

OTIMIZAÇÃO DO MAPEAMENTO GENÉTICO VEGETAL DE POPULAÇÕES DUPLO-HAPLÓIDE VIA SIMULAÇÃO COMPUTACIONAL

OPTIMIZATION OF VEGETAL GENETIC MAPPING OF DOUBLE-HAPLOID POPULATIONS VIA COMPUTATIONAL SIMULATION

Silvan Gomes de BRITO¹; Péricles de Albuquerque Melo FILHO²;
Mairykon Coelho da SILVA³; Eliane Cristina ARCELINO³; Alisson Esdras COUTINHO¹;
Diogo Gonçalves NEDER⁴

1. Mestre em Melhoramento Genético de Plantas, Universidade Federal Rural de Pernambuco - UFRPE, Recife, PE, Brasil. gomesilvapb@hotmail.com; 2. Professor Associado do Departamento de Agronomia, Universidade Federal Rural de Pernambuco - UFRPE, Recife, PE, Brasil; 3. Graduando em Agronomia, Universidade Federal Rural de Pernambuco - UFRPE, Recife, PE, Brasil; 4. Professor Doutor do Departamento de Agroecologia e Agropecuária/CCAA, Universidade Estadual da Paraíba – UEPB, Lagoa Seca, PB, Brasil.

RESUMO: O mapeamento genético é um passo necessário para entender a organização genômica e a relação entre genes e o fenótipo. Um dos principais problemas está em encontrar a ordem, o espaçamento correto dos marcadores em um mapa genético, assim como o número de indivíduos a compor uma população. Deste modo, o objetivo deste estudo foi avaliar o nível de saturação do genoma e o tamanho ideal de populações simulada duplo-haplóide para a construção de mapas de ligação mais confiáveis por meio de simulação computacional. Foram simulados genomas parentais e populações duplo-haplóide considerando marcadores moleculares do tipo dominante, espaçados de forma equidistante a 5, 10 e 20 cM. Os tamanhos das populações geradas foram de 100, 200, 300, 500, 800 e 1000 indivíduos, com dez grupos de ligação cada e 100 repetições por amostra. Procedeu-se a análise de todas as populações geradas obtendo um genoma analisado o qual foi comparado com o genoma simulado inicialmente. Observou-se que o tamanho ideal de populações duplo-haplóide para mapeamento genético foi de no mínimo 200, 500 e 1000 indivíduos para genomas saturados, medianamente saturados e com baixa saturação. Populações de mesmo tamanho tendem a produzir mapas com maior acurácia em níveis de saturação do genoma mais elevados.

PALAVRAS-CHAVE: Grupos de ligação. Nível de saturação. Genoma. Mapa de ligação. Ordem de marcas.

INTRODUÇÃO

Estudos com enfoque molecular são praticados com objetivo de explorar a variabilidade genética no melhoramento de plantas através de técnicas biotecnológicas. Com o advento dos marcadores moleculares, tornou-se possível a construção de mapas de ligação para diversas espécies vegetais.

Na construção de mapas genéticos são utilizados diferentes tipos de populações segregantes derivada do cruzamento entre duas linhas puras (populações F₂, retrocruzamentos, F₁“pseudo-testcross”), assim como, linhagens endogâmicas recombinantes - RILs (*Recombinant Inbred Lines*) derivada de uma população F₂ por meio de sucessivas autofecundações, ou uma população de linhagens duplo-haplóide obtidas a partir da cultura de anteras ou de micrósoros de plantas F₁ (SCHUSTER; CRUZ, 2008) a fim de auxiliar o mapeamento de QTLs (*Quantitative Trait Loci* – Loco de Característica Quantitativa). Por outro lado, têm amplo emprego como ferramenta biotecnológica em estudos de variabilidade genética para o melhoramento das plantas cultivadas

(TEIXEIRA et al., 2004; FRANCESCHINELLI et al., 2006; MELO et al., 2009), permitindo o mapeamento de genes de interesse agrônomico.

