

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

(21) **PI0707579-0 A2**

(22) Data de Depósito: 12/02/2007
(43) Data da Publicação: 10/05/2011
(RPI 2105)



(51) *Int.Cl.:*
C12N 15/85

(54) Título: **SISTEMA DE EXPRESSÃO DE POLINUCLEOTÍDEOS, MÉTODOS DE CONTROLE DA POPULAÇÃO DE UM ORGANISMO EM UM MEIO AMBIENTE NATURAL PARA O MESMO, DE CONTROLE BIOLÓGICO, E DE SEPARAÇÃO DE SEXOS, E, MÉTODO OU CONTROLE BIOLÓGICO OU DA POPULAÇÃO**

(57) Resumo: SISTEMA DE EXPRESSÃO DE POLINUCLEOTÍDEOS, MÉTODOS DE CONTROLE DA POPULAÇÃO DE UM ORGANISMO EM UM MEIO AMBIENTE NATURAL PARA O MESMO, DE CONTROLE BIOLÓGICO, E DE SEPARAÇÃO DE SEXOS, E, METODO OU CONTROLE BIOLÓGICO OU DA POPULAÇÃO. Um sistema de expressão de polinucleotídeos é provido, o qual é capaz de emendar alternativamente os transcritos de RNA de uma sequência polinucleotídica a ser expressada em um organismo.

(30) Prioridade Unionista: 25/10/2006 GB 0621234.4,
10/02/2006 US 11/352177

(73) Titular(es): Oxitec Limited

(72) Inventor(es): Luke Alphey

(74) Procurador(es): Momsen, Leonardos & CIA.

(86) Pedido Internacional: PCT GB2007000488 de 12/02/2007

(87) Publicação Internacional: WO 2007/091099 de 16/08/2007

“SISTEMA DE EXPRESSÃO DE POLINUCLEOTÍDEOS, MÉTODOS DE
CONTROLE DA POPULAÇÃO DE UM ORGANISMO EM UM MEIO
AMBIENTE NATURAL PARA O MESMO, DE CONTROLE
BIOLÓGICO, E DE SEPARAÇÃO DE SEXOS, E, MÉTODO OU
5 CONTROLE BIOLÓGICO OU DA POPULAÇÃO”

Todas as referências incorporadas aqui são por meio disto
incorporadas por referência, salvo de outra forma evidente.

INTRODUÇÃO

A presente invenção refere-se a um sistema de expressão de
10 genes, em combinação com seqüências de controle de emenda, referidas
seqüências de controle provendo um mecanismo para emenda alternativa.

Emenda alternativa envolve a remoção de um ou mais íntrons
e ligação dos éxons de flaqueamento. Esta reação é catalisada pela
espliceossoma, uma máquina macromolecular composta de cinco RNAs e
15 centenas de proteínas (Jurica, M. S. & Moore, M. J. (2003) Mol. Cell 12, 5-
14). Emenda alternativa gera mRNAs múltiplos a partir de um gene único,
deste modo aumentando a diversidade de proteomas (Graveley, B. R. (2001)
Trends Genet. 17, 100-107).

Emenda alternativa também desempenha um papel chave na
20 regulação de expressão de genes em muitos processos relativos ao
desenvolvimento variando de determinação de sexo a apoptose (Black, D. L.
(2003) Annu. Rev. Biochem. 72, 291-336), e defeitos na emenda alternativa
foram ligadas a muitos distúrbios humanos (Cáceres, J. F. & Kornblihtt, A. R.
(2002) Trends Genet. 18, 186-193). Geralmente, emenda alternativa é
25 regulada por proteínas que se associam com pré-mRNA e funcionam ou para
aumentar ou reprimir a capacidade do espliceossoma de reconhecer o(s)
sítio(s) de emenda flaqueando o éxon regulado (Smith, C. W. & Valcarcel, J.
(2000) Trends Biochem. Sci. 25, 381-388).

Se um éxon alternativo particular será incluído ou excluído de

um RNA maduro em cada célula é determinado, como se pensa, pela concentração relativa de um número de reguladores de emenda positivos e negativos e as interações destes fatores com o pré-mRNA e componentes do espliceossoma (Smith, C. W. & Valcarcel, J. (2000) Trends Biochem. Sci. 25, 381-388).

Espliceossomas são complexos grandes de RNA nuclear pequeno e partículas de proteína (snRNPs) que juntam-se com pré-mRNA para realizar a emenda de RNA, removendo os íntrons dos RNAs nucleares eucarióticos, assim produzindo mRNA que é então traduzido para proteína em ribossomas.

Apesar de pelo menos 74% de genes humanos codificarem mRNAs emendados alternativamente (Johnson, J. M., Castle, J., Garrett-Engele, P., Kan, Z., Loerch, P. M., Armour C. D., Santos, R., Schadt, E. E., Stoughton, R. & Shoemaker, D. D. (2003) Science 302, 2141-2144), relativamente poucos reguladores de emenda foram identificados.

SUMÁRIO DA INVENÇÃO

Assim, em um primeiro aspecto, a presente invenção provê um sistema de expressão de polinucleotídeos compreendendo:

pelo menos uma seqüência polinucleotídica heteróloga codificando uma proteína funcional, definida entre um códon de partida e um códon de parada, e/ou polinucleotídeos para RNA interferência (RNAi), a ser expressada em um organismo; ligado operativamente

pelo menos um promotor ligado operativamente ao mesmo; e

pelo menos uma seqüência de controle de emenda que, em cooperação com um espliceossoma, é capaz de (i) mediar a emenda de um transcrito de RNA da seqüência de codificação para produzir um primeiro produto de RNA mensageiro (mRNA) emendado, e (ii) mediar pelo menos uma emenda alternativa de referido transcrito de RNA para produzir um produto de mRNA emendado alternativo;

em que, quando pelo menos uma seqüência polinucleotídica heteróloga codifica uma proteína funcional, pelo menos um dos produtos de mRNA maduro compreendendo uma Matriz de Leitura Aberta (ORF) contínua estendendo-se a partir de referido códon de partida para referido códon de parada, assim definindo uma proteína, que é a referida proteína funcional, ou está relacionada com a referida proteína funcional por pelo menos uma deleção de aminoácido, e que é funcional quando traduzida, e, opcionalmente, sofreu uma modificação pós-translacional;

a mediação sendo selecionada dentre o grupo consistindo de:
10 mediação específica de sexo, mediação específica de estágio, mediação específica de linhagem germinal, mediação específica de tecido, e combinações das mesmas.

O sistema de expressão pode ser DNA ou RNA ou um híbrido ou combinação de ambos. Visa-se que o sistema compreenda ambos ribo- e
15 desoxirribonucleotídeos, isto é porções de DNA e porções de RNA. Estes poderiam corresponder a elementos genéticos diferentes, tal que o sistema é um híbrido de DNA/RNA, com alguns elementos funcionais providos por DNA e outros por RNA.

Preferivelmente, a mediação é em um modo específico de
20 sexo, específica de estágio, específica de linhagem germinal ou específica de tecido. Em particular, a mediação específica de sexo é particularmente preferida. No entanto, também prefere-se que uma combinação destes quatro modos de mediação possa ser utilizada. Prefere-se particularmente que, quando uma combinação destes modos é usada, que isto inclua a mediação
25 específica de sexo. Um exemplo particularmente preferido de tal combinação é uma combinação de mediação específica de sexo, específica de tecido e específica de estágio de emenda alternativa.

O sistema pode ser adaptado para expressão de um gene. Preferivelmente, a seqüência polinucleotídica a ser expressada compreende

uma sequência de codificação para uma proteína ou polipeptídeo, isto é pelo menos um éxon, e preferivelmente 2 ou mais éxons, capazes de codificar um polipeptídeo, tal como uma proteína ou fragmento da mesma.

5 Será entendido que um éxon é qualquer região de DNA dentro de um gene, que está presente em uma molécula de RNA maduro derivado deste gene, em vez de ser emendado da molécula de RNA transcrito. Para genes de codificações de proteína, moléculas de RNA maduro correspondem a moléculas de mRNA maduro, que podem codificar uma ou mais proteínas ou polipeptídeos. Éxons de muitos genes eucarióticos intercalam-se com
10 segmentos de DNA não de codificação.

A pelo menos uma sequência polinucleotídica heteróloga pode codificar uma proteína funcional, definida entre um códon de partida e um códon de parada a ser expressada em um organismo. Alternativamente, ou em
15 adição, a pelo menos uma sequência polinucleotídica heteróloga codifica ou compreende polinucleotídeos para RNA interferência (RNAi), a ser expressado em um organismo.

Estas sequências, a serem expressadas no organismo, também podem ser referidas como as sequências, cuja expressão deve ser regulada no referido organismo.

20 Preferivelmente, a sequência polinucleotídica a ser expressada compreende dois ou mais éxons de codificação, sendo segmentos ou sequências polinucleotídicas que codificam aminoácidos quando traduzidas do mRNA. Preferivelmente, os éxons diferentes são diferencialmente emendados juntos para prover mRNAs alternativos. Preferivelmente,
25 referidos mRNAs emendados alternativos têm potenciais de codificação diferentes, isto é codificam proteínas diferentes ou sequências de polipeptídeos. Assim, a expressão da sequência de codificação é regulada pela emenda alternativa nos modos acima mencionados de mediação.

A sequência polinucleotídica a ser expressada pode

compreender polinucleotídeos de RNA interferência (RNAi). Tais seqüências são capazes de prover, por exemplo, uma ou mais extensões de RNA filamento duplo (dsRNA), preferivelmente na forma de um transcrito primário, que por sua vez é capaz de processar pela enzima de tipo RNA Pol III "Dicer." Tais extensões incluem, por exemplo, extensões de RNA filamento único que podem formar laços, tais como os encontrados em RNA grampo de cabelo curto (shRNA), ou com regiões mais longas que são substancialmente auto-complementares.

Assim, onde o sistema é DNA, os polinucleotídeos para RNA interferência são desoxirribonucleotídeos que, quando transcritos nos ribonucleotídeos pré-RNA, provêm uma extensão de dsRNA, como discutido acima.

Polinucleotídeos para RNA interferência são particularmente preferidos quando referidos polinucleotídeos estão posicionados para minimizar a interferência com emenda alternativa. Isto pode ser realizado por posicionamento distal destes polinucleotídeos das seqüências de controle de emenda alternativa, preferivelmente 3' para as seqüências de controle. Em outra forma de realização preferida, regiões substancialmente auto-complementares podem ser separadas de cada por uma ou mais seqüências de controle de emenda, tais como um íntron, que media q emenda alternativa. Preferivelmente, as regiões auto-complementares são arranjadas como uma série de duas ou mais repetições invertidas, cada repetição invertida separada por seqüência de controle de emenda, preferivelmente um íntron, como definido em outra parte.

Nesta configuração, transcritos emendados alternativamente diferentes podem ter suas regiões auto-complementares substancialmente separadas por diferentes comprimentos de seqüência não auto-complementar no transcrito maduro (emenda pós alternativa). Será notado que regiões que são substancialmente auto-complementares são aqueles que são capazes de

formar grampos de cabelo, por exemplo, como porções da seqüência são capazes de formar pares de bases com outras porções da seqüência. Estas duas porções não têm de ser exatamente complementares mutuamente, porque podem ocorrer alguns desalinhamentos ou tolerância de extensões em cada porção que não forma pares de base umas com as outras. Tais extensões podem não ter um equivalente em outra porção, tal que a simetria é perdida e "protuberâncias" se formam, como é conhecido com complementação de pares de base geralmente.

Em outra forma de realização preferida, um ou mais segmentos de seqüência substancialmente complementares a outra seção do transcrito primário é posicionado, relativo à pelo menos uma seqüência de controle de emenda, de modo que isto não está incluído em todos dos transcritos produzidos por emenda alternativa do transcrito primário. Por esse método, alguns transcritos são produzidos que tendem a produzir dsRNAs enquanto outros não; por mediação da emenda alternativa, por exemplo mediação específica de sexo, mediação específica de estágio, mediação específica de linhagem germinal, mediação específica de tecido, e combinações das mesmas, dsRNA pode ser produzido em um modo específico de sexo, específica de estágio, específica de linhagem germinal ou específica de tecido, ou combinações das mesmas.

O sistema é preferivelmente capaz de expressar pelo menos uma proteína de interesse, isto é referida proteína funcional a ser expressada em um organismo. Referida pelo menos uma proteína de interesse pode ter um efeito terapêutico ou pode, preferivelmente, ser um marcador, por exemplo DsRed, proteína fluorescente verde (GFP) ou um ou mais de seus mutantes ou variantes, ou outros marcadores que são bem conhecidos na técnica.

O mais preferivelmente, a proteína funcional a ser expressada em um organismo tem um efeito letal, deletério ou esterilizante. Onde

referência é feita aqui para um efeito letal, será notado que ela abrange um efeito deletério e esterilizante, tal como um efeito capaz de matar o organismo per se ou sua descendência, ou capaz de reduzir ou destruir a função de determinados tecidos do mesmo, dos quais os tecidos reprodutivos são particularmente preferidos, de modo que o organismo ou sua descendência são estéreis. Assim, alguns efeitos letais, tais como venenos, matarão o organismo ou tecido em uma matriz de tempo curto relativo a seu tempo de vida, enquanto outros podem simplesmente reduzir a capacidade dos organismo de funcionar, por exemplo reprodutivamente.

