



**República Federativa do Brasil**  
Ministério da Indústria, Comércio Exterior  
e Serviços  
Instituto Nacional da Propriedade Industrial

**(11) PI 0916265-8 B1**

**(22) Data do Depósito:** 23/07/2009

**(45) Data de Concessão:** 31/07/2018



\* B R P I 0 9 1 6 2 6 5 B 1 \*

---

**(54) Título:** MÉTODO DE SELECIONAR UMA PLANTA DE SOJA OU GERMOPLASMA COM TOLERÂNCIA OU TOLERÂNCIA MELHORADA A HERBICIDAS QUE INIBEM A FUNÇÃO DE PROTOPORFIRINOGÊNIO OXIDASE, MÉTODO DE RASTREAR SOJA PARA DETERMINAR A PRESENÇA OU AUSÊNCIA DE UM LÓCUS DE TRAÇO QUANTITATIVO ASSOCIADO COM TOLERÂNCIA A HERBICIDAS QUE INIBEM A FUNÇÃO DE PROTOPORFIRINOGÊNIO OXIDASE

**(51) Int.Cl.:** A01H 1/04; A01H 5/10; C12Q 1/68

**(30) Prioridade Unionista:** 23/07/2008 US 61/083,038

**(73) Titular(es):** PIONEER HI-BRED INTERNACIONAL, INC.

**(72) Inventor(es):** DONALD KYLE; MARK D. VOGT

**(85) Data do Início da Fase Nacional:** 21/01/2011

MÉTODO DE SELECIONAR UMA PLANTA DE SOJA OU GERMOPLASMA COM TOLERÂNCIA OU TOLERÂNCIA MELHORADA A HERBICIDAS QUE INIBEM A FUNÇÃO DE PROTOPORFIRINOGENO OXIDASE, MÉTODO DE RASTREAR SOJA PARA DETERMINAR A PRESENÇA OU AUSÊNCIA DE UM LÓCUS DE TRAÇO

5 QUANTITATIVO ASSOCIADO COM TOLERÂNCIA A HERBICIDAS QUE INIBEM A FUNÇÃO DE PROTOPORFIRINOGENO OXIDASE

#### CAMPO DA INVENÇÃO

10 Esta invenção está relacionada, em geral, à detecção das diferenças genéticas entre sojas.

#### FUNDAMENTOS DA INVENÇÃO

15

A soja (*Glycine max* L. Merr.) é um importante cultivo comercial e artigo de investimento na América do Norte e outros lugares. O óleo de soja é um dos óleos comestíveis mais amplamente usados, e soja é usada em todo o mundo tanto

20 na alimentação animal como na produção de alimentos humanos. Adicionalmente, a utilização de soja está se expandindo para aplicações industriais, manufatureiras e farmacêuticas. O manejo de ervas daninhas em campos de soja é importante para maximizar as produções. Um desenvolvimento recente na

25 tecnologia de soja foi o desenvolvimento de variedades de soja tolerantes a herbicidas. Sojas tolerantes a glifosato foram comercialmente introduzidas em 1996 e contabilizaram mais de 85% da área plantada de soja nos Estados Unidos em 2007.

Algumas ervas daninhas estão começando a apresentar tolerância aumentada a glifosato. Essa tolerância aumentada diminui a eficiência da aplicação de glifosato e resulta em menores produtividades para fazendeiros. Como resultado,  
5 existe uma necessidade na arte para variedades de soja que são tolerantes a outra química herbicida.

### **SUMÁRIO DA INVENÇÃO**

10 Esta invenção refere-se, em geral, à detecção de diferenças genéticas entre sojas. Mais particularmente, a invenção refere-se loci de traço quantitativo (QTL) de soja para tolerância a inibidores de protoporfirinogênio oxidase (PPOase), para plantas de soja possuindo esses QTLs, que  
15 mapeiam a uma nova região cromossômica, e a marcadores genéticos que são indicativos de fenótipos associados com tolerância a inibidor de protoporfirinogênio oxidase. Métodos e composições para uso desses marcadores na genotipagem e seleção de soja são também descritos.

20 Um novo método é provido para determinar a presença ou ausência, em germoplasma de soja, de um QTL associado com tolerância a inibidores de protoporfirinogênio oxidase. O traço de tolerância foi encontrado como sendo proximamente  
25 ligado a um número de marcadores moleculares que mapeiam aos grupamentos de ligação L e N. Plantas de soja, sementes, culturas de tecidos, variantes e mutantes tendo tolerância a inibidores de protoporfirinogênio oxidase produzida pelos métodos a seguir são também providas nesta invenção.

De acordo com a presente invenção, o QTL associado com tolerância a inibidores de protoporfirinogênio oxidase mapeia a um grupamento de ligação L e/ou N de soja. Esses QTL podem  
5 estar mapeados por um ou mais marcadores moleculares. Para o grupamento de ligação L, os marcadores incluem SATT495, P10649C-3, SATT182, SATT388, SATT313, SATT613, ou marcadores proximamente ligados a esses. Outros marcadores do grupamento de ligação L também podem ser usados para identificar a  
10 presença ou ausência do gene, incluindo outros marcadores acima do marcador SATT613. Para o grupamento de ligação N, os marcadores incluem Sat\_379, SCT\_195, SATT631, S60167-TB, SATT675, SATT624, SATT080, SATT387, ou marcadores proximamente ligados a esses. Outros marcadores do grupamento de ligação N  
15 também podem ser usados para identificar a presença ou ausência do gene, incluindo outros marcadores acima do marcador SATT387.

A informação aqui descrita em relação ao QTL para  
-20 tolerância a inibidores de protoporfirinogênio oxidase que mapeia ao grupamento de ligação L e/ou N de soja é usada para auxiliar na seleção de plantas de melhoramento, linhas e populações contendo tolerância a inibidores de protoporfirinogênio oxidase para uso na introdução desse traço  
25 em germoplasma de soja de elite, ou germoplasma de superioridade genética comprovada adequado para a liberação da variedade.

É também provido um método para introduzir um QTL de soja associado com tolerância a inibidores de protoporfirinogênio oxidase em germoplasma de soja não tolerante ou germoplasma de soja menos tolerante. De acordo com o método, marcadores de ácidos nucleicos mapeando no QTL são usados para selecionar plantas de soja contendo o QTL. Plantas selecionadas dessa maneira possuem uma alta probabilidade de expressar o traço de tolerância a inibidores de protoporfirinogênio oxidase. Plantas selecionadas dessa forma podem ser usadas em um programa de melhoramento de soja. Através do processo de introdução, o QTL associado com tolerância a inibidores de protoporfirinogênio oxidase é introduzido a partir de plantas identificadas usando seleção assistida por marcador para outras plantas. De acordo com o método, plantas e sementes desejáveis agronomicamente podem ser produzidas contendo o QTL associado com tolerância a inibidores de protoporfirinogênio oxidase a partir do germoplasma contendo o QTL. Fontes de tolerância a inibidores de protoporfirinogênio oxidase são descritas abaixo.

—20

É também provido aqui um método para produzir uma planta de soja adaptada para conferir tolerância a inibidores de protoporfirinogênio oxidase. Primeiro, são selecionadas plantas de soja doadoras para uma linha parental contendo o QTL de tolerância. De acordo com o método, a seleção pode ser conseguida por seleção associada a marcadores de ácido nucleico como aqui explicado. Material vegetal selecionado pode representar, dentre outros, uma linha intercruzada, um híbrido, uma população heterogênea de plantas de soja, ou

simplesmente uma planta individual. De acordo com técnicas bem conhecidas na arte de melhoramento vegetal, essa linha parental doadora é cruzada com uma segunda linha parental. Tipicamente, a segunda linha parental é uma linha de alto rendimento. Esse cruzamento produz uma população de plantas segregantes composta de plantas geneticamente heterogêneas. Plantas da população de plantas segregantes são rastreadas para o QTL de tolerância e estão sujeitas a melhoramento adicional. Esse melhoramento adicional pode incluir, dentre outras técnicas, cruzamentos adicionais com outras linhas, híbridos, cruzamento reverso, ou autocruzamento. O resultado é uma linha de plantas de soja que é tolerante a inibidores de protoporfirinogênio oxidase e também possui outros traços desejáveis de uma ou mais linhas de soja diferentes.

15

Plantas de soja, sementes, culturas de tecidos, variantes e mutantes tendo tolerância a inibidores de protoporfirinogênio oxidase produzidas pelos métodos a seguir são também providas nesta invenção.

20

#### **BREVE DESCRIÇÃO DAS FIGURAS**

A invenção pode ser mais completamente entendida a partir da descrição detalhada a seguir e das figuras que acompanham e a listagem de sequências, que forma uma parte desta aplicação.

25

A Figura 1, Painel A, provê um mapa genético integrado de marcadores de soja no grupamento de ligação L, incluindo o marcador tipo (SSR ou ASH/SNP). As posições no mapa genético

dos marcadores estão indicadas em centiMorgans (cM), tipicamente com a posição zero sendo o primeiro marcador (mais distal) no cromossomo. O Pannel B provê uma tabela listando marcadores genéticos que estão ligados aos marcadores de tolerância a inibidor de protoporfirinogênio oxidase (PPOase) identificados pela presente invenção no grupamento de ligação L. Esses marcadores são do mapa público composto de soja de 18 de junho de 2008 para o grupamento de ligação L.

10 A Figura 2, Pannel A, provê um mapa genético integrado de marcadores de soja no grupamento de ligação N, incluindo o marcador tipo (SSR ou ASH/SNP). As posições dos marcadores no mapa genético estão indicadas em centiMorgans (cM), tipicamente com a posição zero sendo o primeiro marcador (mais distal) no cromossomo. O Pannel B provê uma tabela listando marcadores genéticos que são ligados aos marcadores de tolerância a inibidores de protoporfirinogênio oxidase (PPOase) identificados pela presente invenção no grupamento de ligação N. Esses marcadores são do mapa composto público de soja de 18 de junho de 2008 para o grupamento de ligação N.

25 A Figura 3 provê uma tabela listando marcadores SSR, incluindo aqueles marcadores que demonstraram desequilíbrio de ligação com o fenótipo de tolerância a inibidores de protoporfirinogênio oxidase (PPOase). A tabela provê as sequências dos *primers* de PCR da esquerda e da direita usados na análise de genotipagem de lócus marcador SSR. Também apresentado está a sequência *pigtail* usada na extremidade 5' do *primer* da direita.

A Figura 4 provê uma tabela listando os marcadores SNP que demonstraram desequilíbrio de ligação com o fenótipo de tolerância a inibidor de protoporfirinogênio oxidase (PPOase).

5 A tabela provê as sequências dos *primers* de PCR usados para gerar um *amplicon* contendo SNP, as sondas alelo específicas que foram usadas para identificar o alelo SNP em um teste de hibridização alelo específica (teste ASH).

10 A Figura 5 provê um exemplo de cultivares com fenótipos de tolerância a inibidor de protoporfirinogênio oxidase (PPOase) vastamente diferentes. Amostras de campo são apresentadas, com uma variedade não tolerante à esquerda (círculo branco: atordoada, necrótico) e uma variedade  
15 tolerante à direita (crescimento normal).

A Figura 6 provê um exemplo de cultivares com fenótipos de tolerância a inibidor de protoporfirinogênio oxidase (PPOase) vastamente diferentes. Amostras de estufa são  
20 apresentadas, com uma variedade não tolerante com checagens não tolerantes (seta, lado esquerdo) e tolerantes (lado direito), apresentando plantas tratada em primeiro plano, e plantas não tratadas ao fundo.

## 25 **DESCRIÇÃO DETALHADA**

Deve ser entendido que esta invenção não está limitada a avanços ou exemplos em particular, os quais podem, obviamente, variar. Também deve ser entendido que a terminologia aqui

usada é para propósito de descrever apenas os avanços em particular, e não pretende-se que seja limitante. Como usado nesta especificação e reivindicações em anexo, os termos no singular e formas no singular "um", "uma" e "o", "a", por exemplo, incluem referentes no plural, a menos que o conteúdo claramente dite ao contrário. Assim, por exemplo, referência a "planta", "à planta" ou "uma planta" também inclui uma pluralidade de plantas; também, dependendo do contexto, o uso do termo "planta" pode também incluir progênie geneticamente similar ou idêntica daquela planta; o uso do termo "um ácido nucléico" opcionalmente inclui, como matéria prática, várias cópias daquela molécula de ácido nucléico; similarmente, o termo "sonda", opcionalmente (e tipicamente), engloba várias moléculas sonda similares ou idênticas.

15

Certas definições usadas na especificação são providas abaixo. Também nos exemplos que se seguem, um número de termos é usado. Os termos não especificamente definidos aqui deveriam receber seu significado corriqueiro para aqueles da arte. Para prover um entendimento claro e consistente da especificação e reivindicações, incluindo o escopo a ser dado a tais termos, as seguintes definições são providas:

20

AGRONÔMICO, TRAÇOS AGRONÔMICOS, e PERFORMANCE AGRONÔMICA referem-se aos traços e elementos genéticos de uma dada variedade de planta que contribuem para produzir sobre o curso da estação de crescimento. Traços agronômicos individuais incluem vigor de emergência, vigor vegetativo, tolerância a estresse, resistência ou tolerância a doenças, resistência ou

25

tolerância a herbicidas, ramificação, florada, arranjo de sementes, tamanho de sementes, densidade de sementes, tempo de permanência, habilidade de resistência e similares.

5        ALELO significa qualquer de uma ou mais formas alternativas de uma sequência genética. Em uma célula ou organismo diplóide, os dois alelos de uma dada sequência tipicamente ocupam loci correspondentes em um par de cromossomos homólogos.

10

O termo AMPLIFICAR no contexto de amplificação de ácidos nucleicos é qualquer processo pelo qual cópias adicionais de um ácido nucleico selecionado (ou uma forma transcrita desses) são produzidas. Métodos de amplificação típicos incluem  
15 vários métodos de replicação baseado em polimerase, incluindo a reação de cadeia de polimerase (PCR), métodos mediados por ligase, tais como a cadeia de reação de ligase (LCR) e métodos de amplificação baseados em RNA polimerase (e.g., por transcrição). Um "*amplicon*" é um ácido nucleico amplificado,  
20 e.g., um ácido nucleico que é produzido amplificando-se um ácido nucleico modelo por qualquer método de amplificação disponível (e.g., PCR, LCR, transcrição, ou similares).

Uma LINHA ANCESTRAL é uma linha parental usada como fonte  
25 de genes.

Uma POPULAÇÃO ANCESTRAL é um grupo de ancestrais que contribuíram para o volume da variação genética que foi usada para desenvolver as linhas de elite.

CRUZAMENTO REVERSO é um processo no qual um criador cruza uma variedade de progênie de volta a um dos genótipos parentais uma ou mais vezes.

5 MELHORAMENTO significa a manipulação genética de organismos vivos.

10 O termo SEGMENTO CROMOSSÔMICO designa um par de cromossomos contíguos de DNA genômico que reside *in planta* em um único cromossomo.

15 CULTIVAR e VARIEDADE são usados como sinônimos e significam um grupo de plantas em uma espécie (e.g., *Glycine max*) que compartilham traços genéticos que os separam da forma típica e de outras possíveis variedades naquela espécie. Cultivares de soja são linhas intercruzadas produzidas após diversas gerações de autopolinização. Indivíduos em um cultivar de soja são homogêneos, quase geneticamente idênticos, com a maioria dos loci no estado homozigoto.

20

Uma LINHA DE ELITE é uma linha agronomicamente superior que resultou de vários ciclos de cruzamento e seleção para performance agrônômica superior. Numerosas linhas de elite estão disponíveis para aqueles especialistas na arte de 25 melhoramento de soja.

Uma POPULAÇÃO DE ELITE é uma seleção de indivíduos ou linhas de elite que podem ser usadas para representar o estado

da arte em termos de genótipos agronomicamente superiores de uma dada espécie de cultivo, tal como soja.

5 Um MAPA GENÉTICO é uma descrição de relações de ligação genética entre loci em um ou mais cromossomos ou grupamentos de ligação em uma dada espécie, geralmente representada em uma forma diagramática ou tabular.

10 GENÓTIPO refere-se à constituição genética de uma célula ou organismo.

15 GERMOPLASMA significa o material genético que compreende a base física das qualidades hereditárias de um organismo. Assim como, germoplasma inclui sementes e tecido vivo a partir dos quais novas plantas podem ser cultivadas; ou, outra parte de planta, tal como folha, caule, pólen, ou células, que podem ser cultivadas em uma planta inteira. Fontes de germoplasma provêm fontes de traços genéticos usados por cultivadores de plantas para melhorar cultivares comerciais.

20

Um indivíduo é HOMOZIGOTO se o indivíduo possui apenas um tipo de alelo em um dado locus (e.g., um indivíduo diplóide possui uma cópia do mesmo alelo em um locus para cada dos dois cromossomos homólogos). Um indivíduo é "HETEROZIGOTO" se mais 25 de um tipo de alelos está presente em um dado locus (e.g., um indivíduo diplóide com uma cópia de cada um dos dois alelos diferentes). O termo "HOMOGENEIDADE" indica que membros de um grupo possuem o mesmo genótipo em um ou mais loci específicos. Em contraste, o termo "HETEROGENEIDADE" é usado para indicar

que indivíduos no grupo diferem-se em genótipo em um ou mais loci específicos.

INTRODUÇÃO significa a entrada ou introdução de um gene,  
5 QTL, ou locus traço do genoma de uma planta no genoma de outra planta.

Uma LINHA ou LINHAGEM é um grupo de indivíduos de pais idênticos que são geralmente inter cruzados em algum grau e que  
10 são geralmente homozigotos e homogêneos na maioria dos loci (isogênicos ou quase isogênicos). Uma "SUBLINHA" refere-se a um subgrupo inter cruzado de descendentes que são geneticamente distintos de outros subgrupos similarmente inter cruzados descendentes do mesmo progenitor. Tradicionalmente, uma  
15 sublinha foi derivada inter cruzando-se sementes de uma planta de soja individual selecionada na geração F3 a F5 até que os loci residuais segregantes estejam "fixados" ou homozigotos na maioria ou todos os loci. Variedades comerciais de soja (ou linhas) são tipicamente produzidas agregando-se ("juntando-  
20 se") a progênie autopolinizada de uma única planta F3 a F5 a partir de um cruzamento controlado entre 2 pais geneticamente diferentes. Enquanto a variedade tipicamente se aparenta uniforme, a variedade autopolinizada derivada da planta selecionada eventualmente (e.g., F8) se torna uma mistura de  
25 plantas homozigotas que podem variar em genótipo em qualquer locus que foi heterozigoto na planta F3 a F5 originalmente selecionada. No contexto da invenção, sublinhas baseadas em marcadores que diferem uma das outras baseado em polimorfismo qualitativo no nível de DNA em um ou mais loci marcadores

específicos são derivadas genotipando-se uma amostra de sementes derivada de progênie individual autopolinizada derivada de uma planta F3-F5 selecionada. A amostra de sementes pode ser genotipada diretamente como sementes, ou  
5 como tecido vegetal crescido a partir de uma amostra de semente. Opcionalmente, sementes compartilhando um genótipo em comum no locus especificado (ou loci) são juntadas provendo uma sublinha que é geneticamente homogênea em loci importantes identificados para um traço de interesse (rendimento,  
10 tolerância, etc.).