O processo para se obter uma população de duplo-haplóide é bem mais rápido do que a obtenção de populações RILs (SEMAGN et al., 2010). Para duplo-haplóide é necessária apenas uma geração de recombinação para se alcançar a homozigose, enquanto que para linhas endogâmicas recombinantes são realizados sucessivos ciclos de autofecundação da planta F₁, no mínimo 6 gerações. Populações de duplo-haplóides têm sido utilizadas para o mapeamento de QTLs em várias espécies (FORSTER et al., 2000; RADOEV et al., 2008; WANG et al., 2011; SEYMOUR et al., 2012)

Segundo Silva et al. (2007), a disponibilidade de mapas confiáveis depende de vários fatores tais como o tipo e tamanho da população, bem como tipo e número de marcadores.

A resolução do mapa e a capacidade de se determinar a sequência de marcadores nele estão diretamente relacionadas ao tamanho da população, dada a importância de se estimar de forma adequada as distâncias entre os genes, a fim de estabelecer o ordenamento correto, bem como a formação de

grupos de ligações que reflitam o número básico de cromossomos da espécie (BHERING; CRUZ, 2008).

Ao definir o tipo de população para mapeamento, o conteúdo médio de informação associado à variância podem ser utilizados com critérios para estimar a frequência de recombinação (r) a fim de avaliar a confiabilidade do mapa genético (ROCHA et al., 2004).

Apesar da disponibilidade de vários estudos sobre mapeamento genético, há ainda a necessidade de estudos específicos relacionados com a determinação do número ideal de indivíduos numa dada população.

Com isto, o objetivo do presente trabalho foi determinar o nível de saturação do genoma, bem como identificar o tamanho da população ideal a ser empregado no mapeamento genético de populações duplo-haplóides via simulação computacional.

MATERIAL E MÉTODOS

Para a geração dos dados foi considerada uma espécie vegetal diplóide com $2n = 2x = 20$ cromossomos utilizando o módulo de simulação do aplicativo computacional GQMOL (CRUZ, 2004), o qual permite gerar informações sobre genomas, genótipos de genitores e indivíduos de diferentes tipos de populações.

Foram gerados três genomas com três níveis de saturação: saturado, medianamente saturado e pouco saturado, com intervalos entre marcas adjacentes de 5, 10 e 20 cM, representando um número total de marcas por genoma de 210, 110 e 60, respectivamente. Cada genoma composto por 10 grupos de ligação, com 100 cM em cada grupo e, portanto, com comprimento total de 1000 cM. Foram considerados os tipos de marcadores moleculares dominantes distribuídos de forma equidistante no genoma.

Foram simuladas 100 amostras com 100, 200, 300, 500, 800 e 1000 indivíduos. Após a geração dos dados da população duplo-haplóide, foram aplicados testes de qui-quadrado (χ^2) para verificar a razão de segregação de cada marca em todas as populações geradas. No processo de mapeamento foram utilizadas todas as marcas, mesmo aquelas que não segregaram de acordo com a proporção esperada.

Após a aplicação dos testes de segregação, seguiu-se a etapa da estimação da percentagem de recombinação entre pares de marcas utilizando o método da máxima verossimilhança como descrito por Schuster e Cruz (2004). Na formação do grupo de ligação foi utilizada a propriedade transitiva. Os

critérios utilizados no agrupamento foram a frequência máxima de recombinação (r_{max}) e o LOD mínimo (LOD_{min}), para inferir se dois locos estão ligados, foram utilizados os valores de 30% e 3, respectivamente.

Após a formação dos grupos de ligação empregou-se o método da SARF (*Sum of Adjacent Recombination Fractions*), para determinar a melhor ordem das marcas nos grupos. Após a realização de todos os passos descritos até então, foram formados os grupos de ligação para a população simulada, utilizando como medida de distância a porcentagem de frequência de recombinação. O próximo passo foi à comparação dos grupos de ligação formados para todas as populações simuladas com aqueles grupos de ligação estabelecidos no genoma simulado, como descrito a seguir.

