

ARTIGO DE REVISÃO

MAPAS GENÉTICOS EM PLANTAS⁽¹⁾

MONALISA SAMPAIO CARNEIRO⁽²⁾; MARIA LUCIA CARNEIRO VIEIRA⁽²⁾

RESUMO

Ao lado dos projetos de seqüenciamento e das análises do cariótipo pelas técnicas de hibridização *in situ*, o desenvolvimento de mapas genéticos fundamentados em marcadores de DNA tem propiciado consideráveis avanços à *genômica* de plantas. Esta revisão aborda as premissas básicas utilizadas para o mapeamento genético e suas principais aplicações, especialmente para o melhoramento vegetal. Fundamentos teóricos sobre segregação, recombinação e ligação são considerados e relacionados à construção de mapas genéticos com marcas moleculares. Apresentam-se informações sobre tipos de marcadores, populações de mapeamento, cálculo da frequência de recombinação, distorções da segregação, estabelecimento dos grupos de ligação e da ordenação dos marcadores. Discute-se, também, o uso de mapas de ligação em programas de seleção assistida por marcadores, na clonagem de genes e em estudos sobre sintenia.

Palavras-chave: mapas de ligação, marcadores moleculares, QTL, mapeamento comparativo.

ABSTRACT

GENETIC MAPS IN PLANTS

In addition to genome projects and karyotype analysis by *in situ* hybridization techniques, a major advance in plant genome analysis came from the development of genetic maps based on molecular markers (Figure 1). This review clarifies the basic premises used for genetic mapping and its main applications, especially in plant breeding. The theories of segregation, recombination and linkage are considered and related to the construction of genetic maps based on molecular markers. Information about marker types, population mapping, calculation of the recombination rate, segregation distortion, linkage groups and genetic order determination is presented. Exploitation of linkage mapping for marker assisted selection, gene cloning and synteny comparisons is discussed.

Key words: linkage maps, molecular marker, QTL, comparative mapping.

1. INTRODUÇÃO

Com a redescoberta dos trabalhos de Mendel, no final do século XIX, por Hugo de Vries, Carl Correns e Erick von Tschermak-Seysenegg, realizaram-se inúmeras pesquisas a fim de esclarecer os mecanismos básicos da herança. Esses trabalhos culminaram na formulação da Teoria Cromossômica da Herança por W. Sutton e T. Boveri, em 1902, que estabeleceram a relação entre cromossomos e fatores mendelianos (genes), segundo a qual os cromossomos são elementos celulares portadores dos genes.

No mesmo ano, W. Bateson, E.R. Saunders e R.C. Punnet demonstraram, em ervilha-doce, que as segregações dos caracteres cor da flor e formato do pólen não ocorriam de forma independente (LANDER e WEINBERG, 2000).

Da mesma forma, T.H. Morgan e colaboradores: C.B. Bridges, H.J. Muller e A.H. Sturtevant, em 1910, trabalhando com *Drosophila melanogaster*, observaram distorções da segunda lei de Mendel. Morgan sugeriu, então, que alguns genes estariam situados no mesmo cromossomo e que durante a meiose, ocasionalmente, ocorreriam, entre os homólogos,

⁽¹⁾ Recebido para publicação em 13 de julho de 2001 e aceito em 9 de maio de 2002.

⁽²⁾ Departamento de Genética, Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Caixa Postal 83, 13400-970 Piracicaba (SP). E-mail: mlcvieir@carpa.ciagri.usp.br

trocas de segmentos denominadas *crossing-over* ou permuta. A ocorrência de quiasma, que é a visualização citológica da permuta, dava suporte aos resultados genéticos encontrados por Morgan. Este autor também percebeu que a proporção da progênie recombinante variava bastante, e propôs que estas variações, de algum modo, refletiam a distância entre os genes no cromossomo. Em 1913, A.H. Sturtevant, interpretando dados oriundos da segregação de genes ligados, sugeriu o uso da porcentagem de recombinantes como indicador quantitativo da distância linear entre dois genes na construção de mapas genéticos. Os mapas mostravam que a posição dos genes correspondia à sua ordem linear nos cromossomos. Assim, o conceito de localização dos genes em uma ordem linear passou a ser incorporado à teoria cromossômica da herança (GARDNER e SNUSTAD, 1986; GRIFFITHS et al., 1998).

O mapeamento genético tornou-se, rapidamente, uma poderosa ferramenta para os geneticistas, embora a prova definitiva da associação entre mapas de ligação e cromossomos tenha vindo mais tarde, com os estudos realizados por B. McClintock em cromossomos de milho. CREIGHTON e McCLINTOCK (1931) demonstraram que o *crossing-over* é resultado de troca entre segmentos cromossômicos.

Os primeiros mapas genéticos eram fundamentados em marcadores morfológicos e citológicos. Estes se originaram de *estoques cromossômicos* com aberrações (como aneuploidias, translocações, deleções e inversões), principalmente nas culturas do milho, tomate e ervilha (COE et al., 1988; RICK e YODER, 1988).

Em espécies cujo cariótipo é bem conhecido, é possível atribuir genes a um determinado cromossomo, relacionando aberrações cromossômicas com alterações fenotípicas causadas por esses genes. Por exemplo, em trigo, fumo, tomate e milho analisaram-se características em linhagens portadoras de monossomias ou trissomias para cada um de seus cromossomos e, assim, um grande número de genes pôde ser localizado. Essas análises podem ser refinadas utilizando-se aberrações que afetam apenas uma parte do cromossomo, como a ausência de um dos braços ou a duplicação de um pequeno segmento (GUERRA, 1988).

Em 1928, J.W. Lesley cruzando plantas diplóides e triplóides obteve a primeira série de trissômicos em tomate. Com base na análise desses trissômicos, esse autor foi pioneiro em estudos de mapeamento genético em tomate. Ele atribuiu sete mutantes, quatro no cromossomo A (cromossomo 2) e três em diferentes cromossomos. Mais de 1.200 mutantes foram descritos em tomate, dos quais pelo menos 400 foram localizados nos respectivos cromossomos, e suas distâncias de ligação estabelecidas (QUIROS, 2001). A coleção de estoques cromossômicos de tomate é mantida pelo "National Plant Germplasm

System" e pelo "Tomato Genetics Stock Center", ambos nos Estados Unidos (RICK e YODER, 1988).

Apesar da importância dos estoques cromossômicos em estudos genéticos, estes são restritos às espécies de amplo conhecimento citológico (milho, trigo, tomate, ervilha, aveia). Embora o uso de estoques cromossômicos apresente limitações, pois não permitem estudos de caracteres cuja herança é quantitativa, continuam a ser usados em trabalhos envolvendo vegetais de acentuada importância econômica, como os cereais (SCANLON et al., 1994; SEO e PETERSON, 1995; CHENG et al., 2000; DUBCOVSKY et al., 2000; SEAH et al., 2000) e o tomate (FRARY et al., 1996) visando à atribuição de genes.

Caracteres de herança discreta, cujas classes são facilmente distinguíveis, são bons marcadores morfológicos ou fenotípicos. Características como nanismo, deficiência de clorofila, cor das pétalas e morfologia foliar são usadas como marcadores fenotípicos (COE et al., 1988; RICK e YODER, 1988). Esses caracteres são de avaliação simples e de baixo custo. Entretanto, a disponibilidade de marcadores morfológicos é essencialmente restrita a poucas espécies vegetais para as quais o número de informações genéticas é maior, por exemplo, o milho, o tomate e a ervilha (FERREIRA e GRATTAPAGLIA, 1996; LIU, 1998).