LIGAÇÃO refere-se a um fenômeno caracterizado pelo fato de que alelos no mesmo cromossomo tendem a segregar juntos mais frequentemente do que o esperado do que se sua  
15 transmissão fosse independente. A recombinação genética ocorre em uma frequência assumidamente randômica por todo o genoma. Mapas genéticos são construídos medindo-se a frequência de recombinação entre pares de traços ou marcadores. Quanto mais próximos os traços ou marcadores  
20 reside em relação um ao outro no cromossomo, menos a frequência de recombinação, e maior o grau de ligação. Traços ou marcadores são aqui considerados como sendo ligados se geralmente co-segregam. Uma probabilidade de recombinação por geração de 1/100 é definida como uma distância no mapa de 1,0  
25 centiMorgan (1,0 cM). Por exemplo, em soja, 1 cM se correlaciona, em média, a cerca de 400.000 pares de base (400 Kb).

Os elementos genéticos ou genes localizados em um único segmento cromossômico estão fisicamente ligados. No contexto da presente invenção, os elementos genéticos localizados no segmento cromossômico também estão geneticamente ligados, tipicamente em uma distância de recombinação genética de menos do que ou igual a 50 centimorgans (cM), e.g., cerca de 49, 40, 30, 20, 10, 5, 4, 3, 2, 1, 0,75, 0,5, ou 0,25 cM ou menos. Ou seja, dois elementos genéticos em um único segmento cromossômico passam por recombinação um com o outro durante a meiose em uma frequência de menos de ou igual a cerca de 50%, e.g., cerca de 49%, 40%, 30%, 20%, 10%, 5%, 4%, 3%, 2%, 1%, 0,75%, 0,5%, ou 0,25% ou menos.

GRUPAMENTO DE LIGAÇÃO refere-se a traços ou marcadores que geralmente co-segregam. Um grupamento de ligação geralmente corresponde a uma região cromossômico contendo material genético que codifica os traços ou marcadores.

LÓCUS é um segmento de DNA definido.

Uma LOCALIZAÇÃO DE MAPA é uma localização determinada em um mapa genético relativa a marcadores genéticos ligados onde um marcador especificado pode ser encontrado em uma dada espécie. Marcadores são frequentemente descritos como estando "acima" ou "abaixo" outros marcadores no mesmo grupamento de ligação; um marcador está "acima" de outro marcador se este aparece antes no grupamento de ligação, enquanto um marcador está "abaixo" de outro marcador se este aparece depois no grupamento de ligação.

MAPEAMENTO é o processo de definir as relações de ligação de loci através do uso de marcadores genéticos, populações segregantes para os marcadores, e princípios genéticos padrão  
5 de frequência de recombinação.

MARCADOR MOLECULAR é um ácido nucléico ou sequência de aminoácidos que é suficientemente único para caracterizar um  
lócus específico no genoma. Exemplos incluem Polimorfismos de  
10 Comprimento de Fragmento de Restrição (RFLPs), Repetições de Sequência Única (SSRs), Polimorfismos de Amplificação de Região Alvo (TRAPs), Eletroforese de Isoenzimas, DNAs Polimórficos Randomicamente Amplificados (RAPDs), Reação de Cadeia de Polimerase Arbitrariamente Mediada por Primer (AP-PCR), Amplificação de DNA por Impressão Digital (DAF), Regiões  
15 de Sequências Caracterizadas Amplificadas (SCARs), Polimorfismo de Comprimento de Fragmento Amplificado (AFLPs), e Polimorfismos de Nucleotídeos Simples (SNPs). Adicionalmente, outros tipos de marcadores moleculares são  
20 conhecidos na arte, e traços fenotípicos podem ser também usados como marcadores nos métodos desta invenção. Todos os marcadores são usados para definir um locus específico no genoma de soja. Grandes números desses marcadores foram mapeados. Cada marcador é, portanto, um indicador de um  
25 segmento de DNA específico, tendo uma sequência de nucleotídeos particular. As posições do mapa provêm uma medida das posições relativas de marcadores em particular com respeito um ao outro. Quando diz-se que um traço está ligado a um dado marcador, será entendido que o segmento de DNA

propriamente dito, cuja sequência afeta o traço geralmente co-segrega em relação ao marcador. A localização mais precisa e definitiva de um traço pode ser obtida se marcadores são identificados em ambos os lados do traço. Medindo-se a

5   aparência do(s) marcador(es) na progênie de cruzamentos, a existência do traço pode ser detectada por testes moleculares relativamente simples sem de fato avaliar a aparência do traço propriamente dito, o que pode ser difícil e demorado, pois a

10   avaliação do traço propriamente dita requer cultivar plantas a um estágio onde o traço pode ser expressado. Marcadores moleculares foram amplamente usados para determinar a composição genética em sojas. Shoemaker and Olsen, ((1993) Molecular Linkage Map of Soybean (*Glycine max* L. Merr.). p. 6.131-6.138. In S.J. O'Brien (ed.) Genetic Maps: Locus Maps

15   of Complex Genomes. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, New York.), desenvolveram um mapa de ligação genética molecular que consistiu de 25 grupamentos de ligação com cerca de 365 RFLP, 11 RAPD (DNA polimórfico randomicamente amplificado), três marcadores clássicos, e quatro loci de

20   isoenzimas. Ver também Shoemaker R.C. 1994 RFLP Map of Soybean. P. 299-309 In R.L. Phillips and I.K. Vasil (ed.) DNA-based markers in plants. Kluwer Academic Press Dordrecht, the Netherlands.

25       SELEÇÃO ASSISTIDA POR MARCADOR refere-se ao processo de selecionar um traço desejado ou traços desejados em uma planta ou plantas detectando-se um ou mais marcadores moleculares da planta, onde o marcador molecular está ligado ao traço desejado.

O termo FISICAMENTE LIGADO é usado para indicar que dois loci, e.g., dois loci marcadores, estão fisicamente presentes no mesmo cromossomo. De forma vantajosa, os dois loci estão  
5 localizados em proximidade íntima de forma que a recombinação entre pares de cromossomos homólogos não ocorra entre os dois loci durante a meiose com alta frequência, e.g., de forma que loci ligados co-segregam em pelo menos cerca de 90% do tempo, e.g., 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5%,  
10 99,75%, ou mais do tempo.

O termo PLANTA inclui referência a uma planta inteira imatura ou Madura, incluindo uma planta da qual sementes ou grãos ou anteras foram removidas. Semente ou embrião que irá  
15 produzir a planta também é considerado como sendo a planta.

PARTES DE PLANTA incluem folhas, caules, botões, raízes, pontas de raízes, anteras, semente, grão, embrião, pólen, óvulos, flores, cotilédones, hipocótilos, brotos e espigas, tecidos, células e similares.

20

POLIMORFISMO significa uma mudança ou diferença entre dois ácidos nucléicos relacionados. Um "POLIMORFISMO DE NUCLEOTÍDEO" refere-se a um nucleotídeo que é diferente em uma sequência quando comparado a uma sequência relacionada quando  
25 os dois ácidos nucléicos estão alinhados para máxima correspondência. Um "POLIMORFISMO GENÉTICO DE NUCLEOTÍDEO" refere-se a um nucleotídeo que é diferente em uma sequência quando comparado a uma sequência relacionada quando os dois ácidos nucléicos estão alinhados para máxima correspondência,

onde os dois ácidos nucléicos estão geneticamente relacionados, i.e., homólogos, por exemplo, onde os ácidos nucléicos são isolados a partir de diferentes linhagens de uma planta de soja, ou a partir de diferentes alelos de uma única  
5 linhagem, ou similares.

SONDA significa um polinucleotídeo projetado para ser suficientemente complementar a uma sequência em um ácido nucléico desnaturado para ser sondado e para ser limitado sob  
10 condições de estringência selecionadas.

Marcador RAPD significa DNA marcador polimórfico randomicamente amplificado. Pares de sítios complementares a nucleotídeos simples, octanucleotídeos ou decanucleotídeos  
15 podem existir na orientação correta, próximos o suficiente um do outro para amplificação por PCR. Com alguns decanucleotídeos randomicamente escolhidos, nenhuma sequência é amplificada. Com outros, os produtos de mesmo comprimento são gerados a partir de DNAs de diferentes indivíduos. Ainda  
20 com outros, padrões de bandas não são os mesmos para cada indivíduo em uma população. As bandas variáveis são comumente chamadas de bandas de DNA polimórfico randomicamente amplificado (RAPD).

25 FREQUÊNCIA DE RECOMBINAÇÃO é a frequência de um evento de *crossing over* (recombinação) entre dois loci genéticos. A frequência de recombinação pode ser observada seguindo a segregação de marcadores e/ou traços durante a meiose. No contexto desta invenção, um locus marcador está "associado

com" outro locus marcador ou alguns outros loci (por exemplo, um locus de tolerância), quando os loci relevantes são parte do mesmo grupamento de ligação e estão em desequilíbrio de ligação. Isso ocorre quando o locus marcador e um locus  
5 ligado são encontrados juntos nas plantas de progênie mais frequentemente do que se os dois loci segregassem aleatoriamente. Similarmente, um locus marcador pode estar também associado com um traço, e.g., um locus marcador pode estar "associado com tolerância ou tolerância aumentada"  
10 quando o locus marcador está em desequilíbrio de ligação com o traço.

RFLP significa polimorfismo de comprimento de fragmento de restrição. Marcadores moleculares que ocorrem por causa de  
15 qualquer mudança de sequência no DNA, incluindo uma mudança em uma única mudança, inserção, deleção ou inversão, podem resultar em perda ou ganho de um sítio de reconhecimento de endonuclease de restrição. O tamanho e número de fragmentos gerados por tal enzima é, portanto, alterado. Uma sonda que  
20 hibridiza especificamente a DNA na região de tal alteração pode ser usada para rapidamente e especificamente identificar uma região de DNA que apresenta variação alélica entre duas variedades vegetais. Eletroforese de Isoenzimas e RFLPs foram amplamente usados para determinar a composição genética.

25

AUTOCRUZAMENTO ou AUTOPOLINIZAÇÃO é um processo através do qual um criador cruza progênie híbrida com ela mesma; por exemplo, uma segunda geração híbrida F2 com ela mesma para produzir a progênie designada F2:3.

SNP significa polimorfismo de nucleotídeo simples. SNPs são marcadores genéticos nas quais variações na sequência de DNA que ocorrem em um único nucleotídeo (A, T, C, ou G) na  
5 sequência genômica alterada são mapeadas a sítios no genoma de soja. Várias técnicas para detectar SNPs são conhecidas na arte, incluindo hibridização alelo específica, extensão de *primer*, e seqüenciamento direto.

10 SSR significar repetições de sequências curtas. SSRs são marcadores genéticos baseados em polimorfismos em sequências de nucleotídeos repetidas, tais como microssatélites. Um sistema de marcador baseado em SSRs pode ser altamente informativo na análise de ligação relativa a outro sistema de  
15 marcador no qual múltiplos alelos pode estar presentes. A detecção por PCR é feita por uso de dois *primers* de oligonucleotídeo flanqueando o segmento polimórfico do DNA repetitivo. Ciclos repetidos de desnaturação por calor do DNA seguido por anelamento dos *primers* às suas sequências  
20 complementares em baixas temperaturas, e extensão dos *primers* anelados com DNA polimerase, compreendem a parte principal da metodologia.

TOLERANTE e TOLERÂNCIA referem-se a plantas nas quais  
25 doses maiores de um herbicida são necessárias para produzir efeitos similares àqueles observados em plantas não tolerantes. Plantas tolerantes tipicamente exibem menos lesões necróticas, líticas, cloróticas, ou outras lesões quando

sujeitas ao herbicida em concentrações e taxas tipicamente empregadas pela comunidade agrícola.

PLANTA TRANSGÊNICA refere-se a uma planta que compreende  
5 em suas células um polinucleotídeo heterólogo. Geralmente, o polinucleotídeo heterólogo é estavelmente integrado no genoma de forma que o polinucleotídeo é passado a sucessivas gerações. O polinucleotídeo heterólogo pode ser integrado no  
10 genoma sozinho ou como parte de um cassete de expressão recombinante. TRANSGÊNICO é aqui usado para se referir a qualquer célula, linha celular, calo, tecido, parte de planta ou planta, cujo genótipo foi alterado pela presença de ácido nucléico heterólogo incluindo aqueles organismos transgênicos ou células inicialmente alterados dessa maneira, assim como  
15 aqueles criados por cruzamentos ou propagação assexuada a partir do organismo transgênico ou célula inicial. O termo "transgênico" como aqui usado não engloba a alteração do genoma (cromossômico ou extracromossômico) por métodos convencionais de melhoramento vegetal (e.g., cruzamentos) ou  
20 por eventos naturalmente ocorrentes, tais como fertilização cruzada aleatória, infecção viral não recombinante, transformação bacteriana não recombinante, transposição não recombinante, ou mutação espontânea.

25 Marcador TRAP significa marcador de polimorfismo de amplificação de região alvo. A técnica TRAP emprega um *primer* fixado de sequência conhecida em combinação com um *primer* aleatório para amplificar fragmentos genômicos. As diferenças

em fragmentos entre alelos podem ser detectadas por eletroforese em gel.

O termo VETOR é usado em referência a polinucleotídeo ou  
5 outras moléculas que transferem segmento(s) de ácidos  
nucléicos em uma célula. O termo "veículo" é, às vezes, usado  
intercambiavelmente com "vetor". Um vetor opcionalmente  
compreende partes que mediam a manutenção de vetor e permite  
seu uso pretendido (e.g., sequências necessárias para  
10 replicação, genes conferindo resistência a drogas ou  
antibióticos, um sítio de clonagem múltiplo, elementos  
promotores/intensificadores operacionalmente ligados que  
permitem a expressão de um gene clonado, etc.). Vetores são  
frequentemente derivados a partir de plasmídios,  
15 bacteriófagos, ou vírus vegetais ou animais. Um "vetor de  
clonagem" ou "vetor de transporte" ou "vetor de subclonagem"  
contém partes operacionalmente ligadas que facilitam as etapas  
de subclonagem (e.g., um sítio de clonagem múltiplo contendo  
múltiplos sítios de restrição de endonuclease).

20

O termo PRODUÇÃO refere-se à produtividade por unidade de  
área de um produto vegetal em particular de valor comercial.  
Por exemplo, a produção de soja é comumente medida em  
alqueires de sementes por acre ou toneladas métricas de  
25 sementes por hectare por temporada. A produção é afetada por  
ambos fatores genéticos e ambientais. A produção é a  
culminação final de todos os traços agronômicos.

#### INIBIDORES DE PROTOPORFIRINOGÊNIO OXIDASE

Porfirinas são biologicamente importantes estruturas orgânicas que são encontradas em plantas anexadas a clorofila e pigmentos citocromos. Um intermediário na via de síntese de clorofila e citocromo é protoporfirinogênio IX que é convertido para protoporfirina IX pela protoporfirinogênio oxidase. A inibição da protoporfirinogênio oxidase previne essa conversão e resulta em um aumento de protoporfirinogênio IX no citoplasma da planta. O protoporfirinogênio, então, sofre auto-oxidação não enzimática e se torna protoporfirina IX. Quando a protoporfirina IX citoplasmática é exposta à luz do sol, radicais livres são formados, o que resulta em reações de peroxidação de lipídeos que resulta na morte da planta. Famílias químicas de inibidores de protoporfirinogênio oxidase incluem as famílias de difenil éter, triazolinona, N-fenilftalamida, pirimidindiona e oxadiazola. Existem outras famílias de químicos que também pertencem a esse grupo.

A família de difenil éter é caracterizada por dois anéis de benzeno ligados com uma ponte de éter em um grupamento nitro ligado na posição 4. Exemplos de difenis éteres inibidores de protoporfirinogênio oxidase incluem acifluorfen, fomesafen, oxifluorfen e lactofen. Os éteres difenil são tipicamente considerados por serem herbicidas de contato.

A família de triazolinona é caracterizada por um anel de 5 membros contendo três átomos de nitrogênio (dois dos quais são adjacentes) e dois átomos de carbono, um dos átomos de carbono possui ligação dupla com um átomo de oxigênio e um dos

átomos de nitrogênio está ligado a um anel de benzeno. Exemplos de triazolinonas inibidoras de protoporfirinogênio oxidase incluem sulfentrazone, carfentrazone, e azafeniden.

5        A família de N-fenilftalimida é caracterizada pelo grupo ptalimida, caracterizado pelo fato de que o nitrogênio está ligado a um anel de benzeno. Exemplos de N-fenilftalamidas inibidoras de protoporfirinogênio oxidase incluem flumiclorac e flumioxazin.

10

      A família de oxadiazola é caracterizada por um anel de cinco membros consistindo de dois átomos de nitrogênio adjacentes, dois átomos de carbono, e um átomo de oxigênio ou enxofre. Exemplos de oxadiazolas inibidoras de  
15 protoporfirinogênio oxidase incluem oxadiazon e fluthiacet.

20

      As várias famílias de inibidores de protoporfirinogênio oxidase provêm uma ampla variedade de opções de aplicação. Sulfentrazone, por exemplo, possui uma meia-vida relativamente longa (aproximadamente 280 dias), é conhecida por ter atividade residual no solo e é frequentemente usado como um herbicida de pré-emergência. Carfentrazone possui uma meia-vida consideravelmente mais curta (aproximadamente 4 dias), não possui atividade residual no solo, e é usado como um  
25 herbicida de contato/pós-emergência. A família das pirimidindionas de herbicidas PPO é, ao contrário, uma classe pequena que inclui benzfendizona, butagenacil e saflufenacil. Essa diversidade em características químicas combinada com tolerância a inibidor de protoporfirinogênio oxidase provê aos

fazendeiros uma ampla variedade de opções para o manejo de ervas daninhas.

