

O uso da variância como metodologia alternativa para integração de mapas genéticos

Caio César Salgado⁽¹⁾, Cosme Damião Cruz⁽¹⁾, Moisés Nascimento⁽¹⁾ e Carlos Felipe Sanches Barrera⁽¹⁾

⁽¹⁾Universidade Federal de Viçosa, Departamento de Genética e Melhoramento, Edifício Artur Bernardes, Subsolo, Sala 12, CEP 36570-000 Viçosa, MG. E-mail: caiocesio@yahoo.com.br, cdcruz@ufv.br, moysesnascim@vicosa.ufv.br, feba286@hotmail.com

Resumo – O objetivo deste trabalho foi desenvolver um processo de integração de mapas genéticos, com o uso do inverso da variância, e testar sua eficiência. Foram utilizadas populações simuladas F_2 codominante e de retrocruzamento, com tamanhos populacionais de 100, 150, 200 e 400 indivíduos, tendo-se considerado uma espécie diploide fictícia com $2n = 2x = 2$ cromossomos, com o comprimento total do genoma por grupo de ligação estipulado em 100 cM, 21 marcas por grupo de ligação e marcadores equidistantes em 5 cM. Os genomas foram comparados quanto ao tamanho do grupo de ligação, variância das distâncias entre marcas adjacentes, correlação de Spearman e quanto ao estresse relativo à adequação das distâncias estimadas. Cada genoma simulado foi fragmentado em quatro novos mapas: três com oito marcadores e um com nove marcadores, cada qual com quatro marcadores âncoras. Os mapas foram alinhados, ordenados, integrados e, em seguida, comparados ao mapa de origem. O processo de integração de mapas proposto mostrou-se eficiente. Os mapas gerados tiveram pequena tensão interna em comparação aos mapas dos quais se originaram. A integração de mapas depende do tipo de população utilizada, tamanho da população, tipo de marcador, da frequência de recombinação e da fase de ligação.

Termos para indexação: frequência de recombinação, grupos de ligação, marcador molecular, tensão interna.

Variance utilization as an alternative methodology for integrating genetic maps

Abstract – The objective of this work was to develop an integration process of genetic maps using the inverse variance, and to test its efficiency. Simulated F_2 codominant populations and backcrosses of 100, 150, 200, and 400 plants were used, considering a virtual diploid species with $2n = 2x = 2$ chromosomes, with total genome length per linkage group stipulated in 100 cM, 21 marks per linkage group, and equidistant markers spaced at 5 cM. The genomes were compared for linkage group size, distance variance between adjacent marks, Spearman correlation and stress related to the estimated distances adequation. Each simulated genome was fragmented into four new maps: three with eight makers and one with nine, each containing four anchor markers. The maps were aligned, ordered, integrated and, then, compared with the original map. The proposed map integration process proved to be effective. The generated maps showed a little internal tension in comparison to the original ones. Map integration depends on the population type, population size, marker type, recombination frequency, and on the linkage phase.

Index terms: recombination frequency, linkage groups, molecular marker, internal tension.

Introdução

Mapas genéticos baseados em vários tipos de marcas genéticas têm se tornado disponíveis em abundância nos últimos anos. Programas computacionais bem estabelecidos, tais como Map-Maker, Linkage-1 e GQMOL (Suiter et al., 1983; Lander et al., 1987; Cruz, 2008) são frequentemente utilizados para análise desses dados.

Há pelo menos duas fontes de mapas múltiplos. Primeiro, quando o mapeamento de populações surge a partir de diferentes cruzamentos experimentais com uso de materiais independentes (Sewell et al., 1999;

Lombard & Delourme 2001) e, segundo, quando alelos segregam em ambos os pais, em espécies de fecundação cruzada. Neste caso, frequentemente constroem-se mapas de ligação específicos para cada pai, tal como na estratégia de mapeamento do pseudocruzamento teste (Grattapaglia & Sederoff, 1994; Marques et al., 1998; Testolin et al., 2001). Nesse contexto, é importante que se estudem maneiras eficientes de combinar múltiplos mapas genéticos em um único mapa com grande conteúdo de informação.

Para a integração de mapas distintos, é necessário que eles possuam marcadores em comum: “âncoras”

com as quais se possa interligá-los. Apenas quando um número mínimo de âncoras é disponível, mapas distintos podem ser representados juntos. Porém, ainda que marcadores em comum estejam disponíveis, são comuns os problemas relativos à precisão das estimativas da frequência de recombinação, que variam muito entre conjuntos de dados, e ao tipo de informação utilizada. A dificuldade está em definir os pesos dos tipos de informações que poderiam resultar em um mapa único, mais verossímil, com o mínimo de tensão interna ou, geralmente, que poderiam satisfazer critérios preestabelecidos de avaliação. Para isto, é necessária a utilização de ferramentas estatísticas.