10 Um efeito letal resultante em esterilização é particularmente preferido, como isto permite ao organismo competir no ambiente natural ("no estado selvagem") com organismo de tipo selvagem, mas o inseto estéril não pode então produzir uma descendência viável. Desta forma, a presente invenção realiza um resultado similar às técnicas tais como técnica de inseto estéril (SIT) em insetos, sem os problemas associados com SIT, tal como o custo, perigo ao usuário, e reduzida competitividade do organismo irradiado.

15 Preferivelmente, o sistema compreende pelo menos um mecanismo de retro-alimentação positiva, isto é pelo menos proteína funcional a ser diferencialmente expressada, através de emenda alternativa, e pelo menos um promotor para o mesmo, onde um produto de um gene a ser expressado serve como um fator de controle transcricional positivo para o pelo menos um promotor, e pelo qual o produto, ou a expressão do produto, é controlável. Preferivelmente, um melhorador é associado com o promotor, o produto de gene servindo para melhorar a atividade do promotor através do melhorador. Preferivelmente, o fator de controle é o produto do gene de tTA ou um análogo do mesmo, e onde uma ou mais unidades de operador de tetO é ligada operativamente como promotor e é o melhorador, tTA ou seu análogo servindo para melhorar atividade do promotor via tetO. Prefere-se que proteína funcional codifique o produto de tTAV ou tTAF e preferivelmente, o

promotor está substancialmente inativo na ausência do fator de controle transcripcional positivo. Os promotores apropriados, preferivelmente mínimos, para esse sistema podem ser selecionados dentre: hsp70, um promotor mínimo de P, um promotor mínimo de CMV, um promotor mínimo baseado de Act5C, um fragmento de promotor de BmA3, um fragmento de promotor de hunchback, um promotor de núcleo de Adh, e um promotor mínimo de Act5C, ou combinações dos mesmos.

Em uma forma de realização, a proteína funcional é preferivelmente um fator de indução de apoptose, tal como a proteína de AIF descrita por exemplo em Cande et al (Journal of Cell Science 115, 4727-4734 (2002)) ou homólogos da mesma. Homólogos de AIF são encontrados em mamíferos e até mesmo em invertebrados, incluindo insetos, nematóides, fungos, e plantas, significando que o gene de AIF foi conservado ao longo de todo o reino eucariótico. Também preferido é Hid, o produto de proteína do gene defeituoso de involução de cabeça de *Drosophila melanogaster*, ou Reaper (Rpr), o produto do gene Reaper de *Drosophila*, ou mutantes dos mesmos. Uso de Hid foi descrito por Heinrich e Scott (Proc. Natl Acad. Sci USA 97, 8229-8232 (2000)). Uso de um derivado mutante, Hid^{Ala5} foi descrito por Horn e Wimmer (Nature Biotechnology 21, 64-70 (2003)). Uso de um derivado mutante de Rpr, Rpr^{KR}, é descrito aqui (ver também White et al 1996, Wing et al., 2001, e Olson et al., 2003). Ambos, Rpr e Hid são proteínas pró-apoptóticas, que se ligam a IAP1, como se pensa. IAP1 é uma proteína anti-apoptótica bem conservada. Hid e Rpr são, portanto, esperados como trabalhando através de uma extensa faixa filogenética (Huang et al., 2002, Vernooy et al., 2000) apesar de sua própria seqüência não ser bem conservada.

Também é preferido Nipp1Dm, o homólogo de *Drosophila* de Nippl de mamífero (Parker et al Biochemical Journal 368, 789-797 (2002); Bennett et al, Genetics 164, 235-245 (2003)). Nipp1Dm é outro exemplo de

uma proteína com efeito letal se expressada a um nível apropriado, como seria entendido pelo versado. De fato, muitos outros exemplos de proteínas com um efeito letal serão conhecidos pelo versado na técnica.

5 Também prefere-se que a proteína funcional seja, ela mesma, um transativador transcripcional, tal como o sistema tTAV descrito acima.

10 Prefere-se que o promotor possa ser ativado por condições ambientais, por exemplo a presença ou ausência de um fator particular tal como tetraciclina no sistema tet descrito aqui, tal que a expressão do gene de interesse possa ser facilmente manipulada pelo versado. Alternativamente, um exemplo preferido de promotor apropriado é o promotor de choque térmico hsp 70, permitindo ao usuário controlar a expressão por variação da temperatura ambiente para a qual os hospedeiros estão expostos em um laboratório ou no campo, por exemplo. Outro exemplo preferido de controle de temperatura está descrito em Fryxell e Miller (Journal of Economic Entomology 88, 1221-1232 (1995)).

20 Também preferido como um promotor é o promotor específico de embrião syra (Horn & Wirnmer (2003) de *Drosophila melanogaster*, ou seus homólogos, ou promotores de outros genes específicos de embrião ou ativos de embrião, tais como os do gene de *Drosophila* lento como melaço (slam), ou seus homólogos de outras espécies.

25 Também prefere-se que o sistema compreenda outros fatores a montante, 5', e/ou fatores a jusante, 3', para controlar a expressão. Exemplos incluem melhoradores tais como os melhoradores de gordura corporal dos genes de proteína yolk de *Drosophila*, e os melhoradores de região de homologia (hr) de baculovírus, por exemplo AcMNPV. Também será notado que os produtos de RNA irão incluir UTRs 5' e 3' apropriados, por exemplo.

A sequência de controle de emenda permite um nível adicional de controle de expressão de proteína, em adição ao promotor e/ou melhorador do gene. Por exemplo, a expressão específica de sexo ou de tecido em

embriões de inseto apenas seria extremamente difícil por métodos convencionais. Promotores com esta especificidade são desconhecidos, mesmo em *Drosophila*. No entanto, usando controle combinatorial de acordo com a presente invenção, um promotor específico de embrião, por exemplo srya, pode ser combinado com um sistema de emenda alternativa apropriado.

Prefere-se que qualquer combinação de promotor e mecanismo de emenda alternativa seja considerada. O promotor é preferivelmente específico para uma proteína particular tendo um efeito temporal pequeno ou especial confinado, por exemplo um efeito autônomo celular.

Alternativamente, prefere-se que o promotor possa ser específico para uma classe de proteínas ou uma proteína específica que tem um efeito a longo prazo e/ou sistema extenso, tais como um hormônio, fator de crescimento positivo ou negativo, morfogene ou outra molécula secretada ou sinalizando superfície celular. Isto irá permitir, por exemplo, um padrão de expressão mais amplo de modo que uma combinação de um promotor de morfogene com um mecanismo de emenda alternativa específico de estágio poderia resultar no morfogene sendo expressado apenas uma vez que foi alcançada uma determinada fase do ciclo da vida, mas o efeito do morfogene ainda seria sentido (isto é o morfogene pode ainda agir e ter um efeito) para além desta etapa do ciclo da vida. Exemplos preferidos seriam moléculas de morfogene/sinalização Hedgehog, Wingless/WNTs, TGF β /BMPs, EGF e seus homólogos, moléculas de sinalização conservadas evolucionariamente bem conhecidas.

Prevê-se também que um promotor que é ativado por uma faixa de fatores de proteína, por exemplo transativadores, ou que tem um amplo efeito sistêmico, tais como um hormônio ou morfogene, poderia ser usado em combinação com um mecanismo de emenda alternativa para realizar um controle específico de tecido ou sexo ou controle específico de sexo ou estágio, ou outras combinações de controle específico de estágio,

tecido, linhagem germinal e sexo.

Prevê-se também que mais de um promotor, e opcionalmente um melhorador para o mesmo, pode ser usado no presente sistema, quer como meios alternativos para iniciar a transcrição da mesma proteína ou por virtude
5 do fato de que o sistema genético compreende mais de um sistema de expressão de gene (isto é mais de um gene e seu promotor acompanhante).

Em um outro aspecto, a presente invenção provê um método de transformação, compreendendo expressar duas ou mais moléculas de RNA, derivadas de um transcrito primário simples, ou transcritos primários
10 substancialmente similares, por emenda alternativa, referidas duas ou mais moléculas de RNA preferivelmente codificando diferentes proteínas ou polipeptídeos, em um organismo contatando o organismo com o sistema de expressão e preferivelmente induzindo a expressão do sistema de expressão. Métodos de introdução ou transformação do sistema de gene e indução de
15 expressão são bem conhecidos na técnica com relação ao organismo relevante.

Também são providos organismos (isto é transformantes) transformados pelo presente sistema.

Onde referência a um nucleotídeo ou seqüência de proteína
20 particular é feita, será entendido que isto inclui referência a qualquer mutante ou variante dos mesmos, tendo atividade biológica substancialmente equivalente ao mesmo. Preferivelmente, o mutante ou variante tem pelo menos 85%, preferivelmente pelo menos 90%, preferivelmente pelo menos 95%, preferivelmente pelo menos 99%, preferivelmente pelo menos 99,9%, e
25 o mais preferível pelo menos 99,99% de identidade de seqüência com as seqüências de referência.

As seqüências providas podem tolerar alguma variação de seqüência e ainda emendar corretamente. Existem alguns nucleotídeos conhecidos por serem importantes. Estes são aqueles requeridos para todas

emendas, por exemplo como mostrado na Figura 34 abaixo. O GU inicial e o AG final do íntron são particularmente importantes e, por isso, preferidos, como discutido em outra parte, embora ~5% de íntrons iniciam GC, ao contrário. Esta seqüência de consenso é preferida, embora se aplique para todas as emendas, não especificamente a emenda alternativa. Na Figura 34, $P_u = A$ ou G ; $P_y = C$ ou U .

Preferivelmente, o sistema é ou compreende um plasmídeo. Como mencionado acima, isto pode ser quer DNA, RNA ou uma mistura de ambos. Se o sistema compreende RNA, então pode ser preferível traduzir de modo reverso o RNA em DNA por meio de uma transcriptase reversa. Se transcrição reversa é requerida, então o sistema pode também compreender uma seqüência de codificação para a proteína RT e um promotor apropriado para a mesma. Alternativamente, a RTase e promotor assim pode ser provido em um sistema separado, tais como um vírus. Neste caso, o sistema apenas seria ativado seguindo a infecção com aquele vírus. A necessidade de incluir seqüências cis-atuantes apropriadas para a transcriptase reversa ou RNA polimerase dependente de RNA seria evidente para o versado na técnica.

No entanto, é particularmente preferido que o sistema seja predominantemente DNA e mais preferivelmente consista apenas de DNA, pelo menos com relação às seqüências a serem expressadas no organismo.

Enquanto em algumas formas de realização a pelo menos uma seqüência polinucleotídica heteróloga a ser expressada em um organismo é uma seqüência polinucleotídica de RNA interferência (RNAi), é particularmente preferido que esta seja uma seqüência polinucleotídica capaz de codificar uma proteína funcional. A descrição irá focalizar predominantemente as seqüências polinucleotídicas codificando uma proteína funcional, mas será entendido que ela também refere-se a polinucleotídeos de RNA interferência (RNAi), salvo especificado em contrário.

Será entendido que referência é feita à códons de partida e de

parada entre os quais a seqüência polinucleotídica a ser expressada em um organismo é definida, mas que esta não exclui posicionamento de pelo menos uma seqüência de controle de emenda, elementos da mesma, ou outras seqüências, tais como íntrons, nesta região. De fato, será evidente da presente

5 descrição que a seqüência de controle de emenda, pode, em algumas formas de realização, estar posicionada nesta região.

Ademais, a seqüência de controle de emenda, por exemplo, pode sobrepor com o códon de partida pelo menos, no sentido que o G de ATG pode ser, em algumas formas de realização, o 5' G inicial da seqüência

10 de controle de emenda. Assim, o termo "entre" pode ser considerado como referindo a partir do início (3' para o nucleotídeo inicial, isto é A) do códon de partida, preferivelmente 3' para o segundo nucleotídeo do códon de partida (isto é T), até o lado 5' do primeiro nucleotídeo do códon de parada. Alternativamente, como será evidente por uma simples leitura de uma

15 seqüência polinucleotídica, o códon de parada pode também ser incluído.

A pelo menos uma seqüência polinucleotídica heteróloga a ser expressada em um organismo é uma seqüência heteróloga. Por "heterólogo", seria entendido que ele refere-se a uma seqüência que não seria, no tipo selvagem, normalmente encontrada em associação com, ou ligada a, pelo

20 menos um elemento ou componente da pelo menos uma seqüência de controle de emenda. Por exemplo, onde a seqüência de controle de emenda é derivada de um organismo particular, e o polinucleotídeo heterólogo é uma seqüência de codificação de uma proteína ou polipeptídeo, isto é, é uma seqüência polinucleotídica codificando uma proteína funcional, então a seqüência de

25 codificação poderia ser derivada, em parte ou completamente, de um gene do mesmo organismo, desde que a origem de pelo menos alguma parte da seqüência polinucleotídica transcrita não foi a mesma como a origem da pelo menos uma seqüência de controle de emenda. Alternativamente, a seqüência de codificação poderia ser de um organismo diferente e, nesse contexto,

poderia ser pensado como "exógeno". O polinucleotídeo heterólogo poderia também ser pensado como "recombinante", em que a seqüência de codificação para uma proteína ou polipeptídeo é derivada de locais diferentes, ou dentro do mesmo genoma (isto é o genoma de uma espécie simples ou subespécie) ou de genomas diferentes (isto é genomas de diferentes espécies ou subespécies).

Heterólogo pode referir-se à seqüência diferente da seqüência de controle de emenda e pode, assim, referir-se ao fato de o promotor, e outras seqüências tais como 5' UTR e/ou 3' UTR poderem ser heterólogas à seqüência polinucleotídica a ser expressada no organismo, desde que a referida seqüência polinucleotídica não é encontrada em associação com ou ligada operativamente ao promotor, 5' UTR e/ou 3' UTR, no tipo selvagem, isto é o contexto natural de referida seqüência polinucleotídica, se presente.