#### MARCADORES MOLECULARES E LIGAÇÃO GENÉTICA

5

Marcadores moleculares foram usados para seletivamente melhorar cultivos de soja através do uso de seleção assistida por marcador. Qualquer traço polimórfico detectável pode ser usado como marcador, contanto que seja diferencialmente herdado e exiba desequilíbrio de ligação com um traço fenotípico de interesse. Um número de marcadores de soja foram mapeados e grupamentos de ligação criados, como descrito em Cregan, P.B. *et al.*, "An Integrated Genetic Linkage Map of the Soybean Genome" (1999) *Crop Science* 39:1464-90, e mais recentemente em Choi *et al.*, "A Soybean Transcript Map: Gene Distribution, Haplotype and Single-Nucleotide Polymorphism Analysis" (2007) *Genetics* 176:685-96. Vários marcadores de soja estão publicamente disponíveis no sítio da internet da base de soja afiliada do USDA.

20

A maioria dos traços vegetais de importância agronômica é é poligênica, ao contrário conhecidos como traços quantitativos. Um traço quantitativo é controlado por diversos genes localizados em várias localizações, ou loci, no genoma da planta. Os múltiplos genes possuem um efeito cumulativo que contribui para a variedade de fenótipos observados em vários traços vegetais. Esses genes são referidos como loci de traço quantitativo (QTL). A frequência de recombinação mede a extensão a qual um marcador molecular é

ligado a um QTL. Frequências de recombinação menores, tipicamente medidas em centiMorgans (cM), indicam maior ligação entre o QTL e o marcador molecular. A extensão pela qual duas características são ligadas é frequentemente referida como distância genética. A distância genética é também tipicamente relacionada à distância física entre o marcador e o QTL, contudo, certos fenômenos biológicos (incluindo "*hot spots*" de recombinação) podem afetar a relação entre a distância física e distância genética. Geralmente, a utilidade de um marcador molecular é determinada pela distância genética e física entre o marcador e o traço de interesse selecionável.

O método para determinar a presença ou ausência de um QTL associado com tolerância a inibidores de protoporfirinogênio oxidase em germoplasma de soja, compreende analisar DNA genômico de um germoplasma de soja para a presença de pelo menos um marcador molecular, caracterizado pelo fato de pelo menos um marcador molecular estar ligado ao QTL, e caracterizado pelo fato do QTL mapear a grupamento de ligação L e N principal de soja e estar associado com tolerância a inibidores de protoporfirinogênio oxidase. O termo "está associado com" neste contexto significa que o QTL associado com tolerância a inibidores de protoporfirinogênio oxidase foi encontrado, usando análise assistida por marcador, para estar presente em plantas de soja apresentando tolerância a inibidores de protoporfirinogênio oxidase em testes in vivo como aqui descrito.

Geralmente, marcadores que mapeiam mais próximos do QTL mapeado no grupamento de ligação L e N e associados com tolerância a inibidores de protoporfirinogênio oxidase são superiores a marcadores que mapeiam mais distantes do QTL para uso nesta invenção. Em alguns exemplos, um marcador usado para determinar a presença ou ausência de um QTL mapeando no grupamento de ligação L e/ou N de soja e associado com tolerância a inibidores de protoporfirinogênio oxidase mapeia no grupamento de ligação L de soja é SATT495, P10649C-3, SATT182, SATT388, SATT313, e SATT613 (ou outros marcadores acima do marcador SATT613), e aqueles mapeados no grupamento de ligação N são Sat\_379, SCT\_195, SATT631, S60167-TB, SATT675, SATT624, SATT080, e SATT387 (ou outros marcadores acima de SATT387). Qualquer marcador designado para o grupamento de ligação L e/ou N de soja e ligado a um marcador aqui descrito como associado com tolerância a inibidores de protoporfirinogênio oxidase pode ser usado com a invenção. Geralmente, um marcador ligado está a 50 cM do marcador referência. Informações atualizadas em relação a marcadores designados para os grupamentos de ligação L e N de soja podem ser encontradas no sítio da internet da base de soja do USDA. Adicionalmente, o grupamento de ligação L é agora formalmente referido como cromossomo #19 e o grupamento de ligação N é agora formalmente referido como cromossomo #3.

25

Marcadores flanqueando o QTL associado com tolerância a inibidores de protoporfirinogênio oxidase são usados em processos providos na seleção assistida por marcador. O DNA genômico de germoplasma de soja é tipicamente testado para a

presença de pelo menos dois dos marcadores moleculares anteriores, um marcador em cada lado do QTL. Em alguns exemplos um QTL no grupamento de ligação L é usado. Marcadores úteis no grupamento de ligação L incluem SATT495, P10649C-3, SATT182, SATT388, SATT313, e SATT613, incluindo 5 marcadores acima de SATT613. Marcadores que mapeiam próximos a SATT495, P10649C-3, SATT182, SATT388, SATT313, e SATT613 também podem ser usados. Em alguns exemplos, um QTL no grupamento de ligação N é usado. Marcadores úteis no 10 grupamento de ligação N incluem Sat\_379, SCT\_195, SATT631, S60167-TB, SATT675, SATT624, SATT080, e SATT387, incluindo marcadores acima de SATT387. Marcadores que mapeiam próximos de Sat\_379, SCT\_195, SATT631, S60167-TB, SATT675, SATT624, SATT080, e SATT387 também podem ser usados.

15

Métodos de introduzir tolerância a inibidor de protoporfirinogênio oxidase em germoplasma de soja não tolerante ou menos tolerante são providos. Qualquer método para introduzir QTLs em plantas de soja pode ser usado. Em 20 alguns exemplos, um primeiro germoplasma de soja que contém tolerância a inibidores de protoporfirinogênio oxidase derivado do QTL mapeado ao grupamento de ligação L e/ou N que está associado com tolerância a inibidores de protoporfirinogênio oxidase e um segundo germoplasma de soja 25 que não possui tolerância a inibidores de protoporfirinogênio oxidase derivado do QTL mapeado ao grupamento de ligação L e/ou N são providos. A primeira planta de soja pode ser cruzada com a segunda planta de soja para prover sojas progênie. O rastreamento fenotípico e/ou de marcador é, então,

realizado nas plantas da progênie para determinar a presença de tolerância a inibidores de protoporfirinogênio oxidase derivada do QTL mapeado ao grupamento de ligação L e/ou N. A progênie que testar positivo para a presença de tolerância a inibidores de protoporfirinogênio oxidase derivada do QTL mapeado ao grupamento de ligação L e/ou N pode ser selecionada.

Em alguns exemplos, o rastreamento e seleção são efetuados usando-se seleção assistida por marcador usando qualquer marcador ou combinação de marcadores no grupamento de ligação L e/ou N principal provido. Qualquer método de identificar a presença ou ausência desses marcadores pode ser usado, incluindo, por exemplo, análise de polimorfismo de conformação de hélice simples (SSCP), rastreamento de sequência de excisão de base (BESS), análise RFLP, análise heteroduplex, eletroforese de gel de gradiente de desnaturação, e eletroforese de gradiente de temperatura, PCR alélico, sequenciamento direto por reação de cadeia de ligase, mini sequenciamento, hibridização de ácidos nucléicos, ou detecção tipo micro arranjo.

Sistemas, incluindo sistemas automáticos para selecionar plantas que compreendem um marcador de interesse e/ou para correlacionar a presença do marcador com tolerância são também providos. Esses sistemas podem incluir sondas relevantes para a detecção de lócus marcador, detectores para detectar rótulos nas sondas, elementos para manuseio adequado de fluidos e controladores de temperatura que misturam sondas e modelos

e/ou amplificam modelos, e instruções de sistemas que correlacionam a detecção de rótulo à presença de um locus marcador ou alelo em particular.

5        *Kits* são também providos. Por exemplo, um *kit* pode incluir *primers* apropriados ou sondas para detectar loci marcadores associados a tolerância e instruções no uso dos *primers* ou sondas para detectar os loci marcadores e correlacionar os loci com a tolerância a inibidor de protoporfirinogênio oxidase prevista. Os *kits* podem ainda  
10        incluir materiais de embalagem para empacotar as sondas, *primers* ou instruções, controles, tais como reações de amplificação controles que incluem sondas, *primers* ou ácidos nucleicos modelos para amplificações, marcadores de tamanho  
15        molecular, ou similares.

Fragmentos isolados de ácidos nucleicos compreendendo uma sequência de ácidos nucleicos, codificando tolerância a inibidores de protoporfirinogênio oxidase em soja, são  
20        providos. O fragmento de ácidos nucleicos compreende pelo menos uma porção de ácido nucleico pertencendo ao grupamento de ligação L e/ou N. O fragmento de ácidos nucleicos é capaz de hibridizar sob condições estridentes para ácido nucleico de um cultivar de soja tolerante a inibidores de  
25        protoporfirinogênio oxidase contendo um QTL associado com tolerância a inibidor de protoporfirinogênio oxidase que está localizado no grupamento de ligação L e/ou N principal.

Vetores compreendendo tais fragmentos de ácidos nucleicos, produtos de expressão de tais vetores expressados em um hospedeiro compatível com os mesmos, anticorpos para o produto de expressão (ambos policlonal e monoclonal), e ácido nucleico antisenso em relação ao fragmento de ácidos nucleicos também são providos.

10 Sementes de uma soja produzida cruzando-se uma variedade de soja tendo QTL de tolerância a inibidor de protoporfirinogênio oxidase localizado no grupamento de ligação L e/ou N principal em seu genoma com outra variedade de soja, e progênie dessa, são providas.

#### MARCADORES DE TOLERÂNCIA E ALELOS FAVORÁVEIS

15

Na análise tradicional de ligação, nenhum conhecimento direto da relação física de genes em um cromossomo é necessário. A primeira lei de Mendel é que fatores de pares de características são segregadas, significando que alelos de um traço diplóide se separam em dois gametas e, então, nos descendentes diferentes. A análise clássica de ligação pode ser pensada como uma descrição estatística das frequências relativas de co-segregação de diferentes traços. A análise de ligação, como descrita previamente, é a ferramenta descritiva bem caracterizada de como traços são agrupados juntos baseado na frequência com a qual eles segregam juntos. Porque a distância cromossômica é aproximadamente proporcional à frequência de eventos de recombinação entre traços, existe uma

distância física aproximada que se correlaciona com a frequência de recombinação.

Loci marcadores são traços, e podem ser avaliados de acordo com análise padrão de ligação rastreando-se os loci marcadores durante a segregação. Assim, um cM é igual a uma chance de 1% de que um locus marcador seja separado de outro locus (que pode ser qualquer outro traço, e.g., outro locus marcador, ou outro locus traço que codifica um QTL), devido a recombinação em uma única geração. Os marcadores aqui, e.g., para o grupamento de ligação L: SATT495, P10649C-3, SATT182, SATT388, SATT313, e SATT613 (e outros marcadores acima de SATT613) e para o grupamento de ligação N: Sat\_379, SCT\_195, SATT631, S60167-TB, SATT675, SATT624, SATT080, e SATT387 (e outros marcadores acima de SATT387), foram encontrados como correlacionando-se com tolerância ou tolerância aumentada a inibidores de protoporfirinogênio oxidase em soja. Isso significa que os marcadores são suficientemente próximos a um traço de tolerância que podem ser usados como previsores para o traço de tolerância propriamente dito, usando, por exemplo, seleção assistida por marcador (MAS). Plantas de soja ou germoplasma podem ser selecionados para marcadores ou alelo marcadores que se correlacionam positivamente com tolerância, sem, de fato, criar soja e medir a tolerância ou tolerância aumentada (ou, ao contrário, plantas de soja podem ser contra selecionadas se possuem marcadores que se correlacionam negativamente com tolerância ou tolerância aumentada. MAS é um atalho poderoso para selecionar fenótipos desejados e introduzir traços desejados em cultivares de soja (e.g.,

introduzindo traços desejados em linhas de elite). MAS é facilmente adaptada a métodos de análise molecular de alta produção que podem rastrear rapidamente grandes números de material genético de plantas ou germoplasmas para os marcadores de interesse e é mais eficiente em material de custo do que criar e observar plantas para traços visíveis.

Qualquer marcador que possa estar ligado a um traço de interesse (e.g., no presente caso, um traço de tolerância ou traço de tolerância de aumentada) pode ser usado como marcador para aquele traço. Assim, além dos marcadores aqui descritos, marcadores ligados aos marcadores aqui listados podem ser também usados para prever o traço de tolerância ou traço de tolerância aumentada. Tais marcadores ligados são particularmente úteis quando estão suficientemente próximos de um dado marcador, de forma que apresentem uma baixa frequência de recombinação com o dado marcador. Marcadores proximamente ligados aos marcadores no grupamento de ligação L e/ou grupamento de ligação N são também providos. Marcadores proximamente ligados apresentam uma frequência de recombinação com um dado marcador de cerca de 10% ou menos (o dado marcador está a 10cM do dado marcador). Colocado de outra maneira, loci proximamente ligados co-segregam em pelo menos 90% do tempo.

25

Loci marcadores são especialmente úteis quando estão proximamente ligados a um loci alvo (e.g., QTL para tolerância, ou, alternativamente, simplesmente outros loci marcadores, tais como aqueles aqui identificados, que estão

ligados a tal QTL) para o qual estão sendo usados como marcadores. Um marcador mais proximamente ligado a um locus alvo é um melhor indicador para o locus alvo (devido à frequência reduzida de recombinação entre o locus alvo e o  
5 marcador). Assim, em um exemplo, loci proximamente ligados, tais como um locus marcador e um segundo locus (e.g., um dado marcador ou um QTL) apresentam uma frequência de recombinação interlocus de cerca de 10% ou menos, cerca de 9% ou menos, cerca de 8% ou menos, cerca de 7% ou menos, cerca de 6% ou  
—10 menos, cerca de 5% ou menos, cerca de 4% ou menos, cerca de 3% ou menos, ou cerca de 2% ou menos. Em alguns exemplos, os loci relevantes (e.g., um locus marcador e um locus alvo, tal como um QTL) apresentam uma frequência de recombinação de cerca de 1% ou menos, e.g., cerca de 0,75% ou menos, cerca de  
15 0,5% ou menos, ou cerca de 0,25% ou menos. Assim, os loci estão a cerca de 10 cM, 9 cM, 8 cM, 7 cM, 6 cM, 5 cM, 4 cM, 3 cM, 2cM, 1cM, 0,75 cM, 0,5 cM ou 0,25 cM ou menos distantes um do outro. Colocado de outra maneira, dois loci que estão localizados no mesmo cromossomo, e em uma tal distância que a  
—20 recombinação entre os dois loci ocorre em uma frequência de não menos de que 10% (e.g., cerca de 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, 0,75%, 0,5%, 0,25%, ou menos), são ditos pode serem proximais um do outro.

25 Quando referindo-se à relação entre dois elementos genéticos, tais como um elemento genético contribuindo para tolerância e um marcador proximal, a fase de "pareamento" de ligação indica o estado onde o alelo "favorável" no locus de tolerância está fisicamente associado na mesma banda do

cromossomo, como o alelo "favorável" do respectivo locus marcador ligado. Na fase de pareamento, ambos alelos favoráveis são herdados pela progênie que herda aquela banda do cromossomo. Na fase de "repulsão" de ligação, o alelo "favorável" no locus de interesse (e.g., um QTL para tolerância) estão fisicamente ligado com um alelo "não favorável" no locus marcador proximal, e os dois alelos "favoráveis" não são herdados juntos (i.e., os dois loci são "desconectados" um com o outro).

-10

Opcionalmente, um, dois, três ou mais alelo(s) favorável(is) são identificados na, ou introduzidos na planta. Vários alelos marcadores podem ser selecionados ou contra-selecionados durante MAS. São identificadas plantas ou germoplasma que possuem pelo menos um alelo favorável destes que se correlaciona positivamente com tolerância ou tolerância aumentada. Contudo, é útil para propósitos de exclusão durante o cruzamento identificar alelos que se correlacionam negativamente com tolerância, para eliminar tais plantas ou germoplasma de subseqüentes rodadas de cruzamento.

-20

A identificação de alelos marcadores favoráveis é específica de germoplasma. A determinação de quais alelos marcadores se correlacionam com tolerância (ou não tolerância) é determinada para o germoplasma em particular em estudo. Um especialista reconhece que métodos para identificar os alelos favoráveis são rotina e bem conhecidos e, adicionalmente, que a identificação e uso de tais alelos favoráveis está bem dentro do escopo da invenção.

*Primers* ou oligos de amplificação para amplificar loci marcadores e sondas marcadoras adequadas para detectar loci marcadores ou para genotipar alelos SNP são providos.

5 Opcionalmente, outras sequências para qualquer dos lados dos dados *primers* podem ser usadas no lugar dos dados *primers*, contanto que os *primers* possam amplificar uma região que inclui o alelo a ser detectado. Ainda, será ponderado que a sonda precisa a ser usada para a detecção pode variar, e.g.,  
10 qualquer sonda que possa identificar a região de um *amplicon* marcador a ser detectada pode ser substituída por aqueles exemplos aqui providos. A configuração dos *primers* de amplificação de sondas de detecção pode, é claro, variar. Assim, a invenção não está limitada aos *primers* e sondas aqui  
15 citados especificamente.

Em alguns exemplos, a presença de loci marcadores é diretamente detectada em DNA genômico não amplificado realizando-se um *Southern blot* em uma amostra de DNA genômico  
20 usando sondas para os loci marcadores. Procedimentos para realizar *Southern blotting*, amplificação (PCR, LCR, ou similares) e vários outros métodos de detecção de ácidos nucléicos estão bem estabelecidos e ensinados, e.g., em Sambrook et al., Molecular Cloning - A Laboratory Manual (3d  
25 ed.), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 2000 ("Sambrook"); Current Protocols in Molecular Biology, F.M. Ausubel et al., eds., Current Protocols, a joint venture entre Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (supplemented through 2002)

("Ausubel")) e PCR Protocols A Guide to Methods and Applications (Innis et al. eds) Academic Press Inc. San Diego, CA (1990) (Innis). Detalhes adicionais em relação à detecção de ácidos nucleicos em plantas podem ser encontrados em, e.g.,  
5 Plant Molecular Biology (1993) Croy (ed.) BIOS Scientific Publishers, Inc.

Sondas de detecção separadas podem ser também omitidas em métodos de amplificação/detecção, e.g., realizando-se uma  
10 reação de amplificação em tempo real que detecte a formação de produtos por modificação do primer de amplificação relevante com a incorporação em um produto, incorporação de nucleotídeos rotulados em um *amplicon*, ou monitorar mudanças nas propriedades moleculares de rotação de *amplicons* se comparado  
15 a precursores não amplificados (e.g., por polarização de fluorescência).