A integração de mapas genéticos tem sido extensivamente utilizada na pesquisa, o que pode ser verificado pela quantidade de artigos publicados na última década com essa metodologia, aplicada a diferentes espécies vegetais (Freyre et al., 1998; Sewell et al., 1999; Peleman et al., 2000; Lombard & Delourme, 2001; Ulloa et al., 2002; Paran et al., 2004; Yan et al., 2005; Doligez et al., 2006) ou animais (Brown et al., 1998).

A integração de mapas derivados de várias populações apresenta diversas vantagens, como o aumento da saturação do mapa e da cobertura do genoma, além de aumentar a precisão da ordem das marcas comuns. Mapas integrados também são usados em estudos comparativos entre espécies (Beavis & Grant, 1991), para evidenciar rearranjos cromossômicos entre grupos de ligação homólogos (Dirlewanger et al., 2004; Pelgas et al., 2006; Nicolas et al., 2007), e para detecção de QTL em múltiplas populações (Symonds et al., 2005; Blanc et al., 2006).

Assim, não há dúvida de que processos de integração ou construção de mapas genéticos, com alta densidade de marcadores moleculares, são importantes para qualquer programa de melhoramento genético. Entretanto, há uma carência de métodos e mesmo de programas computacionais que atendam à demanda e ao fluxo de informações gerado pelos laboratórios.

Um mapa genético é essencialmente um arranjo linear de marcadores, ordenados a partir de valores de recombinação. O método proposto baseia-se nesse princípio e tem, como principal ferramenta para integração, a análise multiloco, assim como no método de Stam (1993). Em contraste com JoinMap, que estima a informação sobre a recombinação em um dado cruzamento, a partir dos valores limites

de detecção (LOD), e então combina estimativas entre os cruzamentos, assumindo uma distribuição binomial, o método proposto considera a informação de recombinação com base nos valores de variância da distância entre marcas adjacentes. Esses valores dependem ou são influenciados pelo tipo de população utilizado, tamanho da população, tipo de marcador, da frequência de recombinação e da fase de ligação.

O objetivo deste trabalho foi desenvolver um processo de integração de mapas genéticos, com o uso do inverso da variância, e testar sua eficiência.

Material e Métodos

Na integração de mapas genéticos construídos com diferentes técnicas, é indispensável que entre os grupos de ligação exista ao menos um marcador âncora (marcador comum aos grupos de ligação). De maneira geral, a integração envolve quatro etapas: identificação dos marcadores âncoras; obtenção de um mapa alinhado, sem enfatizar distâncias e nem o posicionamento mais apropriado; estabelecimento do melhor ordenamento e geração do mapa ordenado; e obtenção do mapa integrado efetivo, com ordenamento e medidas de distâncias apropriadamente calculadas.

Assim, para o processo de alinhamento de mapas proposto neste trabalho, instituiu-se uma marca como ponto referencial e, com base nas informações dos grupos de ligação, todas as demais foram posicionadas em relação à primeira. Nesse mapa alinhado, estavam presentes todas as marcas, inclusive as consideradas como âncoras que se apresentavam repetidas vezes no grupo de ligação, tendo-se em vista a variação de seu posicionamento relativo nos diferentes grupos.

O ordenamento das marcas consistiu da simples organização do mapa alinhado. Para as marcas âncoras, um posicionamento único foi definido a partir da distância média entre elas, e as demais marcas foram estabelecidas nos grupos de ligação originais. Assim, todas as marcas passaram a estar presentes no mapa ordenado, inclusive os marcadores âncoras que, ao contrário do grupo alinhado, foram representados uma única vez.

Tendo-se ordenado e alinhado os diferentes mapas genéticos, processou-se a integração dos mapas genéticos pela análise multiloco, cujas funções de mapeamento exercem papel fundamental no resultado obtido. A adoção de uma, entre as diferentes funções

de mapeamento, é dependente das pressuposições a respeito da distribuição da permuta, do grau de interferência e do comprimento do segmento cromossômico considerado. O método de integração de mapas genéticos derivados de diferentes cruzamentos foi implementado no programa de análises genômicas GQMOL (Cruz, 2008).