Será entendido que heterólogo também aplica-se a seqüências "designer" ou híbridas que não são derivadas de um organismo particular mas são baseadas em um número de componentes de organismos diferentes, como isto iria atender à exigência de que a seqüência e pelo menos um componente da seqüência de controle de emenda não estão ligados ou encontrados em associação no tipo selvagem, mesmo se uma parte ou elemento da seqüência híbrida for assim encontrada, enquanto pelo menos uma parte ou elemento não é. Preferivelmente, uma porção de pelo menos 50 nucleotídeos da seqüência híbrida não é encontrada em associação com pelo menos um componente da seqüência de controle de emenda, mais preferivelmente 200 nucleotídeos e o mais preferível, 500 nucleotídeos.

Também será entendido que versões sintéticas de seqüências de ocorrência natural são visadas. Tais seqüências sintéticas são também consideradas como heterólogas, a menos que elas sejam de seqüência idêntica a uma seqüência que seria, no tipo selvagem ou contexto natural, normalmente encontrada em associação com, ou ligada a, pelo menos um

elemento ou componente da pelo menos uma seqüência de controle de emenda.

Isto aplica-se igualmente para onde o polinucleotídeo heterólogo é um polinucleotídeo de RNA interferência.

5 Em uma forma de realização, onde a seqüência polinucleotídica a ser expressada compreende uma seqüência de codificação de uma proteína ou polipeptídeo, será entendido que referência a expressão em um organismo refere-se à provisão de uma ou mais seqüências de RNAs transcritos, preferivelmente mRNAs maduros, mas isto pode, preferivelmente, 10 também referir-se aos polipeptídeos traduzidos no referido organismo.

RT-PCR, que demonstra a presença de um transcrito, não de uma proteína, pode ser usado para identificar as seqüências de RNAs transcritos. Isto também é particularmente útil quando a própria proteína não é traduzida ou não é funcional ou não identificável por anticorpos criados 15 contra a proteína de ocorrência natural ou de tipo selvagem, devido a RNAi, modificação pós-translacional ou duplicação distorcida.

Em outra forma de realização, onde a seqüência polinucleotídica a ser expressada compreende polinucleotídeos de RNA interferência, também será entendido que a referência à expressão em um 20 organismo refere-se à interação dos polinucleotídeos de RNA interferência, ou transcritos do mesmo, na via de RNAi, por exemplo ligação de Dicer ou formação de RNA interferência pequeno (siRNA). De fato, é particularmente preferido que os polinucleotídeos de RNA interferência compreendam seqüências de siRNA e são, por esta razão, preferivelmente de 20-25 25 nucleotídeos de comprimento, especialmente onde o organismo é mamífero.

Em insetos e nematóides especialmente, prevê-se prover porção de dsRNA, por exemplo por formação de grampo de cabelo, que pode então ser processada pelo Sistema Dicer. Células de mamíferos geralmente produzem uma retro-alimentação de interferon contra as seqüências de

dsRNA longas, assim para células de mamíferos é mais comum prover seqüências menores, tais como siRNAs. Seqüências anti-sentido ou seqüências tendo homologia a microRNAs que são moléculas de RNA de ocorrência natural marcando os 3' UTRs de proteína são também visadas como seqüências de RNAi de acordo com uma forma de realização da presente invenção.

Cada seqüência de controle de emenda no sistema compreende pelo menos um sítio aceitador de emenda e pelo menos um sítio doador de emenda. O número de sítios doadores e aceitadores pode variar, dependendo do número de segmentos da seqüência que devem ser emendados juntos. Preferivelmente, sítios de ramificação estão incluídos em cada seqüência de controle de emenda. Um sítio de ramificação é a seqüência em que o doador de emenda está inicialmente ligado, ver figura 32, que mostra que a emenda ocorrendo em dois estágios, em que o éxon 5' é separado e então ligado ao éxon 3'.

Referindo-se à referida figura, o A é o único nucleotídeo essencial, e está, por esta razão, preferivelmente incluído. Sem estar limitado pela teoria, acredita-se que emenda de pré-mRNA prossegue através de um intermediário lariat, como é feito na auto-emenda de grupo II. Primeiro, clivagem ocorre na junção 5' - algumas vezes chamada de sítio doador de emenda. O fosfato na extremidade 5' do íntron então torna-se ligado ao 2' OH de uma adenina de aproximadamente 25 nucleotídeos a montante da extremidade 3' do íntron, que é algumas vezes chamado de sítio aceitador. Este resíduo A é chamado de ponto de ramificação. A próxima etapa é que a clivagem ocorre na junção de emenda 3' e o fosfato 5' do éxon a jusante é ligado ao 3'OH do éxon a montante.

Prefere-se particularmente que a maneira ou mecanismo de emenda alternativa seja específico de sexo. Preferivelmente, a seqüência de controle de emenda é derivada de um íntron tra. No entanto, é particularmente

preferido que o mecanismo de emenda alternativa seja derivado do Cctra do gene *transformer* de mosca-do-mediterrâneo, ou de outro ortólogo ou homólogo do gene *transformer* de *Drosophila*, preferivelmente de *C. rosa*, ou *B. zonata* especialmente um derivado de um mosca tefrítida da fruta.

5 Também prefere-se que a seqüência de controle de emenda seja derivada do mecanismo de emenda alternativa do Gene Actin-4, em particular que de *Aedes spp.* e mais preferível de AaActin-4, que é um gene de *Aedes/Stegomyia aegypti* que mostra emenda específica de tecido, estágio e de sexo.

10 Preferivelmente, emenda alternativa, particularmente a mediada por Actin-4, pode adicionar seqüências que influenciam a tradução ou estabilidade de RNA, por exemplo.

Também prefere-se que o mecanismo de emenda compreenda pelo menos um fragmento do gene *doublesex* (dsx), preferivelmente o derivado de *Drosophila*, *B. mori*, lagarta-rosada, traça das maçãs e pêras, ou um mosquito, em particular *A. gambiae* ou especialmente *A. aegypti*.

15 Prefere-se que a seqüência de controle de emenda e a seqüência polinucleotídica heteróloga codificando uma proteína funcional, definidas entre um códon de partida e um códon de parada, e/ou polinucleotídeos para RNA interferência (RNAi), a serem expressadas em um organismo, sejam providas na forma de uma construção de minigene ou um éxon de cassete.

20 Isto é particularmente preferido quando a seqüência de controle de emenda é derivada de dsx (preferivelmente minigene 1 como descrito nos exemplos e representado em SEQ ID NO. 149 (éxons estão presentes nas posições 1-135, 1311-2446 e 3900-4389 de SEQ ID NO. 149) que foi incluída na construção LA3491) ou Actin-4.

Os exemplos particularmente preferidos da presente invenção são providos nos exemplos, e podem ser selecionados dentre o grupo

consistindo dos plasmídeos ou construções, em particular qualquer um dos de acordo com qualquer uma das figuras 19-31, especialmente qualquer um dos plasmídeos mostrados nas Figuras 16-18, 22- 24, 26-32, 49, 52-55, e 61-69, e/ou SEQ ID NOs 46-48, 50-56, 143-145 e 151-162.

5 Preferivelmente, a proteína funcional a ser expressada em um organismo é tTAV, tTAV2 ou tTAV3.

Outras proteínas a serem expressadas no organismo são, ou como evidente previstas, em combinação com a referida proteína funcional, preferivelmente um gene letal como discutido em outra parte.

10 Um ORF contínuo pode também ser pensado como um ORF ininterrupto, isto é uma seqüência polinucleotídica em mRNA maduro, que não inclui nucleotídeos não de codificação, por exemplo aqueles tendo o potencial para serem traduzidos em aminoácidos. Nesta definição, prefere-se que o códon de parada não seja incluído.

15 Em algumas formas de realização, a pelo menos uma seqüência de controle de emenda regula a emenda alternativa por meio de ambos, nucleotídeos intrônicos e exônicos. No entanto, em uma forma de realização, é particularmente preferido que a pelo menos uma seqüência de controle de emenda seja uma seqüência de controle de emenda intrônica. Em
20 outras palavras, prefere-se que a pelo menos uma seqüência de controle de emenda seja substancialmente derivada de polinucleotídeos que formam parte de um íntron e são deste modo excisados do transcrito primário por emenda, de modo que estes nucleotídeos não são conservados na seqüência de mRNA maduro.

25 Assim, seqüências intrônicas podem ser pensadas como distintas de seqüências "exônicas", que são conservadas na molécula de RNA (pós emenda) processada. Onde a molécula de RNA processada codifica uma seqüência de proteína ou polipeptídeo, e é capaz de ser traduzida, isto é tem a estrutura correta e modificações tais como uma terminação, e um sinal de

poliadenilação, por exemplo, é conhecido como mRNA maduro ou processado e algumas das seqüências exônicas então codificam para aminoácidos, quando traduzidos.

Será entendido que em emenda alternativa, seqüências podem ser intrônicas sob algumas circunstâncias (isto é em algumas emendas alternativas variantes), mas exônicas sob outras circunstâncias (isto é em outras variantes). Assim, a pelo menos uma seqüência de controle de emenda da presente invenção é preferivelmente substancialmente derivada de polinucleotídeos que formam parte de um íntron em pelo menos uma variante de emenda alternativa, isto é, ou no primeiro produto de mRNA emendado ou em pelo menos um produto de mRNA alternativamente emendado. Assim, íntrons ou seqüências intrônicas podem ser vistos como emendados em pelo menos um transcrito ou tipo de transcrito.

Por exemplo, considerando o íntron tra de *C. capitata* (íntron Cctra), que é um exemplo particularmente preferido de uma pelo menos uma seqüência de controle de emenda de acordo com a presente invenção. De acordo com figura 2 A de Pane et al, reproduzida como figura 33, todos os 8 dos sítios de ligação de Tra/Tra2 putativos realçados estão na seqüência intrônica no sentido que eles estão em porções da seqüência emendada no F1 de transcrito mas por outro lado 6 dentre os 8 são exônicos no sentido que eles estão em éxons que estão incluídos ou retidos em quer M1 ou M2 de transcrito, ou ambos. Assim, estes sítios de ligação de Tra/Tra2 são intrônicos no presente sentido como eles são capazes de controlar a emenda alternativa, mas são emendados fora, isto é não presentes, em pelo menos uma variante de emenda alternativa, isto é pelo menos um mRNA que foi emendado em uma maneira alternativa de pré-RNA.

Em emenda "normal" (não alternativa) e em emenda alternativa, íntrons são geralmente removidos do pré-RNA para formar um mRNA emendado, que pode então ser traduzido em um polipeptídeo, tais

como uma proteína ou fragmento de proteína, tendo uma seqüência de aminoácido. Assim, será prontamente evidente para um versado como determinar aquelas seqüências do presente sistema que devem ser considerados intrônicas, em vez de exônicas.

5 Será, como evidente, notado que somente uma parte de um mRNA é De fato traduzida, isto é, tipicamente a parte entre o códon de partida e o códon de parada, embora será entendido que algumas vezes múltiplas partidas e paradas estão presentes. Assim, quando a referência é feita aqui para a tradução de uma seqüência de mRNA, será notado que isto é
10 referente à tradução da porção partindo no primeiro nucleotídeo do códon de partida e terminando após o último nucleotídeo antes da partida do códon de parada, que pode ser considerado como a porção de codificação.

 Como mencionado acima, seqüências exônicas podem estar envolvidas na mediação do controle de emenda alternativa, mas prefere-se
15 que pelo menos algumas seqüências intrônicas de controle estejam envolvidas na mediação da emenda alternativa. Em outras palavras, o sistema de expressão de gene da presente invenção pode também incluir seqüências de controle de emenda presentes em éxons, enquanto ocorre algum envolvimento intrônico de controle. Exemplos particularmente preferidos destes são as
20 seqüências de controle de emenda derivadas de ou contendo elementos do gene dsx, onde, sem estar vinculado pela teoria, pensa-se que as seqüências exônicas auxiliam no mecanismo de emenda alternativa.

 Assim, em algumas formas de realização, a pelo menos uma seqüência de controle de emenda não compreende seqüência exônica e será
25 entendido que isto é previsto por definições usadas para descrever a presente invenção. Assim, como será evidente, é possível para alguns nucleotídeos serem englobados dentro da definição de pelo menos uma seqüência de controle de emenda e também dentro da definição de uma seqüência polinucleotídica codificando uma proteína funcional. Em outras palavras, a

definição destes elementos pode se sobrepor, tal que determinados nucleotídeos podem ser cobertos pela definição de mais de um elemento.

No entanto, o versado reconhecerá que isto não é incomum em biologia molecular, como nucleotídeos podem freqüentemente realizar mais de um papel. Por exemplo, na presente invenção, um nucleotídeo pode formar parte de um seqüência de codificação de uma proteína funcional, mas poderia também formar parte de uma seqüência reconhecida e ligada por um fator de emenda, um exemplo de que a proteína TRA ou complexo TRA/TRA, como discutido em outra parte. Não é incomum que, por exemplo, alguns vírus
10 terem um genoma altamente concentrado onde a mesma extensão de polinucleotídeos pode codificar para duas ou mesmo três proteínas diferentes, cada lida em uma matriz diferente.

Como evidente, pode também ser que a seqüência de controle de emenda ou seqüências sejam unicamente intrônicas, isto é sem nenhuma
15 influência exônica. De fato, isto é particularmente preferido.