Tipicamente, marcadores moleculares são detectados por qualquer método estabelecido disponível, incluindo, sem  
20 limitação, hibridização alelo específica (ASH) ou outros métodos para detectar polimorfismos de nucleotídeos simples (SNP), detecção polimorfismo de comprimento de fragmento amplificado (AFLP), detecção de sequência amplificada variável, detecção de DNA polimórfico randomicamente  
25 amplificado (RAPD), detecção de polimorfismo de comprimento de fragmento de restrição (RFLP), detecção de replicação de sequência auto-sustentável, detecção de repetições de sequências simples (SSR), detecção de polimorfismos de conformação de fita simples (SSCP), detecção de isoenzimas

marcadoras, ou similares. Enquanto os marcadores aqui  
providos como exemplos nas tabelas são tanto marcadores SSR ou  
SNP (ASH), quaisquer dos tipos de marcadores mencionados  
anteriormente podem ser empregados para identificar segmentos  
5 cromossômicos englobando um elemento genético que contribua  
para desempenho agrônomo superior (e.g., tolerância ou  
tolerância aumentada).

10 Em outro exemplo, a presença ou ausência de um marcador  
molecular é determinada por sequenciamento de nucleotídeo da  
região marcadora polimórfica. Esse método é prontamente  
adaptado a análise de alta produção assim como são outros  
métodos notados acima, e.g., usando métodos de sequenciamento  
de alta produção, tais como, sequenciamento por hibridização.

15

Em geral, a maioria dos marcadores genéticos confia em  
uma ou mais propriedades de ácidos nucléicos para sua  
detecção. Por exemplo, algumas técnicas para detectar  
marcadores genéticos utilizam hibridização de um ácido  
20 nucléico sonda para ácidos nucléicos correspondendo ao  
marcador genético (e.g., ácidos nucléicos amplificados  
produzidos usando DNA genômico de soja como modelo). Formatos  
de hibridização, incluindo, mas não se limitando a, fase de  
solução, fase sólida, fase mista, ou testes de hibridização *in*  
25 *situ* são úteis para a detecção de alelos. Um guia extensivo  
para a hibridização de ácidos nucléicos é encontrado em  
Tijssen (1993) Laboratory Techniques in Biochemistry and  
Molecular Biology--Hybridization with Nucleic Acid Probes

Elsevier, New York, assim como em Sambrook, Berger and Ausubel.

Por exemplo, marcadores que compreendem polimorfismos de comprimento de fragmento de restrição (RFLP) são detectados, e.g., hibridizando-se uma sonda que é tipicamente um subfragmento (ou um oligonucleotídeo sintético correspondendo a um subfragmento) do ácido nucléico a ser detectado para DNA genômico digerido por restrição. A enzima de restrição é selecionada para prover fragmentos de restrição de pelo menos dois comprimentos alternativos (ou polimórficos) em diferentes indivíduos ou populações. Determinar uma ou mais enzimas de restrição que produzam fragmentos informativos para cada cruzamento é um procedimento simples. Após a separação por comprimento em uma matriz apropriada (e.g., agarose, poliacrilamida, etc.) e transferir para uma membrana (e.g., nitrocelulose, nylon, etc.), a sonda rotulada é hibridizada sob condições que resultam em ligação de equilíbrio da sonda ao alvo seguido por remoção de sonda em excesso por lavagem.

—20

Sondas de ácidos nucléicos para os loci marcadores podem ser clonadas e/ou sintetizadas. Qualquer rótulo adequado pode ser usado com uma sonda. Rótulos adequados para uso com sondas de ácidos nucléicos incluem, por exemplo, qualquer composição detectável por meios espectroscópicos, radioisotópicos, fotoquímicos, bioquímicos, imunoquímicos, elétricos, óticos ou químicos. Rótulos úteis incluem biotina para coloração com estreptavidina conjugada rotulada, anéis magnéticos, pigmentos fluorescentes, radio-rótulos, enzimas, e

rótulos colorimétricos. Outros rótulos incluem ligantes que se ligam a anticorpos rotulados com fluoróforos, agentes químicos luminescentes, e enzimas. Uma sonda pode também constituir *primers* de PCR radio-rotulados que são usados para gerar um *amplicon* radio-rotulado. Métodos e reagentes para para rotular ácidos nucleicos e estratégias de detecção correspondentes podem ser encontradas, e.g., em Haugland (1996) Handbook of Fluorescent Probes and Research Chemicals Sixth Edition by Molecular Probes, Inc. (Eugene OR); ou Haugland (2001) Handbook of Fluorescent Probes and Research Chemicals Eighth Edition by Molecular Probes, Inc. (Eugene OR).

#### Métodos de Detecção baseados em Amplificação

PCR, RT-PCR e LCR estão particularmente em amplo uso como métodos de amplificação e amplificação-detecção para amplificar ácidos nucleicos de interesse (e.g., aqueles compreendendo loci marcadores), facilitar a detecção dos marcadores. Detalhes em relação ao uso desses e outros métodos de amplificação podem ser encontrados em qualquer um de uma variedade de textos padrão, incluindo, e.g., Sambrook, Ausubel, Berger and Croy, *supra*. Vários textos de biologia disponíveis também possuem discussões extensas em relação a PCR e métodos de amplificação relacionados. Qualquer RNA pode ser convertido em uma fita dupla de DNA adequada para digestão por restrição, expansão por PCR e sequenciamento usando transcriptase reversa e um polimerase ("PCR de transcrição

reversa, ou "RT-PCR"). Ver também Ausubel, Sambrook and Berger, *supra*.

### Métodos de Amplificação/ Detecção em Tempo Real

5

Em um aspecto, PCR ou LCR em tempo real é realizado nas misturas de amplificação aqui descritas, e.g., usando dispositivos de sinalização moleculares ou sondas TaqMan™. Um dispositivo de sinalização molecular (MB) é um

10 oligonucleotídeo ou ácido nucléico peptídico (PNA) que, sob condições de hibridização apropriadas, se auto-hibridiza para formar um ramo e uma estrutura em *loop*. O MB possui um rótulo e um repressor no terminal do oligonucleotídeo ou PNA; assim, sob condições que permitem hibridização intra-molecular, o

15 rótulo é tipicamente reprimido (ou ao menos alterado em sua fluorescência) pelo repressor. Sob condições onde o MB não apresenta hibridização intra-molecular (e.g., quando ligando a um ácido nucléico alvo, e.g., a uma região de um *amplicon* durante a amplificação), o rótulo de MB é "desreprimido" e o

20 sinal detectado. Métodos padrão de fazer e usar MBs são bem conhecidos e MBs e reagentes estão comercialmente disponíveis Ver também, e.g., Leone et al. (1995) "Molecular beacon probes combined with amplification by NASBA enable homogenous real-time detection of RNA." Nucleic Acids Res. 26:2150-2155; Tyagi

25 and Kramer (1996) "Molecular beacons: probes that fluoresce upon hybridization" Nature Biotechnology 14:303-308; Blok and Kramer (1997) "Amplifiable hybridization probes containing a molecular switch" Mol Cell Probes 11:187-194; Hsuih et al. (1997) "Novel, ligation-dependent PCR assay for detection of

hepatitis C in serum" J Clin Microbiol 34:501-507; Kostrikis et al. (1998) "Molecular beacons: spectral genotyping of human alleles" Science 279:1228-1229; Sokol et al. (1998) "Real time detection of DNA:RNA hybridization in living cells" Proc. Natl. Acad. Sci. U.S.A. 95:11538-11543; Tyagi et al. (1998) "Multicolor molecular beacons for allele discrimination" Nature Biotechnology 16:49-53; Bonnet et al. (1999) "Thermodynamic basis of the chemical specificity of structured DNA probes" Proc. Natl. Acad. Sci. U.S.A. 96:6171-6176; Fang et al. (1999) "Designing a novel molecular beacon for surface-immobilized DNA hybridization studies" J. Am. Chem. Soc. 121:2921-2922; Marras et al. (1999) "Multiplex detection of single-nucleotide variation using molecular beacons" Genet. Anal. Biomol. Eng. 14:151-156; e Vet et al. (1999) "Multiplex detection of four pathogenic retroviruses using molecular beacons" Proc. Natl. Acad. Sci. U.S.A. 96:6394-6399. Ver também, e.g., Patente US 5.925.517 (20 de julho de 1999) para Tyagi et al. intitulada "Detectably labeled dual conformation oligonucleotide probes, assays and kits;" Patente US 6.150.097 para Tyagi et al. (21 de novembro de 2000) intitulada "Nucleic acid detection probes having non-FRET fluorescence quenching and kits and assays including such probes" e Patente US 6.037.130 para Tyagi et al. (14 de março de 2000), intitulada "Wavelength-shifting probes and primers and their use in assays and kits."

A detecção e quantificação por PCR usando sondas de oligonucleotídeo duplamente rotuladas fluorogênicas pode ser feita, por exemplo, usando sondas "TaqMan™". Essas sondas são

compostas de oligodesoxinucleotídeos curtos (e.g., 20-25 bases) que são rotulados com dois tipos diferentes de pigmentos fluorescentes. No terminal 5' de cada sonda está um pigmento repórter, e no terminal 3' de cada sonda um pigmento repressor é encontrado. A sequência de oligonucleotídeos sonda é complementar a uma sequência alvo interna presente em um *amplicon* de PCR. Quando a sonda está intacta, ocorre transferência de energia entre os dois fluoróforos e a emissão do repórter é reprimida pelo repressor por FRET. Durante a fase de extensão do PCR, a sonda é clivada por atividade de nuclease 5' da polimerase usada na reação, liberando, portanto, o repórter do repressor de oligonucleotídeo e produzindo um aumento na intensidade de emissão do repórter. Conseqüentemente, sondas TaqMan™ são oligonucleotídeos que possuem um rótulo e um repressor, onde o rótulo é liberado durante a amplificação pela ação de exonuclease da polimerase usada na amplificação. Isso provê uma medida em tempo real de amplificação durante a síntese. Uma variedade de reagentes TaqMan™ são comercialmente disponíveis, e.g., da Applied Biosystems (Division Headquarters em Foster City, CA), assim como de uma variedade de vendedores especializados, tais como Biosearch Technologies (e.g., sondas repressoras *black hole*).

Detalhes adicionais em relação a sequências amplificadas variáveis, SSR, AFLP, ASH, SNPs e marcadores de isoenzimas

Sequências amplificadas variáveis referem-se a sequências amplificadas do genoma de planta, que exibem alta variabilidade de resíduos de ácidos nucléicos entre membros da

mesma espécie. Todos os organismos possuem seqüências genômicas variáveis e cada organismo (à exceção de um clone) possui um diferente grupo de seqüências variáveis. Uma vez identificadas, a presença de seqüências variáveis específicas pode ser usada para prever traços fenotípicos. Tipicamente, o DNA da planta serve como modelo para a amplificação com *primers* que flanqueiam uma seqüência variável de DNA. A seqüência variável é amplificada e, então, seqüenciada.

Alternativamente, replicação de seqüências auto-sustentável pode ser usada para identificar marcadores genéticos. Replicação de seqüências auto-sustentável refere-se a um método de amplificação de ácidos nucleicos usando seqüências de ácidos nucleicos alvo que são replicadas exponencialmente *in vitro* sob condições substancialmente isotérmicas usando três atividades enzimáticas envolvidas em replicação retroviral: (1) transcriptase reversa, (2) Rnase H, e (3) uma RNA polimerase dependente de DNA (Guatelli et al. (1990) Proc Natl Acad Sci USA 87:1874). Imitando a estratégia retroviral de replicação de RNA por meios de intermediários de DNAC, essa reação acumula cópias de DNAC e RNA do alvo original.

Polimorfismos de comprimento de fragmento amplificado (AFLP) podem ser também usados como marcadores genéticos (Vos et al. (1995) Nucl Acids Res 23:4407). A frase "polimorfismo de comprimento de fragmento amplificado" refere-se a fragmentos de restrição selecionados que são amplificados antes ou após a clivagem por uma endonuclease de restrição. A

etapa de amplificação permite uma detecção mais fácil de fragmentos de restrição específicos. AFLP permite a detecção de grandes números de marcadores polimórficos e foi usada para mapeamento genético em plantas (Becker *et al.* (1995) Mol Gen Genet 249:65; and Meksem *et al.* (1995) Mol Gen Genet 249:74).

Hibridização alelo-específica (ASH) pode ser usada para identificar os marcadores genéticos da invenção. A tecnologia ASH é baseada em anelamento estável de um oligonucleotídeo sonda curto, de fita simples, para um ácido nucléico alvo de fita simples completamente complementar. A detecção é através de um rótulo isotópico ou não-isotópico anexado à sonda.

Para cada polimorfismo, duas ou mais sondas ASH diferentes são projetadas para ter seqüências de DNA idênticas, exceto nos nucleotídeos polimórficos. Cada sonda terá homologia exata com uma seqüência alélica de forma que a amplitude de sondas possa distinguir todas as seqüências alélicas alternativas conhecidas. Cada sonda é hibridizada ao DNA alvo. Com o desenho de sonda e condições de hibridização apropriados, uma falta de correspondência de uma única base entre a sonda e o DNA alvo irá prevenir a hibridização. Dessa maneira, apenas uma das sondas alternativas irá hibridizar a uma amostra alvo que é homozigota ou homogênea para um alelo. Amostras que são heterozigotas ou heterogêneas para dois alelos irão hibridizar a ambas as sondas alternativas..

Marcadores ASH são usados como marcadores dominantes onde a presença ou ausência de apenas um alelo é determinada da

hibridização ou falta de hibridização por apenas uma sonda. O alelo alternativo pode ser inferido da falta de hibridização. A sonda ASH e moléculas alvo são, opcionalmente, RNA ou DNA; as moléculas alvo são de qualquer comprimento de nucleotídeos além da sequência que é complementar à sonda; a sonda é projetada para hibridizar com qualquer fita de um DNA alvo; a sonda varia em tamanho para se conformar a várias condições estridentes de hibridização, etc.

PCR permite que a sequência alvo para ASH seja amplificada a partir de baixas concentrações de ácido nucléico em volumes relativamente pequenos. De outra maneira, a sequência alvo de DNA genômico é digerida com uma endonuclease de restrição e separada por tamanho por eletroforese de gel.. Hibridizações ocorrem tipicamente com a sequência alvo ligada à superfície de uma membrana ou, como descrito na Patente U.S. 5.468.613, a sequência sonda ASH pode estar ligada a uma membrana.

Em um exemplo, dados de ASH são obtidos, tipicamente, amplificando-se fragmentos de ácidos nucléicos (*amplicons*) de DNA genômico usando PCR, transferindo o DNA *amplicon* alvo a uma membrana em um formato *dot-blot*, hibridizando uma sonda de oligonucleotídeo rotulada ao *amplicon* alvo, e observando os pontos de hibridização por auto-radiografia.

Polimorfismos de nucleotídeo simples (SNP) são marcadores que consistem de uma sequência compartilhada diferenciada na base de um único nucleotídeo. Tipicamente, essa distinção é

detectada por padrões de migração diferencial de um *amplicon* compreendendo o SNP em e.g., um gel de acrilamida. Contudo, modos alternativos de detecção, tais como hibridização, e.g., ASH, ou análise RFLP são também apropriados.

5

Marcadores de isoenzimas podem ser aqui empregados como marcadores genéticos, e.g., para rastrear marcadores que não os marcadores de tolerância, ou para rastrear marcadores de isoenzimas ligados aos marcadores daqui. Isoenzimas são formas múltiplas de enzimas que diferem um do outro em seus aminoácidos e, portanto, suas seqüências de ácidos nucleicos. Isoenzimas são formas múltiplas de enzimas que diferem umas das outras em sua seqüência de aminoácidos e, portanto, sua seqüência de ácidos nucleicos. Algumas isoenzimas são enzimas multiméricas contendo sub-unidades levemente diferentes. Outras isoenzimas são tanto multiméricas ou monoméricas, mas foram clivadas a partir da pró-enzima em diferentes sítios na seqüência de aminoácidos. Isoenzimas podem ser caracterizadas e analisadas no nível de proteína, ou alternativamente, isoenzimas que diferem no nível de ácido nucleico podem ser determinadas. Em tais casos, qualquer dos métodos baseados em ácidos nucleicos aqui descritos pode ser usado para analisar marcadores de isoenzimas.

## 25 Métodos de Síntese de Sondas/*Primers*

Em geral, métodos sintéticos para fazer oligonucleotídeos, incluindo sondas, *primers*, dispositivos sinalizadores moleculares, PNAs, LNAs (ácidos nucleicos

trancados), etc., são bem conhecidos. Por exemplo, oligonucleotídeos podem ser sintetizados quimicamente de acordo com o método de fosforamidita triéster em fase sólida descrito por Beaucage and Caruthers (1981), Tetrahedron Letts, 22(20):1859-1862, e.g., usando um sintetizador automático comercialmente disponível, e.g., como descrito em Needham-VanDevanter et al. (1984) Nucleic Acids Res., 12:6159-6168. Oligonucleotídeos, incluindo oligonucleotídeos modificados podem ser solicitados de uma variedade de fontes comerciais conhecidas de especialistas. Existem vários fornecedores comerciais de serviços de oligo síntese e, dessa forma, essa é uma tecnologia amplamente acessível. Qualquer ácido nucléico pode ser pedido de qualquer de uma variedade de fontes comerciais, tais como The Midland Certified Reagent Company (mcrc@oligos.com), The Great American Gene Company (genco.com), ExpressGen Inc. (expressgen.com), Operon Technologies Inc. (Alameda, CA) e várias outras. Similarmente, PNAs podem ser solicitados sob medida de qualquer de uma variedade de fontes comerciais, tais como PeptidoGenic (pkim@ccnet.com), HTI Bio-products, inc. (htibio.com), BMA Biomedicals Ltd (U.K.), Bio. Synthesis, Inc., e várias outras.

#### Deteccção de marcadores *in silico*

25

Em avanços alternativos, métodos *in silico* podem ser usados para detectar os loci marcadores de interesse. Por exemplo, a sequência de um ácido nucléico compreendendo o locus marcador de interesse pode ser armazenada em um

computador. A sequência lócus marcadora desejada ou sua homóloga pode ser identificada usando um algoritmo de busca de ácidos nucleicos apropriado como provido, por exemplo, em tais programas prontamente disponíveis como BLAST, ou até simples  
5 processadores de palavras.