Os estudos realizados com marcadores morfológicos em muito contribuíram para a elucidação do fundamento teórico da análise de ligação gênica e para a construção das primeiras versões de mapas genéticos (KNAPP, 1991). Atualmente, têm sido úteis nos casos em que os grupos de ligação já estão associados a determinados cromossomos, podendo ser também utilizados para ancorar grupos de ligação construídos com base em marcadores moleculares a cromossomos específicos.

No início da década de 60, as isoenzimas foram descobertas (MARKERT e MÖLLER, 1959) e passaram a ser usadas como marcadores bioquímicos, permitindo a construção de mapas genéticos, potencialmente, em todas as espécies de plantas (GARVIN e WEEDEN, 1994). Os marcadores isoenzimáticos apresentam um controle genético normalmente simples, em que a ação alélica é, na maioria dos casos, codominante.

Considerando-se que o número total de locos isoenzimáticos detectado seja superior a 100 (MURPHY et al., 1990), são raros os trabalhos que utilizam mais de 30 locos. Esse número é muito baixo para fins de mapeamento, principalmente quando se pretende obter ampla cobertura do genoma (RICK e YODER, 1988; FERREIRA e GRATTAPAGLIA, 1996).

A partir da década de 80, com o advento dos marcadores de DNA, o mapeamento genético tornou-se, efetivamente, ilimitado a todas as

espécies (BOSTEIN et al., 1980). Mapas genéticos que levaram décadas para serem desenvolvidos em virtude da restrita disponibilidade de marcadores morfológicos e isoenzimáticos foram amplamente saturados com o uso de marcas moleculares. Assim ocorreu em milho (HELENTJARIS et al., 1986), arroz (MCCOUCH et al., 1988), trigo (CHAO et al., 1989), batata (GEBHARDT et al., 1989) e tomate (TANKSLEY et al., 1992). Várias espécies para as quais os estudos de herança são restritos, e tampouco foi descrita ligação gênica entre qualquer tipo de marcador, tiveram mapas genéticos rapidamente desenvolvidos, como o eucalipto (GRATTAPAGLIA e SEDEROFF, 1994), kiwi (TESTOLIN et al., 2001) e maracujá (CARNEIRO et al., 2002).

2. PREMISSAS BÁSICAS DA CONSTRUÇÃO DE MAPAS DE LIGAÇÃO

A grande disponibilidade de marcadores neutros, altamente polimórficos, cuja herança não sofre influência ambiental, aliada a procedimentos estatísticos complexos, tem permitido a construção de mapas de ligação para a maioria das espécies vegetais de interesse agrônomo, até mesmo para aquelas de longo ciclo de vida, como as florestais e as frutíferas. A construção de mapas de ligação tem como base a análise de segregação de centenas de marcadores e, por isso, é computadorizada. Programas como "Mapmaker" (LANDER et al., 1987), "Linkage 1" (SUITER et al., 1983), "Gmendel" (LIU e KNAPP, 1992) foram desenvolvidos para auxiliar na análise genética dos dados visando à construção de mapas genéticos e estão disponíveis na internet (<http://linkage.rockefeller.edu/>).

Premissas importantes devem ser consideradas na construção de um mapa genético, a saber:

2.1. Seleção da população apropriada

A seleção da população para mapeamento envolve a escolha de genitores e a determinação do tipo de cruzamento (F_2 , retrocruzamento, etc), sendo considerada uma etapa crítica para o sucesso da construção do mapa (STAUB et al., 1996). No entanto, independentemente dessa escolha, duas condições básicas devem ser atendidas: o máximo de polimorfismo entre os genitores, e que sejam produzidas gerações em que os locos estejam em desequilíbrio de ligação (PATERSON et al., 1991; TANKSLEY, 1993).

Os genitores escolhidos para gerar a população de mapeamento devem ser contrastantes (até mesmo quanto à expressão fenotípica para caracteres de interesse agrônomo), já que o número de marcas no mapa representa o número

de caracteres contrastantes nos genitores. Em casos cujo polimorfismo é baixo, pode-se realizar cruzamentos interespecíficos para obter maior número de locos informativos (TANKSLEY et al., 1992). No entanto, como salientam STAUB et al. (1996), a supressão da recombinação entre genomas distantes e a ocorrência de alelos deletérios podem provocar distorções das proporções mendelianas esperadas na descendência, quebrando as premissas assumidas na análise estatística, afetando o cálculo da distância e a ordem dos marcadores nos grupos de ligação.

As gerações F_2 ou aquelas oriundas de retrocruzamentos, cruzamentos entre heterozigotos e linhagens duplo-haplóides são as populações mais usadas na construção de mapas de ligação (TANKSLEY, 1993; LYNCH e WALSH, 1998). Conforme salienta COELHO (2000), determinados genótipos (indivíduos F_2) são portadores de dois gametas informativos, enquanto em outras populações (retrocruzamento), cada indivíduo carrega somente um gameta informativo, sendo o outro complemento não informativo. Assim, de modo aproximado, um indivíduo F_2 fornece duas vezes mais informações que um indivíduo oriundo de uma geração de retrocruzamento. Esta relação é aproximada porque nem todos os genótipos de F_2 são informativos (PATERSON et al., 1991; LIU, 1998).

A escolha do tipo de população mais apropriada ao mapeamento está relacionada com o tipo de marcador empregado (TANKSLEY et al., 1988a). O máximo de informação genética é alcançado quando se usa uma população F_2 e marcadores codominantes (STAUB et al., 1996).

O desequilíbrio de ligação entre os locos segregantes é atribuído à ligação física entre os locos, e originado do desequilíbrio gamético de ligação decorrente, por sua vez, da redução da frequência de recombinação entre genes situados em regiões próximas ao longo de determinado cromossomo (WEIR, 1996; COELHO, 2000). Cabe ressaltar que o desequilíbrio não deve ser causado por seleção ou deriva genética. O desequilíbrio de ligação, decorrente da ligação física entre os locos, atinge seu ponto máximo nas populações derivadas de cruzamentos controlados e, como consequência, a capacidade de detectar a ligação também é máxima (TANKSLEY, 1993; LIU, 1998).

O número de genótipos identificados estabelece o nível máximo de resolução que pode ser alcançado com número ilimitado de marcadores no mapa genético. Assim, frequências de recombinação abaixo de 10% apenas podem ser obtidas pela avaliação de mais de dez gametas passíveis de recombinação e, portanto, informativos (COELHO, 2000).

2.2. Tipo de marcador genético

Os marcadores moleculares são um recurso muito útil em análises genéticas. Há grande quantidade de mapas genéticos disponíveis, tanto para espécies animais como vegetais, e um número crescente de genes têm sido detectados com auxílio desses mapas. Vários tipos de marcadores podem ser utilizados na construção de mapas (Quadro 1), diferindo em alguns aspectos: número de locos que pode ser detectado, grau de polimorfismo entre e dentro de acessos e características de dominância (MALIEPAARD et al., 1997; JOSHI et al., 1999). COELHO (2000) ressalta a importância de se considerar o conteúdo informativo dos marcadores de natureza codominante (RFLPs e SSRs) em relação aos marcadores dominantes (RAPDs e AFLPs), além de aspectos como custo, tempo necessário para a realização das avaliações, e dificuldades práticas inerentes à execução de cada técnica.

A influência do número de marcadores genéticos está relacionada até quanto a resolução pré-estabelecida pelo número de genótipos identificados será, de fato, satisfeita. Tanto o grau de cobertura do genoma como a densidade do mapa a serem construídos são funções diretas do número de marcadores (COELHO, 2000). LIU (1998) e LYNCH e WALSH (1998) apresentam procedimentos estatísticos para estimar o número adequado de marcadores na construção de mapas genéticos.