As variáveis analisadas foram: número de grupos de ligação obtidos; tamanho dos grupos de ligação; distâncias médias entre marcas adjacentes nos grupos de ligação; variâncias das distâncias entre marcas adjacentes nos grupos de ligação; estresse (expressa o grau de concordância dos valores de distância entre cada par de marcas adjacentes nos grupos de ligação, simulados em relação às distâncias nos respectivos pares de marcas no genoma de referência); inversão de posição dos marcadores, verificada pela correlação de Spearman. Nas análises apresentadas, foram utilizadas apenas as repetições em que houve recuperação dos três grupos de ligação no mapeamento genético.

Foram comparadas pelo teste de Scott-Knott (1974), a 5% de probabilidade (erro tipo I) as médias das variáveis: tamanho do grupo de ligação, distância média de marcas adjacentes, variância e estresse para cada grupo de ligação obtido para vários tamanhos de população. Também foram comparadas as médias gerais (médias de todos os grupos de ligação), para cada tamanho de população dentro de cada nível de saturação do genoma.

Todas as análises foram realizadas utilizando o programa ONEMAP (MARGARIDO et al., 2007) e as comparações por meio do programa computacional R (R DEVELOPMENT CORE TEAM, 2011).

RESULTADOS E DISCUSSÃO

Vários tipos de populações controladas vêm sendo utilizado na construção de mapas de ligação cada vez mais saturados em busca de uma maior confiança no ordenamento de marcadores moleculares.

A veracidade das informações do mapa depende do tipo de população, do tipo de marcador molecular e principalmente do número de indivíduos considerados na obtenção da porcentagem de recombinação entre pares de marcas (SALGADO et al., 2011) além de técnicas de análises estatísticas e computacionais para estimativa de ligação e de distâncias entre marcadores, bem como seu ordenamento nos cromossomos.

A qualidade de um mapa pode ser avaliada pela confiança no agrupamento e ordenamento dos marcadores, assim como a proporção do genoma representado pelo comprimento e a densidade total do mapa, tamanho dos intervalos entre os marcadores. Por outro lado, à medida que aumenta a cobertura de um mapa genético, o número de grupos de ligação aproxima-se do número de cromossomos.

Era esperada a formação de 10 grupos de ligação para os três níveis de saturação. Considerando o número de grupos de ligação formados para as 100 amostras analisadas, observou-se um número diferente do esperado para tamanhos de população de 100, 200, 300, 500 e 800 indivíduos nos três níveis de saturação (Tabela 1). É possível visualizar que as populações com 100 indivíduos apresentaram menor número de repetições com recuperação dos 10 grupos de ligação, com valores de 78, 42 e 4 para os níveis de saturação de 5, 10 e 20 cM, respectivamente.

Este fato pode ser explicado pelo pequeno número de indivíduos utilizado na constituição da população de duplo-haplóide não ter sido eficiente na localização de ligação entre marcas, o que acabou não representando muito bem a diversidade de gametas produzidos pelos parentais. Silva (2005), estudando o mapeamento em populações RILs, detectou número de grupos de ligação acima do esperado em populações com tamanho inferior a 100 indivíduos e atribuiu à divisão de grupos de ligação em consequência da não detecção de ligação entre marcas onde certamente existia. Deste modo, independente do nível de saturação, o uso de tamanho da população com 100 indivíduos é inadequado para mapeamento em populações duplo-haplóide.

Conforme verificado, o número de grupos de ligação recuperados tende a ser igual ao número de grupos de ligação do genoma simulado à medida que o tamanho da população e a densidade do mapa genético aumentam. Assim, sugere-se que são necessárias populações com tamanho mínimo de 200, 500 e 1000 indivíduos, para construção de

mapas de ligação em populações duplo-haplóide nos níveis de saturação de 5, 10 e 20 cM, respectivamente.