Na integração de mapas genéticos, é importante, também, que se considere a disponibilidade de matrizes de distância, em porcentagem de recombinação, entre cada par de marcas moleculares em cada mapa. Como ilustração, foi considerada a integração de grupos de ligação resultantes de três diferentes mapas, em que as matrizes de distâncias entre pares de marcas estavam disponíveis e denominadas de D1, D2 e D3. O ideal é que haja informação da distância entre todos os pares de marcas. Entretanto, deve-se ressaltar a possibilidade de reconstituir D1, D2, D3, a partir das informações dos mapas em que apenas valores de distâncias de marcas consecutivas foram diretamente mensurados. Nesse caso, a distância entre pares de marcas não consecutivas é obtida de forma inferencial e, portanto, grande quantidade de erro é incorporada à matriz, em razão do desconhecimento da verdadeira taxa de recombinação e pelo desconhecimento da taxa de interferência entre regiões adjacentes.

Sempre haverá diferentes medidas de distância originadas dos grupos de ligação a serem integradas entre pares de marcadores âncoras. É possível obter estimativas multiloco da distância entre os pares de âncoras, para aquele ordenamento específico estabelecido no grupo de ligação ordenado. Assim, o mapa pode apresentar o posicionamento dessas marcas.

Para obtenção do vetor m , com as distâncias apropriadas entre marcadores âncoras, por meio da análise multiloco, é utilizado o sistema de equações em que se considera uma matriz P , cujos elementos dependem de pesos (L_{ij}) a serem atribuídos às distâncias observadas, e outra Q , cujos elementos dependem de uma função linear das distâncias (M_{ij}) apropriadamente ponderadas pelos respectivos pesos. Assim, $Pm = Q$:

$$P = \begin{pmatrix} & a_{LC} \\ b_{LC} & \end{pmatrix}_{n \times n-1} \quad Q = \begin{pmatrix} q_1 \\ \dots \\ q_{n-1} \end{pmatrix}_{n \times 1} \quad m = \begin{pmatrix} m_1 \\ \dots \\ m_{n-1} \end{pmatrix}_{n-1 \times 1}$$

em que: n é o número de marcas moleculares; a_{LC} é o elemento da diagonal ou acima da diagonal da matriz

$$P - \text{ dado por } a_{LC} = \sum_{i=1}^L \sum_{j>c}^n L_{ij} \quad (L \leq C);$$

b_{LC} é o elemento abaixo da diagonal da matriz

$$P - \text{ dado por } b_{LC} = \sum_{i=1}^n \sum_{j=1}^c L_{ij} \quad (L > C);$$

q_i , elemento da i -ésima linha do vetor

$$Q - \text{ dado por } q_i = \sum_{i=1}^L \sum_{j>L}^n L_{ij} M_{ij}.$$

Nessas expressões, L_{ij} é o peso atribuído à informação do valor da distância m_{ij} , admitido neste estudo como sendo o inverso da variância de m_{ij} . A variância depende do tipo de população, do tipo de marcador, da fase de ligação e do próprio valor de m_{ij} .

Se a matriz D tiver sido reconstituída a partir do mapa genético, esta análise será inócua, pois reconstituirá exatamente o mapa fornecido para a função de mapeamento utilizado.

Entretanto, para obtenção do mapa consenso, o valor de distância entre pares de marcadores pode ser mais bem estimado não só pela informação dos elementos de uma particular matriz D , mas de todas elas tomadas em conjunto. Assim, tem-se: $P1m = Q1$, para o mapa 1; $P2m = Q2$, para o mapa 2; $Pcm = Qc$, para o c -ésimo mapa de ligação. De forma conjunta, pode-se escrever: $[P1 + P2 + \dots + Pc]m = [Q1 + Q2 + \dots + Qc]$, em que: $Pcm = Q1c$ e $m = Pc-1 Qc$.

Para o estudo do processo de integração, utilizou-se o programa de análises genômicas GQMOL, para simular uma espécie diploide fictícia com $2n = 2x = 2$ cromossomos (Figura 1), tomada como referência, cujo

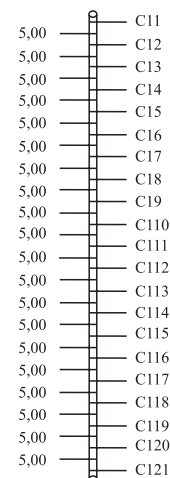


Figura 1. Grupo de ligação 1 do genoma analisado com saturação de 21 marcas equidistantes 5 cM.

comprimento total do genoma, por grupo de ligação, foi estipulado em 100 cM. Foi gerado o genoma com nível de saturação de 21 marcas moleculares (ou 5 cM de intervalo) por grupo de ligação. O genoma foi composto por apenas um grupo de ligação, com comprimento total de 100 cM.