2.3. Verificação da distorção da segregação mendeliana

Como mencionado, a construção de um mapa genético é feita com dados oriundos de populações segregantes, ou seja, retrocruzamentos, geração F_2 , entre outras. Espera-se um padrão mendeliano típico para cada uma delas, que é pressuposto nos modelos probabilísticos (BEARZOTI, 2000). A ausência do padrão mendeliano é chamada distorção da segregação mendeliana, sendo causada, por exemplo,

pelo efeito de amostragem ou viabilidade diferencial de gametas. O critério mais imediato de verificação da distorção mendeliana é o teste de aderência (χ^2), cuja hipótese H_0 não admite a ocorrência de distorções (FREGENE et al., 1997; LESPINASSE et al., 2000).

Segundo LIU (1998), os locos que apresentam distorção mendeliana afetam os testes estatísticos usados para detectar a ligação podendo gerar falsos positivos. É recomendado, preferencialmente, o descarte dos locos que apresentam distorções da segregação mendeliana para não comprometer a qualidade do mapa (YAO et al., 1999; BEARZOTI, 2000). Na construção do mapa de ligação de *Cryptomeria japonica*, 50% dos marcadores apresentaram desvios na segregação esperada, não sendo utilizados nas análises de ligação. Os autores atribuíram esse resultado à sobreposição de fragmentos de AFLP, de tamanhos bastante semelhantes, o que provavelmente dificultou a sua correta identificação (NIKAIDO et al., 2000).

No entanto, há relatos do uso de marcadores com distorções mendelianas na construção de mapas de ligação. Mecanismos de seleção pré- ou pós-zigóticos explicam, mormente, tais desvios (KUBISIAK et al., 1995; CONNER et al., 1997; HEUSDEN et al., 2000).

2.4. Cálculo da frequência de recombinação entre pares de marcadores e funções de mapeamento

A mensuração da frequência de gametas recombinantes entre dois locos constitui o procedimento a partir do qual se realiza o mapeamento (STAUB et al., 1996). Diversos tratamentos estatísticos podem ser usados para estimar a frequência de recombinantes, entre eles, o método dos momentos, o método dos quadrados mínimos e o da máxima verossimilhança. De modo geral, pelo método dos momentos, obtêm-se os parâmetros igualando-se os valores observados às suas expectativas, em termos de esperanças matemáticas. Pelo método dos quadrados mínimos,

Quadro 1. Comparação entre os principais tipos de marcadores moleculares utilizados na construção de mapas de ligação em plantas

Características	Isoenzimas	RFLPs	RAPDs	SSRs	AFLPs
Ação gênica	Codominante	Codominante	Dominante	Dodominante	Dominante
Número de alelos/loco	2 a 4	Multialélico	2	Multialélico	2
Disponibilidade de marcadores no genoma	20 a 50	Ilimitada	Ilimitada	Ilimitada	Ilimitada
Nível de polimorfismo	Baixo	Alto	Alto	Muito alto	Muito alto
Necessidade de procedimentos prévios	Não há	Construção de bibliotecas genômicas ou de cDNAs	Não há	Desenho de primers	Não há
Concentração de DNA	-	10 mg	25 ng	50 ng	500 ng
Marcação radioativa	-	Sim/Não	Não	Não	Sim/Não
Reprodutibilidade	Alta	Alta	Média	Alta	Média
Investimento	baixo	Alto	Médio	Alto	Médio

obtêm-se as estimativas para os parâmetros de modo que a soma de quadrados dos desvios entre os valores observados e os estimados seja mínima, enquanto pelo método de máxima verossimilhança obtêm-se as estimativas de modo que os valores observados representem o caso mais provável de ocorrência (WEIR, 1996; LIU, 1998).

Embora a frequência de recombinantes indique a medida de distância entre dois locos ao longo de determinado cromossomo, o ideal é que as distâncias usadas na construção de mapas genéticos sejam totalmente aditivas, de modo que, ao atribuir novos locos (marcadores) ao mapa, não seja necessário ajustar as distâncias já estabelecidas. Entretanto, as frequências de recombinação não são aditivas; são, portanto, inadequadas como medida direta de distâncias entre dois locos a serem mapeados (STAUB et al., 1996; LYNCH e WALSH, 1998).

Para ilustrar essa questão, LYNCH e WALSH (1998) tomam como exemplo três locos que supostamente estão na ordem A, B e C, cujas probabilidades de ocorrência de recombinação são de r_{AB} , r_{AC} e r_{BC} . Assim, a probabilidade de ocorrência de recombinação corresponde à probabilidade de ocorrência de um número ímpar de *crossing-overs* entre dois locos, enquanto $(1-r)$ indica a probabilidade de ocorrência de um número par (inclusive zero) de *crossing-overs* entre os dois locos. Dessa forma, existem duas maneiras de ocorrer um número ímpar de permutações no fragmento A-C: quando se tem um número ímpar de *crossing-overs* no fragmento A-B, e par no fragmento B-C, ou quando se tem um número par em A-B, e ímpar em B-C. Admitindo-se as ocorrências de permuta como eventos independentes entre si (ausência de interferência) e utilizando as regras básicas de probabilidade, tem-se as equações a seguir:

$$r_{AC} = r_{AB} (1 - r_{BC}) + (1 - r_{AB}) r_{BC} \quad (1)$$

$$r_{AC} = r_{AB} + r_{BC} - 2 r_{AB} r_{BC} \quad (2)$$

A expressão (2) é conhecida como fórmula de Trow, e demonstra o caráter não aditivo das frequências de recombinação. Obviamente, em casos cujo segmento cromossômico é pequeno, as frequências de recombinação podem ser consideradas aditivas, desde que os valores de r_1 e r_2 sejam baixos, de modo que o produto $r_1 r_2$ se torne desprezível (LYNCH e WALSH, 1998).

Segundo BAILEY (1961), uma unidade de medida de distância que é, ao mesmo tempo, função da frequência de recombinantes e possui caráter aditivo, equivale ao número de eventos de recombinação que ocorre entre dois locos (marcadores) em uma geração. De acordo com WEIR (1996), uma unidade de medida de distância genética de mapa é definida como igual à distância entre dois locos ao longo da qual se espera

que ocorra uma permuta por gameta, por geração. Essa medida é denominada Morgan (M), sendo expressa em centiMorgans (cM).

Para estabelecer a relação entre as distâncias genéticas entre dois locos no mapa e a frequência de gametas recombinantes criaram-se as funções de mapeamento (LYNCH e WALSH, 1998). Um grande número de funções de mapeamento tem sido desenvolvido (HALDANE, 1919; KOSAMBI, 1944; CARTER e FALCONER, 1951; RAO et al., 1977; FELSENSTEIN, 1979; KARLIN, 1982), sendo as funções de Haldane e Kosambi as mais utilizadas no mapeamento genético de vegetais (STAUB et al., 1996).

A função de Haldane admite que a ocorrência de permuta se dá de modo independente (ausência de interferência), seguindo uma distribuição de Poisson. Ao contrário, a função de Kosambi admite a ocorrência de permutas próximas como eventos não independentes (presença de interferência). Assim, esta função assume a interferência completa entre regiões arbitrariamente próximas, sendo decrescente para locos mais distantes, e igual a zero para locos independentes (WEIR, 1996; LIU, 1998).

Outro aspecto importante, é que não existe correlação entre distância física (número de pares de bases) e distância de mapa, ou seja, existem regiões cromossômicas de apenas algumas dezenas de milhares de pares de bases (distância física pequena) em que a probabilidade de recombinação é alta (regiões-alvo de recombinação - *hotspots*). Ao mesmo tempo, há regiões de alguns milhões de pares de bases (distância física grande) onde a recombinação é praticamente suprimida e a distância entre os locos, mínima - regiões centroméricas e teloméricas (TANKSLEY et al., 1992; FERREIRA e GRATTAPAGLIA, 1996; JIANG et al., 1997; SALIBA-COLOMBANI et al., 2000).