Em algumas formas de realização, prefere-se que a pelo menos uma seqüência de controle de emenda seja capaz de ser removida do pré-RNA, por emenda. Preferivelmente, a pelo menos uma seqüência de controle de emenda não resulta em uma armação em pelo menos uma variante de
20 emenda. Preferivelmente isto é uma variante de emenda codificando uma proteína funcional de comprimento completo. Em outras palavras, pelo menos a uma seqüência de controle de emenda preferivelmente não media a remoção de nucleotídeos que formam parte, ou foram planejados para formar parte de, uma seqüência polinucleotídica codificando uma proteína funcional, definida
25 entre um códon de partida e um códon de parada, e/ou polinucleotídeos de RNA interferência (RNAi), a serem expressados em um organismo. Por isso, entende-se que nucleotídeos que são excisados por emenda, em pelo menos uma variante de emenda, não são nucleotídeos que codificam aminoácidos na forma de tipo selvagem da proteína ou gene. Uma ou mais variantes de

emendas podem ter referidos nucleotídeos excisados, mas pelo menos uma variante deve reter estes nucleotídeos, de modo que uma armação não induzida em pelo menos uma variante. Estes nucleotídeos removidos são aqueles que são removidos em adição às seqüências que são normalmente emendadas tais como o íntron.

No entanto, em vista do acima, prevê-se também que as variantes de emenda diferentes podem resultar na mesma seqüência sendo lida em matrizes diferentes.

Interação da pelo menos uma seqüência de controle de emenda com maquinário de emenda celular, por exemplo o espliceossoma, conduz a ou media a remoção de uma série de, preferivelmente, pelo menos 50 nucleotídeos consecutivos do transcrito primário e ligação (emenda) junto de seqüências nucleotídicas que não foram consecutivas no transcrito primário (porque elas, ou seu complemento se a seqüência anti-sentido for considerada, não foram consecutivas na seqüência de modelo original do qual o transcrito primário foi transcrito). Referida série de pelo menos 50 nucleotídeos consecutivos compreende um íntron. Esta mediação atua preferivelmente em um modo específico de sexo, específica de estágio, específica de linhagem germinal ou específica de tecido, ou combinação das mesmas, tal que os transcritos primários equivalentes em sexos, estágios, tipos de tecido, diferentes etc, tendem a remover íntrons de diferente tamanho ou seqüência, ou em alguns casos podem remover um íntron em um caso mas não no outro. Este fenômeno, a remoção de íntrons de tamanho diferente ou seqüência em circunstâncias diferentes, ou a remoção diferencial de íntrons de um tamanho ou seqüência dado, em circunstâncias diferentes, é conhecido como emenda alternativa. Emenda alternativa é um fenômeno bem conhecido na natureza, e muitos exemplos são conhecidos, ver acima.

Em algumas formas de realização, a pelo menos uma seqüência de controle de emenda é associada com uma matriz de leitura

aberta heteróloga tal que, em pelo menos uma variante de emenda, a matriz de leitura aberta heteróloga é separada, por exemplo por um códon de parada ou armação, enquanto em pelo menos uma variante de emenda alternativa a matriz de leitura aberta heteróloga não é separada. Transcritos do segundo tipo codificam ou potencialmente codificam uma proteína funcional, considerando que aqueles do primeiro tipo codificam uma proteína com função alterada, separada ou mesmo sem função, atividade ou estabilidade relativa àqueles de segundo tipo.

5 Geralmente, será evidente para o versado na técnica que a
10 matriz de leitura aberta heteróloga pode ser, ela mesma, um compósito ou fusão de seqüências de várias fontes. Emenda para produzir uma proteína funcional pode ainda produzir uma proteína alterada relativa à matriz de leitura aberta heteróloga de protótipo, por exemplo se o íntron emendado
15 alternativamente inserido incluir seqüência que é exônica em todas as formas de emenda alternativa, e por esta razão conservada em mRNAs maduros do segundo tipo. No entanto, é particularmente preferido que pelo menos um transcrito remova todas, ou substancialmente todas, das seqüências emendadas alternativamente inseridas, tal que a matriz de leitura aberta heteróloga seja restaurada, ou substancialmente restaurada, para a forma
20 intacta, com poucas ou nenhuma seqüência endogenamente associada com o íntron restante no mRNA maduro. Endógeno é usado aqui em contraste a heterólogo, então será entendido que este refere-se a uma seqüência que seria, no tipo selvagem, normalmente encontrada em associação com, ou ligada a, pelo menos um elemento ou componente da pelo menos uma seqüência de
25 controle de emenda.

Alternativamente, um ou mais transcritos podem remover nucleotídeos adicionais, de modo que a matriz de leitura aberta heteróloga é separada, não pela inserção de nucleotídeos extras (por exemplo códon de parada ou armação, mas também seqüência potencialmente de codificação

que separa a função), mas sim por deleção de nucleotídeos da matriz de leitura aberta heteróloga, por exemplo de tal forma a induzir uma armação. Uma ou mais variantes de emenda podem ter nucleotídeos excisados, mas pelo menos uma variante precisa conservar estes nucleotídeos, de modo que

5 uma armação não é induzida em pelo menos uma variante. Estes nucleotídeos removidos são aqueles que são removidos em adição às seqüências que são normalmente emendadas tais como o íntron, onde uma seqüência intrônica pode ser considerada como uma que forma parte de um íntron em pelo menos uma variante de emenda alternativa do análogo natural.

10 Quando nucleotídeos exônicos devem ser removidos, então estes devem ser removidos em múltiplos de três, se é desejado evitar uma armação, mas como um nucleotídeo único ou múltiplos de dois (que não são também múltiplos de três) se é desejado induzir uma armação. Será notado que se apenas um ou alguns múltiplos de dois nucleotídeos são removidos,

15 então isto poderia conduzir a uma seqüência de proteína completamente diferente sendo codificada em ou em torno da junção de emenda do mRNA.

Este é particularmente o caso em uma forma de realização do sistema onde éxon de cassetes são usados para interromper uma matriz de leitura aberta em algumas variantes de emenda mas não outras, tais como em,

20 por exemplo, tra, especialmente Cctra.

Em outra forma de realização preferida da presente invenção, toda ou parte de uma matriz de leitura aberta está em um éxon de cassete, por exemplo algumas formas de realização de Dsx derivadas de *Aedes*, são providas com, por exemplo, uma região de codificação de tTAV em um éxon

25 de cassete que está apenas presente em variantes de emendas específicas de fêmeas.

Onde a mediação de emenda alternativa é específica de sexo, prefere-se que a variante de emenda codificando uma proteína funcional a ser expressada em um organismo seja a variante de emenda F1, isto é uma

variante de emenda encontrada apenas ou predominantemente em fêmeas, e preferivelmente é a variante a mais abundante encontrada em fêmeas, apesar da mesma não ser essencial. Correspondentemente para configurações onde toda ou parte de uma matriz de leitura aberta funcional está em um éxon de cassete, prefere-se que este éxon de cassete esteja incluído em transcritos encontrados apenas ou predominantemente em fêmeas, e preferivelmente tais transcritos são, individualmente ou em combinação, as variantes as mais abundantes encontradas em fêmeas, apesar disto não ser essencial.

Em uma forma de realização preferida, seqüências são incluídas em uma seqüência híbrida ou recombinante ou construção que são derivadas de seqüências intrônicas de ocorrência natural que são elas próprias sujeitas a emenda alternativa, em seu contexto nativo ou original. Por esta razão, uma seqüência intrônica pode ser considerada como uma que forma parte de íntron em pelo menos uma variante de emenda alternativa do análogo natural. Assim, seqüências correspondentes a extensões contíguas únicas de seqüência intrônica de ocorrência natural são previstas, mas também híbridos de tais seqüências, incluindo híbridos de duas seqüências intrônicas de ocorrência natural diferentes, e também seqüências com deleções ou inserções relativas a extensões contíguas únicas de seqüência intrônica de ocorrência natural, e híbridos das mesmas. Referidas seqüências derivadas de seqüências intrônicas de ocorrência natural poderão ser associadas, na invenção, com seqüências não fazendo parte, elas mesmas, de qualquer íntron de ocorrência natural. Se tais seqüências forem transcritas, e preferivelmente retidas no RNA maduro em pelo menos uma variante de emenda, elas podem ser então consideradas exônicas.

Também será notado que a referência a uma "armação" poderia também referir-se a uma codificação direta de um códon de parada, que também é provável de conduzir a uma proteína não funcionando como seria uma ruptura da seqüência de mRNA emendado causada por inserção ou

deleção de nucleotídeos. Produção de variantes de emenda diferentes de duas ou mais proteínas diferentes ou seqüências de polipeptídeo de função diferencial é também prevista, em adição à produção de duas ou mais proteínas diferentes ou seqüências de polipeptídeo das quais uma ou mais não tem função prevista ou discernível. Também é prevista a produção de variantes de emenda diferentes de duas ou mais proteínas diferentes ou seqüências de polipeptídeo de função similar, mas diferindo a localização subcelular, estabilidade ou capacidade de se ligar a ou se associar com outras proteínas ou ácidos nucléicos.

10 Preferivelmente, a pelo menos uma seqüência de controle de emenda é intrônica e compreende em sua extremidade 5' um nucleotídeo de guanina (G). Em outras palavras, nucleotídeo 5' da seqüência de controle de emenda, 3' para o sítio doador de emenda, e preferivelmente na superfície ou junção do éxon com seqüência de controle de emenda, é Guanina (G), no pré-
15 RNA, ou C em uma seqüência de DNA anti-sentido correspondendo à mesma.

Além disso, o nucleotídeo adjacente (3' para referida G) é preferivelmente Citosina (C) no pré-RNA, ou uma G correspondente em uma seqüência de DNA, mas é mais preferível Uracila (U) no pré-RNA, ou um A
20 correspondente na seqüência anti-sentido de DNA. Assim, os dois nucleotídeos 5' da seqüência de controle de emenda são preferivelmente 5'GT com relação ao filamento sentido de DNA, 5'-GU no transcrito primário.

Preferivelmente, pelo menos uma seqüência de controle de emenda intrônica também compreende em sua extremidade 3' um nucleotídeo
25 de guanina 3' e preferivelmente AG-3' na junção do sítio aceitador de emenda com éxon, por exemplo, ver figura 34.

Preferivelmente, a seqüência de flanqueamento 5' para o sítio doador de emenda no sistema compreende 5'-TG, de modo que a seqüência pode ser representada 3' de seqüência de controle de emenda-** de 5'-TG-*,

onde * representa o sítio doador de emenda e ** representa o sítio aceitador de emenda.

Preferivelmente, a seqüência de controle de emenda é também flanqueada em seu lado 3' por um nucleotídeo de G, e mais preferível por nucleotídeos de GT, tal que a seqüência poderia ser representada como: 5'-TG- *- seqüência de controle de emenda ** -GT-3' *. Será notado que esta é a seqüência de DNA de filamento sentido (TG). Assim, o pré-RNA transcrito irá ler UG por exemplo, onde U substitui T.

Derivados de Guanina ou Timina tendo a mesma função são também visados.

Prefere-se particularmente que a emenda seja específica de sexo e ainda mediada ou controlada ligando a proteína TRA ou complexo de proteína TRA/TRA2, ou homólogos da mesma. Em insetos, por exemplo, a proteína TRA é diferencialmente expressada em sexos diferentes. Em particular, a proteína TRA é conhecida como estando presente basicamente em fêmeas, por esta razão, media a emenda alternativa de tal forma que um seqüência de codificação é expressada em um modo específico de sexo, isto é, que em alguns casos uma proteína é expressada apenas em fêmeas ou em um nível muito maior em fêmeas do que em machos ou, alternativamente, em outros casos uma proteína é expressada apenas em machos, ou em um nível muito maior em machos do que em fêmeas. Enquanto prefere-se que a proteína seja expressada apenas em machos, é particularmente preferido que a proteína seja expressada apenas em fêmeas, No entanto, o mecanismo para realizar esta emenda alternativa específica de sexo mediada pela proteína TRA ou pelo complexo TRA/TRA2 é conhecido e é discutido, por exemplo, em Pane et al (Development 129, 3715-3725 (2002)).

Preferivelmente, a pelo menos uma seqüência de controle de emenda compreende, e mais preferivelmente consiste de, o íntron tra derivado do gene tra de *Ceratitis capitata* (Cctrd), que tem uma região alternativamente

emendada. No transcrito F1, como ilustrado pela figura 33 (figura 2A de Pane et al (2002) supra), este é o primeiro íntron. Homólogos do gene tra em outras espécies, tais como *Bactrocera oleae*, *Ceratitis rosa*, *Bactrocera zonata* e *Drosophila melanogaster* também tem região alternativamente emendada em um local similar dentro da seqüência de codificação de tra. Íntrons tras derivados destes insetos também são particularmente preferidos.

O padrão de emenda em Cctra em particular é bem conservado, com aqueles transcritos encontrados em machos contendo material exônico adicional relativo ao transcrito de F1, tal que estes transcritos não codificam proteína funcional de comprimento completo. Por contraste, o transcrito de F1 não codifica proteína Tra de comprimento completo; este transcrito é substancialmente específico de fêmea na maior parte dos estágios do ciclo da vida, ainda que seja especulado que embriões muito prematuros de ambos os sexos podem conter uma quantidade pequena deste transcrito. Os requerentes descreveram a seqüência emendada do transcrito de F1, mas não os transcritos específicos de machos ou não específicos de sexo, como o íntron tra, ou mesmo o íntron F1 de tra. Assim a versão desta seqüência encontrada no gene CCtra é o íntron Cctra.