Para o estudo dos mapas integrados e comprovação da funcionalidade da metodologia proposta, foram utilizadas populações F_2 codominante e de retrocruzamento com 100, 150, 200 e 400 indivíduos, com 21 marcas por grupo de ligação e marcadores equidistantes em 5 cM,

em um total de quatro simulações para F_2 codominante e quatro para retrocruzamentos (Figura 2 A, B). Todas as populações foram simuladas pelo programa de análise molecular GQMOL (Cruz, 2008).

Os genomas foram comparados quanto ao tamanho do grupo de ligação, variância das distâncias entre marcas adjacentes, correlação de Spearman e quanto ao estresse. Cada genoma simulado (população F_2 e de retrocruzamento) foi fragmentado em quatro novos mapas: três com oito marcadores e um com nove, cada qual com quatro marcadores âncoras. Os mapas

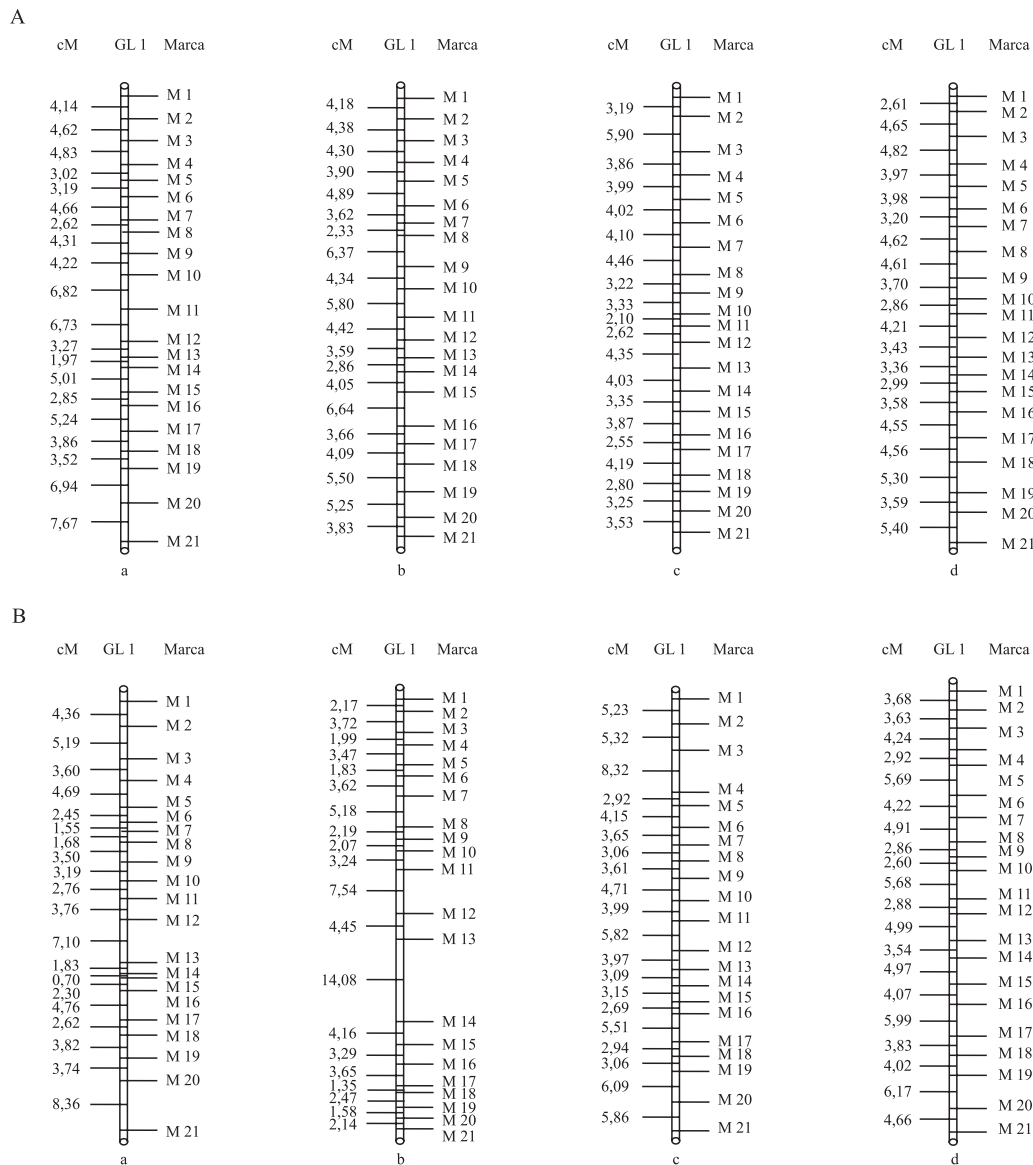


Figura 2. Genomas estudados no processo de integração de mapas: A, populações F_2 codominantes com tamanhos de 100 (a), 150 (b), 200 (c) e 400 (d) indivíduos; B, populações de retrocruzamentos com tamanhos de 100 (a), 150 (b), 200 (c) e 400 (d) indivíduos.

foram alinhados, ordenados, integrados e, em seguida, comparados ao mapa de origem.