#### PRIMERS DE AMPLIFICAÇÃO PARA DETECÇÃO DE MARCADORES

Em alguns exemplos, marcadores moleculares são detectados  
10 usando um método de detecção adequado baseado em PCR, onde o tamanho ou sequência do *amplicon* de PCR é indicativo da ausência ou presença do marcador (e.g., um alelo marcador em particular). Nesses tipos de métodos, *primers* de PCR são hibridizados às regiões conservadas flanqueando a região  
15 polimórfica marcadora. *Primers* adequados podem ser projetados usando qualquer método adequado. Não se pretende que a invenção esteja limitada a qualquer *primer* ou par de *primers* em particular. Por exemplo, os *primers* podem ser projetados usando qualquer *software* adequado, tal como LASERGENE®.

Em alguns exemplos, os *primers* da invenção são radio-  
rotulados, ou rotulados por qualquer modo adequado (e.g., usando uma marca fluorescente não-radioativa), para permitir rápida visualização dos *amplicons* de diferentes tamanhos  
25 seguindo uma reação de amplificação sem qualquer etapa de rotulação adicional ou etapa de visualização. Em algumas incorporações, os *primers* não são rotulados, e os *amplicons* são visualizados seguindo seu tamanho de resolução, e.g., seguindo eletroforese de gel de agarose. Em alguns exemplos,

coloração por brometo de etídio dos *amplicons* de PCR seguindo resolução de tamanho permite a visualização de *amplicons* de diferentes tamanhos.

5 Os *primers* usados para amplificar os loci e alelos marcadores daqui não estão limitados a amplificar toda a região do *locus* relevante. Em alguns exemplos, a amplificação de marcador produz um *amplicon* de pelo menos 20 nucleotídeos em comprimento, ou alternativamente, pelo menos 50  
10 nucleotídeos em comprimento, ou alternativamente, pelo menos 100 nucleotídeos em comprimento, ou alternativamente, pelo menos 200 nucleotídeos em comprimento, ou até e incluindo o comprimento total do *amplicon*.

## 15 SELEÇÃO ASSISTIDA POR MARCADOR E MELHORAMENTO DE PLANTAS

Uma motivação primária para o desenvolvimento de marcadores moleculares em espécies de cultivo é o potencial para eficiência aumentada no melhoramento de plantas através  
20 de seleção assistida por marcador (MAS). Marcadores genéticos são usados para identificar plantas que contêm um *genótipo* desejado em um ou mais loci, e que são esperados para transferir o *genótipo* desejado, junto com um *fenótipo* desejado para sua *progênie*. Marcadores genéticos podem ser usados para  
25 identificar plantas que contêm um *genótipo* desejado em um *locus*, ou em vários loci ligados ou não-ligados (e.g., um *haplótipo*), e que seriam esperadas de transferir o *genótipo* desejado, junto com um *fenótipo* desejado a sua *progênie*. Meios para identificar plantas, particularmente plantas de

soja, que são tolerantes, ou que exibem tolerância aumentada a inibidores de protoporfirinogênio oxidase são providos, por exemplo, indentificando plantas tendo loci marcadores especificados e.g., para o grupamento de ligação L: SATT495, 5 P10649C-3, SATT182, SATT388, SATT313, and SATT613 (e outros marcadores acima de SATT613) e/ou para o grupamento de ligação N: Sat\_379, SCT\_195, SATT631, S60167-TB, SATT675, SATT624, SATT080, e SATT387 (e outros marcadores acima de SATT387). Similarmente, identificando-se plantas sem o locus marcador 10 desejado, não tolerantes ou menos tolerantes, plantas podem ser identificadas, e, e.g., eliminadas de cruzamentos subsequentes. Similarmente, esses loci marcadores pode ser introduzidos em qualquer base genômica desejada, germoplasma, planta, linhagem, variedade, etc., como parte de um programa 15 de reprodução por MAS total projetado para intensificar a produção de soja.

Em geral, a aplicação de MAS usa a identificação de uma população de plantas tolerantes e mapeamento genético do traço 20 de tolerância. Loci polimórficos na vizinhança do traço de tolerância mapeado são escolhidos como potenciais marcadores de tolerância. Tipicamente, um locus marcador mais próximo do locus de tolerância é um marcador preferido. Análise de ligação é, então, usada para determina qual sequência de alelo 25 marcador polimórfica demonstra uma probabilidade estatística de co-segregação com o fenótipo tolerante (assim, um "alelo marcador de tolerância"). Após a identificação de um alelo marcador para co-segregação com o alelo de tolerância, é possível usar esse marcador para rastreamento rápido e acurado

de linhas de plantas para o alelo de tolerância sem a necessidade de se criar plantas através de seu ciclo de vida e esperar avaliações fenotípicas, e, adicionalmente, permite a seleção genética para o alelo de tolerância em particular, mesmo quando a identidade molecular do QTL de tolerância propriamente dito é anônima. Amostras de tecidos podem ser retiradas, por exemplo, da primeira folha da planta e rastreadas com o marcador molecular apropriado, e em dias fica determinado qual progênie irá avançar. Marcadores ligados também removem o impacto de fatores ambientais que podem frequentemente influenciar a expressão fenotípica.

Após um fenótipo desejado (e.g., tolerância a inibidores de protoporfirinogênio oxidase) e um locus marcador cromossômico polimórfico serem determinado para co-segregarem, o locus marcador polimórfico pode ser usado para selecionar alelos marcadores que segregam com o fenótipo de tolerância desejado. Esse processo geral é tipicamente chamado de seleção assistida por marcador (MAS). Resumidamente, um ácido nucléico correspondendo ao ácido nucléico marcador é detectado em uma amostra biológica de uma planta a ser selecionada. Essa detecção pode tomar a forma de hibridização de um ácido nucléico sonda a um alelo marcador ou *amplicon* deste, e.g., usando hibridização alelo específica, análise *Southern*, análise *northern*, hibridização *in situ*, hibridização de *primers* seguida por amplificação por PCR de uma região do marcador, ou similares. Após a presença (ou ausência) de um marcador em particular na amostra biológica ser verificada, a

planta é selecionada, e.g., usada para fazer plantas descendentes por reprodução seletiva.

Criadores de plantas de soja desejam combinações de loci  
5 de tolerância com genes para alto rendimento e outros traços  
desejados para desenvolver variedades melhoradas de soja.  
Rastrear grandes números de amostras por métodos não  
moleculares (e.g., avaliação de traços em plantas de soja)  
pode ser caro, consumir tempo, e não ser confiável. O uso dos  
10 marcadores polimórficos aqui descritos, quando geneticamente  
ligados a loci de tolerância, provê um método efetivo para  
selecionar variedades resistentes em programas de  
melhoramento. Por exemplo, uma vantagem da seleção assistida  
por marcador sobre avaliações em campo para tolerância é que  
15 MAS pode ser feita em qualquer época do ano, independente da  
temporada de cultivo. Além do mais, efeitos ambientais são  
largamente irrelevantes para a seleção assistida por marcador.

Quando uma população está segregando para múltiplos loci  
20 afetando um ou múltiplos traços, e.g., loci múltiplos  
envolvidos em tolerância, ou loci múltiplos, cada qual  
envolvido com tolerância ou tolerância a diferentes doenças, a  
eficiência de MAS comparada ao rastreamento fenotípico se  
torna ainda maior, porque todos os loci podem ser avaliados no  
25 laboratório juntos a partir de uma única amostra de DNA. No  
presente exemplo, para o grupamento de ligação L: SATT495,  
P10649C-3, SATT182, SATT388, SATT313, e SATT613 (ou outros  
marcadores acima de SATT613) e para o grupamento de ligação N:  
Sat\_379, SCT\_195, SATT631, S60167-TB, SATT675, SATT624,

SATT080, e SATT387 (ou outros marcadores acima de SATT387) marcadores podem ser testados simultaneamente ou sequencialmente em uma única amostra ou população de amostras.

5        Outro uso de MAS no melhoramento de plantas é prestar assistência à recuperação dos genótipos parentais recorrentes por cruzamento reverso. Cruzamento reverso é o processo de cruzar uma progênie de volta a um de seus pais ou linhagens parentais. Cruzamento reverso é usualmente feito para o  
10   propósito de introduzir um ou alguns loci de um pai doador (e.g., um pai compreendendo loci marcadores de tolerância desejados) em antecedente genético desejado do pai recorrente (e.g., uma linhagem de soja, de outra maneira, de alta produção). Quanto mais ciclos de cruzamento reverso são  
15   feitos, maior é a contribuição genética do pai recorrente à variedade introduzida resultante. Isso é frequentemente necessário, porque plantas tolerantes podem ser, de outra forma, indesejáveis, e.g., devido à baixa produção, baixa fecundidade, ou similar. Em contraste, linhagens que são o  
-20   resultado de programas intensivos de reprodução podem ter excelente produção, fecundidade ou similar, meramente sendo deficiente em um traço desejado, tal como tolerância a inibidores de protoporfirinogênio oxidase.

25        A presença e/ou ausência de um marcador genético ou alelo em particular, e.g., para o grupamento de ligação L: SATT495, P10649C-3, SATT182, SATT388, SATT313, e SATT613 (incluindo marcadores acima de SATT613) e para o grupamento de ligação N: Sat\_379, SCT\_195, SATT631, S60167-TB, SATT675, SATT624,

SATT080, e SATT387 (incluindo marcadores acima de SATT387) no genoma de uma planta exibindo um traço fenotípico preferido é feita por qualquer método aqui anotado. Se os ácidos nucleicos da planta são positivos para um alelo marcador genético desejado, a planta pode ser auto-fertilizada para criar uma linhagem de cruzamento verdadeira com o mesmo genótipo, ou pode ser cruzada com uma planta com o mesmo marcador ou com outras características desejadas para criar uma geração híbrida cruzada sexualmente.

-10

#### Introdução de Alelos Favoráveis - Cruzamento Eficiente de Marcadores de Tolerância em Outras Linhagens

Uma aplicação de MAS é usar os marcadores de tolerância ou tolerância aumentada para aumentar a eficiência de uma introdução ou esforço de cruzamento reverso visado quando introduzindo um QTL de tolerância em um antecedente desejado (tipicamente com alta produção). No cruzamento reverso de marcadores específicos assistido por marcador (e QTL associados) de uma fonte doadora, e.g., para um antecedente genético de elite ou exótico, seleciona-se entre a progênie de cruzamentos reversos para o traço doador.

-20

Assim, os marcadores e métodos podem ser utilizados para guiar a seleção assistida por marcador ou cruzamento de variedades de soja como o complemento desejado (grupo) de formas alélicas de segmentos cromossômicos associados com performance agrônômica superior (tolerância, junto com qualquer outro marcador disponível para produção, tolerância a

doenças, etc.). Qualquer dos alelos marcadores descritos pode ser introduzido em uma linhagem de soja via introdução, por cruzamento tradicional (ou introduzido via transformação, ou ambos) para produzir uma planta de soja performance agrônômica superior. O número de alelos associados com tolerância que podem ser introduzidos, ou estar presente em uma planta de soja da presente invenção, varia de 1 ao número de alelos aqui descritos, cada número inteiro dos quais é aqui incorporados como se explicitamente citados.

10

Métodos de fazer uma planta de soja descendente e essas plantas de soja descendentes tendo tolerância a inibidores de PPO são providos. Esses métodos compreendem cruzar uma primeira planta de soja parental com uma segunda planta de soja e cultivar a planta de soja fêmea sob condições de cultivo para produzir plantas de soja descendentes. Tal progênie de planta de soja pode ser testada para alelos associados com tolerância e, dessa forma, a progênie desejada ser selecionada. Tais plantas descendentes ou sementes podem ser vendidas comercialmente para produção de soja, usadas para alimentação, processadas para obter um constituinte desejado da soja, ou ainda utilizadas em subseqüentes rodadas de cruzamento. Ao menos uma da primeira ou segunda plantas de soja é uma planta de soja compreendendo pelo menos uma das formas alélicas dos marcadores providos, de forma que a progênie seja capaz de herdar o alelo.

20

25

A herdabilidade do alelo de tolerância desejado pode ser rastreada, tal como a partir de linhagens progenitoras ou

descendentes no pedigree de plantas de soja alvo, de forma que o número de gerações separando as plantas de soja sendo sujeito aos métodos da presente invenção será geralmente de 1 a 20, comumente 1 a 5, e tipicamente 1, 2, ou 3 gerações de separação, e muito frequentemente um descendente direto ou pai direto da planta de soja será o objeto do método (*i.e.*, 1 geração de separação).

#### MÉTODOS PARA IDENTIFICAR PLANTAS DE SOJA TOLERANTES A INIBIDOR DE PROTOPORFIRINOGÊNIO

Cultivadores de plantas experientes podem reconhecer plantas de soja tolerantes no campo, e podem selecionar os indivíduos ou populações tolerantes para propósitos de melhoramento ou para propagação. Neste contexto, o cultivador reconhece plantas de soja tolerantes e não tolerantes.

O rastreamento e seleção podem ser também efetuados expondo as plantas contendo dito germoplasma progênie a inibidores de protoporfirinogênio oxidase em um teste e selecionando aquelas plantas apresentando tolerância a inibidores de protoporfirinogênio oxidase como contendo germoplasma de soja no qual o germoplasma tendo tolerância a inibidores de protoporfirinogênio oxidase derivado do QTL mapeado no grupamento de ligação L e/ou N foi introduzido. O teste *in vivo* pode ser qualquer teste conhecido na arte, *e.g.*, Taylor-Lovell *et al.* (2001) Weed Tech 15:95-102.

Contudo, tolerância de plantas é um espectro fenotípico consistindo de extremos de alta tolerância até não tolerância com um contínuo de fenótipos de tolerância intermediários. A avaliação desses fenótipos intermediários usando testes reproduzíveis são de valor para cientistas que procuram identificar loci genéticos que acarretam tolerância, conduzir seleção assistida por marcador para a população tolerante, e para técnicas de introdução para criar um traço de tolerância em uma linha de soja de elite, por exemplo. Descrever o contínuo de tolerância pode ser feito usando qualquer sistema de pontuação conhecido ou derivado desse, incluindo os sistemas de pontuação descritos nos Exemplos 1-4.

#### SISTEMAS DE DETECÇÃO/CORRELAÇÃO AUTOMÁTICOS

Em alguns exemplos, os métodos incluem um sistema automático para detectar marcadores e ou correlacionar os marcadores com um fenótipo desejado (e.g., tolerância). Dessa forma, um sistema típico pode incluir um grupo de sondas marcadoras ou *primers* configurados para detectar pelo menos um alelo favorável de um ou mais loci marcadores associados com tolerância ou tolerância aumentada a inibidores de protoporfirinogênio oxidase. Essas sondas ou *primers* são configuradas para detectar os alelos marcadores notados nas tabelas e exemplos daqui, e.g., usando qualquer formato de detecção de alelo, e.g., detecção baseada em arranjo de fase sólida ou líquida, detecção baseada em amostra microfluida, etc.

Em alguns exemplos, marcadores envolvendo o grupamento de ligação L são usados. Em alguns exemplos, um marcador proximalmente ligado ao locus marcador de SATT495, P10649C-3, SATT182, SATT388, SATT313, e SATT613 é usado, e a grupo de  
5 sondas é configurado para detectar o(s) marcador(es) proximalmente ligado(s). Em alguns exemplos, o locus marcador é SATT495, P10649C-3, SATT182, SATT388, SATT313, e SATT613 (ou outro marcador acima de SATT613) e o grupo de sondas é configurado para detectar o locus. Similarmente, alelos de  
10 SATT495, P10649C-3, SATT182, SATT388, SATT313, e SATT613 podem ser detectados.

Em alguns exemplos, marcadores envolvendo o grupamento de ligação N são usados. Em alguns exemplos, um marcador  
15 proximalmente ligado ao locus marcador de Sat\_379, SCT\_195, SATT631, S60167-TB, SATT675, SATT624, SATT080, e SATT387 (ou outro marcador acima de SATT387) é usado, e o grupo de sondas é configurado para detectar o(s) marcador(es) proximalmente ligado(s). Em alguns exemplos, o locus marcador é Sat\_379,  
20 SCT\_195, SATT631, S60167-TB, SATT675, SATT624, SATT080, e SATT387 e o grupo de sondas é configurado para detectar o locus. Similarmente, alelos de Sat\_379, SCT\_195, SATT631, S60167-TB, SATT675, SATT624, SATT080, e SATT387 podem ser detectados.

25

O sistema típico inclui um detector que está configurado para detectar um ou mais sinais resultantes do grupo de sondas marcadoras ou *primers*, ou *amplicon* desse, identificando, dessa forma, a presença ou ausência do alelo. Uma ampla variedade de

aparatos de detecção de sinais está disponível, incluindo tubos de foto multiplicadores, espectrofotômetros, arranjos CCD, arranjos e escaners de arranjos, detectores de rastreamento, fototubos e fotodiodos, estações de

5 microscópios, galvo-escaners, dispositivos de detecção de amplificação de ácidos nucleicos microfluida e similares. A configuração precisa do detector dependerá, em parte, do tipo de rótulo usado para detectar o alelo marcador, assim como a instrumentação que é mais convenientemente obtida para o

10 usuário. Detectores que detectem fluorescência, fosforescência, radioatividade, pH, carga, absorbância, luminescência, temperatura, magnetismo ou similares podem ser usados. Exemplos de detectores típicos incluem detectores de luz (e.g., fluorescência) ou detectores de radioatividade.

15 Por exemplo, a detecção de uma emissão de luz (e.g., uma emissão de fluorescência) ou outro rótulo de sonda é indicativa da presença ou ausência de um alelo marcador. A detecção fluorescente é geralmente usada para a detecção de ácidos nucleicos amplificados (contudo, operações acima e/ou

20 abaixo podem ser também efetuadas em *amplicons*, as quais podem envolver outros métodos de detecção). Em geral, o detector detecta uma ou mais emissões de rótulos (e.g., luz) de um rótulo de sonda, o que é indicativo da presença ou ausência de um alelo marcador.

25

O(s) detector(es) opcionalmente monitora um ou uma pluralidade de sinais de uma reação de amplificação. Por exemplo, o detector pode monitorar sinais óticos que

correspondem a resultados de testes de amplificação em "tempo real".