Colaborando com esses resultados, os valores da correlação de Spearman iguais à unidade nos 10 grupos de ligação obtidos no mapeamento das populações segregantes para os 3 níveis de saturação, indicam que a ordem das marcas não foi alterada em relação à ordem previamente estabelecida no genoma o qual foi utilizado para a geração das populações (Tabela 1). Nota-se que apesar das diferenças no número de repetições analisadas para cada tamanho de população, não foi verificada inversão das marcas, o que evidencia boa qualidade dos dados analisados. Entretanto, caso os valores de correlação de Spearman obtidos fossem menores que a unidade à indicação seria que a ordem das marcas nos grupos de ligação obtidos no mapeamento da população segregante teria sido alterada em relação à ordem do genoma de referência.

O tamanho de cada grupo de ligação esperado era de 100 cM, no entanto, se observa uma pequena variação no comprimento do genoma para os três níveis de saturação (Tabela 1). Quando comparado com o tamanho do grupo de ligação simulado, as diferenças são de 1,422 e 4,056 cM para o menor e maior valor, respectivamente. Apesar de todos os valores estarem situados acima do inicialmente esperado (100 cM), as médias para os tamanhos populacionais avaliados se aproximam do valor simulado independente do nível de saturação, não diferindo estatisticamente entre si pelo teste de Scott-Knott realizado com $P < 0,05$. Estes resultados estão de acordo com os encontrados por Bhering e Cruz (2008). Os autores ressaltaram que não se pode fazer referência de um determinado tamanho da população, sendo necessária uma maneira adicional para observar o comportamento do comprimento médio dos grupos de ligação.

Comportamento semelhante é apresentado pela distância média entre marcas para o genoma da população simulada. Os valores esperados eram de 5, 10 e 20 cM, contudo, verifica-se que para este último, as distâncias médias foram variantes, de modo que para 5 e 10 cM, estatisticamente não houve diferença significativa entre as distâncias (Tabela 1). De modo geral, as distâncias médias entre os marcadores se aproximam dos valores esperados a medida em que o tamanho da população aumenta, assim como o tamanho médio dos grupos de ligação.

Tabela 1. Valores médios de correlação, tamanho, distância, variância e estresse dos grupos de ligação em função do tamanho de população e nível de saturação do genoma em populações duplo-haplóide utilizando marcadores.

SG	TP	NRA	Correlação	Tamanho	Distância	Variância	Estresse
5	100	78	1	104,056 a	5,203 a	5,066 a	46,060 a
	200	100	1	102,949 a	5,147 a	2,353 b	36,258 b
	300	100	1	103,490 a	5,174 a	1,571 b	32,500 c
	500	100	1	102,794 a	5,140 a	0,963 b	29,821 d
	800	100	1	103,255 a	5,163 a	0,623 b	27,795 e
	1000	100	1	103,112 a	5,156 a	0,506 b	26,722 e
10	100	42	1	103,879 a	10,388 a	9,816 a	28,248 a
	200	93	1	102,721 a	10,272 a	4,931 b	19,705 b
	300	99	1	103,145 a	10,314 a	3233 c	18,138 b
	500	100	1	103,459 a	10,346 a	2,011 d	15,191 c
	800	100	1	103,268 a	10,327 a	1,248 d	13,374 c
	1000	100	1	103,211 a	10,321 a	1,017 d	12,835 c
20	100	4	1	103,828 a	21,021 a	17,914 a	23,995 a
	200	70	1	101,422 a	20,300 b	9,386 b	12,926 b
	300	83	1	101,503 a	20,311 b	7,328 c	12,330 b
	500	95	1	101,967 a	20,393 b	5,030 d	10,455 b
	800	99	1	101,824 a	20,365 b	3,030 e	8,417 c
	1000	100	1	101,560 a	20,312 b	2,483 e	7,187 c

⁽¹⁾SG = Saturação do genoma; ⁽²⁾TP = Tamanho da população; ⁽³⁾NRA = Número de repetições avaliadas. Médias seguidas por letras iguais nas colunas, dentro de cada nível de saturação do genoma, não diferem entre si pelo teste de Scott Knott a 5% de probabilidade ($P < 0,05$).