Resultados e Discussão

Pelo processo de integração de mapas proposto, foi possível gerar um mapa com todos os marcadores com a melhor ordem, cujas distâncias entre marcadores âncoras e não âncoras são estimadas por análise multiloco. A veracidade das informações do mapa depende do tipo de população, do tipo de marca molecular e, principalmente, do número de indivíduos considerados na obtenção da porcentagem de recombinação entre pares de marcas.

Na Tabela 1, encontram-se os tipos e tamanhos de mapas utilizados no processo de integração, bem como a distância média de cada mapa, a variância, correlação de Spearman e o estresse avaliado em relação ao mapa de origem.

As correlações de Spearman se mostraram constantes em todos os mapas integrados, tanto nas populações F_2 codominantes como nas de retrocruzamento, e não apresentaram valores diferentes da unidade em nenhum dos mapas integrados, em comparação aos mapas dos quais se originaram. Esse resultado pode ser considerado um bom indicativo de confiabilidade do processo de integração dos mapas. Silfverberg-Dilworth et al. (2006) relataram que

alguns casos de inversão na ordem das marcas podem ocorrer quando as marcas estão estreitamente ligadas. O problema na identificação de inversões, em um mapa integrado, é como identificá-las, uma vez que não há um mapa referência. Quando se trabalha com espécies de fecundação cruzada, tal como na estratégia de mapeamento do pseudocruzamento, as comparações podem ser feitas com os mapas individuais construídos para cada pai, como a realizada por Doligez et al. (2006).

Diferenças na construção de mapas genéticos, pelos diferentes programas, são relatadas na literatura. Sewell et al. (1999), ao trabalhar com mapas genéticos de pinus, relataram que os mapas obtidos pelo JoinMap são ligeiramente diferentes dos obtidos pelo Mapmaker. Qi et al. (1996) relataram diferenças em mapas de ligação, construídos para cevada, e as atribuíram à maneira como cada programa calcula as distâncias, quando a interferência real difere da assumida.

A maioria dos relatos encontrados na literatura diz respeito apenas à utilização de dois programas que realizam a integração de mapas, sem uma explicação clara da metodologia utilizada por esses aplicativos. O programa mais citado na literatura, para a integração de mapas, é o JoinMap (Stam, 1993). São utilizados outros, como o descrito por Peleman et al. (2000) e Paran et al. (2004), com auxílio do programa Int-Map. Mais recentemente, Hu et al. (2004) apresentaram

Tabela 1. Tipos e tamanhos dos mapas utilizados no processo de integração, distância média entre marcas, variância, correlação de Spearman e estresse avaliados em relação ao mapa de origem.

Número de indivíduos	Tipo de mapa	População	Tamanho (cM)	Distância média (cM)	Variância	r	Estresse
100	Integrado	F_2 codominante	99,5810	4,9790	4,4891	1	16,4315
100	Ordenado	F_2 codominante	89,4225	4,4711	2,4412	1	0,4237
150	Integrado	F_2 codominante	97,1799	4,8590	1,6399	1	13,3624
150	Ordenado	F_2 codominante	88,0000	4,4000	1,1766	1	0,0
200	Integrado	F_2 codominante	78,9327	3,9466	1,0607	1	11,1122
200	Ordenado	F_2 codominante	72,7100	3,6355	0,7025	1	0,0
400	Integrado	F_2 codominante	87,9047	4,3952	1,0570	1	13,3325
400	Ordenado	F_2 codominante	79,9900	3,9995	0,6434	1	0,0
100	Integrado	Retrocruzamento	78,8587	3,9429	4,2338	1	10,6942
100	Ordenado	Retrocruzamento	71,9600	3,5980	3,4173	1	0,0
150	Integrado	Retrocruzamento	82,6401	4,1320	3,3211	1	15,2943
150	Ordenado	Retrocruzamento	74,1900	3,7095	8,0796	1	0,4861
200	Integrado	Retrocruzamento	97,7246	4,8862	3,0727	1	17,6581
200	Ordenado	Retrocruzamento	87,1400	4,3570	2,1534	1	0,0
400	Integrado	Retrocruzamento	99,3229	4,9624	2,2587	1	51,9245
400	Ordenado	Retrocruzamento	80,8900	4,4939	1,9859	1	47,4830