Instruções de sistema que correlacionam a presença ou  
5 ausência do alelo favorável com a tolerância prevista são  
também um aspecto da invenção. Por exemplo, as instruções  
podem incluir ao menos uma tabela de consulta que inclui uma  
correlação entre a presença ou ausência dos alelos favoráveis  
e a tolerância ou tolerância aumentada prevista. A forma  
10 precisa das instruções pode variar dependendo dos componentes  
do sistema, e.g., podem estar presentes como *software* de  
sistema em uma ou mais unidades integradas do sistema (e.g.,  
um microprocessador, computador ou meio legível em  
computador), ou podem estar presentes em uma ou mais unidades  
15 (e.g., computadores ou meio legível em computador)  
operacionalmente combinadas ao detector. Como notado em um  
exemplo típico, as instruções de sistema incluem ao menos uma  
tabela de consulta que inclui uma correlação entre a presença  
ou ausência dos alelos favoráveis e a tolerância ou tolerância  
20 aumentada prevista. As instruções também incluem,  
tipicamente, instruções provendo uma interface de usuário com  
o sistema, e.g., para permitir um usuário visualizar os  
resultados de uma análise de amostras e incluir parâmetros no  
sistema.

25

O sistema tipicamente inclui componentes para armazenar  
ou transmitir dados legíveis em computador representando ou  
designando os alelos detectados pelos métodos da presente  
invenção, e.g., em um sistema automático. O meio legível em

computador pode incluir memória cache, principal, e de armazenamento e/ou outros componentes de armazenamento de dados eletrônicos (discos rígidos, *drives* de disquetes, drives de armazenamento, etc.) para armazenamento de código de computador. Dados representando alelos detectados pelo método da presente invenção podem se transmitidos, também, eletronicamente, oticamente, magneticamente em um sinal de dados de computador incorporado em um meio de transmissão em uma rede, tal como *intranet* ou *internet* ou combinações destas.

10 O sistema pode transmitir, também ou alternativamente, dados via *wireless*, IR, ou outras alternativas de transmissão disponíveis.

Durante a operação, o sistema tipicamente compreende uma amostra que está para ser analisada, tal como um tecido vegetal, ou material isolado do tecido, tal como DNA genômico, DNA genômico amplificado, DNAC, DNAC amplificado, RNA, RNA amplificado, ou similar.

20 A frase "sistema de detecção/ correlação de alelos" refere-se a um sistema no qual dados entrando em um computador correspondem a objetos físicos ou processos externos ao computador, e.g., um alelo marcador, e um processo que, em um computador, causa uma transformação física dos sinais de entrada a diferentes sinais de saída. Em outras palavras, o dado inserido, e.g., amplificação de um alelo marcador em particular é transformado em dado de saída, e.g., a identificação da forma alélica de um segmento cromossômico. O processo no computador é um grupo de instruções, ou

"programa", pelo qual sinais positivos de amplificação ou hibridização são reconhecidos pelo sistema integrado e atribuídos às amostras individuais como um genótipo. Programas adicionais correlacionam a identidade de amostras individuais com valores fenotípicos ou alelos marcadores, e.g., métodos estatísticos. Além disso, existem numerosos programas, e.g., C/C++, para computação, programas Delphi e/ou Java para interfaces GUI, e ferramentas de produtividade (e.g., Microsoft Excel e/ou SigmaPlot) para fazer gráficos ou criar

5 tabelas de consulta de correlações alelo-traço relevantes. Outras ferramentas de *software* úteis no contexto dos sistemas integrados da invenção incluem pacotes estatísticos, tais como SAS, Genstat, Matlab, Mathematica, e S-Plus e pacotes de modelagem genética, tal como QU-GENE. Adicionalmente,

10

15 linguagens de programação adicionais, tais como *visual basic* são também adequadamente empregadas nos sistemas integrados.

Por exemplo, valores de alelos marcadores de tolerância designados a uma população de progênie descendendo de cruzamentos entre linhagens de elites são gravados em um meio legível em computador, estabelecendo, dessa forma, uma base de dados correspondendo a alelos de tolerância com identificadores particulares para membros da população da progênie. Qualquer arquivo ou pasta, seja feita sob encomenda

20

25 ou comercialmente disponível (e.g., da Oracle ou Sybase) adequado para gravar dados em um meio legível em computador é aceitável como uma base de dados no contexto da presente invenção. Dados em relação a genótipos para um ou mais marcadores moleculares, e.g., ASH, SSR, RFLP, RAPD, AFLP, SNP,

marcadores de isoenzimas ou outros marcadores, como aqui descrito acima, são similarmente gravados em uma base de dados acessível por computador. Opcionalmente, dados de marcadores são obtidos usando um sistema integrado que automatiza um ou  
5 mais aspectos do teste (ou testes) usado para determinar genótipos marcadores. Em tal sistema, dados de entrada correspondendo a genótipos para marcadores moleculares são revezados de um detector, e.g., uma ordenação, um escaner, um CCD, ou outro dispositivo de detecção diretamente para  
10 arquivos em um meio legível em computador acessível à unidade central de processamento. Um grupo de instruções de sistema (tipicamente incorporadas em um ou mais programas) codificando as correlações entre tolerância e os alelos da invenção é, então, executado pelo dispositivo computacional para  
15 identificar correlações entre alelos marcadores e traços fenótipos previstos.

Tipicamente, o sistema inclui, também, um dispositivo de inserção de dados pelo usuário, tal como um teclado, um *mouse*,  
20 uma tela de toque, ou similares, para, e.g., selecionar arquivos, corrigir dados, revisar tabelas de informações de marcadores, etc., e um dispositivo de saída (e.g., um monitor, um impressora, etc.) para visualizar ou recuperar o produto das análises estatísticas.

25

Sistemas integrados compreendendo um computador ou meio legível em computador compreendendo grupos de arquivos e/ou uma base de dados com ao menos um grupo de dados que correspondam aos alelos marcadores daqui são aqui providos.

Os sistemas incluem, também, uma interface de usuário permitindo um usuário visualizar seletivamente uma ou mais dessas bases de dados. Além disso, *softwares* padrão de manipulação de textos, tais como *softwares* processadores de palavras (e.g., Microsoft Word™ ou Corel WordPerfect™) e *softwares* de base de dados ou planilhas (e.g., *software* de planilhas, tais como Microsoft Excel™, Corel Quattro Pro™, ou programas de bases de dados, tais como Microsoft Access™ ou Paradox™) podem ser usados em conjunção com a interface de usuário (e.g., um GUI e um sistema operacional padrão, tal como um sistema Windows, Macintosh, Unix ou Linux) para manipular séries de caracteres correspondendo aos alelos ou outros aspectos da base de dados.

Os sistemas opcionalmente incluem componentes para manipulação de amostras, e.g., dispositivos robóticos incorporados. Por exemplo, uma armadura robótica para controle de líquidos para transferir soluções (e.g., extratos de células vegetais) de uma fonte a um destino, e.g., de uma placa de microtítulo para um substrato de ordenação, é opcionalmente operacionalmente ligada ao computador (ou a um computador adicional no sistema integrado). Um dispositivo de inserção para inserir dados no computador para controlar transferências de líquidos de alta produção pela armadura robótica para controle de líquido e, opcionalmente, controlar a transferência pela armadura ao suporte sólido é comumente um aspecto do sistema integrado. Vários tais sistemas robóticos automáticos de manuseio de fluídos são comercialmente disponíveis. Por exemplo, uma variedade de sistemas

automáticos são disponíveis da Caliper Technologies (Hopkinton, MA), os quais utilizam vários sistemas Zymate, e que tipicamente incluem, e.g., módulos robóticos e de manuseio de fluídos. Similarmente, o robô comum ORCA®, que é usado em  
5 uma variedade de sistemas laboratoriais, e.g., para manipulação de bandejas de microtítulo, também está comercialmente disponível, e.g., da Beckman Coulter, Inc. (Fullerton, CA). Como uma alternativa para robótica convencional, sistemas de microfluídos para realizar o  
10 manuseio e detecção de fluídos são também amplamente disponíveis, e.g., da Caliper Technologies Corp. (Hopkinton, MA) e Agilent technologies (Palo Alto, CA).

Sistemas para análise de marcador molecular podem incluir  
15 um computador digital com um ou mais *softwares* de controle de líquido, *softwares* de análise de imagem para analisar dados de rótulos de marcadores, *softwares* de interpretação de dados, armadura robótica para controle de líquido para transferir soluções de uma fonte para um destino operacionalmente ligado  
20 ao computador, um dispositivo de inserção (e.g., um teclado de computador) para entrada de dados no computador para controlar a transferência de líquidos pela armadura robótica para controle de líquido e, opcionalmente, um escaner de imagem para digitalizar sinais de rótulos das sondas hibridizadas  
25 rotuladas, e.g., para marcadores em um suporte sólido operacionalmente ligados ao computador. O escaner de imagem interage com a o software de análise de imagens para prover uma medição de, e.g., intensidade de rótulo de sonda de ácido nucléico em hibridização para uma população amostra de ácidos

nucléicos ordenada (e.g., compreendendo um ou mais marcadores), onde a medição de intensidade de rótulo da sonda é interpretada pelo *software* de interpretação de dados para mostrar se, e em qual grau, a sonda rotulada hibridiza a um

5 ácido nucléico marcador (e.g., um alelo marcador amplificado). Os dados derivados dessa forma são, então, correlacionados com a identidade de amostra, para determinar a identidade de uma planta com um genótipo(s) em particular para marcadores ou

17 alelos em particular, e.g., para facilitar a seleção assistida por marcador de plantas de soja com formas alélicas favoráveis de segmentos cromossômicos envolvidos na performance

10 agrônômica (e.g., tolerância ou tolerância aumentada).

Imagens óticas, e.g., padrões de hibridização

15 visualizados (e, opcionalmente, gravados) por uma câmera ou outro dispositivo de gravação (e.g., um fotodiodo e dispositivo de armazenamento de dados) são opcionalmente processados em qualquer dos avanços aqui, e.g., digitalizando a imagem e/ou armazenando e analisando a imagem em um

20 computador. Uma variedade de equipamentos periféricos e *softwares* comercialmente disponíveis estão disponíveis para digitalizar, armazenar e analisar um vídeo digitalizado ou imagem ótica digitalizada, e.g., usando computadores PC (máquinas baseadas em Intel x86 ou *pentium chip-compatible*

25 DOS™, OS2™ WINDOWS™, WINDOWS NT™ ou WINDOWS95™), computadores baseados em MACINTOSH™, LINUX, ou UNIX (e.g., SUN™ work station).

Também incluídos nesta invenção são plantas de soja produzidas por qualquer dos métodos precedentes.

A presente invenção é ilustrada pelos exemplos a seguir. As descrições precedente e subsequente da presente invenção e dos vários avanços não pretendem limitar a invenção, mas, ao contrário, ser ilustrativa da mesma. Assim, será entendido que a invenção não está limitada aos detalhes específicos desses exemplos.

#### **EXEMPLOS**

**Exemplo 1: Identificação de Linhas de Soja Sensíveis e Tolerantes a Sulfentrazone - bioensaio de rastreamento de herbicida e diagnóstico intergrupo baseado em associação de marcador**

Sulfentrazone é um inibidor de PPO é o ingrediente ativo no herbicida Authority®. Authority® 75DF (FMC Corp., Philadelphia, PA, USA) é uma formulação de ingrediente 75% ativa de sulfentrazone contendo nenhum outro ingrediente ativo.

#### Parte 1: Bioensaio de Herbicida

Cento e dezesseis (116) linhas de soja de elite foram rastreadas para tolerância a sulfentrazone usando o seguinte protocolo. Sementes de variedades de soja com qualidade de

sementes adequadas, tendo mais do que 85% de taxa de germinação foram usadas.

**DESENHO E REPLICAÇÃO:** Após o plantio, entradas foram preparadas em um desenho de bloco completo randomizado, bloqueado por replicação. Foram usadas três replicações por experimento. Uma ou mais de variedades de checagem bem estabelecidas foram incluídas no experimento, tais como linhas de checagem do setor público.

10

Checagem não tolerante: Pioneer 9692, Asgrow A4715

Checagem tolerante: Pioneer 9584, Syngenta S5960

As condições de crescimento foram como a seguir (estufa/câmara de crescimento): fotoperíodo de 16 hr a 85° F (w/ 75° ajuste noturno). A iluminação é crítica para o sucesso do rastreamento como afirmado a seguir.

**MÉTODO DE RASTREAMENTO:** Potes plásticos de quatro polegadas foram preenchidos com um solo de pote universal de alta qualidade. As entradas foram plantadas em uma profundidade de 1 polegada em uma taxa de 5 sementes/pote. Uma tarja plástica com um código de barras foi usada para identificar cada entrada. Após o plantio, os potes foram deixados descansando na estufa de um dia para o outro para se aclimatar ao solo e melhorar a germinação. No dia seguinte, uma solução de herbicida sulfentrazone foi lentamente pulverizada em cada pote e deixada mesmo infiltrar através de todo o perfil do solo. Isso assegurou que cada semente foi exposta a uma

quantidade igual de sulfentrazone. Potes foram posicionados em bandejas de alumínio e colocados em uma estufa ou câmara de crescimento sob condições de luminosidade de alta intensidade com densidade de fluxo de fótons fotossintéticos (PPFD) de pelo menos 500  $\mu\text{mol}/\text{m}^2/\text{s}$ . Condições de luminosidade adequadas foram necessárias para esse rastreamento devido à natureza do inibidor de PPO usado. Os pots foram levemente enxaguados de forma que o herbicida não vazasse do perfil de solo. Após a emergência da soja, os pots foram enxaguados mantendo-se as bandejas de alumínio preenchidas com  $\frac{3}{4}$ " de água sob cada pote.

**SOLUÇÃO DE HERBICIDA:**

- A) Misturar uma solução de estoque de 0,926 g Authority® 75DF (FMC Corp.), cuidadosamente dissolvida em 1000 ml de água.
- B) Misturar 10 ml de SOLUÇÃO DE ESTOQUE em 1000 ml de água para criar a solução final.
- C) Pulverizar 100 ml de SOLUÇÃO FINAL em cada pote.

**GRAVAÇÃO DE DADOS:** 10-14 dias após o tratamento, as plantas estavam prontas para serem pontuadas. Todas as pontuações foram baseadas em uma comparação às checagens e avaliadas como a seguir:

- 9 = Equivalente, ou melhor quando comparado com a checagem tolerante
- 7 = Muito pouco dano ou resposta notados.
- 5 = Resposta ou danos intermediários
- 3 = Dano acentuado, incluindo necrose foliar e atordoamento

1 = Dano severo, incluindo atordoamento severo e necrose; equivalente ou pior quando comparado com a checagem não tolerante

5        Das 116 linhagens de soja rastreadas, 102 apresentaram pelo menos alguma tolerância a herbicidas baseados em sulfentrazone e 11 apresentaram alta sensibilidade. Uma referência relevante para este protocolo seria: Dayan *et al.* (1997) 'Soybean (*Glycine max*) cultivar differences in response  
10 to sulfentrazone' Weed Science 45:634-641.

## Parte 2: Análise Intergrupo

15        Uma análise de "Distribuição de Frequência Intergrupo de Alelos" foi conduzida usando o *software* GeneFlow™ versão 7.0. Uma análise de distribuição de frequência intergrupo de alelos provê um método para encontrar distribuições não aleatórias de alelos entre dois grupos fenotípicos.

20        Durante o processamento, uma tabela de contingência de frequências de alelos foi construída e, a partir dela, foram calculadas a estatística G e a probabilidade. A estatística G foi ajustada usando o fator de correção de William. O valor de probabilidade foi ajustado para levar em consideração o fato  
25 de que múltiplos testes estão sendo feitos (assim, existe uma taxa esperada de falsos positivos). A probabilidade ajustada é proporcional à probabilidade de que as diferenças da distribuição de alelos observada entre as duas classes ocorreriam por acaso sozinhas. Quanto menor o valor da

probabilidade, maior a chance de que o fenótipo de tolerância e o marcador co-segrem. Uma discussão mais completa da derivação dos valores de probabilidade pode ser encontrada na documentação do software GeneFlow™ versão 7.0. Ver também

5 Sokal and Rolf (1981), *Biometry: The Principles and Practices of Statistics in Biological Research*, 2nd ed., San Francisco, W. H. Freeman and Co.

A lógica é a de que marcadores com distribuições de

10 alelos significativamente diferentes entre os grupos tolerantes e não tolerantes (i.e., distribuições não aleatórias) devem estar associados com o traço e podem ser usados para separá-los para propósitos seleção assistida por marcador de linhas de soja com tolerância ou não tolerância a

15 inibidores de protoporfirinogênio oxidase previamente não caracterizados. A presente análise examinou um locus marcador em um momento e determinou se a distribuição de alelos no grupo tolerante é significativamente diferente da distribuição de alelos no grupo não tolerante. Uma distribuição de alelos

20 significativamente diferente é uma indicação de que o marcador está ligado a um locus que é associado com tolerância ou não tolerância a inibidores de protoporfirinogênio oxidase. Nesta análise, probabilidades não ajustadas menores do que um são consideradas significativas (o marcador e o fenótipo

25 apresentam desequilíbrio de ligação), e probabilidades ajustadas menores do que aproximadamente 0,05 são consideradas altamente significativas. Classes de alelos representadas por menos de 5 observações em ambos os grupos não foram incluídas

na análise estatística. Nesta análise, 1043 loci marcadores tinham observações suficientes para a análise.

Esta análise compara a pontuação fenotípica das plantas com os genótipos em vários loci. Este tipo de análise intergrupo nem gera nem necessita de qualquer dado de mapeamento. Subsequentemente, dados de mapeamento (por exemplo, um mapa genético composto de soja) são relevantes já que múltiplos marcadores significativos que também estão geneticamente ligados podem ser considerados como evidência colaborativa de que uma dada região cromossômica está associada com o traço de interesse.

### Resultados

15

A tabela 1, abaixo, provê uma listagem dos marcadores de soja que demonstraram desequilíbrio de ligação com o fenótipo de inibidor de tolerância a protoporfirinogênio oxidase. Foram usados 1043 marcadores nesta análise. Também indicados naquela tabela estão os cromossomos nos quais os marcadores estão localizados e sua posição aproximada no mapa relativa a outros marcadores conhecidos, dada em cM, com a posição zero sendo o primeiro (mais distal) marcador conhecido no início do cromossomo. Essas posições no mapa não são absolutas, e representam uma estimativa da posição no mapa. As probabilidades estatísticas de que o alelo marcador e o fenótipo de tolerância estejam segregando independentemente estão refletidas nos valores de Probabilidade Ajustada. Dos 584 loci estudados em 38 linhas de soja sensíveis e 160 linhas

de soja tolerantes, QTLs em Lg L e em Lg N foram altamente significativos, como mostrado na tabela abaixo.