Assim o gene tra é regulado em parte por emenda alternativa específica de sexo, enquanto seu produto chave, a proteína Tra, é ela própria envolvida em emenda alternativa. Em insetos, emenda alternativa específica de sexo mediada pela proteína Tra, ou um complexo compreendendo as proteínas TRA e TRA2, incluem seqüências de controle de emenda de Dípteros derivadas do gene *doublesex* (dsx) e também o próprio íntron tra, embora o mesmo excluiria o íntron tra de *Drosophila* (Dmtra), que é principalmente mediado pelo produto de gene Sxl em *Drosophila*, em vez de TRA ou complexo TRA/TRA2.

Fora de *Drosophila*, o produto de gene Sxl não é diferencialmente expressado em sexos diferentes. Não se pensa que Sxl atue

na mediação de emenda alternativa específica de sexo em insetos não drosófilos.

Exemplos da proteína TRA que liga aos sítios de proteína de ligação (as seqüências nucleotídicas especificamente reconhecidas pela proteína Tra) no íntron tra são preferivelmente de Díptera, preferivelmente da família *Tephritidae*, mais preferivelmente dos gêneros *Ceratitis*, *Anastrepha* ou *Bactrocera*. No entanto, prevê-se também que outros dípteros, tais como drosófilos ou mosquitos de várias formas discutidas acima, também são capazes de prover a proteína TRA ou homólogos da mesma, que são capazes de se ligar aos sítios apropriados nas seqüências de controle de emenda derivadas do gene *dsx*, o gene *tra* ou o íntron tra, isto é o íntron tra alternativamente emendado completamente removido no transcrito de F1, mesmo naqueles casos, tais como *Drosophila*, onde o gene *tra* natural (*Dmtra*) não é, ele mesmo, regulado por proteína Tra. Em algumas formas de realização, o "íntron tra" pode ser definido como uma seqüência de controle de emenda onde emenda alternativa do transcrito de RNA é regulada por TRA, por exemplo sua ligação, sozinha ou em combinação (isto é quando complexada) com TRA2. Isto exclui o íntron tra de *Drosophila*.

Prefere-se particularmente que as seqüências de controle de emenda sejam derivadas do íntron tra. Referido íntron tra pode ser derivado, como discutido em outra parte, de *Ceratitis*, *Anastrepha* ou *Bactrocera*. O íntron tra de *Ceratitis capitata* do gene *transformer* foi inicialmente caracterizado por Pane et al (2002), supra. No entanto, será notado que homólogos existem em outras espécies, e podem ser facilmente identificados em referidas espécies e também em seus vários gêneros. Assim, quando referência é feita a tra será notado que isto também refere-se aos homólogos de tra em outras espécies, especialmente em espécies de *Ceratitis*, *Anastrapha* ou *Bactrocera*.

Por "derivado" será entendido que, usando a referência ao

íntron tra, isto refere-se a seqüências que aproximam-se a ou replicam exatamente o íntron tra, como descrito na técnica, nesse caso por Pane et al (2002), supra. No entanto, será notado que, como estas são seqüências intrônicas, alguns nucleotídeos podem ser adicionados ou substituídos sem uma perda substancial na função.

Exemplos preferidos disso incluem o íntron dsx, preferivelmente provido na forma de um minigene. Nesse exemplo, pode ser preferível deletar, como feito nos exemplos, quantidades relativamente grandes de íntrons alternativamente emendados, por exemplo 90% ou mais de um íntron em alguns casos, enquanto ainda conservando a função de emenda alternativa. Assim, apesar de deleções grandes estarem previstas, prevê-se também que inserções, substituições ou deleções mesmo menores, por exemplo de nucleotídeo único, também são preferidas.

O comprimento exato da seqüência de controle de emenda derivado do íntron tra não é essencial, desde que seja capaz de mediar a emenda alternativa. A este relação, pensa-se que cerca de 55 a 60 nucleotídeos é o comprimento mínimo de um íntron tra modificado, embora o íntron tra de tipo selvagem (variante de emenda F1) de *C. capitata* é na região de 1345 nucleotídeos de comprimento.

Prefere-se particularmente que a seqüência de 1345 ntd de comprimento completo de Cctra seja usada.

Como com todas as seqüências nucleotídicas discutidas aqui, prefere-se que um certo grau de homologia de seqüência seja prevista, salvo de outra forma evidente. Assim, prefere-se que a seqüência de controle de emenda tenha pelo menos 80% de homologia de seqüência com a referência SEQ ID NO., preferivelmente pelo menos 80% de homologia de seqüência com a referência SEQ ID NO., preferivelmente pelo menos 80% de homologia de seqüência com a referência SEQ ID NO., mais preferivelmente pelo menos 90% de homologia de seqüência com a referência SEQ ID NO.,

mais preferivelmente pelo menos 95% de homologia de seqüência com a referência SEQ ID NO., ainda mais preferivelmente pelo menos 99% homologia de seqüência com a referência SEQ ID NO., e a mais preferível, pelo menos 99.9% de homologia de seqüência com a referência SEQ ID NO.

5 Um algoritmo apropriado tal como BLAST pode ser usado para determinar homologia de seqüência. Se quantidades grandes de seqüência são deletadas do tipo selvagem, então a comparação de seqüência pode ser sobre o comprimento completo do tipo selvagem ou sobre seqüências alinhadas de homologia similar.

10 No entanto, será entendido que apesar da homologia de seqüência acima, determinados elementos, em particular os nucleotídeos de flaqueamento e sítio de ramificação de emenda devem ser conservados, para funcionamento eficiente do sistema. Em outras palavras, enquanto porções podem ser deletadas ou, de outro modo, alteradas, funcionalidade ou atividade de emenda alternativa, a pelo menos 30%, preferivelmente 50%,
15 preferivelmente 70%, mais preferivelmente 90%, e, a mais preferível, 95% comparada ao tipo selvagem deve ser conservada. Isto pode ser aumentado do tipo selvagem, bem como, por engenharia apropriada dos sítios que ligam fatores de emenda alternativa ou interação com spliceossoma, por exemplo.

20 Em particular, prefere-se que onde a seqüência de controle de emenda compreende um íntron tra modificado, ela compreenda pelo menos 20 a 40 pares de base de 5' e, preferivelmente, assim a extremidade 3' de referido íntron. Além disso, prefere-se que pelo menos 3 ou 4 e, mais preferido, pelo menos 5, preferivelmente 6, mais preferivelmente 7 e, mais
25 preferido, todos os 8 dos 8 domínios de ligação de TRA putativos do íntron tra de *C. capitata*, como ensinado por Pane et al (2002), ou homólogos disso, sejam providos. Como evidente, se mais de tais sítios forem descobertos em devido curso, então é visado que a seqüência de controle de emenda poderia incluir mais do que 8 sítios. De fato, é previsto que mais do que 8 sítios

possam ser engenheirados na sequência de controle de emenda e que a emenda alternativa possa ser regulada desta forma, especialmente se alguns sítios estiverem ligados com afinidades diferentes levando um resultado de emendas alternativas diferentes.

5 Uma sequência de consenso par os domínios de ligação de TRA putativos do íntron tra de *C. capitata* é dada abaixo como SEQ ID NO 1, uma sequência de DNA, embora o RNA correspondente equivalente é também preferido.

10 As seqüências de consenso preferidas são 1: TCWWCRATCAACA (SEQ ID NO. 1), onde W = A ou T e R = A ou G.

Considerações similares aplicam-se a *doublesex*, onde a sequência de consenso da proteína TRA é também aquela dada em SEQ ID NO. 1 como um complexo de proteína compreendendo as 'proteínas Tra e TRA2 é um regulador chave de emenda alternativa de *doublesex*, como é de 15 homólogos tra (embora não sejam os homólogos tra encontrados em drosofilídeos).

Como mencionado acima, as seqüências de controle de emenda são preferivelmente derivadas do íntron tra, preferivelmente da família *Tephritidae*. Prefere-se particularmente que o íntron tra seja derivado 20 de *B. zonata* ou, preferivelmente, de outras moscas de frutas não drosofilídeos. No entanto, é particularmente preferido que o íntron tra seja derivado dos gêneros de *Ceratitis*, em particular *C. rosa* e, o mais preferível, *C. capitata*. Estes são mais vastamente conhecidos como as moscas-de-natal e do-mediterrâneo, respectivamente.

25 Com relação ao íntron tra derivado de *B. zonata*, os requerentes demonstraram que isto pode levar a uma emenda alternativa específica de sexo em mosca mexicana das frutas transgênica (*Anastrepha ludens*) e em mosca-do-mediterrâneo transgênica (*C. capitata*). Os requerentes também demonstraram que uma variedade de proteínas pode ser expressada

em um modo específico de sexo através de emenda alternativa, incluindo tTAV 3 e Rpr.

Em relação ao íntron tra derivado de *C. rosa*, os requerentes demonstraram com sucesso emenda alternativa em um modo específico de sexo de um transgene em mosca-do-mediterrâneo.

Com relação ao íntron tra derivado de *C. capitata* (mosca-do-mediterrâneo), os requerentes demonstraram que este pode mediar a emenda específica de sexo em mosca-do-mediterrâneo transgênica, e outras tefritidas, e outras tefritidas tais como *A. ludens* (mosca mexicana das frutas). Não apenas isso, os requerentes demonstraram que este íntron pode trabalhar com sucesso através de toda faixa de insetos e, em particular, dípteros. De fato, os requerentes demonstraram que o íntron tra de *C. capitata* (referido como Cctra) pode prover emenda alternativa específica de sexo em *Drosophila* transgênica, que não é uma tefritida, e também no mosquito *Aedes aegypti*. Embora mosquitos sejam dípteros, eles divergiram de drosófilos e tefritidos cerca de 250 milhões de anos atrás e, por esta razão, são parentes bem mais distantes que os drosofilídeos são para as tefritidas, para os quais o tempo de divergência foi estimado como 120-150 milhões de anos. Assim, isto mostra uma ampla aplicabilidade da presente invenção através de uma extensa faixa de insetos.

No que diz relação a seqüências de controle de emenda derivadas do íntron dsx, também os requerentes demonstraram que isto pode ser usado para emendar alternativamente, em um modo específico de sexo, em uma ampla faixa de insetos. Conseqüentemente, é particularmente preferido que o dsx seja derivado de *Bombyx mori* (bicho da seda), *Pectinophora gossypiella* (lagarta rosada) *Pectinophora gossypiella*, *Cydia pomonella* (traça das maçãs e pêras), *Drosophila*, e mosquitos tais como *Anopheles* sp., por exemplo *A. gambiae*. Mosquitos particularmente preferidos incluem *Stegomyia* spp., particularmente *S. aegypti* (também

conhecido como *Aedes aegypti*).

De fato, em *A. aegypti*, os requerentes demonstraram um número considerável de construções de DNA, que são capazes de prover emenda alternativa específica de sexo.

5 Será notado que o sistema ou construção seja preferivelmente administrado como um plasmídeo, mas geralmente testado após integrar no genoma. Administração pode ocorrer por métodos conhecidos na técnica, tais como parenteral, intravenosa, intramuscular, oral, transdérmica, liberada através de uma membrana mucosa, e assim por diante. Injeção em embriões é particularmente preferida. O plasmídeo pode ser linearizado antes ou durante 10 administração, e nem todos dentre os plasmídeos podem ser integrados no genoma. Onde apenas parte do plasmídeo é integrada no genoma, prefere-se que esta parte inclua a pelo menos uma seqüência de controle de emenda capaz de mediar a emenda alternativa.

15 Preferivelmente, o sistema de expressão de polinucleotídeos é um sistema genético letal dominante recombinante, cujo efeito letal é condicional. Condições apropriadas incluem temperatura, de modo que o sistema é expressado a uma temperatura mas não, ou em menor grau, em outra temperatura, por exemplo. O sistema genético letal pode atuar em 20 células ou tecidos específicos ou impor seu efeito em todo o organismo. Sistemas que não são estritamente letais mas impõem um custo substancial à aptidão física são também previstos, por exemplo levando à a cegueira, aptidão ao voo (para organismos que poderiam normalmente voar), ou esterilidade. Sistemas que interferem com a determinação de sexo são 25 também previstos, por exemplo transformando ou tendendo a transformar todo ou parte de organismo de um tipo sexual a outro. Será entendido que todos tais sistemas e conseqüências são englobados pelo termo letal como usado aqui. Similarmente, "matar", e termos similares referem-se à expressão efetiva do sistema letal e assim a imposição de um fenótipo deletério ou de

deformidade sexual, por exemplo morte.

5 Mais preferivelmente, o sistema de expressão de polinucleotídeo é um sistema genético letal dominante recombinante, cujo efeito letal é condicional e não é expressado sobre condições permissivas requerendo a presença de uma substância que está ausente do ambiente natural do organismo, tal que o efeito letal do sistema letal ocorre no ambiente natural do organismo.

10 Em outras palavras, as seqüências de codificações codificam um ligado letal a um sistema tal como o sistema tet descrito em WO 01/39599 e/ou WO2005/012534.

15 De fato, prefere-se que a expressão de referido gene letal ocorra sob o controle de uma proteína de transativador reprimível. Também prefere-se que o gene cuja expressão é regulada por emenda alternativa codifique uma proteína de transativador tal como tTA. Isto não é incompatível com a proteína regulada ser uma letal. De fato, é particularmente preferido que ela seja os dois. A este respeito, os requerentes preferem particularmente que o sistema inclua um sistema de retro-alimentação positivo como ensinado em WO2005/012534.