2.5. Estabelecimento dos grupos de ligação

Do ponto de vista biológico, os grupos de ligação são definidos como grupos de genes (marcadores) cujos locos estão localizados no mesmo cromossomo. Estatisticamente, grupo de ligação pode ser definido como um conjunto de locos herdados juntos (LIU, 1998).

Para demonstrar que dois locos estão ligados, podem-se utilizar as razões de verossimilhança em que no numerador é disposta a função de verossimilhança considerando $r = 0,5$ (ausência de ligação) e, no denominador, a função que possui a estimativa de máxima verossimilhança de r . Para facilitar a interpretação, essa razão é colocada no logaritmo com base 10, sendo este termo definido como *LOD score*. Por exemplo, um LOD igual a 3 indica que a ocorrência de ligação é mil vezes mais provável que a de segregação independente (BEARZOTI, 2000). O estabelecimento dos grupos de ligação deve ser feito com o *LOD score* variando de 4

a 6. Quando existe um grande número de marcadores, LODs com valores baixos (menores que 4) podem ocasionar ligações espúrias, enquanto o aumento nos valores de LOD, acima de 6, poderia resultar na fragmentação dos grupos de ligação.

Para separar os grupos de ligação, realizaram-se testes considerando duas marcas (locos) de cada vez, admitindo a hipótese de que a frequência entre ambas é igual a 0,5 (H_0 : segregação independente). Outro procedimento estatístico usado para detectar a ligação (desequilíbrio de ligação), além da máxima verossimilhança, é o teste de aderência (χ^2) (BEARZOTI, 2000).

O teste do qui-quadrado (χ^2) baseia-se na comparação das frequências observadas das diversas classes genotípicas com aquelas esperadas sob condição de segregação independente. Esse teste permite a estimativa da probabilidade de ocorrência dos desvios, dada a condição de independência. Normalmente, utiliza-se como ponto crítico para a rejeição da hipótese de independência (H_0) probabilidade de 5%. Isso significa rejeitar H_0 e, portanto, assumir que os locos não estão segregando independentemente, quando a probabilidade de ocorrência dos desvios encontrados, dada a condição de independência, for menor que 5% (LIU, 1998; BEARZOTI, 2000).

2.6. Determinação da ordem dos marcadores nos grupos de ligação

Uma vez conhecidas as frações de recombinação entre cada par de marcadores, e já discriminados os grupos de ligação, deve-se determinar a melhor ordem para os marcadores dentro de cada grupo. Historicamente, a ordenação de locos tem sido um processo que visa minimizar o número de *crossing-overs*.

WEIR (1996) sugere a ordenação de três locos, A, B, e C, feita pela soma dos coeficientes de recombinação adjacentes (*sar* - *sum of adjacent recombination coefficients*). Esse processo propõe que a ordem correta entre os locos seja aquela que possibilite o mínimo valor de *sar*. O valor de *sar* sob a ordem incorreta será sempre maior que a correta. Esse método pode ser empregado para qualquer número de locos. No entanto, o uso dos valores obtidos pelo critério *sar* para determinar a ordem correta pode não ser adequado quando a análise envolve um número elevado de locos. O número de ordens possíveis para m locos é igual a $m!/2$. Assim, se tivermos $m = 30$, serão 13.263×10^{32} combinações possíveis. Por isso, esse procedimento é considerado impraticável, tornando-se necessários outros métodos de ordenação (WEIR, 1996; LIU, 1998).

Existem vários métodos para esse propósito: a) delimitação rápida em cadeia (DOERGE, 1996); b) seriação (BUETOW e CHAKRAVARTI, 1987a,b); c) "simulated annealing" (KIRKPATRICK et al., 1983) e d) ramos e conexões (THOMPSON, 1987).

Um dos enfoques alternativos para se trabalhar com muitos locos é o da delimitação rápida em cadeia (DRC, DOERGE, 1996). Ilustra essa situação, a clássica parábola do "caixeiro viajante" que precisa percorrer várias cidades e deseja minimizar a distância. O problema aumenta sua complexidade à medida que aumenta o número de cidades envolvidas. Por exemplo, se o viajante fosse percorrer 40 cidades, seria possível encontrar $8,16 \times 10^{47}$ rotas (LIU, 1998). A DRC consiste, com base na matriz de recombinações de todos os pares de marcas, na obtenção de uma ordem preliminar para os locos. Em seguida, tentam-se inversões sucessivas em triplas marcas, a fim de minimizar a soma das recombinações adjacentes (*sar*) (BEARZOTI, 2000).

A seriação é um método simples, no qual um conjunto de regras é proposto com base nas frações de recombinação entre dois locos. A principal desvantagem desse método é não assegurar a melhor ordenação, sobretudo nos casos em que as matrizes de frações de recombinantes não são monotônicas (LIU, 1998).

O "simulated annealing" é um dos métodos que analisa o maior número de ordens possíveis. Esse método foi construído com base na termodinâmica da cristalização dos líquidos. De acordo com essa linguagem, quando um líquido é resfriado muito lentamente, ele se cristaliza em um estado mínimo de energia (E). Assim, a ordem dos locos, considerada como E_1 , apresenta um determinado valor de *sar*. À medida que esta ordem sofre permutações, uma nova ordem (energia) é alcançada (E_2), com um novo valor de *sar*. Se $E_2 < E_1$, então a nova ordem será aceita. No entanto, se $E_2 > E_1$, a nova ordem é aceita com dada probabilidade, P. Ao final de sucessivas permutações, a melhor ordem pode ser dada quando o estado mínimo de energia é alcançado com determinada probabilidade. No entanto, não existe garantia de que a melhor ordem será encontrada (WEIR, 1996).

O método de ramos e conexões é firmado na estrutura das árvores. Os ramos da árvore correspondem aos locos, e as conexões são pontos alternativos para a inserção de novos locos. Supondo os dois primeiros locos como C e D, o próximo loco, B, pode ser inserido em uma das três posições, correspondendo às seguintes ordens BCD, CBD e CDB. Se um novo loco fosse inserido, loco F, quatro diferentes posições para cada uma das três ordens anteriormente descritas seriam possíveis (FBCD, BFCD, BCFD, BCDF, FCBD, CFBD, CBF, CDBF, FCDB, CFDB, CDFB e CDBF). O número de recombinantes é calculado para cada ramo. Se o número de recombinantes para ordem BCD é menor que aquele encontrado para a ordem CBD ou CDB, por exemplo, então, não é necessário considerar as demais outras ordens subsequentes, sendo o ramo descartado da análise. Esse método garante encontrar a melhor ordem dos locos, mesmo que não se tenha avaliado todas as ordens possíveis (WEIR, 1996).

3. APLICAÇÕES

O desenvolvimento de mapas genéticos é considerado uma das aplicações de maior impacto da tecnologia de marcadores moleculares. Presta-se às análises filogenéticas, e potencialmente, ao melhoramento de plantas. Diante desses avanços, procedimentos como a localização e o mapeamento de QTL's (*Quantitative Trait Loci*), estudos de sintenia e clonagem de genes com base em mapas genéticos passaram a representar complementos importantes a serem considerados em programas de melhoramento das mais variadas espécies vegetais (LEE, 1995).

Uma das aplicações mais importantes dos mapas genéticos é a localização de genes que controlam características de importância econômica, como produção de grãos, altura da planta, teor de proteína, resistência a doenças, que resultam da ação cumulativa de um conjunto de genes. Essas características são denominadas poligênicas, quantitativas ou de herança complexa, e os locos que as controlam são chamados de QTL's (TANKSLEY, 1993).