Tabela 1: Resultados da Análise Intergrupos para os marcadores

5 LgL e LgN

| Lócus     | Teste | Chrom# | Posição | Valor de G | df | Probabil. (G) | Probabil. ajustada |
|-----------|-------|--------|---------|------------|----|---------------|--------------------|
| S00224-1  | GW    | L      | 12,03   | 89,87      | -1 | 0             | 0                  |
| P10649C-3 | ASH   | L      | 3,6     | 86,01      | -1 | 0             | 0                  |
| SATT523   | SSR   | L      | 32,4    | 24,02      | -1 | 0,000001      | 0,000592           |
| S60167-TB | SSR   | N      | 26      | 62,35      | -1 | 0             | 0                  |
| P5467A-1  | ASH   | N      | 25      | 16,25      | -1 | 0,000056      | 0,032192           |
| P5467A-2  | ASH   | N      | 25      | 16,2       | -1 | 0,000057      | 0,032731           |

A Tabela 2 abaixo apresenta a distribuição de alelos entre 101 linhas tolerantes e 32/33 linhas não tolerantes analisadas.

10 Linhas exibindo tolerância estão indicadas na primeira coluna como "TOL," e linha exibindo não tolerância estão indicadas na primeira coluna como "NON." Chamadas de marcador para o lócus P10649C-3 e o lócus S60167-TB lócus estavam disponíveis para 132 e 63 das linhas respectivamente.

15

Tabela 2: Distribuição de alelos

| <b>Fenótipo</b> | <b>alelo<br/>P10649C-3<br/>LG-L</b> | <b>alelo<br/>S60167-TB<br/>LG-N</b> |
|-----------------|-------------------------------------|-------------------------------------|
| TOL             | 1                                   | 1                                   |
| TOL             | 1                                   | 1                                   |
| TOL             | 1                                   |                                     |
| TOL             | 1                                   |                                     |
| TOL             | 1                                   |                                     |
| TOL             | 1                                   | 2                                   |
| TOL             | 1                                   |                                     |
| TOL             | 1                                   |                                     |
| TOL             | 1                                   | 1                                   |
| TOL             | 1                                   |                                     |

|     |   |   |
|-----|---|---|
| TOL | 1 | 1 |
| TOL | 1 |   |
| TOL | 1 | 1 |
| TOL | 1 | 1 |
| TOL | 1 | 1 |
| TOL | 1 | 1 |
| TOL | 1 |   |
| TOL | 1 | 1 |
| TOL | 1 |   |
| TOL | 1 |   |
| TOL | 1 | 1 |
| TOL | 1 | 1 |
| TOL | 1 |   |
| TOL | 1 | 1 |
| TOL | 1 | 1 |
| TOL | 1 | 1 |
| TOL | 1 | 1 |
| TOL | 1 |   |
| TOL | 1 | 1 |
| TOL | 1 | 1 |
| TOL | 1 |   |
| TOL | 1 |   |
| TOL | 1 | 1 |
| TOL | 1 | 1 |
| TOL | 1 |   |
| TOL | 1 | 1 |
| TOL | 1 |   |
| TOL | 1 |   |
| TOL | 1 | 1 |
| TOL | 1 | 1 |
| TOL | 1 | 1 |
| TOL | 1 | 1 |
| TOL | 1 |   |
| TOL | 1 | 1 |
| TOL | 1 | 1 |
| TOL | 1 |   |
| TOL | 1 | 1 |
| TOL | 1 | 1 |
| TOL | 1 |   |
| TOL | 1 | 1 |
| TOL | 1 | 1 |
| TOL | 1 | 1 |
| TOL | 1 |   |
| TOL | 1 | 1 |
| TOL | 1 | 1 |
| TOL | 1 | 1 |
| TOL | 1 |   |
| TOL | 1 | 1 |
| TOL | 1 | 1 |
| TOL | 1 | 1 |
| TOL | 1 |   |
| TOL | 1 |   |



|     |     |     |
|-----|-----|-----|
| NON | 3   |     |
| NON | 1   | 1_2 |
| NON | 3   | 2   |
| NON | 2   | 1_2 |
| NON | 3   | 2   |
| NON | 2   |     |
| NON | 2   | 2   |
| NON | 2   | 2   |
| NON | 1   | 1   |
| NON | 2_3 |     |
| NON | 3   | 2   |
| NON | 3   | 2   |
| NON | 2_3 |     |
| NON | 3   |     |
| NON | 3   |     |
| NON | 1   | 1   |
| NON | 2   |     |
| NON | 3   |     |
| NON | 3   | 2   |
| NON | 3   |     |
| NON | 2   | 2   |
| NON | 3   | 2   |
| NON | 1_3 | 2   |
| NON | 3   |     |
| NON | 2   |     |
| NON | 1   |     |
| NON | 3   |     |

A distribuição não aleatória de alelos entre os grupos de plantas tolerantes e não tolerantes nos loci marcadores na Tabela 2 é uma boa evidência de que um QTL influenciando tolerância a inibidores de protoporfirinogênio oxidase está ligado a esses loci marcadores.

**Exemplo 2: Previsão e confirmação de seleção baseada por marcador para resposta a químicos de PPO em um grupo de linhas de soja públicas diversificadas.**

Dados de haplótipos marcadores para um grupo de 17 linhas de soja públicas diversificadas foram determinados para dois

QTL identificados no Exemplo 1 para marcadores moleculares do Grupamento de ligação L P10649C-3 (posição aproximada 3,6) e S00224-1 (posição aproximada 12,0). A resposta dessas linhas ao herbicida sulfentrazone foi publicada por Hulting et al.

5 (Soybean (Glycine max (L.) Merr.) cultivar tolerance to sulfentrazone. 2001 Science Direct, Vol. 20(8): 679-683). A resposta fenotípica foi relatada como um índice de redução de crescimento: altura de planta e ferimentos visíveis como expressado como uma porcentagem da plotagem de checagem de  
 10 cada cultivar. Dados para o haplótipo marcador no Grupamento de ligação L e os resultados do bioensaio de herbicida estão apresentados na Tabela 3. O uso do diagnóstico molecular P10649C-3 (QTL ligado no Grupamento de ligação L, posição aproximada 3,6) para este grupo de linhas de soja fenotipadas  
 15 é 92% preditivo para tolerância a sulfentrazone quando o ferimento é ajustado a 39% ou menor GRI e é 100% preditivo para não tolerância a sulfentrazone quando o ferimento é ajustado a 40% ou maior GRI. O uso do marcador S00224-1 (posição aproximada 12,0) para este grupo de linhas de soja é  
 20 88% preditivo para tolerância a sulfentrazone quando o ferimento é ajustado a 39% ou menor GRI e é 100% preditivo para não tolerância a sulfentrazone quando o ferimento é ajustado a 40% ou mais GRI.

25 Tabela 3. Haplótipo marcador no/próximo QTL no Grupamento de ligação L para resposta a herbicida PPO (sulfentrazone) e medida fenotípica de resposta de cultivo, expressada em termos de Índice de Redução de Crescimento, para cultivares de soja (itens em *itálico* indicam desvio do esperado)

|                                                                                                                                                                                                                                          |                                                                        | Grupamento de ligação L<br>QTLs |                                   |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------|---------------------------------|-----------------------------------|
|                                                                                                                                                                                                                                          |                                                                        | Posição 3,6                     | Posição 12,0                      |
| <b>Cultivar</b>                                                                                                                                                                                                                          | <b>Índice de Redução<br/>de Crescimento *</b>                          | <b>P10649C-3</b>                | <b>S00224-1</b>                   |
| PI88788                                                                                                                                                                                                                                  | 2                                                                      | 1,1                             | 3,3                               |
| Richland                                                                                                                                                                                                                                 | 4                                                                      | 1,1                             | 3,3                               |
| Lincoln                                                                                                                                                                                                                                  | 5                                                                      | 1,1                             | 3,3                               |
| PI180501                                                                                                                                                                                                                                 | 8                                                                      | 1,1                             | 3,3                               |
| Illini                                                                                                                                                                                                                                   | 8                                                                      | 1,1                             | 3,3                               |
| S100                                                                                                                                                                                                                                     | 8                                                                      | 1,1                             | 3,3                               |
| Mukden                                                                                                                                                                                                                                   | 8                                                                      | 1,1                             | 3,3                               |
| Arksoy                                                                                                                                                                                                                                   | 10                                                                     | 1,1                             | 3,3                               |
| Capital                                                                                                                                                                                                                                  | 10                                                                     | 1,1                             | 3,3                               |
| Haberlandt                                                                                                                                                                                                                               | 10                                                                     | 3,3                             | 2,2                               |
| Ral soy                                                                                                                                                                                                                                  | 13                                                                     | 1,1                             | 2,3                               |
| Dunfield                                                                                                                                                                                                                                 | 16                                                                     | 1,1                             | 3,3                               |
| Peking                                                                                                                                                                                                                                   | 22                                                                     | 1,1                             | 3,3                               |
| Roanoke                                                                                                                                                                                                                                  | 40                                                                     | 3,3                             | 2,2                               |
| Ogden                                                                                                                                                                                                                                    | 42                                                                     | 3,3                             | 2,2                               |
| Hutcheson                                                                                                                                                                                                                                | 46                                                                     | 3,3                             | 2,2                               |
| Ransom                                                                                                                                                                                                                                   | 52                                                                     | 3,3                             | 2,2                               |
|                                                                                                                                                                                                                                          |                                                                        |                                 |                                   |
|                                                                                                                                                                                                                                          | <b>Percentual de<br/>acurácia de carga<br/>de chamada de<br/>alelo</b> |                                 |                                   |
|                                                                                                                                                                                                                                          | <b>Tolerante</b>                                                       | <b>(alelos 1) 24/26 = 92%</b>   | <b>(alelo 3)<br/>23/26 = 88%</b>  |
|                                                                                                                                                                                                                                          | <b>Não tolerante</b>                                                   | <b>(alelo 3) 8/8 = 100%</b>     | <b>(alelo 2) =<br/>8/8 = 100%</b> |
|                                                                                                                                                                                                                                          |                                                                        |                                 |                                   |
|                                                                                                                                                                                                                                          |                                                                        |                                 |                                   |
| *índice de redução de crescimento (altura de planta e ferimentos visíveis expressos como uma porcentagem de plotagem de chcagem de cada cultivar); Aplicação de sulfentrazona pré-emergência a 0,28 kg ai/ha, de Hulting, et al. (supra) |                                                                        |                                 |                                   |

Dados de haplótipos para um grupo de 15 linhas de soja públicas diversificadas foram determinados para dois QTL

5 identificados no Exemplo 1 para o Grupamento de ligação N marcador molecular S60167 (posição aproximada 26,0). A resposta dessas 15 linhas para herbicida sulfentrazona foi determinada e publicada por Hulting et al. (Soybean (Glycine

max (L.) Merr.) cultivar tolerance to sulfentrazone. 2001 Science Direct, Vol. 20(8): 679-683). A resposta fenotípica foi relatada como um índice de redução de crescimento: altura de planta e ferimentos visíveis como expressado como uma porcentagem da plotagem de checagem de cada cultivar. Dados para o haplótipo marcador no Grupamento de ligação N e os resultados do bioensaio de herbicida são apresentados na Tabela 4. O cultivar Ral soy é heterozigoto para o marcador S60167. O uso do marcador S60167 para esse grupo de linhas de soja fenotipadas é 88% preditivo para tolerância a sulfentrazona quando o ferimento é ajustado a 39% ou menos GRI e é 100% preditivo para tolerância a sulfentrazona quando o ferimento é ajustado a 40% ou maior GRI.

Tabela 4. Haplótipo marcador no/próximo QTL no Grupamento de ligação N para resposta a herbicida PPO (sulfentrazona) e medida fenotípica de resposta de cultivo, expressada em termos de Índice de Redução de Crescimento, para cultivares de soja (itens em itálico indicam desvio do esperado).

|                 |                                              | Grupamento de ligação<br>N QTL |
|-----------------|----------------------------------------------|--------------------------------|
|                 |                                              | Posição 26                     |
| <b>Cultivar</b> | <b>Índice de redução de<br/>crescimento*</b> | <b>S60167-TB</b>               |
| PI88788         | 2                                            | 1,1                            |
| Richland        | 4                                            | 1,1                            |
| Lincoln         | 5                                            | 1,1                            |
| Illini          | 8                                            | 1,1                            |
| S100            | 8                                            | 1,1                            |
| Mukden          | 8                                            | 1,1                            |
| Arksoy          | 10                                           | 1,1                            |
| Haberlandt      | 10                                           | 1,1                            |
| Ral soy         | 13                                           | 1,2                            |
| Dunfield        | 16                                           | 1,1                            |
| CNS             | 20                                           | 2,2                            |
| Peking          | 22                                           | 1,1                            |

|           |                                                                    |                              |
|-----------|--------------------------------------------------------------------|------------------------------|
| Roanoke   | 40                                                                 | 2,2                          |
| Ogden     | 42                                                                 | 2,2                          |
| Hutcheson | 46                                                                 | 2,2                          |
|           |                                                                    |                              |
|           | <b>Percentual de acurácia<br/>de carga de chamada de<br/>alelo</b> |                              |
|           | <b>Tolerante</b>                                                   | <b>(alelo 1) 21/24 = 88%</b> |
|           | <b>Não tolerante</b>                                               | <b>(alelo 2) 6/6 = 100%</b>  |

**Exemplo 3: Previsão e confirmação de seleção baseada em marcador para resposta a químicos PPO em um grupo de linhas de soja comerciais**

5

Dados de haplótipos marcadores para um grupo de 7 linhas de soja comerciais foi determinado para dois QTL identificados no exemplo anterior para os marcadores moleculares do Grupamento de ligação L P10649C-3 (posição 3,6) e S00224-1 (posição 12,0). A resposta dessas linhas a herbicida sulfentrazone foi determinada pelo método usado no Exemplo 1. Além disso, a mesma escala foi usada para pontuação, de forma que:

- 15        9 = Equivalente, ou melhor quando comparado com a  
          checagem tolerante
- 7 = Muito pouco dano ou resposta notados.
- 5 = Resposta ou danos intermediários
- 3 = Dano acentuado, incluindo necrose foliar e  
20        atordoamento
- 1 = Dano severo, incluindo atordoamento severo e necrose;  
          equivalente ou pior quando comparado com a checagem não  
          tolerante

Dados para o haplótipo marcador no Grupamento de ligação L e os resultados do bioensaio de herbicida são apresentados na Tabela 5. O uso de um dos dois ou ambos marcadores para esse grupo de linhas de soja fenotipadas é 100% preditivo tanto para tolerância (pontuação de 7 ou 9) e não tolerância (pontuação de 1 para a checagem não tolerante).

Tabela 5: Previsão e confirmação de seleção baseada em marcador em QTL para o grupamento de ligação L para resposta a químicos PPO (sulfentrazone) em um grupo de variedades de soja comerciais.

|                  |                                                 | Posição 3,6      | Posição 12,0    |
|------------------|-------------------------------------------------|------------------|-----------------|
| <b>Variedade</b> | <b>Pontuação de ferimento por sulfentrazone</b> | <b>P10649C-3</b> | <b>S00224-1</b> |
| 93B41            | 9                                               | 1,1              | 3,3             |
| 93B82            | 9                                               | 1,1              | 3,3             |
| 9281             | 9                                               | 1,1              | 3,3             |
| 9584             | 9                                               | 1,1              | 3,3             |
| 92B52            | 7                                               | 1,1              | 3,3             |
| 92B61            | 7                                               | 1,1              | 3,3             |
| 9692             | 1                                               | 3,3              | 2,2             |

**Exemplo 4: Previsão e confirmação de seleção baseada em marcador para resposta a químicos PPO (sulfentrazone) em dez linhas de um grupo de linhas de soja fenotipadas na Universidade de Illinois (University of Illinois)**

Uma comparação para a previsibilidade de marcador a resposta a PPO foi conduzida. O experimento de bioensaio de herbicida usado está descrito em *Phytoxic Response and Yield of Soybean (Glycine max) Varieties Treated with Sulfentrazone ou Flumioxazin* (Taylor-Lovell et al., 2001 Weed Technology

15:96-102). Dados fenotípicos foram tomados a partir da Tabela 2 da publicação para aquelas variedades para as quais dados de marcador *in-house* estavam disponíveis. Pontuação fenotípica e dados de haplótipos para um grupo de 10 linhas de soja (1 pública 9 comerciais) nas regiões cromossômicas ao redor do QTL para o Grupamento de ligação L estão apresentados na Tabela 6. A pontuação fenotípica foi determinada como percentual de ferimentos, o que é definido como taxas de ferimentos visíveis incluindo atordoamento, clorose, e sintomas de bronzeamento (0 = nenhum ferimento; 100 = morte completa) com aplicação em campo de 448 g ai/ha. As taxas foram obtidas 12 dias após o tratamento. O uso do marcador P10649C (QTL ligado no Grupamento de ligação L, posição aproximada 3,6, chamada de alelo 1) para esse grupo de linhas de soja fenotipadas é 100% preditivo para tolerância (chamada de alelo 1) para sulfentrazone quando os ferimentos são de 21% ou menos e é 100% preditiva de não tolerância (chamada de alelo 2 ou 3) para sulfentrazone quando os ferimentos são 43% ou mais. A previsibilidade do marcador S00224-1 é também 100% acurada para tolerância (alelo 3) e não tolerância (alelo 2) para este grupo de material.

Tabela 6. Haplótipo marcador no ou próximo ao QTL no Grupamento de ligação L para resposta a herbicida PPO (sulfentrazone) e medida fenotípica de ferimento de cultivo

|           |                                          | Posição 3,6 | Posição 12,0 |
|-----------|------------------------------------------|-------------|--------------|
| Variedade | Pontuação de ferimento por sulfentrazone | P10649C-3   | S00224-1     |
| P9584     | 5                                        | 1,1         | 3,3          |
| P9671     | 5                                        | 1,1         | 3,3          |

|               |                                              |                                 |                           |
|---------------|----------------------------------------------|---------------------------------|---------------------------|
| P9151         | 8                                            | 1,1                             | 3,3                       |
| P9306         | 15                                           | 1,1                             | 3,3                       |
| Elgin         | 18                                           | 1,1                             | 3,3                       |
| P9282         | 19                                           | 1,1                             | 3,3                       |
| P9352         | 21                                           | 1,1                             | 3,3                       |
| P9362         | 43                                           | 2,2                             | 2,2                       |
| 91B01         | 58                                           | 3,3                             | 2,2                       |
| P9552         | 61                                           | 3,3                             | 2,2                       |
| LSD<br>(0,05) | 8                                            |                                 |                           |
|               |                                              |                                 |                           |
|               | <b>acurácia de carga de chamada de alelo</b> |                                 |                           |
|               | <b>Tolerante</b>                             | (alelos 1 ou 2)<br>14/14 = 100% | (alelo 3)<br>14/14 = 100% |
|               | <b>Não tolerante</b>                         | (alelo 3) 8/8 =<br>100%         | (alelo 2) = 8/8<br>= 100% |

**Exemplo 5: Fotográficas de resposta de variedades de soja (variedades de checagem tolerantes e não tolerantes) a ferimentos por sulfentrazone em bioensaios no campo e em estufa/câmara de crescimento**

Germoplasma conhecido não tolerante (*i.e.*, Pioneer variety 9692, Asgrow variedade A4715) e tolerante (*i.e.*, Pioneer variedade 9584, Syngenta variedade S5960) pode exibir severas diferenças em sintomas quando condições de campo conduzem ao dano e quando condições de laboratório para bioensaios são otimizadas para propósitos de seleção. As figuras 5 e 6 apresentam essas diferenças em fenótipo. A Figura 5 apresenta uma amostra de campo, com uma variedade não tolerante na esquerda (atordoada, necrótica) e variedade tolerante na direita (crescimento normal). A Figura 6 apresenta uma amostra de estufa, com variedades de checagem não tolerante (lado esquerdo) e tolerante (lado direito), tratadas no primeiro plano, não tratadas no plano de fundo.