20 Preferivelmente, o efeito letal do sistema letal dominante é condicionalmente supressivo.

25 Organismos apropriados sobre os quais o presente sistema pode ser usado incluem mamíferos tais como camundongos, ratos e animais de criação. Também são preferidos peixes, tais como salmão e truta. Plantas também são preferidas, mas é particularmente preferido que o organismo hospedeiro seja um inseto, preferivelmente um díptero ou tefrítidos. Preferivelmente, o organismo não é um humano, preferivelmente não é um mamífero, preferivelmente não é um pássaro, preferivelmente é um invertebrado, preferivelmente é um artrópode.

Em particular, prefere-se que o inseto seja da Ordem Díptera,

especialmente dípteros superiores e particularmente ele é uma mosca tefrítida das frutas, preferivelmente mosca-do-mediterrâneo (*Ceratitis capitata*), preferivelmente mosca mexicana das frutas (*Anastrepha ludens*), preferivelmente mosca oriental das frutas (*Bactrocera dorsalis*), mosca da azeitona (*Bactrocera oleae*), mosca do melão (*Bactrocera cucurbitae*), mosca-de-natal (*Ceratitis rosa*), mosca da cereja (*Rhagoletis cerasi*), mosca das frutas de Queensland (*Bactrocera tyroni*), mosca do pêssgo (*Bactrocera zonata*), mosca das frutas do Caribe (*Anastrepha suspensa*) ou mosca das frutas da Índia Ocidental (*Anastrepha obliqua*). Também é particularmente preferido que o organismo hospedeiro seja um mosquito, preferivelmente dos gêneros *Stegomyia*, *Aedes*, *Anopheles* ou *Culex*. São particularmente preferidos *Stegomyia aegyptae*, também conhecidos como *Aedes aegypti*, *Stegomyia albopicta* (também conhecidos como *Aedes albopictus*), *Anopheles stephensi*, *Anopheles albimanus* e *Anopheles gambiae*.

Dentro dos Dípteros, outro grupo preferido é *Calliphoridae*, particularmente o verme-parafuso do Novo Mundo (*Cochliomyia hominivorax*), verme-parafuso do Velho Mundo (*Chrysomya bezziana*) e mosca australiana da ovelha (*Lucilia cuprina*). Lepidópteros e Coleópteros também são preferidos, especialmente traças, incluindo traça das maçãs e pêras (*Cydia pomonella*), e o bicho da seda (*Bombyx mori*), a lagarta rosada (*Pectinophora gossypiella*), a lagarta da couve (*Plutella xylostella*), a mariposa (*Lymantria dispar*), a traça de laranja-da-baía (*Amyelois transitella*), a lagarta mineira do pessegueiro (*Anarsia lineatella*), e a broca da haste do arroz (*Tryporyza incertulas*), as traças noctuídas, especialmente *Heliothiane*.

Dentre os Coleópteros, o besouro japonês (*Popilla japonica*), o besouro franjado branco (*Graphognatus spp.*), o gorgulho do algodão (*Anthonomous grandis*), a lagarta da raiz do milho (*Diabrotica spp*), e escaravelho da batateira (*Leptinotarsa decemlineata*) são particularmente preferidos.

Preferivelmente, o inseto não é um drosofilídeo, especialmente

Dm. Assim, em algumas formas de realização, expressão em drosofilídeos, especialmente Dm é excluída. Em outras formas de realização, a seqüência de controle de emenda não é derivada do íntron tra de um drosofilídeo, especialmente Dm.

5 Prefere-se que a expressão da seqüência polinucleotídica heteróloga conduza a uma conseqüência fenotípica no organismo. Prefere-se particularmente que a proteína funcional não seja beta-galactosidase, mas possa estar associada com marcadores visíveis (incluindo fluorescência), viabilidade, fertilidade, fecundidade, conveniência, aptidão ao vôo, visão, e
10 diferenças comportamentais. Será notado, como evidente, que, em algumas formas de realização, os sistemas de expressão são tipicamente condicionais, com o fenótipo sendo expressado apenas sobre algumas condições, por exemplo restritivas.

 Em um outro aspecto, provê-se também um método de
15 controle populacional de um organismo em um ambiente natural para o mesmo, compreendendo:

 i) reproduzir um lote do organismo,
 o organismo transportando um sistema de expressão de gene compreendendo um sistema de acordo com a presente invenção que é um
20 sistema genético letal dominante,

 ii) distribuir os referidos animais de lote no meio ambiente em um local para controle populacional; e

 iii) obter um controle da população através da letalidade em estágio prematuro por expressão do sistema letal na descendência que resulta
25 da inter-reprodução dos referidos indivíduos no lote com os indivíduos do sexo oposto da população selvagem.

 Preferivelmente, a letalidade de estágio prematuro é embriônica ou antes de maturidade sexual, preferivelmente no início de desenvolvimento, mais preferível, nos estágios prematuros da vida larval ou

embriônica.

Preferivelmente, o efeito letal do sistema letal é condicional e ocorre no referido ambiente natural através da expressão de um gene letal,

5 a expressão de referido gene letal estando sob controle de uma proteína de transativados reprimível, a referida reprodução ocorrendo sob condições permissivas na presença de uma substância, a substância estando ausente do referido ambiente natural e capaz de reprimir referido transativador.

10 Preferivelmente, o efeito letal é expressado nos embriões de referida descendência. Preferivelmente, o organismo é um animal multicelular invertebrado ou é como discutido em outra parte.

Também provê-se um método de controle biológico, compreendendo:

15 i) reproduzir um lote de organismos machos e fêmeas transformados com o sistema de expressão de acordo com a presente invenção sob condições permissivas, permitindo a sobrevivência dos machos e fêmeas, para dar um agente de controle biológico de sexo dual;

20 ii) opcionalmente antes da próxima etapa impor ou permitir condições restritivas para causar a morte de indivíduos de um sexo e, assim, prover um agente de controle biológico de sexo único compreendendo indivíduos do outro sexo transportando o sistema genético letal condicional;

iii) liberar o agente de controle biológico de sexo dual ou de sexo único no meio ambiente em um locus para controle biológico; e

25 iv) obter controle biológico através da expressão do sistema genético na descendência resultante de inter-reprodução dos indivíduos do agente de controle biológico com indivíduos do sexo oposto da população selvagem.

Preferivelmente, ocorre separação de sexos antes da distribuição de organismo por expressão de um sistema genético letal

específico de sexo.

Preferivelmente, o efeito letal resulta em matar mais de 90% da classe alvo da progênie de acasalamentos entre os organismo liberados e a população selvagem.

5 Também provê-se um método de separação de sexos compreendendo:

i) reproduzir um lote de organismos machos e fêmeas transformados com o sistema de expressão de gene sob condições permissivas ou restritivas, permitindo a sobrevivência dos machos e fêmeas; e

10 ii) remover as condições permissivas ou restritivas para induzir o efeito letal do gene letal em um sexo e não em outro por emenda alternativa específica de sexo do gene letal.

15 Preferivelmente, o efeito letal resulta em matar mais do que 90% da classe alvo da progênie de acasalamentos entre organismos liberados e a população selvagem.

Também provê-se um método de controle biológico ou populacional compreendendo;

20 i) reproduzir um lote de organismo machos e fêmeas transformados com o sistema de expressão de genes sob condições permissivas ou restritivas, permitindo a sobrevivência dos machos e fêmeas;

ii) remover as condições permissivas ou restritivas para induzir o efeito letal do gene letal em um sexo e não em outro por emenda alternativa específica de sexo do gene letal para obter uma separação de sexos;

25 iii) esterilizar ou parcialmente esterilizar os indivíduos separados e

iv) realizar referido controle através de da liberação dos indivíduos estéreis ou parcialmente estéreis separados no meio ambiente natural do organismo.

Preferivelmente, a esterilização é realizada através do uso e

radiação ionizante. Geralmente, no entanto, métodos evitando irradiação, como usado na técnica de inseto estéril (SIT) são especialmente preferidos e têm muitas vantagens de custo e saúde sobre os métodos associados com ou seguidos pelo uso de radiação.

5 Também provê-se um método para eliminar seletivamente fêmeas de uma população. O equivalente para machos também é previsto.

Métodos de separação de sexo são enormemente importantes comercialmente em, por exemplo, bichos da seda, onde os machos produzem seda em maior e melhor quantidade do que as fêmeas. Assim, métodos de
10 separação de sexo que eliminam as fêmeas e, em particular bichos da seda fêmeas são particularmente preferidos.

Prevê-se também que a proteína funcional pode ser uma expressada diferencialmente, mas detectável em mais de um variante de emenda e preferivelmente, por esta razão, em ambos sexos, por exemplo. Tais
15 exemplos incluem uma proteína fluorescente, tais como eGFP, CopGFP e DsRed2. Esta pode ser usada em um método de separação ou escolha de sexo não letal, de modo que se pode separar os dois tipos sem matar qualquer um deles.

Os requerentes verificaram surpreendentemente que o
20 posicionamento da seqüência de controle de emenda pode ser alterado e melhores resultados obtidos. Preferivelmente, a seqüência de controle de emenda é a "primeira" seqüência de controle de emenda, quando lida a partir do promotor, na direção de 5' a 3'. Os requerentes verificaram que em determinadas construções com um íntron em 5' UTR do sistema este fato
25 conduzida a reduzidos níveis ou expressão de proteína alternativamente emendada mediada pela seqüência de controle de emenda da presente invenção.

Preferivelmente, a seqüência de controle de emenda é 3' para o códon de partida. Preferivelmente, a seqüência de controle de emenda está

inserida dentro do primeiro éxon, isto é a extensão de sequência imediatamente 3' para o sítio de partida de transcrição. Será entendido que tais termos podem referir-se à sequência de DNA que codifica o transcrito, ou para o próprio transcrito de RNA.

5 Onde a sequência de controle de emenda é 3' para o códon de partida, prefere-se que seja também 5' para o primeiro códon de parada na matriz (que é 3' para e em matriz com o códon de partida), de modo que a emenda alternativa produz transcritos que codificam proteína ou sequências de polipeptídeos diferentes. Assim em uma forma de realização preferida, a
10 construção ou sequência polinucleotídica compreende os seguintes elementos em ordem de 5' a 3', com relação ao filamento sentido ou transcrito primário: partida de transcrição, partida de tradução, íntron capaz de emenda alternativa, sequência de codificação de toda ou parte de uma proteína, códon de parada.

15 A sequência de controle de emenda pode ser definida como preferivelmente até e incluindo 5' G (GT/C) seu equivalente 3' G, especialmente em tra, mas, como mencionado acima, pode incluir algumas sequências exônicas e, por esta razão, poderia incluir o nucleotídeo o mais 3' (último) do éxon (isto é G).

20 Prefere-se particularmente que a sequência de controle de emenda seja imediatamente adjacente, na direção de 3', do códon de partida, de modo que G de ATG está 5' para partida (extremidade 5') da sequência de controle de emenda. Isto é particularmente vantajoso como permite a G do códon de partida de ATG ser a sequência de flanqueamento de 5' G para a
25 sequência de controle de emenda.

Alternativamente, a sequência de controle de emenda é 3' para o códon de partida mas dentro de 1000 bp exônico, preferivelmente 500 bp exônico, preferivelmente 300 bp exônico, preferivelmente 200 bp exônico, preferivelmente 150 bp exônico, preferivelmente 100 bp exônico, mais

preferivelmente 75 bp exônico, mais preferivelmente 50 bp exônico, mais preferivelmente 30 bp exônico, mais preferivelmente 20 bp exônico, e, o mais preferível, 10 ou mesmo 5, 4, 3, 2, ou 1 bp exônico.

5 A presente invenção é um aperfeiçoamento no sistema definido como LA1188 em WO2005/012534. Esse plasmídeo tem um número de defeitos, sendo o principal que os nucleotídeos exônicos foram excisados com o íntron Ccra usado aí, assim resultando em uma armação induzida no transcrito.

10 Especificamente, em adição a seqüência derivada de Ccra (o íntron Ccra), 4 nucleotídeos de seqüência de tTAV foram removidos no transcrito específico de fêmea. Por esta razão, embora vários transcritos emendados alternativamente foram produzidos, incluindo um transcrito específico de fêmea, nenhum foi capaz de codificar a proteína tTAV funcional. Por esta razão, esta construção não foi capaz de prover expressão
15 específica de sexo de proteína tTAV funcional.

Desde que emenda não foi direcionada à seqüência de doador de emenda (5'-GT...) normalmente usada no íntron Ccra, claramente esta construção não conteve todas das seqüências regulatórias necessárias para direcionar emenda na forma do íntron Ccra em "seu contexto nativo" No
20 entanto, isto realça outro tecido. Provavelmente a única coisa faltando foi o TG...GT de flanqueamento, do qual é possível que apenas importava 5' G.

Uma benefício chave da presente invenção é, em particular em relação a tra, que as exigências das seqüências exônicas são tão mínimas (por exemplo 2 nucleotídeos em cada extremidade) que elas podem facilmente ser
25 projetadas em maiores seqüências de codificação, usando a redundância no código genético. Assim, os nucleotídeos exônicos "extras" podem ser tanto parte de seqüência de proteína heteróloga como a seqüência de flanqueamento do íntron em seu contexto nativo ao mesmo tempo.