A ligação entre um marcador genético e um QTL foi primeiramente demonstrada por SAX (1923). No entanto, os fundamentos da teoria do mapeamento de QTL's foram entendidos a partir do trabalho de THODAY (1961). Esse autor sugeriu que se um oligogene (herança complexa) estiver ligado a um monogene (herança simples), os efeitos fenotípicos do oligogene podem ser indiretamente estudados com base nos efeitos do gene vizinho. Com o advento dos marcadores moleculares, o número de associações entre esses marcadores e caracteres de herança oligênica de importância econômica foi ampliado significativamente. O grande número de marcadores distribuídos pelo genoma permite uma amostragem mais completa na busca de ligação gênica com o loco de interesse (TANKSLEY, 1993).

O mapeamento de QTL's possibilita mensurar o número de locos quantitativos envolvidos na herança complexa, bem como suas localizações cromossômicas, modo de ação gênica (aditividade, dominância, heterose e epistasia), além de possibilitar a decomposição da interação genótipos por ambientes ao nível de cada QTL (FERREIRA e GRATTAPAGLIA, 1996; AUSTIN e LEE, 1998). A capacidade de detecção de um QTL é função da magnitude do seu efeito sobre a característica, do tamanho da população segregante avaliada, da frequência de recombinação entre o marcador e o QTL, bem como da herdabilidade da característica (PATERSON et al., 1991; TANKSLEY, 1993).

Uma vez detectados os marcadores ligados a uma determinada característica de interesse, é possível selecionar os indivíduos com base no fenótipo do marcador, sem que haja a necessidade de avaliar o

fenótipo da característica. Essa abordagem é conhecida como seleção assistida por marcadores (SAM). É importante, principalmente, nos casos em que a característica de interesse é de avaliação difícil e cara, nos estudos de espécies perenes de ciclo longo, e para características fenotípicas de herdabilidade baixa, o que pode significar redução no tempo e nos custos de programas de melhoramento (SOUZA, 2001).

Os exemplos de SAM ainda são escassos na literatura, em especial porque as informações sobre programas de melhoramento de empresas privadas, que são hoje, possivelmente, as principais executoras da SAM, não estão disponíveis (MILLACH, 1998).

No entanto, alguns exemplos de SAM têm sido descritos: transferência de genes favoráveis de acessos silvestres para a espécie cultivada visando ao aumento do tamanho do fruto em tomateiro (TANKSLEY et al., 1996) e dois locos de resistência à ferrugem em cevada (TOOJINDA et al., 1998).

Dentre os métodos propostos para a localização de QTL's encontram-se os modelos lineares (ANOVA e regressão linear), mapeamento por intervalo ("interval mapping" ou "flanking-marker analysis"), mapeamento por intervalo composto ("composite interval mapping") (ZENG, 1993; 1994; JANSSEN, 1993) e mapeamento multiponto ("multipoint mapping") (KEARSEY e HYNE, 1994).

Os procedimentos mais simples para detecção de QTL's são os modelos lineares que analisam, separadamente, a diferença entre os valores fenotípicos médios para cada marcador (EDWARDS et al., 1987), não sendo necessária a construção de mapas de ligação. Como limitação, há o fato de quanto maior for a distância entre o QTL e o marcador, menor será a probabilidade de detectá-lo em vista da possibilidade de recombinação entre o gene marcador e o QTL. Além disso, a magnitude do QTL, pelo mesmo motivo, é normalmente subestimada. Outro inconveniente é que não é possível determinar a posição do QTL no genoma (LYNCH e WALSH, 1998).

O método de mapeamento com base em intervalo ("interval mapping") foi proposto por LANDER e BOTSTEIN (1989) e tem sido alternativa para aumentar o poder de detecção das associações, permitindo estimar o efeito e a posição dos QTL's. Esse método tem como base informações de segregação de pares de marcadores adjacentes como unidades de análise, utilizando o método de máxima verossimilhança para estimar a frequência de recombinantes e a magnitude do efeito do QTL no intervalo entre dois marcadores ligados (LYNCH e WALSH, 1998). Uma desvantagem atribuída ao mapeamento fundamentado em intervalo é a de que os outros QTL's fora do intervalo são ignorados, o que pode resultar em duas consequências principais. Toda a variação genética devida a esses outros QTL's é residual, o que pode diminuir a precisão das

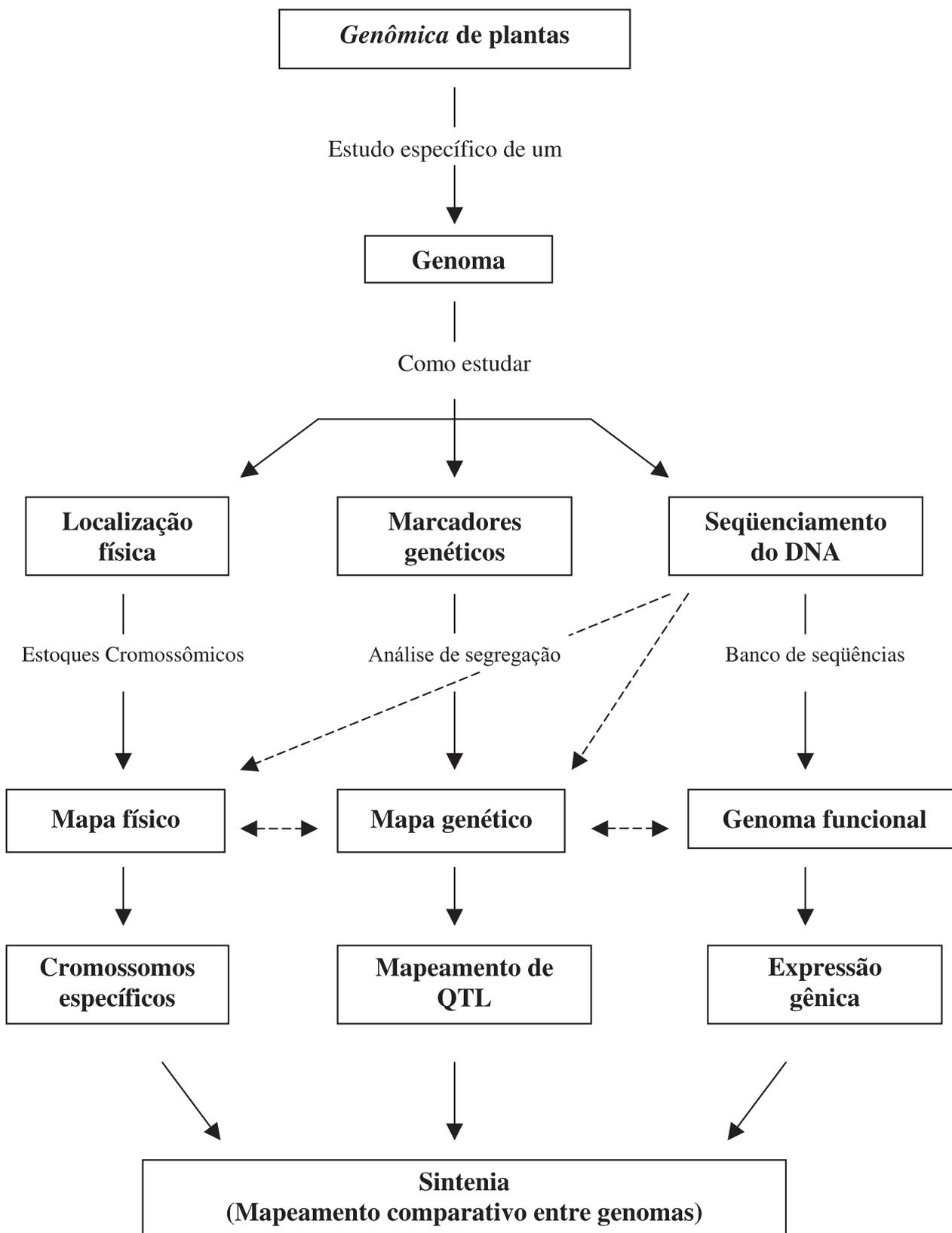


Figura 1. Diferentes abordagens para o estudo do genoma em plantas e suas inter-relações.

estimativas e o poder dos testes. Além disso, eventuais QTL's que estejam ligados ao intervalo em questão interferem no processo de estimação, levando, em casos extremos, a declarar a existência de um QTL no intervalo quando na realidade não há nenhum (BEARZOTI, 2000).