Esta aproximação e a inalterabilidade das estimativas nos níveis de saturação mais elevados podem ser explicadas em termos de precisão das estimativas de recombinação em populações com maiores tamanhos amostrais e do nível de resolução para alta saturação (GOOD GOD, 2008). Isso sugere que populações com maior número de indivíduos permitem uma melhor amostragem dos eventos de permuta dos gametas, contribuindo para estimativas mais precisas e acuradas.

A variância das estimativas das distâncias entre os marcadores também pode ser tomada como medida de análise da acurácia de mapas mais fidedignos.

Ainda na Tabela 1, é possível visualizar o comportamento da variância da distância média entre marcadores nos três níveis de saturação do genoma simulado. Como era de se esperar, os valores referentes à variância estão de acordo com as estimativas de tamanho e distâncias médias entre os marcadores.

Ao analisar os níveis de saturação em separado, observa-se uma diminuição da variância na medida em que o tamanho da amostra é ampliado. Esta redução é evidenciada pelas diferenças significativas dadas pelo teste de médias,

em que as menores médias estão associadas às maiores populações, portanto, levando à maior precisão no mapeamento genético. Seja por exemplo, o nível de saturação de 5 cM, os valores da variância correspondem à 5,066 e 0,506 para os tamanhos de população com 100 e 1000 indivíduos, respectivamente.

Quando se considera um mesmo tamanho de população nos diferentes níveis de saturação, nota-se que o valor da variância será menor tanto quanto menor for o valor do nível de saturação do genoma (Tabela 1). Isto pode ser melhor entendido, quando se analisa, por exemplo, o tamanho da população de 100 indivíduos nos níveis de saturação de 5, 10 e 20 cM, onde os valores da variância são 5,066, 9,816 e 17,914, respectivamente. Alguns autores também observaram diminuição nos valores da variância na medida em que o tamanho da população e o nível de saturação aumentaram, em estudos relacionados ao mapeamento de populações via simulação (SILVA, 2005; BARROS, 2007; BHERING; CRUZ, 2008; SALGADO, 2008). Portanto, quanto menor o valor de variância mais ajustados estarão os valores do real, indicando uma boa recuperação do genoma, bem como maior uniformidade na distribuição dos marcadores dentro dos grupos de ligação.

Colaborando com estes resultados, verifica-se que o estresse expressa o grau de concordância dos valores de variância da distância média entre cada par de marcadores nos grupos de ligação simulados em relação às distâncias nos respectivos pares de marcadores no genoma de referência.

De modo semelhante ao que ocorre com a variância, os valores do estresse tendem a diminuir à medida que o tamanho da população avaliada aumenta independente do nível de saturação do genoma (Tabela 1). A redução dos valores de estresse com o aumento do tamanho das populações dentro do nível de saturação do genoma é evidenciada pelas diferenças significativas quando comparadas pelo teste de Scott-Knot ($P < 0,05$), de modo que, os maiores valores associados às menores populações, diferem significativamente das menores médias associadas às maiores populações para cada nível de saturação. Com essa diminuição, pode-se inferir que a qualidade dos dados analisados será mais confiável quanto maior for o número de indivíduos na população.

Ferreira et al. (2006) verificaram que os valores referentes ao estresse também diminuía com o aumento no tamanho da população. Por fim, estes resultados indicam que populações com

tamanho de mínimo 200, 500 e 1000 indivíduos, é suficiente para se obter mapas de ligação com número de grupos de ligação esperado, ordenamento correto das marcas, tamanho médio de grupos de ligação e distância média entre marcas coerentes.

CONCLUSÕES

O tamanho ideal de populações duplo-haplóide para mapeamento genético deve ser de 200, 500 e 1000 indivíduos quando os níveis de saturação do genoma são de 5, 10 e 20 cM, respectivamente.