Conseqüentemente, o íntron Ccra em LA1118 foi +132bp 3'

para o G do Códon de partida ATG (para o último nucleotídeo exônico). De fato, embora o íntron Cctra em LA1118 seja o primeiro íntron lido na direção de 5' a 3' do códon de partida ATG, ele não é o "primeiro" íntron quando lido na direção de 5' a 3' do promotor. De fato, é o segundo íntron, como não há mais um íntron (derivado de gene *Adh* de *Drosophila melanogaster*) a montante do Códon de partida de ATG. Esta informação está incluída na Tabela 3.

Será entendido que onde referência é feita a códon de partida de ATGs ou G de flanqueamento, ou seqüências de 5'- TG...GT-3', que isto está relacionado com a seqüência de DNA, mas isto também abrange a seqüência de DNA anti-sentido correspondente e, igualmente, a seqüência de RNA correspondente.

Descrição das Seqüências da presente invenção

- SEQ ID NO. 1 seqüência de consenso tra
- SEQ ID NO. 2 Seqüência de flanqueamento 5' de LA3097
- SEQ ED NO. 3 Seqüência de flanqueamento 3' de LA3097
- SEQ ID NO. 4 iniciador 688 iel-transcr
- SEQ ID NO. 5 iniciador 790 Aedsx-m-r2
- SEQ ID NO. 6 iniciador 761 Aedsx-fem-r
- SEQ ID NO. 7 iniciador AedsxRI
- SEQ ID NO. 8 Seqüência de consenso de Pane et al
- SEQ ID NO. 9 seqüência de consenso de Scali et al 2005
- SEQ ID NOS. 10 - 33 e 107 - 138 seqüências de consenso de sítios de ligação Tra/Tra2 putativos deduzidos de *Drosophila* (ver Tabela 2).
- SEQ ID NO. 34: Matriz de leitura aberta de tTAV
- SEQ ID NO. 35: Seqüência de proteína de tTAV
- SEQ ID NO. 36: Matriz de leitura aberta de tTAV2
- SEQ ID NO. 37: Seqüência de proteína de tTAV2
- SEQ ID NO. 38: Matriz de leitura aberta de tTAV3

SEQ ID NO. 39: Sequência de proteína de tTAV3

SEQ ID NO. 40: Fragmento 1 de sequência específica de fêmea dsx de lagarta rosada

5 SEQ ID NO. 41: fragmento 2 de sequência específica de fêmea dsx de lagarta rosada (PBW, *Pectinophora gossypiella*)

SEQ ID NO. 42: sequência específica de macho de dsx de lagarta rosada (PBW, *Pectinophora gossypiella*)

10 SEQ ID NO. 43: Sequência de gene parcial dsx de *Aedes aegypti*. Toda sequência exônica é incluída, mas apenas sequência intrônica parcial- ver figuras 47 e 48 para anotação.

15 SEQ ID NO. 44: sequência de gene de fêmea dsx de traça das maçãs e pêras (*Cydia pomonella*): inclui uma extensão de nucleotídeos desconhecidos, preferivelmente além de 100, preferivelmente inferior a 50, mais preferivelmente inferior a 20, mais preferivelmente inferior 10, e, o mais preferível, inferior a 5.

SEQ ID NO. 45: sequência de macho dsx de traça das maçãs e pêras (*Cydia pomonella*).

SEQ ID NO. 46: sequência de construção de pLA3435/plasmídeo dsx de *Bombyx mori*.

20 SEQ ID NO. 47: sequência de construção de pLA3359 dsx de *Anopheles gambiae*.

SEQ ID NO. 48: sequência de construção de pLA3433 Agdsx (*Anopheles gambiae*) com éxon 2 incluído.

25 SEQ ID NO. 49: sequência de construção de pLA1118 - íntron Cctra

SEQ ID NO. 50: sequência de construção de pLA3077 - íntron Cctra-tTAV.

SEQ ID NO. 51 : sequência de construção de pLA3097 - íntron Cctra-tTAV.

SEQ ID NO. 52: sequência de pLA3233 - íntron Ccra - tTAV2.

SEQ ID NO 53: sequência de construção de pLA3014 - íntron Ccra - ubiquitina - ReaperKR.

5 SEQ ID NO. 54: sequência de construção de pLA3166 - íntron Ccra -ubiquitina - ReaperKR.

SEQ BD NO. 55: sequência de pLA3376 - íntron Bztra - ReaperKR -- íntron Bztra - tTAV3.

10 SEQ ED NO. 56: sequência de construção de pLA3242 - íntron Crtra - ReaperKR.

SEQ ID NO. 57: Sequência parcial de um transcrito de macho gerado em *Drosophila melanogaster* de transformantes de LA3077 que diferem para a sequência gerada em linhagens de LA3077 de mosca-do-mediterrâneo. Esta sequência corresponde ao transcrito M3 representado na
15 figura 36.

SEQ ID NO. 58: Sequência parcial de homólogo tra de *Bactrocera zonata*. Sequência de íntron prevista para ser emendada em um transcrito específico de fêmea tra de *B. zonata* (+3 a +970bp em sequência). Nucleotídeo exônicos de flaqueamento estão nas posições 1-2 e 971-972,
20 isto é nas extremidades de 5' e 3' da sequência intrônica. De fato, deve-se observar que a sequência intrônica é flanqueada em sua extremidade 5' por um nucleotídeo de guanina, que se pensa ser crítico de uma saída limpa do íntron.

SEQ ID NO 59: Sequência parcial de homólogo tra de
25 *Ceratitis rosa*. Sequência de íntron prevista para ser emendada em um transcrito específico de fêmea tra de *C. rosa* (+3 a 1311 bp em sequência).

Nucleotídeo exônicos de flaqueamento são apresentados nas posições 1-2 e 1312-3. Novamente, observa-se que sequência intrônica é flanqueada em sua extremidade de 5' por um nucleotídeo de guanina, que se

pensa ser crítico de uma saída limpa do íntron.

SEQ ID NOS. 60-70: Iniciadores como referido nas figuras 44-46 e 50-51.

5 SEQ ID NO. 71 : fragmento 3 específico de fêmea dsx de lagarta rosada (PBW, *Pectinophora gossypiella*).

SEQ ID NO. 72: Matriz de leitura aberta de ubiquitina de *Drosophila melanogaster*.

SEQ ID NO. 73: Seqüência de proteína de ubiquitina de *Drosophila melanogaster*.

10 SEQ ID NOS. 74-105 são iniciadores como discutido acima nos exemplos.

SEQ ID NO. 106 é seqüência nucleotídica de LA1172, incluindo espinha dorsal de plasmídeo.

SEQ ID NOS 107-138 estão descritas acima.

15 SEQ ID NO. 139 iniciador de HSP

SEQ ID NO. 140 iniciador de VP 16

SEQ ID NO.141 iniciador de Agéxon1F

SEQ ID NO. 142 iniciador TETRR1

SEQ ID NO. 143 seqüência de plasmídeo LA3576

20 SEQ ID NO. 144 seqüência de plasmídeo LA3582

SEQ ID NO. 145 seqüência de plasmídeo LA3596

SEQ ID NO. 146 PBW - dsx (Fig 6)

SEQ ID NO. 147 bombyx - dsx (Fig 6)

SEQ ID NO. 148 traça das maçãs e pêras - dsx (Fig 6)

25 SEQ ID NO. 149 DSX Minigene 1 de construção LA3491

SEQ ID NO. 150 DSX Minigene 2 de construção LA3534

SEQ ID NO. 151 seqüência completa de plasmídeo LA3619

SEQ ID NO. 152 seqüência completa de plasmídeo LA3612

SEQ ID NO. 153 seqüência de plasmídeo LA3491

SEQ ID NO. 154 seqüência de plasmídeo LA3515

SEQ ID NO. 155 seqüência de plasmídeo LA3545

SEQ ID NO. 156 seqüência de plasmídeo LA3604

SEQ ID NO. 157 seqüência de plasmídeo LA3646

5 SEQ ID NO. 158 seqüência de plasmídeo LA3054

SEQ ID NO. 159 seqüência de plasmídeo LA3056

SEQ ID NO. 160 seqüência de plasmídeo LA3488

SEQ ID NO. 162 seqüência de plasmídeo LA3570

10 A invenção será agora descrita por referência aos seguintes,
exemplos não limitantes.

EXEMPLOS

Transformer

Exemplo 1-íntron tra de *Ceratitis capitata*

Os requerentes prepararam uma inserção de um cassete de
15 íntron Cctra em uma matriz de leitura aberta sintética (ORP). Duas versões
desta emenda corretamente em mosca-do-mediterrâneo, em outras palavras a
emenda do cassete de íntron Cctra fielmente recapitula o que seria
normalmente feito no contexto do gene de Cctra endógeno. Isto foi para
produzir 3 variantes de emendas 3 (principais ou sozinhas) em fêmeas, uma
20 das quais é específica de fêmea (chamada F1), enquanto as outras duas de
ambos, machos e fêmeas (chamadas M1 e M2). Uma vez que cada um dos
transcritos não específicos de sexo contêm material exônico adicional com
códon de parada, os requerentes também realizaram isto de modo que apenas
a variante de emenda de fêmea produz proteína funcional.

25 Cada uma destas construções (LA3077 e LA3097) tem o
íntron Cctra flanqueado por TG e GT (para dar 5'...TG\íntron\GT...3'. Uma
construção antiga, que não funciona perfeitamente, é LA1188. LA1118 é
completamente bem caracterizada - emenda é exatamente como acima exceto
que 4 nucleotídeos adicionais são removidos. O íntron está no contexto

5'...TGGCAC|íntron|GT...3'; emenda remove 4 bases adicionais, isto é 5'...TG|GCAC|íntron|G7...3' (figura 33).

Em todos os casos o íntron é invariável, e é simplesmente a sequência de íntron Cctra completa. Como é normal para íntrons, ele começa com GT e termina com AG. Quase todos os íntrons começam com GT, assim, o uso de GC alternativo raro em LA1118 é surpreendente [íntrons GC-AG são uma alternativa conhecida - em um estudo em grande escala, 0,5% de todos íntrons foram notificados para uso de GC-AG (Burset et al., 2001), embora isto possa ser subestimado, particularmente de íntrons alternativamente emendados, dos quais talvez 5% poderiam utilizar GC-AG (Thanaraj e Clark, 2001)] .

Análise de RT-PCR foi efetuada em LA3077, (uma construção de retro-alimentação positiva com o íntron Cctra na matriz de leitura aberta de tTAV). Moscas adultas transformadas de ambos os sexos foram criadas em dieta substancialmente livre de tetraciclina ("sem tetraciclina") por 7 dias. Moscas foram então coletadas para extração de RNA e RT PCR usando iniciadores (HSP- SEQ ID NO. 104 e VP16 SEQ ID NO. 105) foram usados para analisar o padrão de emenda do íntron CcTRA (Figura 34). Em duas amostras de fêmea os requerentes encontraram o padrão de emenda correto do Cctra (776 bp, correspondendo a remoção precisa do íntron Cctra) e não notaram tal banda em machos.

Os requerentes encontraram que LA3077 e LA3097 correspondentemente deram uma letalidade específica de fêmea repressível. LA3077 foi testado fenotipicamente através de cruzamento de moscas heterozigóticas de LA3077 para tipo selvagem, com e sem tetraciclina. Letalidade de fêmea variou de 50 a 70%. LA3097 (uma versão modificada de LA3077 por meio da qual o íntron Cctra imediatamente segue o códon de partida em tTAV ORF), demonstrou um nível muito maior de letalidade específica de fêmea, atingindo um máximo de 100% (Figura 35). O íntron

Cctra foi também inserido em tTAV2 na mesma posição como LA3097, na construção LA3233, e isto deu um resultado fenotípico similar como LA3097 (Figura 35).

Os requerentes também prepararam transformantes de LA3077 em *Drosophila*. Fenotipicamente, a construção funciona perfeitamente, isto é, que ela é letal específica de fêmea altamente efetiva. No entanto, sequenciamento das variantes de emenda de uma destas inserções demonstrou que a emenda desta construção em *Drosophila* não é igual como é na mosca-do-mediterrâneo (SEQ ID NO. 57). O transcrito crítico, aquele específico de fêmea, é o mesmo em ambos, mas em pelo menos um dos transcritos não específicos de sexo é diferente. Ele incorpora ainda seqüência exônica extra, com códons de parada, mas a junção de emendas não são exatamente iguais (Figura 36). Esta observação é extremamente importante na medida em que demonstra que esse método (regulação de expressão de gene por uso de íntrons alternativamente emendados) pode ser bastante utilizado em toda uma faixa filogenética.

Um teste simples para determinar se um regulador de emenda exônica ainda não caracterizado (tais como melhoradores e supressores) pode estar modificando a função do íntron emendado alternativamente, poderiam incluir a realização da construção e introdução no tecido alvo, e então examinar seu padrão de emenda. Em muitos casos isto não vai requerer a transformação de linhagem germinal, assim, o teste pode ser bastante rápido, por exemplo por expressão transiente em células de cultura de tecido apropriadas ou in vivo. Por exemplo, teste in vivo em insetos poderia ser realizado liberando o DNA por microinjeção. No entanto, como o versado apreciará, microinjeção junto com eletroporação, ou eletroporação, transformação química, métodos balísticos, por exemplo, foram todos utilizados em vários contextos e tais métodos de introdução de plasmídeo e expressão de proteína a partir dos mesmos são bem conhecidos na técnica.