uma metodologia baseada em uma função de verossimilhança conjunta, para auxiliar na construção de mapas genéticos integrados. Estudos de simulação mostram que a diferença desta técnica, em comparação à metodologia proposta por Stam (1993), é pequena quando marcas codominantes são usadas, mas que a segunda é melhor, principalmente quando marcadores dominantes ou uma mistura de marcadores dominantes e codominantes são usados. Isso ocorre porque, com a metodologia baseada na função de verossimilhança, é possível encontrar os pesos ótimos entre as diferentes classes de dados, enquanto o método proposto por Stam (1993) não prediz a informação de cruzamentos que envolvem marcadores dominantes.

Doligez et al. (2006) integraram diferentes mapas de ligação, em uva, com o programa Carthagene e dados provenientes de família de irmãos completos, com tamanhos de 96, 45, 112, 139 e 153 indivíduos. O total do comprimento do mapa integrado de leitura foi de 1.485 cM, com média de distância inter-locus de 6,2 cM. Esses mapas foram integrados com o programa JoinMap, e os mapas integrados nos dois programas foram comparados. Esses autores observaram diferenças entre eles, principalmente inversões entre algumas marcas.

Para as populações F_2 codominantes, pelos mapas integrados, observou-se que quanto maior foi a população, maior foi a tendência de aproximação do tamanho do grupo de ligação do mapa integrado ao tamanho do mapa original. O mesmo ocorreu com a distância média das marcas. Quanto à variância, pôde-se observar queda com o aumento da população. O mesmo fato não ocorreu com o estresse, que decresceu na população simulada com marcas codominantes.

Segundo Soller & Beckmann (1983), quando marcadores codominantes estão disponíveis, análises baseadas em gerações F_2 serão mais úteis do que aquelas baseadas em gerações de retrocruzamento, já que fornecem informações tanto em relação à dominância quanto ao efeito maior do QTL identificado. Isso também pode ser uma explicação para a necessidade de maior tamanho de população, para a obtenção de mapas mais confiáveis em populações de retrocruzamentos.

Para as populações de retrocruzamento, observou-se aumento do número de indivíduos, com o aumento no tamanho dos grupos de ligação e das distâncias médias entre as marcas, o que não era esperado. Essas medidas se afastaram das observadas nos mapas originais, o que pode ser verificado pelo aumento do estresse com

o aumento da população, mas se aproximaram das medidas do mapa original que era de 100 cM e com marcadores equidistantes a 5 cM.

Outro resultado a ser destacado refere-se à construção dos mapas ordenados que, em quase todas as situações, tiveram estresse zero ou bem próximo de zero. Ocorreu uma única exceção com a população de retrocruzamento de tamanho 400, que apresentou estresse elevado. O estresse zero ou muito diminuto, em relação ao mapa dos quais os mapas foram originados, significa que eles eram exatamente iguais. Este fato pode levar a conclusões equivocadas, pois se poderia pensar que os mapas ordenados seriam mais indicados do que a integração dos mapas genéticos. Isto, no entanto, não é uma realidade, pois quando se obtém o mapa integrado efetivo, deve-se levar em conta que ele foi submetido a uma análise multiloco, ou seja, a ordem, bem como as distâncias entre as marcas, foram analisadas em conformidade com todas as marcas presentes no grupo de ligação.

Programas como MapMaker e Carthagene não informam qual metodologia utilizam para integração de mapas genéticos, e os artigos que fazem uso desses programas apenas informam quais os comandos utilizados para se obter o mapa integrado. Hu et al. (2004) avaliaram um método de integração de mapas baseado em estimativas de verossimilhança, em que uma estimativa da função de verossimilhança comum, que combina informações entre todos os cruzamentos, é proposta para se obter uma estimativa comum de recombinação. O grande problema da utilização do método proposto por Hu et al. (2004) é que eles não apresentaram um programa computacional específico para a construção de mapas genéticos integrados, o que restringe substancialmente a utilização dessa metodologia. Segundo os autores, no entanto, a metodologia proposta se mostrou, por simulações, mais eficiente do que a proposta por Stam (1993). Porém, para o emprego da metodologia de Stam (1993), somente a fase de ligação mais provável pode ser usada. Isto a tornaria mais eficiente com tamanho maior de progênies, em que a fase incorreta pode ter probabilidade muito menor que a fase correta, e ter pequeno efeito na probabilidade total. Entretanto, quando o tamanho da progênie é pequeno, ao escolher-se a fase mais provável, pode-se introduzir um viés em razão dos efeitos do pequeno tamanho da amostra, pela incongruência da posição das marcas (Daryl et al., 2004).