A fim de resolver esses problemas, JANSEN (1993) e ZENG (1993, 1994) propuseram o mapeamento por intervalo composto, método em que se consideram os QTL's fora do intervalo em questão por um modelo de regressão linear múltipla. Além de proporcionar maior resolução no mapeamento, permite o uso de dados obtidos em vários ambientes, para avaliar os efeitos da interação QTL's *vs.* Ambientes.

Outra abordagem mais recente é o mapeamento por intervalo múltiplo, proposto por KAO et al. (1999), que trouxe considerável aprimoramento no mapeamento de QTL's, pois incorpora parâmetros de *epistasia* ao modelo. Em presença de epistasia, esse método apresenta algumas vantagens potenciais como maior eficiência e precisão na identificação de QTL's; seus efeitos são estimados sem viés e, ao contrário de outros modelos, pode aumentar a eficiência da seleção assistida, considerando os efeitos epistáticos (BEARZOTI, 2000).

O mapeamento comparativo (*comparative mapping* ou *synteny mapping*) constitui-se em outra importante utilização dos mapas. A comparação das estruturas genômicas de diferentes espécies, do ponto de vista de homologia de genes, conservação de distâncias e da ordem de ligação nos cromossomos contribui para o entendimento sobre a evolução dos genomas (TANKSLEY et al., 1988b; KIANIAN e QUIROS, 1992; AHN e TANKSLEY, 1993; KELLER e FEUILLET, 2000). O mapeamento comparativo é também estratégia para obtenção de mapa único de referência para a maioria das espécies vegetais cultivadas (MOORE et al., 1995; SEWELL et al., 1999), pelo menos ao nível de famílias taxonômicas.

A clonagem de genes com base em mapas genéticos (*map-based cloning* e *positional cloning*) constitui-se na mais audaciosa aplicação dos mapas genéticos. Nessa abordagem, o efeito fenotípico do gene é verificado anteriormente avaliando-se a característica de interesse na progênie, e sua posição cromossômica aproximada é determinada pelo mapeamento. Esse mapeamento detalhado, de alta resolução, consiste em identificar marcadores próximos a um gene, sendo essa resolução dependente do número de marcadores e da disponibilidade de grande número de indivíduos na progênie, o que aumenta a probabilidade de se encontrar eventos de recombinação no segmento de interesse (FERREIRA e GRATTAPAGLIA, 1996; LIU, 1998).

Com o desenvolvimento dos chamados *projetos genoma* (Figura 1), é possível associar marcas mapeadas a genes, simplesmente pelo seqüen-

ciamento dos fragmentos mapeados, os quais servem como ponto de partida para a comparação com seqüências depositadas nos bancos, facilitando a identificação e clonagem desses genes. Além disso, os fragmentos mapeados em cada um dos grupos de ligação podem servir como sondas em ensaios de hibridação *in situ* contra metáfases mitóticas, possibilitando a atribuição dos grupos de ligação aos cromossomos das espécies em estudo.

AGRADECIMENTOS

Os autores agradecem ao Professor Giancarlo Conde Xavier de Oliveira, do Departamento de Genética da ESALQ/USP, pela leitura do trabalho e sugestões apresentadas.

REFERÊNCIAS BIBLIOGRÁFICAS

- AHN, S.; TANKSLEY, S.D. Comparative linkage maps of the rice and maize genomes. *Proceedings of the National Academic of Sciences of the United States of America*, Washington, v.90, n.17, p.7980-7984, 1993.
- AUSTIN, D.F.; LEE, M. Detection of quantitative trait loci for grain yield and yield components in maize across generations in stress and nonstress environments. *Crop Science*, Madison, v.38, p.1296-1308, 1998.
- BAILEY, N.T.J. *Introduction to the mathematical theory of genetic linkage*. London: Oxford University Press, 1961. 421p.
- BEARZOTI, E. Mapeamento de QTL. In: PINHEIRO, J.B., CARNEIRO, I.F (Eds.). *Análise de QTL no melhoramento de plantas*. Goiânia: FUNAPE, 2000. p.63-223.
- BOSTEIN, D.; WHITE, R.L.; SKOLNICK, M.; DAVIS, R.W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal Human Genetics*, Chicago, v.32, n.3, p.314-331, 1980.
- BUETOW, K. H.; CHAKRAVARTI, A. Multipoint gene mapping using seriation. I. General methods. *American Journal Human Genetics*, Chicago, v.41, n.2, p.180-188, 1987a.
- BUETOW, K.H.; CHAKRAVARTI, A. Multipoint gene mapping using seriation. I. Analysis of simulated and empirical data. *American Journal Human Genetics*, Chicago, v.41, n.2, p.189-201, 1987b.
- CARNEIRO, M.S.; CAMARGO, L.E.A.; COELHO, A.S.G.; VENCOSKY, R.; LEITE-JÚNIOR, R.P.; STENZEL, N.M.C.; VIEIRA, M.L.C. RAPD-based genetic linkage maps of yellow passion fruit (*Passiflora edulis* Sims. f. *flavicarpa* Deg.). *Genome*, Ottawa, v.45, p.670-678, 2002.
- CARTER, T.C.; FALCONER, S.D. Stocks for detecting linkage in the mouse and the theory of their design. *Journal of Genetics*, Bangalore, v.50, n.2, p.307-323, 1951.

