares de meio-irmãos não existem quando apenas dois alelos estão presentes em um *locus*. Enquanto o tipo VI se torna o mais presente, quando o *locus* tem quatro ou mais alelos, os tipos I e II têm a sua porcentagem reduzida drasticamente com o aumento do polimorfismo do *locus*. Esta variação no percentual dos diferentes tipos de pares-de-irmãos, quando o *locus* é mais polimórfico, acarreta redução na porcentagem de pares-de-irmãos não-informativos, conforme observado na Tabela 6.

Assim, considerando-se que, quanto maior for o número de pares de indivíduos informativos ($\hat{\pi} \neq 0,25$), maior será a probabilidade de se detectar e mapear um QTL localizado entre dois marcadores (ROBERTSON, 1973), deduz-se, pelos resultados apresentados, que os *locus* marcadores polimórficos com quatro ou mais alelos são os mais eficientes para se alcançarem estes objetivos. Da mesma forma, embora haja perda de informações quando os genótipos dos pais são desconhecidos parcial ou totalmente, devido ao procedimento sugerido para se estimar a proporção de alelos idênticos por descendência nestas situações, verificou-se que a porcentagem de pares-de-irmãos não-informativos se manteve ou mesmo decresceu, quando o *locus* marcador é bastante polimórfico (seis ou mais alelos).

Tabela 4 - Proporção ($\hat{\pi}$) de alelos idênticos por descendência, quando os genótipos das mães são desconhecidosTable 4 - Proportion ($\hat{\pi}$) of alleles IBD when dam genotypes are unknown

Tipo de par-de-irmãos <i>Sib pair type</i>	Genótipo <i>Genotype</i>		$\hat{\pi}$	
	do par <i>of the pair</i>	do pai em comum <i>of the common parent</i>	4 alelos <i>4 alleles</i>	≥ 6 alelos ≥ 6 alleles
I	$A_i A_i - A_i A_i$	$A_i A_i$	0,25	0,25
	$A_x A_x - A_x A_x$	$A_i A_i$	0,00	0,00
	$A_i A_i - A_i A_i$	$A_i A_j$	0,50	0,50
	$A_j A_j - A_j A_j$	$A_i A_j$	0,50	0,50
II	$A_i A_i - A_j A_j$	$A \bullet A \bullet^2$	0,00	0,00
III	$A_i A_i - A_i A_j$	$A_i A_i$	0,25	0,25
	$A_i A_i - A_i A_j$	$A_i A_j$	0,29	0,29
	$A_j A_j - A_i A_j$			
	$A_i A_i - A_i A_x$	$A_i A_j$	0,44	0,47
	$A_j A_j - A_i A_x$			
	$A_x A_x - A_x A_i$			
	$A_x A_x - A_x A_j$	$A_i A_j$	0,00	0,00
	$A_x A_x - A_x A_x$			
IV	$A_i A_i - A_j A_x$	$A \bullet A \bullet$	0,00	0,00
V	$A_i A_j - A_i A_j$	$A_i A_j / A_i A_i$	0,25	0,25
	$A_i A_x - A_i A_x$	$A_i A_j$	0,49	0,48
	$A_j A_x - A_j A_x$			
	$A_x A_x - A_x A_x$	$A_i A_j$	0,00	0,00
VI	$A_i A_j - A_i A_x$	$A_i A_i$	0,25	0,25
	$A_i A_j - A_j A_x$			
	$A_i A_j - A_j A_x$	$A_i A_j$	0,25	0,29
	$A_i A_x - A_i A_x$			
	$A_j A_x - A_j A_x$			
	$A_i A_x - A_j A_x$	$A_i A_j$	0,00	0,00
VII	$A_i A_j - A_k A_l$	$A \bullet A \bullet$	0,00	0,00

¹ A_x = qualquer alelo diferente do alelo do pai em comum.
any allele different from the allele of the common parent.

² $A \bullet$ = qualquer um dos alelos (i, j, k, l, ..., x).
anyone alleles (i, j, k, l, ..., x).

Tabela 5 - Proporções ($\hat{\pi}_{(hom)}$, $\hat{\pi}_{(het)}$, $\hat{\pi}$) estimadas de alelos IBD, quando ambos os genótipos dos pais são desconhecidosTable 5 - Estimated proportions ($\hat{\pi}_{(hom)}$, $\hat{\pi}_{(het)}$, $\hat{\pi}$) of alleles IBD when genotypes of parents are unknown

Par de meio-irmãos <i>Half-sibs pair</i>		$\hat{\pi}_{(hom)}$	$\hat{\pi}_{(het)}$	$\hat{\pi}$	
Tipo <i>Type</i>	Genótipo <i>Genotype</i>				
I	$A_i A_i - A_i - A_i$	0,25	0,5	0,44 ¹	0,48 ²
II	$A_i A_i - A_j - A_j$	0	0	0	0
III	$A_i A_i - A_i - A_j$	0,25	0,39	0,36	0,38
IV	$A_i A_i - A_j - A_k$	0	0	0	0
V	$A_i A_j - A_i - A_j$	0,25	0,38	0,35	0,37
VI	$A_i A_j - A_i - A_k$	0,25	0,29	0,28	0,29
VII	$A_i A_j - A_k - A_l$	0	0	0	0

¹ $\pi_i = 0,25$ e $n = 4$.