Conclusões

1. O processo proposto de integração de mapa é eficiente e gera mapas com pequena tensão interna, em comparação aos mapas dos quais se originaram.

2. O aumento do tamanho da população de retrocruzamento e F₂ codominante é eficiente na recuperação do tamanho dos grupos de ligação e da distância média entre as marcas originais, além de reduzir a variância.

3. A integração de mapas depende do tipo de população utilizada, tamanho da população, tipo de marcador utilizado, da frequência de recombinação e da fase de ligação.

Referências

- BEAVIS, W.D.; GRANT, D. A Linkage map based on information from four F₂ populations of maize (*Zea Mays* L.). **Theoretical and Applied Genetics**, v.82, p.636-644, 1991.
- BLANC, G.; CHARCOSSET, A.; MARGIN, B.; GALLAIS, A.; MOREAU, L. Connected populations for detecting quantitative trait loci and testing for epistasis: an application in maize. **Theoretical and Applied Genetics**, v.113, p.206-224, 2006.
- BROWN, D.M.; MATISE, T.C.; KOIKE, G.; SIMON, J.S.; WINER, E.S.; ZANGEN, S.; MCLAUGHLIN, M.G.; SHOIZAWA, M.; ATKINSON, O.S.; HUDSON JUNIOR, J.R.; CHAKRAVARTI, A.; LANDER, E.S.; JACOB, H. J. An integrated genetic linkage map of the laboratory rat. **Mammalian Genome**, v.9, p.521-530, 1998.
- CRUZ, C.D. **Programa GQMOL**: programa para análise de genética quantitativa e molecular. Viçosa: UFV, 2008.
- DARYL, J.S.; PETER, I.; EDWARDS, K. A high-density microsatellite consensus map for bred wheat (*Triticum aestivum* L.). **Theoretical and Applied Genetics**, v.109, p.1105-1114, 2004.
- DIRLEWANGER, E.; GRAZIANO E.; JOOBEUR T.; GARRIGA-CALDERE, F.; CASSON, P.; HOWARD, W.; ARUS, P. Comparative mapping and marker-assisted selection in Rosaceae fruit crops. **Proceedings of the National Academy of Sciences of the United States of America**, v.101, p.9891-9896, 2004.
- DOLIGEZ, A.; ADAM-BLONDON, A.F.; CIPRIANI, G.; DI GASPERO, G.; LAUCOU, V.; MERDINOGLU, D.; MEREDITH, C.P.; RIAZ, S.; ROUX, C.; THIS, P. An integrated SSR map of grapevine base on five mapping populations. **Theoretical and Applied Genetics**, v.113, p.369-382, 2006.
- FREYRE, R.; SKROCH, P.W.; GEFFROY, V.; ADAM-BLONDON, A.-F.; SHIRMOHAMADALI, A.; JOHNSON, W.C.; LLACA, V.; NODARI, R.O.; PEREIRA, P.A.; TSAI, S.M.; TOHME, J.; DRON, M.; NEINHUIS J.; VALLEJOS, C.E.; GEPTS, P. Towards an integrated linkage map of common bean. 4. Development of a core linkage map and alignment of RFLP maps. **Theoretical and Applied Genetics**, v.97, p.847-856, 1998.
- GRATTAPAGLIA, D.; SEDEROFF, R. Genetic linkage maps of *Eucalyptus grandis* and *E. urophylla* using a pseudo-testcross mapping strategy and RAPD markers. **Genetics**, v.137, p.1121-1137, 1994.
- HU, X.S.; GOODWILLIE, C.; RITLAND, K.M. Joining genetic linkage maps using a joint likelihood function. **Theoretical and Applied Genetics**, v.109, p.996-1004, 2004.
- LANDER, E.S.; GREEN, P.; ABRAHANSON, J.; BARLOW, A.; DALY, M.J.; LINCOLN, S.E.; NEWBURG, L. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. **Genomics**, v.1, p.174-81, 1987.
- LOMBARD, V.; DELOURNE, R. A consensus linkage map for rapeseed (*Brassica napus* L.): construction and integration of three individual maps from DH populations. **Theoretical and Applied Genetics**, v.103, p.491-507, 2001.
- MARQUES, C.M.; ARAUJO, J.A.; FERREIRA, J.G.; WHETTEN, R.; O'MALLEY D.M.; LIU, B.-H.; SEDEROFF, R. AFLP genetic maps of *Eucalyptus globulus* and *E. tereticornis*. **Theoretical and Applied Genetics**, v.96, p.727-737, 1998.
- NICOLAS, S.D.; MIGNON, G.L.; EBER, F.; CORITON, O.; MONOD, H.; CLOUET, V.; HUTEAU, V.; LOSTANLEN, A.; DELOURME, R.; CHALHOUB, B.; RYDER, C.D.; CHEVRE, A.M.; JENCZEWSKI, E. Homologous recombination plays a major role in chromosome rearrangements that occur during meiosis of *Brassica napus* Haploids. **Genetics**, v.175, p.487-503, 2007.
- PARAN, I.; VAN DER VOORT, J.R.; LEFEBVRE, V.; JAHN, M.; LANDRY, L.; SCHRIEK, M.V.; TANYOLAC, B.; CARANTA, C.; CHAIN, A.B.; LIVINGSTONE, K.; PALLOIX, A.; PELEMAN, J. An integrated genetic linkage map of pepper (*Capsicum* spp.). **Molecular Breeding**, v.13, p.251-261, 2004.
- PELEMAN, J.; WIJK, R.V.; OEVEREN, J.V.; SCHAİK, R.V. Linkage map integration: an integrate genetic map of *Zea mays* L. In: PLANT AND ANIMAL GENOME CONFERENCE, 8., 2000, San Diego. **Proceedings**. [S.l: s.n.], 2001. 624p.
- PELGAS, B.; BEAUSEIGLE, S.; ACHERÉ, V.; JEANDROZ, S.; BOUSQUET, J.; ISABEL, N. Comparative genome mapping among *Picea glauca*, *P. mariana* x *P. rubens* and *P. abies*, and correspondence with other Pinaceae. **Theoretical and Applied Genetics**, v.113, p.1371-1393, 2006.
- QI, X.; STAM, P.; LINDHOUT, P. Comparason and integration of four barley genetic maps. **Genome**, v.39, p.379-394, 1996.
- SEWELL, M.M.; SHERMAN, B.K.; NEALE, D.B. A consensus map for loblolly pine (*Pinus taeda* L.). I. Construction and integration of individual linkage maps from two outbreed three-generation pedigrees. **Genetics**, v.151, p.321-330, 1999.
- SILFVERBERG-DILWORTH, E.; MATASCI, C.L.; VAN DE WEG, W.E.; VAN KAAUWEN, M.P.W.; WALSER, M.; KODDE, L.P.; SOGLIO, V.; GIANFRANCESCHI, L.; DUREL, C.E.; COSTA, F.; YAMAMOTO, T.; KOLLER, B.; GESSLER, C.; PATOCCHI, A. Microsatellite markers spanning the apple (*Malus x domestica* Borkh). **Tree Genetics and Genomes**, v.2, p.202-224, 2006.