- CHAO, S.; SHARP, P.J.; WORLAND, A.J.; WARHAM, E.J.; KOEBNER, R.M.D.; GALE, M.D. RFLP-based genetic maps of wheat homoeologous group 7 chromosomes. *Theoretical Applied Genetics*, New York v.78, p.495-504, 1989.
- CHENG, Z.K.; YAN, H.H.; YU, H.X.; QIAN, Q.; YI, C.D.; GU, M.H.; ZHU, L.H. Fast assignment of DNA sequences to individual chromosome arms based on dosage effects from a set of rice telotrisomics. *Acta Botanica Sinica*, Beijing, v.42, n.7, p.708-711, 2000.
- COE, H.E.; NEUFFER, M.G.; HOISINGTON, D.A. The genetics of corn. In: SPRAGUE, G. F.; DUDLEY, J.W. (Eds.). *Corn and corn improvement*, 1988. p.81-237.
- COELHO, A.S.G. Considerações gerais sobre a análise de QTL's. In: PINHEIRO, J.B.; CARNEIRO, I.F. (Eds.). *Análise de QTL no melhoramento de plantas*. Goiânia: FUNAPE, 2000. p.1-36.
- CONNER, P.J.; BROWN, S.K.; WEEDEN, N.F. Randomly amplified polymorphic DNA-based genetic linkage maps of three apple cultivars. *Journal of the American Society for Horticultural Science*, Alexandria, v.122, n.3, p.350-359, 1997.
- CREIGHTON, H.B.; McCLINTOCK, B. A correlation of cytological and genetical crossing-over in *Zea mays*. *Proceedings of the National Academic of Sciences of the United States of America*, Washington, v.17, p.492-497, 1931.
- DOERGE, R. Constructing genetic maps by rapid chain delineation. *Journal of Quantitative Trait Loci*, v.2, p.121-132, 1996.
- DUBCOVSKY, J.; TRANQUILLI G.; KHAN, I.A.; PFLUGER, L.A.; SUAREZ, E; ROUSSET, M.; DVORAK, J. Comparisons of recombination frequencies in hybrids involving telocentric and bibrachial wheat chromosomes. *Theoretical and Applied Genetics*, New York, v.100, n.2, p.308-314, 2000.
- EDWARDS, M.D.; STUBER, C.W.; WENDEL, J.F. Molecular-marker-facilitated investigations of quantitative trait loci in maize. Numbers, genomic distribution and types of gene action. *Genetics*, Baltimore, v.116, p.113-125, 1987.
- FELSENSTEIN, J. A mathematically tractable family of genetic mapping functions with different amount of interference. *Genetics*, Baltimore, v.91, n.4, p.769-775, 1979.
- FERREIRA, M.E.; GRATTAPAGLIA, D. *Introdução ao uso de marcadores moleculares em análise genética*. Brasília: EMBRAPA/CENARGEM, 1996. 220p.
- FRARY, A.; PRESTING, G.G.; TANKSLEY, S.D. Molecular mapping of the centromeres of tomato chromosomes 7 and 9. *Molecular General Genetics*, New York, v.250, n.3, p.295-304, 1996.
- FREGENE, M.; ANGEL, F.; GOMEZ, R.; RODRIGUEZ, F.; CHAVARRIAGA, P; ROCA, W.; TOHME, J.; BONIERBALE, M. A molecular genetic map of cassava (*Manihot esculenta* Crantz). *Theoretical and Applied Genetics*, New York, v.95, n.3, p.431-441, 1997.
- GARDNER, E. J.; SNUSTAD, P. *Genética*. Rio de Janeiro: Interamericana, 1986. 497p.
- GARVIN, D.F.; WEEDEN, N.F. Genetic-linkage between isozyme, morphological, and DNA markers in tepary bean. *Journal of Heredity*, Cary, v.85, n.4, p.273-278, 1994.
- GEBHARDT, C.; RITTER, E.; DEBENER, T.; SCHACHTSCHABEL, U.; WALKEMEIER, B.; UHRIG, H.; SALAMINI, F. RFLP analysis and linkage mapping in *Solanum tuberosum*. *Theoretical and Applied Genetics*, New York, v.78, p.65-75, 1989.
- GRATTAPAGLIA, D., SEDEROFF, R. Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-tescross mapping strategy and RAPD markers. *Genetics*, Baltimore, v.137, p.1121-1137, 1994.
- GRIFFITHS, A.J.F.; MILLER, J.H.; SUZUKI, D.T.; LEWONTIN, R.C.; GELBART, W.M. *Introdução à genética*. Rio de Janeiro: Guanabara, 1998. 856p.
- GUERRA, M.S. *Introdução à citogenética geral*. Rio de Janeiro: Guanabara, 1988. 140p.
- HALDANE, J.B.S. The combination of linkage values, and the calculation of distance between the loci of linked factors. *Journal of Genetics*, Bangarole, v.8, p.299-309, 1919.
- HELENTJARIS, T.; SLOCUM, M.; WRIGHT, S.; SHAEFER, A.; NIENHUIS, J. Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. *Theoretical Applied Genetics*, New York, v.72, p.761-769, 1986.
- HEUSDEN, A.W. van; OOIJEN, J.W. van; VRIELINK-VAN GINKEL, R.; VERBEEK W.H.J.; WIETSMA, W.A.; KIK, C. A genetic map of an interspecific cross in *Allium* based on amplified fragment length polymorphism (AFLP (TM)) markers. *Theoretical and Applied Genetics*, New York, v.100, n.1, p.118-126, 2000.
- JANSEN, R.C. Interval mapping of multiple quantitative trait loci. *Genetics*, Baltimore, v.135, n.1, p.205-211, 1993.
- JIANG, C.; LEWIS, M.E.; SINK, K.C. Combined RAPD and RFLP molecular linkage map of asparagus. *Genome*, Ottawa, v.40, n.1, p.69-76, 1997.
- JOSHI, S.P.; RANJEKAR, P.K.; GUPTA, V.S. Molecular markers in plant genome analysis. *Current Science*, Bangarole, v.77, n.2, p.230-240, 1999.
- KAO, C. H.; ZENG, Z.B.; TEASDALE, R.D. Multiple interval mapping for quantitative trait loci. *Genetics*, Baltimore, v.152, p.1203-1216, 1999.
- KARLIN, S. Theoretical aspects of genetic map functions in recombination processes. In: CHAKRAVARTI, A. (Ed.). *Human population genetics: the Pittsburgh symposium*. New York: van Nostrand Reinhold, 1982. p.245-286.

- KEARSEY, M.J.; HYNE, V. QTL analysis: a simple marker regression approach. *Theoretical and Applied Genetics*, New York, v.3, n.6, p.698-702, 1994.
- KELLER, B.; FEUILLET, C. Colinearity and gene density in grass genomes. *Trends in Plant Science*, London, v.5, n.6, p.246-251, 2000.
- KIANIAN, S.F.; QUIROS, C.F.; Generation of a *Brassica oleracea* composite RFLP map: linkage arrangements among various populations and evolutionary implications. *Theoretical and Applied Genetics*, New York, v.84, n.5/6, p.544-554, 1992.
- KIRKPATRICK, S.; GELATT, C.D.; VECCHI, M.P. Optimization by simulated annealing. *Science*, Washington, v.220, n.4598, p.671-680, 1983.
- KNAPP, S.J. Mapping quantitative trait loci using molecular makers: multilocus estimators of backcross, recombinant inbreed, and doubled haploid parameters. *Theoretical and Applied Genetic*, New York, v.81, p.333-338, 1991.
- KOSAMBI, D.D. The estimation of map distance from recombination values. *Annuaire of Eugenetics.*, v.12, p.172-175, 1944.
- KUBISIAK, T.C.; NELSON, C.D.; NANCE, W.L.; STINE, M. RAPD linkage mapping in a longleaf pine x slash pine F₁ family. *Theoretical and Applied Genetics*, New York, v.90, n.7/8, p.1119-1127, 1995.
- LANDER, E.S.; BOTSTEIN, D. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, Baltimore, v.121, n.1, p.185-199, 1989.
- LANDER, E.S.; GREEN, P.; ABRAHAMSON, J.; BARLOW, A.; DALY, M.J.; LINCOLN, E.E.; NEWBURG, L. Mapmaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics*, v.1, p.174-181, 1987.
- LANDER, E.S.; WEINBERG, R.A. Genomics: journey to the century of biology. *Science*, Washington, v.287, n.5459, p.1777-1782, 2000.
- LEE, M. DNA markers and plant breeding programs. *Advances in Agronomy*, San Diego, v.55, p.265-344, 1995.
- LESPINASSE, D.; RODIER-GOUD, M.; GRIVET, L.; LECONTE, A.; LEGNATE, H.; SEGUIN, M. A saturated genetic linkage map of rubber tree (*Hevea* spp.) based on RFLP, AFLP, microsatellite, and isozyme markers. *Theoretical and Applied Genetics*, New York, v.100, n.1, p.127-138, 2000.
- LIU, B.H. *Statistical genomics*. New York: CRC, 1998. 610p.
- LIU, B.H.; KNAPP, S.J. GMENDEL: a program for Mendelian segregation and linkage analysis of individual or multiple progeny populations using log-likelihood ratios. *Journal of Heredity*, Cary, v.81, n.5, p.407-418, 1992.
- LYNCH, M.; WALSH, B. *Genetics and analysis of quantitative traits*. Sunderland: Sinauer Associates, 1998. 978p.
- MALIEPAARD, C.; JANSEN, J.; van OOIJEN, J.W. Linkages analysis in a full-sibs family of an outbreeding plant species: overview and consequences for applications. *Genetics Research*, New York, v.70, p.237-250, 1997.
- MARKERT, C.L.; MOLLER, F. Multiple forms of enzyme: tissues, ontogenetic, and species specific patterns. *Proceedings of the National Academic of Sciences of the United States of America*, Washington, v.45, p.753-763, 1959.
- McCOUCH, S.R.; KOCHERT, G.; YU, Z.H.; WANG, Z.Y.; KHUSH, G.S.; COFFMAN, W.R.; TANKSLEY, S.D. Molecular mapping of rice chromosomes. *Theoretical Applied Genetics*, New York, v.76, p.815-829, 1988.
- MILACH, S.C.K. Marcadores de DNA: Aplicações no melhoramento de plantas. *Biotecnologia*, v.5, p.14-17, 1998.
- MOORE, G.; DEVOS, K.M.; WANG, Z.; GALE, M.D. Cereal Genome Evolution: Grasses, line up and from a circle. *Current Biology*, London, v.5, n.7, p.737-749, 1995.
- MURPHY, R.W.; SITES, J.W.; BUTH, D.G.; HAUFLE, C.H. Proteins I: isoenzyme electrophoresis. In: HILLIS, D.M.; MORITZ, C. (Eds.). Sunderland: Sinauer Associates, 1990. p.45-126.
- NIKAIIDO, A.M.; UJINO, T.; IWATA, H.; YOSHIMURA, K.; YOSHIMURA, H.; SUYAMA, Y.; MURAI, M.; NAGASAKA, K.; TSUMURA, Y. AFLP and CAPS: linkage maps of *Cryptomeria japonica*. *Theoretical and Applied Genetics*, New York, v.100, n.6, p.825-831, 2000.
- PATERSON, A.H.; TANKSLEY, S.D.; SORRELLS, M.E. DNA markers in plant improvement. *Advances in Agronomy*, San Diego, v.46, p.39-89, 1991.
- QUIROS, C.F. Solanacea: Tomato (*Lycopersicon esculentum*). Disponível em: <<http://veghome.ucdavis.edu/classes/vc221/tomato/tomsum.html>>. Acesso em 12 de julho de 2001.
- RAO, D.C.; MORTON, N.E.; LINDSTEN, J.; HULTEN, M.; YEE, S. A mapping function for man. *Humboldt Heredity*, Basel, v.27, n.2, p.99-104, 1977.
- RICK, C.M.; YODER, J.I. Classical and molecular genetics of tomato highlights and perspectives. *Annual Review of Genetics*, Palo Alto, v.22, p.281-300, 1988.
- SALIBA-COLOMBANI, V.; CAUSSE, M.; GERVAIS, L.; PHILOUZE, J. Efficiency of RFLP, RAPD, and AFLP markers for the construction of an intraspecific map of the tomato genome. *Genome*, Ottawa, v.43, n.1, p.29-40, 2000.
- SAX, K. The association of size differences with seed coat pattern and pigmentation in *Phaseolus vulgaris*. *Genetics*, Baltimore, v.8, p.552-560, 1923.