² $\pi_i = 0,10$ e $n = 10$.

Tabela 6 - Porcentagem de pares de meio-irmãos com um ou dois locus de marcadores ligados entre si com $\hat{\pi} = 0,25$
 Table 6 - Percentage of half-sib pairs with one or two marker loci linked together with $\hat{\pi} = 0,25$

Situações ¹ Situations	Porcentagem desconhecida de genótipo das mães Percentage of dam genotype unknown											
	0				50				100			
	N ² =2	4	6	10	2	4	6	10	2	4	6	10
$\hat{\pi}_1$ e $\hat{\pi}_2 = 0,25$	15,5	7,3	6,1	2,7	27,5	8,7	5,0	1,5	39,3	9,0	2,9	1,3
$\hat{\pi}_1=0,25$ e $\hat{\pi}_2=0,0$	48,5	40,3	36,8	27,1	45,6	38,8	34,9	22,5	46,0	44,6	30,1	11,4
$\hat{\pi}_2=0,25$ e $\hat{\pi}_1=0,0$												
$\hat{\pi}_1$ e $\hat{\pi}_2 \neq 0,25$	36,0	52,4	57,1	70,2	26,9	52,5	60,1	76,0	14,7	46,4	67,0	87,3

¹ $\hat{\pi}_1$ e $\hat{\pi}_2$ são as proporções de alelos IBD em dois locus 1 e 2, ligados entre si por uma distância de 20 cM.

$\hat{\pi}_1$ e $\hat{\pi}_2$ are proportion of alleles IBD in two loci 1 and 2, linked together with 20 cM interval between them.

² Número de alelos, de mesma frequência, em cada locus do marcador.

Number of alleles of same frequency in each marker locus.

Tabela 7 - Porcentagem dos diferentes tipos de pares de meio-irmãos, quando os genótipos de ambos os pais e das progênie são conhecidos

Table 7 - Percentage of different half-sibs pairs type when genotypes of both parents and offspring are known

Tipo de pares de meio-irmãos Half-sibs pairs type	Número de alelos Number of alleles			
	2	4	6	10
I	20,1	4,9	2,4	1,0
II	6,3	2,2	1,1	0,3
III	48,0	27,3	19,0	9,3
IV	-	8,7	9,0	7,1
V	25,6	16,7	12,1	9,0
VI	-	35,9	43,7	43,5
VII	-	4,3	12,7	29,0

Conclusões

Estes resultados sugerem que os procedimentos adotados nesta pesquisa podem levar à maior eficiência, quando da utilização de pares-de-irmãos para se detectarem ou mapearem QTLs nas situações em que os genótipos dos pais são desconhecidos parcial ou totalmente.

É importante, todavia, que se possam testar estes procedimentos por intermédio de simulações nas quais deverão ser avaliados outros fatores que afetem diretamente a eficiência de se detectarem QTLs e a precisão com que são estimados.

Referências Bibliográficas

- AMOS, C.I., ELSTON, R.C., WILSON, A.F. et al. 1989. A more powerfull robust sib-pair test of linkage for quantitative traits. *Genet. Epidemiology*, 6:435-449.
- FULKER, D.W., CARDON, L.R. 1994. A sib-pair approach to interval mapping of quantitative trait loci. *Am. J. Hum. Genet.*, 54:1092-110.
- GOLDGAR, D.E. 1990. Multipoint analysis of human quantitative genetic variation. *Am. J. Hum. Genet.*, 47:957-967.
- HASEMAN, J.M., ELSTON, R.C. 1972. The investigation of linkage between a quantitative trait and a marker locus. *Behav. Genet.*, 2:3-19.
- KEMPTHORNE, O. 1957. *An introduction to genetics statistics*. New York: John Wiley. 545p.
- ROBERTSON, A. 1973. Linkage between marker loci and those affecting a quantitative trait. *Behav. Genet.*, 3:389-391.
- SCHORK, N.J. 1993. Extended multipoint identical-by-descent analysis of human quantitative traits: efficiency, power, and modeling considerations. *Am. J. Hum. Genet.*, 53:1386-1393.
- XU, S., ATCHLEY, R.W. 1995. A random model approach to interval mapping of quantitative trait loci. *Genetics*, 141:1189-1197.

Recebido em: 13/07/99

Aceito em: 10/09/99