Será aparente àqueles especialistas na arte que não pretende-se que a invenção esteja limitada a tais avanços ou mecanismos ilustrativos, e que modificações podem ser feitas sem sair do escopo ou espírito da invenção, como definido pelas reivindicações em anexo. Pretende-se que tais modificações e variações óbvias seja incluídas no escopo da presente invenção, como definido nas reivindicações em anexo. As reivindicações devem cobrir os componentes e etapas reivindicados em qualquer sequência que seja efetiva para encontrar os objetivos aí pretendidos, a menos que o contexto especificamente indique o contrário.

Todas as publicações aqui referidas são aqui incorporadas por referência para o propósito citado na mesma extensão como se cada uma fosse especificamente e individualmente indicada a ser aqui incorporada por referência.

**REIVINDICAÇÕES**

1. Método de selecionar uma planta de soja ou germoplasma com tolerância ou tolerância melhorada a herbicidas que inibem a função de protoporfirinogênio oxidase, caracterizado por compreender:

a) performar um bioensaio de herbicida em uma planta de soja ou germoplasma;

b) avaliar a tolerância ou tolerância melhorada das plantas de soja ou germoplasmas a herbicidas que inibem a função de protoporfirinogênio oxidase;

c) detectar na planta de soja ou germoplasma pelo menos um alelo de um locus marcador que é associado com a tolerância ou tolerância melhorada a herbicidas que inibem a função de protoporfirinogênio oxidase, através da análise comparativa da pontuação fenotípica das plantas em que o um ou mais locus marcador é um locus marcador localizado em (i) uma distância de recombinação genética menor ou igual a 10 cM do marcador P10649C-3 no grupamento de ligação L com sequências primer SEQ ID NO 5 e SEQ ID NO 6 e com sondas SEQ ID NO 7 e SEQ ID NO 8; (ii) uma distância de recombinação genética menor ou igual a 10 cM do marcador S00224-1 no grupamento de ligação L com sequências primer SEQ ID NO 10 e SEQ ID NO 11 e com sondas SEQ ID NO 12 e SEQ ID NO 13; (iii) uma distância de recombinação genética menor ou igual a 1 cM do marcador SATT523 no grupamento de ligação L com

sequências primer SEQ ID NO 3 e SEQ ID NO 4; (iv) uma distância de recombinação genética menor ou igual a 10 cM do marcador S60167-TB no grupamento de ligação com sequências primer SEQ ID NO 1 e SEQ ID NO 2; (v) uma distância de recombinação genética menor ou igual a 1 cM do marcador P546A-1 no grupamento de ligação N com sequência primer SEQ ID NO 14 e com sonda SEQ ID NO 16; (vi) uma distância de recombinação genética menor ou igual a 1 cM do marcador P5467A-2 no grupamento de ligação com sequência primer SEQ ID NO 15 e com sonda SEQ ID NO 17; e

d) selecionar a planta de soja ou germoplasma compreendendo pelo menos um alelo de um ou mais loci marcador, selecionando, dessa forma, uma planta de soja com tolerância ou tolerância melhorada a herbicidas que inibem a função de protoporfirinogênio oxidase.

2. Método de acordo com a reivindicação 1, caracterizado pelo fato de que o dito locus de traço quantitativo é localizado em um intervalo cromossômico definido por e incluindo os marcadores P10649C-3 (SEQ ID 5-9) e S00224-1 (SEQ ID 10-13).

3. Método de acordo com a reivindicação 1, caracterizado pelo fato de que o herbicida é selecionado a partir do grupo consistindo de difeniléteres, N-fenilptalamidas, oxadiazola e triazolinonas.

4. Método de acordo com a reivindicação 3, caracterizado pelo fato de que o herbicida é selecionado a partir do grupo consistindo de sulfentrazone, carfentrazone-etil, acifluorfen, lactofen, fomesafen, flumioxazina, flumiclorac-pentil e oxifluorfen.

5. Método de rastrear soja para determinar a presença ou ausência de um locus de traço quantitativo associado com tolerância a herbicidas que inibem a função de protoporfirinogênio oxidase, caracterizado por compreender:

- a. rastrear a soja para um ácido nucléico que se associa com tolerância a herbicidas que inibem a função de protoporfirinogênio oxidase e que se localizam a (i) uma distância de recombinação genética menor ou igual a 10 cM do marcador P10649C-3 no grupamento de ligação L com sequências primer SEQ ID NO 5 e SEQ ID NO 6 e com sondas SEQ ID NO 7 e SEQ ID NO 8; (ii) uma distância de recombinação genética menor ou igual a 10 cM do marcador S00224-1 no grupamento de ligação com sequências primer SEQ ID NO 10 e SEQ ID NO 11 e com sondas SEQ ID NO 12 e SEQ ID NO 13; (iii) uma distância de recombinação genética menor ou igual a 1 cM do marcador SATT523 no grupamento de ligação L com sequências primer SEQ ID NO; 3 e SEQ ID NO 4; (iv) uma distância de recombinação genética menor ou igual a 10 cM do marcador S60167-TB no grupamento de ligação com sequências primer SEQ ID NO 1 e SEQ ID NO 2; (v) uma distância de recombinação genética menor ou igual a 10

cM do marcador P546A-1 no grupamento de ligação N com sequência primer SEQ ID NO 14 e com sonda SEQ ID NO 16; ou (vi) uma distância de recombinação genética menor ou igual a 10 cM do marcador P5467A-2 no grupamento de ligação com sequência primer SEQ ID NO 15 e com sonda SEQ ID NO 17; e

- b. determinar se a soja compreende dito ácido nucléico, determinando, dessa maneira, a presença ou ausência de um locus de traço quantitativo associado com tolerância a herbicidas que inibem a função de protoporfirinogênio oxidase.

FIGURA 1A

| Marcador  | Grupo Ligante | Posição | Tipo |
|-----------|---------------|---------|------|
| SATT495   | L             | 0.00    | SSR  |
| SATT723   | L             | 0.44    | SSR  |
| Sat 408   | L             | 1.00    | SSR  |
| Sat 301   | L             | 10.31   | SSR  |
| SATT446   | L             | 11.13   | SSR  |
| P10649C-3 | L             | 12.5    | ASH  |
| SATT232   | L             | 12.55   | SSR  |
| SATT182   | L             | 13.90   | SSR  |
| SATT238   | L             | 19.41   | SSR  |
| Sat 071   | L             | 20.04   | SSR  |
| SATT388   | L             | 21.61   | SSR  |
| SATT497   | L             | 26.06   | SSR  |
| SATT313   | L             | 27.35   | SSR  |
| SATT143   | L             | 28.16   | SSR  |
| Sat 397   | L             | 28.26   | SSR  |
| SATT418   | L             | 28.57   | SSR  |
| Sat 134   | L             | 28.66   | SSR  |
| SATT652   | L             | 28.67   | SSR  |
| SATT711   | L             | 28.67   | SSR  |
| Sat 187   | L             | 28.68   | SSR  |
| Sat 195   | L             | 28.68   | SSR  |
| Sat 388   | L             | 28.71   | SSR  |
| SATT694   | L             | 28.71   | SSR  |
| SATT398   | L             | 28.90   | SSR  |
| Sat 191   | L             | 29.19   | SSR  |
| Sat 405   | L             | 29.40   | SSR  |
| Sat 320   | L             | 29.74   | SSR  |
| SATT523   | L             | 30.18   | SSR  |
| SATT278   | L             | 30.34   | SSR  |
| SATT613   | L             | 32.64   | SSR  |

FIGURA 1B

| Marcadores Ligados |          |         |         |          |          |
|--------------------|----------|---------|---------|----------|----------|
| SATT495            | SATT723  | Sat 408 | A169 1  | EV2 1    | Sle3 4s  |
| BLT010 2           | BLT007 1 | SATT232 | Sat 301 | SATT446  | SATT182  |
| R176 1             | JUBC090  | SATT238 | Sat 071 | BLT039 1 | Bng071 1 |
| SATT388            | A264 1   | RGA 7   | RGA7    | SATT523  | Sat 134  |
| LbA                | i8 2     | A450 2  | A106 1  | Sat 405  | SATT143  |
| B124 2             | A459 1   | SATT398 | SATT694 | Sat 195  | Sat 388  |
| SATT652            | SATT711  | Sat 187 | SATT418 | SATT278  | Sat 397  |
| Sat 191            | Sat 320  | O109 1  | A204 2  | SATT497  | G214 17  |
| SATT313            | B164 1   | G214 16 | SATT613 | A023 1   | SATT284  |
| AW508247           | SATT462  | L050 7  | E014 1  | A071 5   | B046 1   |
| L1                 | B162 2   |         |         |          |          |

FIGURA 2A

| Marcador  | Grupo Ligante | Posição | Tipo |
|-----------|---------------|---------|------|
| Sat 379   | N             | 2.70    | SSR  |
| SCT 195   | N             | 4.63    | SSR  |
| SATT631   | N             | 20.31   | SSR  |
| SATT159   | N             | 21.76   | SSR  |
| SATT009   | N             | 22.42   | SSR  |
| SATT641   | N             | 23.32   | SSR  |
| Sat 186   | N             | 23.91   | SSR  |
| SATT152   | N             | 24.79   | SSR  |
| S60167-TB | N             | 26.00   | SSR  |
| SATT530   | N             | 27.41   | SSR  |
| SATT675   | N             | 28.92   | SSR  |
| SATT683   | N             | 28.96   | SSR  |
| SATT624   | N             | 29.55   | SSR  |
| SATT393   | N             | 29.64   | SSR  |
| SATT125   | N             | 30.30   | SSR  |
| SATT485   | N             | 30.30   | SSR  |
| SATT584   | N             | 30.64   | SSR  |
| Sat 166   | N             | 31.09   | SSR  |
| Sat 084   | N             | 31.36   | SSR  |
| Sat 275   | N             | 33.84   | SSR  |
| Sat 208   | N             | 35.48   | SSR  |
| Sat 280   | N             | 36.82   | SSR  |
| SATT080   | N             | 38.56   | SSR  |
| Sat 266   | N             | 40.02   | SSR  |
| SATT387   | N             | 45.03   | SSR  |

FIGURA 2B

| Marcadores Ligados |         |          |         |          |
|--------------------|---------|----------|---------|----------|
| SCT 195            | Sat 379 | A071 10  | A071 3  | A071 4   |
| A071 6             | AC telo | Rps7     | R022 1  | L050 12  |
| SATT152            | OP N03  | BLT004 1 | SATT631 | SATT159  |
| SATT009            | SATT641 | RGA6a    | Sat 186 | A071 2   |
| K418 1             | Rps1    | K395 2   | OPAC12b | SATT530  |
| gc34 2             | SATT683 | SATT675  | A280 1  | SATT624  |
| A426 2             | Sle 003 | i4 2     | Sat 084 | SATT393  |
| SATT584            | SATT485 | Sat 166  | Sat 208 | BLT049 1 |
| Bng095 2           | OP F13  | SATT125  | Sat 275 | Sle2 3   |
| Mng456 1           | RGA6b   | OP U09b  | mO128 1 | Sat 280  |
| SATT080            | Sat 266 | L103 1   | B162 1  | peG488 2 |
| SATT387            | Rpg4    | Sat 236  | Sat 033 |          |

FIGURA 3

| Marcador Nome  | Sequência do Primer esquerdo                | Sequência do Primer direito                  | "Pigtail" |
|----------------|---------------------------------------------|----------------------------------------------|-----------|
| S60167-TB LG-N | TTATTGAGGTGGGCAAGGTGTG<br>(SEQ ID NO. 1)    | CATGAACGTCTGGTGGTTGAACA<br>(SEQ ID NO. 2)    | GTTTCTT   |
| SATT523 LG-L   | GCGATTCTTCCTTGAAGAATTTCTG<br>(SEQ ID NO. 3) | GCGCTTTTTCGGCTGTTATTTTAACT<br>(SEQ ID NO. 4) | GTTTCTT   |

FIGURA 4

| Marcador Nome  | Primers de PCR                                                                                                              | Sondas de Alelos                                                                                                                                            |
|----------------|-----------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------|
| P10649C-3 LG-L | Primer Seq 1 (SEQ ID NO. 5):<br>GAGGGCTATGTTTCTTCTCCAGATGTGAG<br>Primer Seq 2 (SEQ ID NO. 6):<br>AAGGTCGGCTTGGTGGTTAAAGGCAG | Sonda de Alelo 1 (SEQ ID NO. 7):<br>TCATcTgTGATAA<br>Sonda de Alelo 2 (SEQ ID NO. 8):<br>TCATgTgTGATAA<br>Sonda de Alelo 3 (SEQ ID NO. 9):<br>TCATcTcTGATAA |
| S00224-1 LG-L  | Primer Seq 1 (F) (SEQ ID NO. 10):<br>CTGGACCTACCCGGGATGAAAA<br>Primer Seq 2 (R) (SEQ ID NO. 11):<br>TCTTCCTCTCCCTTCTCTCGC   | Sonda de Alelo 1 (PF1) (SEQ ID NO. 12):<br>CGCGAcTCTCCTC<br>Sonda de Alelo 2 (PV1) (SEQ ID NO. 13):<br>CGCGAgTCTCCTC                                        |
| P5467-1 LG-N   | Primer Seq 1 (SEQ ID NO. 14):<br>TCCCAGGTTAGATTTTCTGAACGAAGA<br>Primer Seq 2 (SEQ ID NO. 15):<br>CATCAGCACAAAAGGGCATCCTCA   | Sonda de Alelo 1 (SEQ ID NO. 16):<br>CACTCCTTAAGgTAAT<br>Sonda de Alelo 2 (SEQ ID NO. 17):<br>CACTCCTTAAGaTAAT                                              |

FIGURA 5

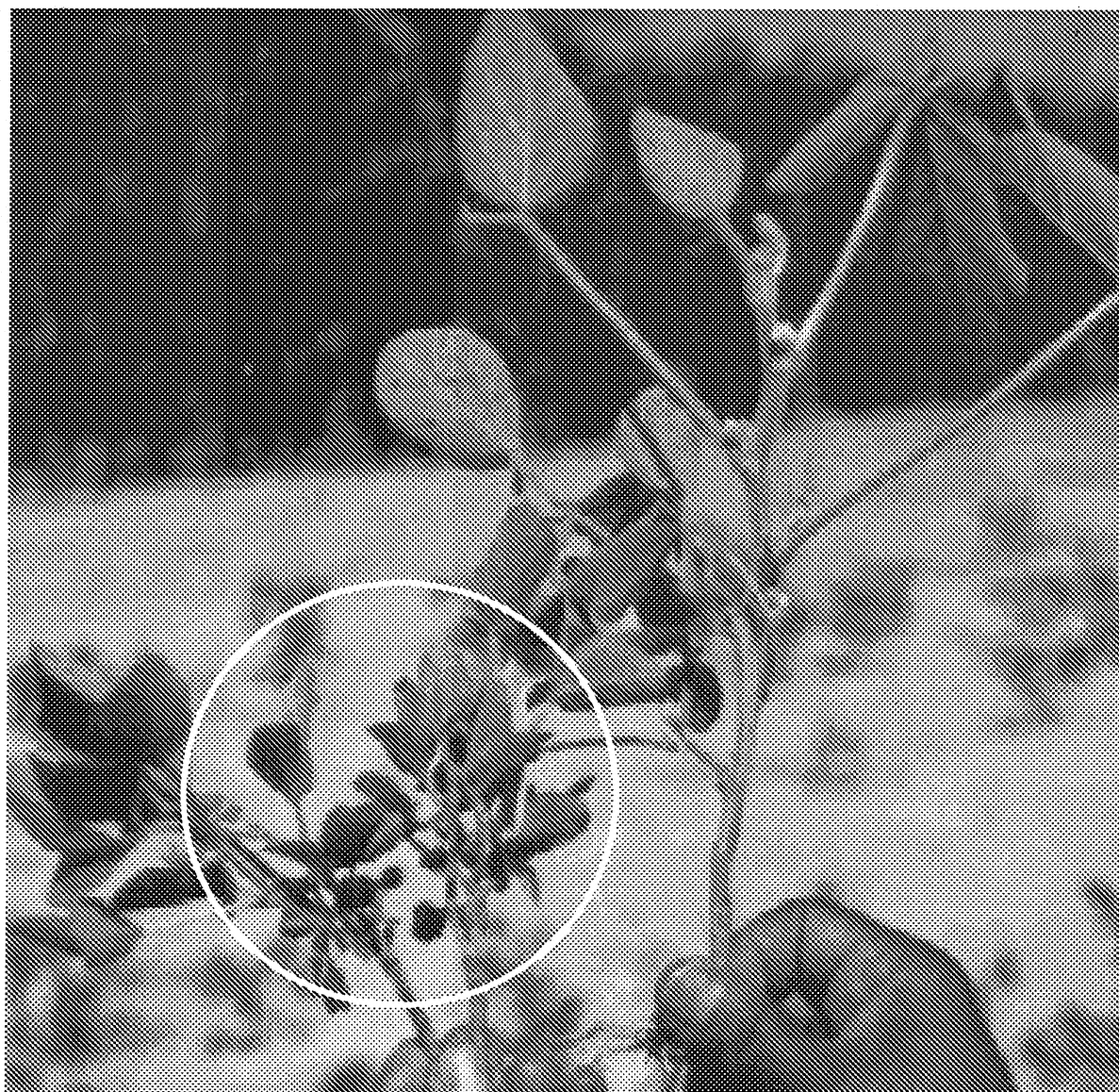


FIGURA 6



Não tolerante