Níveis de saturação mais elevados proporcionam mapas mais precisos, mais confiáveis em populações de mesmo tamanho.

O aumento no tamanho da população permitiu a obtenção de mapas de ligação com maior acurácia independente do nível de saturação.

AGRADECIMENTOS

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES, pela concessão da bolsa de estudos.

ABSTRACT: Genetic mapping is a necessary step to understand the genomic organization and the relationship between genes and phenotypes. A major problem is to find the order, the correct spacing of the markers in a genetic map, and the number of individuals to compose a population. Thus, the objective of this study was to evaluate the saturation level of the genome and the optimal size of simulated double-haploid populations for the construction reliable linkage maps by means of computer simulation. Parental genomes and double-haploid populations were simulated considering dominant molecular markers, spaced equidistantly at 5, 10 and 20 cM. The sizes of the generated populations were 100, 200, 300, 500, 800 and 1000 individuals, with ten linking groups and 100 replicates per sample. It was proceeded the analysis of all generated population obtaining a genome which was compared with the first simulated genome. It was observed that the optimal size of double-haploid populations for genetic mapping has been at least 200, 500 and 1000 individuals for saturated genomes, medium unsaturated and low saturation. Populations of the same size tend to produce maps with greater accuracy in higher levels of genome saturation.

KEYWORDS: Linkage groups. Saturation level. Genome. Linkage map. Order marks.

REFERÊNCIAS

- BARROS, W. S. **Genotipagem seletiva e outras estratégias de amostragem no mapeamento genético e na detecção de QTLs em populações F₂ simuladas**. 2007. 158p. Tese (Doutorado) - Universidade Federal de Viçosa, Viçosa.
- BHERING, L. L.; CRUZ, C. D. Tamanho da população ideal para mapeamento genético em famílias de irmãos completos. **Pesquisa Agropecuária Brasileira**, Brasília, v. 43, p. 379-385, 2008.
- CRUZ, C. D. **Programa para análises de dados moleculares e quantitativos – Gqmol**. Viçosa: UFV, 2004.
- FERREIRA, A.; SILVA, M. F.; SILVA, L. C.; CRUZ, C. D. Estimating the effects of population size and type on the accuracy of genetic maps. **Genetics and Molecular Biology**, Ribeirão Preto, v. 29, p. 187-192, 2006.

FORSTER, B. P.; ELLIS, R. P.; THOMAS, W. T. B.; NEWTON, A. C.; TUBEROSA, R.; THIS, D.; EL-ENEIN, R. A.; BAHRI, M. H.; SALEM, M. B. The development and application of molecular markers for abiotic stress tolerance in barley. **Journal of Experimental Botany**, Oxford, v. 51, p. 19-27, 2000. Disponível em: <<http://jxb.oxfordjournals.org/content/51/342/19.abstract>>. Acesso em: 17 mai. 2012.

FRANCESCHINELLI, E. V.; JACOBI, C. M.; DRUMMOND, M. G.; RESENDE, M. F. S. The genetic diversity of two Brazilian *Vellozia* (Velloziaceae) with different patterns of spatial distribution and pollination biology. **Annals of Botany**, London, v. 97, p. 585-592, 2006. Disponível em: <<http://aob.oxfordjournals.org/content/97/4/585>>. Acesso em: 14 fev. 2012.

GOOD GOD, P. I. V. **Mapeamento genético em famílias de meio irmãos por simulação computacional**. 2008. 114p. Tese (Doutorado) – Universidade Federal de Viçosa, Viçosa.

MARGARIDO, G. R. A.; SOUZA, A. P.; GARCIA, A. A. F. OneMap: software for genetic mapping in outcrossing species. **Hereditas**, Lund, v. 144, p. 78-79, 2007. Disponível em: <<http://onlinelibrary.wiley.com/doi/10.1111/j.2007.0018-0661.02000.x/pdf>>. Acesso em: 20 mar. 2012.