- SOLLER, M.; BECKMANN, J.S. Genetic polymorphism in varietal identification and genetic improvement. **Theoretical and Applied Genetics**, v.67, p.25-33, 1983.
- SUITER, K.A.; WENDEL, J.F.; CASE, J.S. LINKAGE-1: a PASCAL computer program for the detection and analysis of genetic linkage. **Journal of Heredity**, v.74, p.203-204, 1983.
- STAM, P. Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. **The Plant Journal**, v.3, p.739-744, 1993.
- SYMONDS, V.V.; GODOY, A.V.; ALCONADA, T.; BOTTO, J.F.; JUENGER, T.E.; CASAL, J.J.; LLOYD, A.M. Mapping quantitative trait loci in multiple populations of *Arabidopsis thaliana* identifies natural allelic variation for trichome density. **Genetics**, v.169, p.1649-1658, 2005.
- TESTOLIN, R.; HUANG, W.G.; LAIN, O.; MESSINA, R.; VECCHIONE, A.; CIPRIANI, G. A kiwifruit (*Actinidia* spp.) linkage map based on microsatellites and integrated with AFLP markers. **Theoretical and Applied Genetics**, v.103, p.30-36, 2001.
- ULLOA, M.; MEREDITH JUNIOR, W.R.; SHAPPLEY, Z.W.; KAPHLER, A.L. RFLP genetic linkage maps from F_{2,3} populations and a joinmap of *Gossypium hirsutum* L. **Theoretical and Applied Genetics**, v.104, p.200-208, 2002.
- YAN, Z.; DENNEBOOM, C.; HATTENDORF, A.; DOLSTRA, O.; DEBENER, T.; STAM, P.; VISSER, P.B. Construction of an integrated map of rose with AFLP, SSR, PK, RGA, RFLP, SCAR and morphological markers. **Theoretical and Applied Genetics**, v.110, p.766-777, 2005.

Recebido em 9 de setembro de 2010 e aprovado em 20 de outubro de 2010