- SCANLON, M.J.; STINARD, P.S.; JAMES, M.G.; MYERS, A.M.; ROBERTSON, D.S. Genetic-analysis of 63 mutations affecting maize kernel development isolated from mutator stocks. *Genetics*, Baltimore, v.136, n.1, p.281-294, 1994.
- SEAH, S.; SPIELMEYER, W.; JAHIER, J.; SIVASITHAMPARAM, K.; LAGUDAH, E.S. Resistance gene analogs within an introgressed chromosomal segment derived from *Triticum ventricosum* that confers resistance to nematode and rust pathogens in wheat. *Molecular Plant-Microbe Interactions*, Saint Paul, v.13, n.3, p.334-341, 2000.
- SEO B.S.; PETERSON P.A. A transposable element in diverse corn lines, ubiquitous (uq)-allelism test. *Theoretical and Applied Genetics*, New York, v.90, n.7/8, p.1188-1197, 1995.
- 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 outbred three-generation pedigrees. *Genetics*, Baltimore, v.151, n.1, p.321-330, 1999.
- SOUZA, A.P. Biologia molecular aplicada ao melhoramento. In: NASS, L.L.; VALOIS, A. C.C.; MELO, I.S.; VALADARES-INGLIS, M.C. *Recursos Genéticos e melhoramento*. Fundação MT, 2001. p.939-965.
- STAUB, J.E, SERQUEN, F.C., GUPTA, M. Genetic markers, map construction, and their application in plant breeding. *HortScience*, Alexandria, v.31, n.5, p.729-740, 1996.
- SUITER, K.A.; WENDELL, J.F.; CASE, J.S. Linkage 1: a pascal computer-program for the detection and analysis of genetic-linkage. *Journal of Heredity*, Cary, v.74, n.3, p.203-204, 1983.
- TANKSLEY, S.D. Mapping polygenes. *Annual Review of Genetics*, Palo Alto, v.27, p.205-233, 1993.
- TANKSLEY, S.D.; BERNATZKY, R.; LAPITAN, N.L.; PRINCE, J.P. Conservation of gene repertoire but not gene order in pepper and tomato. *Proceedings of the National Academic of Sciences of the United States of America*, Washington, v.85, n.17, p.6419-6423, 1988b.
- TANKSLEY, S.D.; GANAL, M.W.; PRINCE, J.P.; VICENTE, M.C.; BONIERBALE, M.W.; BROUN, P.; FULTON, T.M.; GIOVANNONI, J.J.; GRANDILLO, S.; MARTIN, G.B.; MESSEGUER, R.; MILLER, J.C.; MILLER, L.; PATERSON, A.H.; PINEDA, O.; RODER, M.S.; WING, R.A.; WU, W.; YOUNG, N.D. High density molecular linkage maps of the tomato and potato genomes. *Genetics*, Baltimore, v.132, n.4, p.1141-1160, 1992.
- TANKSLEY, S.D.; GRANDILLO, S.; FULTON, T.M.; ZAMIR, D.; ESHED, Y.; PETIARD, V.; LOPEZ, J.; BECKBUNN, T. Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium*. *Theoretical and Applied Genetics*, New York, v.92, p.213-224, 1996.
- TANKSLEY, S.D.; MILLER, J.; PATERSON, A.; BERNATSKY, R. *Molecular mapping of plant chromosomes*. In: GUSTAFSON, J.P.; APPELS, R. (Eds.). New York: Plenum Press, 1988a. p.157-173.
- 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*, New York, v.103, p.30-36, 2001.
- THODAY, J.M. Location of polygenes. *Nature*, v.191, p.368-370, 1961.
- THOMPSON, E.A. Crossover counts and likelihood in multipoint linkage analysis. *IMA- Journal of Mathematical Applied Medicine Biology*, Oxford, v.4, n.2, p.93-108, 1987.
- TOOJINDA, T.; BAIRD, E.; BOOTH, A.; BROERS, L.; HAYES, P.; POWELL, W.; THOMAS, W.; VIVAR, H.; YOUNG, G. Introgression of quantitative trait loci (QTLs) determining stripe rust resistance in barley: an example of marker-assisted line development with limited resources. *Theoretical and Applied Genetics*, New York, v.96, p.123-131, 1998.
- WEIR, B. *Genetic data analysis*. 2.ed. Sunderland: Sinauer Associates, 1996. 447p.
- YAO, Y.; BAN, M.; BRANDLE, J. A genetic linkage map for *Stevia rebaudiana*. *Genome*, Ottawa, v.42, n.4, p.657-661, 1999.
- ZENG, Z.B. Precision mapping of quantitative trait loci. *Genetics*, Baltimore, v.136, n.4, p.492-496, 1994.
- ZENG, Z.B. Theoretical basis for separation of multiple linked gene effects in mapping of quantitative trait loci. *Proceedings of the National Academic of Sciences of the United States of America*, Washington, v.90, n.23, p.10972-10976, 1993.