MELO, R. A.; MENEZES, D.; RESENDE, L. V.; WANDERLEY JÚNIOR, L. J. G.; SANTOS, V. F.; MESQUITA, J. C. P.; MAGALHÃES, A. G. Variabilidade genética em progênes de meios-irmãos de coentro. **Horticultura Brasileira**, Brasília, v. 27, p. 325-329, 2009.

R DEVELOPMENT CORE TEAM. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Disponível em: www.R-project.org. Acesso em: 10 nov. 2011.

RADOEV, M.; BECKER, H. C.; ECKER, W. Genetic analysis of heterosis for yield and yield components in rapeseed (*Brassica napus* L.) by quantitative trait locus mapping. **Genetics**, Austin, v. 179, p. 1547-1558, 2008.

ROCHA, R. B.; CRUZ, C. D.; BARROS, W. S.; FERREIRA, F. M.; ARAÚJO, E. F. Comparisons of segregating populations for genetic mapping. **Crop Breeding and Applied Biotechnology**, Viçosa, v. 4, p. 408-415, 2004.

SALGADO, C. C. **Integração de mapas genéticos**. 2008. 142p. Dissertação (Mestrado) - Universidade Federal de Viçosa, Viçosa.

SALGADO, C. C.; CRUZ, C. D.; NASCIMENTO, M. BARRERA, C. F. S. O uso da variância como metodologia alternativa para integração de mapas genéticos. **Pesquisa Agropecuária Brasileira**, Brasília, v. 46, p. 66-73, 2011.

SCHUSTER, I.; CRUZ, C. D. Estatística genômica aplicada a populações derivadas de cruzamentos controlados. Viçosa: UFV, 2008. 568p.

SEMAGN, K.; BJORNSTAD, A.; XU, Y. The genetic dissection of quantitative traits in crops. **Electronic Journal of Biotechnology**, v. 13, 2010. Chile, Disponível em <<http://dx.doi.org/10.2225/vol13-issue5-fulltext-14>>. Acesso em: 20 mai. 2012.

SEYMOUR, D. K.; FILIAULT, D. L.; HENRY, I. M.; MONSON-MILLER, J.; RAVI, M.; PANGA, A.; COMAI, L.; CHAN, S. W. L.; MALOOF, J. N. Rapid creation of *Arabidopsis* doubled haploid lines for quantitative trait locus mapping. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 109, p. 4227-4232, 2012. Disponível em: <<http://www.pnas.org/content/early/2012/02/23/1117277109.short>>. Acesso em: 03 mai. 2012.

SILVA, L. C. **Simulação do tamanho da população e da saturação do genoma para mapeamento genético de RILs**. 2005. 132p. Dissertação (Mestrado) - Universidade Federal de Viçosa, Viçosa.

SILVA, L. C.; L. C.; CRUZ, C. D.; MOREIRA, M. A.; BARROS, E. G. Simulation of population size and genome saturation level for genetic mapping of recombinant inbred lines (RILs). **Genetics and Molecular Biology**, Ribeirão Preto, v. 30, p. 1101-1108, 2007.

TEIXEIRA, H.; VIEIRA, M. G. G. C.; MACHADO, J. C. Marcadores RAPD na análise da diversidade genética de isolados de *Acremonium strictum*. **Fitopatologia Brasileira**, Brasília, v. 29, p. 651-655, 2004.

WANG, N.; QIAN, W.; SUPPANZ, I.; WEI, L.; MAO, B.; LONG, Y.; MENG, J.; MULLER, A. E.; JUNG, C. Flowering time variation in oilseed rape (*Brassica napus* L.) is associated with allelic variation in the FRIGIDA homologue BnaA.FRI.a. **Journal of Experimental Botany**, Oxford, v. 62, p. 5641-5658, 2011. Disponível em: <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3223056/>>. Acesso em: 17 mai. 2012.