Os requerentes também realizaram recentemente e obtiveram transgênicos com, o íntron Ccra em um gene diferente (LA3014) (todos os exemplos acima são tTAV). LA3014 contém uma fusão ubiquitina - Reaper^{KR} a jusante de um íntron Ccra. Dados fenotípicos (Figura 35) mostram que LA3014 mosca-do-mediterrâneo transgênica resultou em letalidade específica de fêmea repressível. Análise de RT-PCR em RNA extraído de machos e fêmeas adultos criados sem tetraciclina, usando iniciadores (HSP, SEQ ID NO 74) e ReaperKR (SEQ ID NO. 75), demonstra que emenda correta estava ocorrendo em fêmeas (banda de 508bp) e nenhuma destas bandas foi encontrada em machos (Figura 37). LA3166 é outra construção com o íntron Ccra localizado dentro da região de codificação de ubiquitina fusionada ao Reaper^{KR}, mas localizada em uma posição diferente em ubiquitina. LA3166 também produz um efeito letal específico de fêmea repressível dominante na mosca-do-mediterrâneo (Figura 35).

Os requerentes também realizaram recentemente e obtiveram transgênicos com construções baseadas em Ccra com "apenas íntron" com íntron em um gene diferente (todos os exemplos acima estão em tTAV ou uma de suas variantes, isto é tTAV2 ou tTAV3). Estas construções funcionam como previsto. Este é um resultado importante, deste modo mostrando que não ocorrem seqüências exônicas essenciais em Ccra que os requerentes simplesmente duplicaram (em função, se não necessariamente em seqüência) por acaso, em tTAV. Também os requerentes realizaram construções ubi-rpr^{KR} deste tipo (LA3014 e LA3166), o que também valida o método de fusão de ubiquitina descrito acima.

A fim de demonstrar a faixa filogenética do íntron Ccra, os requerentes geraram LA3097 e LA3233 de *Anastrepha ludens* transgênico. LA3097 e LA3233 foram selecionados para injeção em *Anastrepha ludens* como demonstraram a melhor letalidade específica de fêmea em *Ceratitis capitata* (ver exemplo 13). Dados fenotípicos foram gerados de 4 linhagens de

LA3097 independentes e 1 linhagem de LA3233 (ver figura 38). Letalidade específica de fêmea foi geralmente pouco menor em *Anastrepha ludens* quando comparada a *C. capitata* mas atingiu 100% em uma linhagem.

Anastrepha ludens transformada com LA3097 e criada com tetraciclina até eclosão foram isolados e mantidos sem tetraciclina durante 7 dias. RNA foi então extraído e análise de RT-PCR foi realizada usando iniciadores HSP (SEQ ID NO. 76) e TETRR1 (SEQ ID NO. 77). O padrão correto de emenda específico de fêmea (como F1) foi observado a partir de RNA isolado em fêmeas (348bp) mas não de machos demonstrando a função do íntron Ccra em uma espécie diferente (figura 39)

A banda do macho mais brilhante e banda específica de fêmea foram purificadas e precipitadas para sequenciamento. O transcrito específico de fêmea foi verificado como sendo corretamente emendado nas fêmeas de mosca mexicana das frutas como esperado para LA3097:

LA3097: AGCCACCATG|GT...íntron...AG | GTCAGCCGCC

As duas seqüências de flanqueamento acima são SEQ ID NOS. 2 e 3, respectivamente.

Exemplo 2: Íntron tra de *Bactocera zonata*

Os requerentes isolaram o íntron tra de *Bactocera zonata* (*B. zonata*) (SEQ ID NO. 58) usando iniciadores ROSA1 (SEQ ID NO. 78), ROSA2 (SEQ ID NO. 79), e ROSA3 (SEQ ID NO. 80).

Estas seqüências de iniciadores foram designadas com base em seqüência de codificação conservada de *Ceratitis capitata* e homólogos tra de *Bactrocera oleae*. Usando ROSA2 e ROSA3 ou ROSA1 e ROSA3 como iniciadores, o íntron tra e sua região de codificação de flanqueamento foram amplificados de DNA genômico de *Bactrocera zonata*. Então os requerentes usaram estes produtos de PCR como um modelo e amplificaram o fragmento de íntron tra para fazer a construção LA3376 (figura 31 e SEQ ID NO. 55). Os iniciadores (BZNHE- SEQ ID NO. 81 e BZR-SEQ ID NO. 82) foram

usados para fazer as construções; estes iniciadores contêm seqüências adicionais para propósitos de clonagem. O íntron Bztra em LA3376 é clonado em ORF de tTAV3 e também de Reaper^{KR}. Transformantes de mosca-do-mediterrâneo foram gerados e RNA extraído das moscas macho e fêmea.

5 RT-PCR foi então realizado em ambos as emendas Reaper^{KR} (HB- SEQ ID NO. 83) e ReaperKR SEQ ID NO. 84) e tTAV3 (SRY- SEQ ID NO. 85) e AV3F- SEQ ID NO. 86). Os fragmentos esperados de 200bp para Reaper^{KR} e 670 bp para tTAV3, correspondendo a emenda em um padrão equivalente ao transcrito de F1 de Cctra (Pane et al, 2002), foram gerados em
10 fêmeas (figura 40).

Exemplo 3: Isolamento e emenda do íntron tra de *Ceratitis rosa* (*C. rosa*, mosca-de-natal).

Iniciadores ROSA2 (SEQ ID NO. 87) e ROSA3 (SEQ ID NO. 88) foram projetados baseados na seqüência de codificação conservada de
15 *Ceratitis capitata* e *Bactrocera oleae*. Usando ROSA2 e ROSA3 como iniciadores, o íntron tra e sua região de codificação de flanqueamento foram amplificados de DNA genômico de *Ceratitis rosa* (SEQ ID NO. 59). Então os requerentes usaram os produtos de PCR como um modelo e amplificaram o fragmento de íntron tra para fazer as construções. Os iniciadores (CRNHE-
20 SEQ ID NO 89 e CRR SEQ ID NO 90) foram usados durante a construção de LA3242 (SEQ ID NO. 56 e figura 32. LA3242 contém íntron de *C. rosa* na extremidade 5' de ORF de ReaperKR. Embriões de *Ceratitis capitata* foram injetados com DNA de LA3242, embriões injetados foram criados até idade adulta em uma dieta substancialmente livre de tetraciclina. RNA foi extraído
25 de machos e fêmeas adultos; isto foi usado como um modelo de RT PCR usando iniciadores HB (SEQ ID NO. 91) e ReaperKR (SEQ ID NO. 92). A banda de emenda específica de fêmea esperada (200bp), correspondendo a emenda no padrão equivalente ao do transcrito F1 de Cctra, foi observada em fêmeas e não em machos (figura 41).

Doublesex

Exemplo 4 - dsx *Bombyx mori* em PBW

A seqüência de um homólogo de *Bombyx mori* (bicho-da-seda) de Dsx de *Drosophila* (Bmdsx) foi previamente descrita e um produto de emenda específico de macho e de fêmea foi identificado (Suzuki et al, 2001).
5 Ambos machos e fêmeas usaram o mesmo 3' polyA, e há dois éxons específicos de fêmea. Um trabalho sugeriu que a emenda específica de sexo não é dependente em tra/tra2, em outras palavras mesmo se o padrão parece o mesmo, o mecanismo subjacente pode ser diferente (Suzuki et al., 2001),
10 embora seus dados, principalmente a falta de sítios de ligação tra-tra2 reconhecíveis, no entanto, não sejam atraentes. Além disso, uma construção de minigene dsx de *B. mori* (contendo seqüência exônica e seqüência intrônica truncada) foi transformada no *B. mori* e os transformantes de linhagem germinal mostram emenda específica de sexo (Funaguma et al.,
15 2005).

Os requerentes geraram um minigene de Bmdsx com base na seqüência usada no trabalho de Funaguma et al, com algumas mudanças significantes, e o injetaram na lagarta rosada para averiguar se se poderia obter emenda específica de sexo em uma espécie divergente. A construção de
20 minigene, gerada pelos requerentes, não incluía éxon 1, que está presente em ambos, machos e fêmeas. Além disso, os requerentes removeram o íntron entre éxon 3 e 4 (os dois éxons específicos de fêmea), incluíram uma seqüência heteróloga (contendo sítios de clonagem múltiplos, MCS), usaram a seqüência de melhorador/promotor de Hr5-IE1 do baculovírus AcNPV e
25 usaram uma seqüência de terminação transcripcional de 3' derivada de SV40 (ver figura 42 para um esquema). Os fragmentos de éxon / íntron de flaqueamento individuais usados foram amplificados e recombinados juntos por PCR e ligados em uma construção transportando um fragmento de promotor melhorador Hr5/ IE1 e SV40 3' UTR (figura 22 e SEQ ID NO. 22).

A3435 foi injetado nos embriões da lagarta rosada (*Pectinophora gossypiella*). As larvas do primeiro instar foram coletadas após 5-7 dias e analisadas individualmente por RT-PCR (usando iniciadores IEL transcr- SEQ ID NO. 93 e SV40-RT-P2- SEQ ID NO. 94) para determinar se BMdsx pode sofrer emenda específica de macho e fêmea (figura 43). A análise dos requerentes detectou a banda específica de macho (prevista para ser 442bp) em 4 amostras (Trilhas 1, 2, 3 e 4) e a banda específica de fêmea (prevista para ser 612bp) em 1 amostra (Trilha 5).

A emenda correta de dsx *B. mori* em PBW demonstra que os requerentes podem realizar (de fato realizaram) a expressão específica de sexo de uma seqüência heteróloga (aqui, MCS) em um lepidóptero utilizando um sistema de emenda alternativa. Além disso, uma vez que este sistema de emenda foi derivado de uma espécie heteróloga, isto sugere que tais construções poderiam funcionar sobre uma extensa faixa filogenética. No entanto, a identificação de sistemas de emenda alternativa nas espécies de interesse é também prevista, e métodos para identificar tais sistemas de emenda alternativa são providos aqui ou serão conhecidos do versado na técnica. Ao prover um MCS no exemplo dos requerentes (ver figura 42), a expressão de uma seqüência de interesse, por exemplo uma região de codificação de uma proteína de interesse poderia prontamente ser conseguida inserindo-se a referida seqüência. Se a referida seqüência codificou uma proteína apropriada, um fenótipo específico de sexo, por exemplo letalidade específica de sexo condicional, poderia assim ser introduzido, por exemplo em uma lagarta rosada.

Exemplo 5: Isolamento de dsx de traça das maçãs e pêras

O gene dsx da traça das maçãs e pêras (*Cydia pomonella*) foi isolado efetuando 3' RACE usando iniciadores que foram baseados em alinhamentos de seqüência de *B. oleae*, *B. tyroni*, *C. capitata*, *D. melanogaster*, *B. mori*, e *A. gambiae*. RNA foi isolado de uma traça das maçãs

e pêras, macho e fêmea, e 3' RACE, para gerar cDNA, foi realizado usando o iniciador de TT7T25 (SEQ ID NO. 95).

5 PCR foi realizado usando os iniciadores ds1c (SEQ ID NO. 96) e TT7 (SEQ ID NO. 97). Dois ciclos de PCR aninhado foram então realizados no produto do primeiro PCR usando os iniciadores codling2a (SEQ ID NO. 98) e TT7 (SEQ ID NO. 99) e o produto do segundo ciclo de PCR usando Codling2b (SEQ ID NO. 100) e TT7. As seqüências específicas de macho e fêmea isoladas compartilham a similaridade de seqüência a homólogos dsx previamente isolados (Macho-SEQ ID NO. 43 e Fêmea- SEQ ID NO. 42).

Exemplo 6: Isolamento de PBW dsx

O gene dsx da lagarta rosada foi isolado realizando 3' RACE usando iniciadores que foram baseados em alinhamentos de seqüência de *B. oleae*, *B. tyroni*, *C. capitata*, *D. melanogaster*, *B. mori*, e *A. gambiae*. RNA foi isolado de uma traça das maçãs e pêras, macho e fêmea, e 3' RACE, para gerar cDNA, foi realizado usando TT7T25 (seqüência definida aqui). PCR foi realizado usando os iniciadores Pbwdsx2 (SEQ ID NO. 101) e TT7 (SEQ ID NO. 102). PCR aninhado foi então realizado no produto do primeiro PCR usando os iniciadores Pbwdsx3 (SEQ ID NO. 103) e TT7. Três seqüências específicas de fêmea foram isoladas: PBWdsx-F1 (SEQ ID NO. 40), PBWdsx-F2 (figura 10), e PBWdsx-F3 (SEQ ID NO. 71) e uma seqüência específica de macho (SEQ ID NO. 42). As seqüências específicas de macho e fêmea isoladas compartilham similaridade de seqüência a homólogos dsx isolados previamente.

25 Exemplo 7: dsx em *Anopheles gambiae*

A seqüência do gene dsx de *Anopheles gambiae* foi previamente descrita (Scali et al 2005). No entanto, quando os requerentes tentaram repetir o trabalho descrito, eles verificaram que existiam algumas diferenças na emenda que ocorre. Quando os requerentes tentaram repetir a

amplificação do transcrito específico de fêmea usando iniciadores projetados da sequência de mRNA (Acesso; AY903308 de sequência de codificação de fêmea e AY903307 de sequência de codificação de macho), a amplificação falhou. No entanto, quando Scali e colaboradores mostraram que havia um
 5 éxon compartilhado, o que não tinha sido previamente descrito, os requerentes projetaram iniciadores para amplificar o transcrito e gene *dsx* completo. Usando estes iniciadores e iniciadores projetados a partir da sequência de DNA genômico (Acesso; GI: 19611767) os requerentes verificaram que a emenda do transcrito de fêmea é diferente daquele descrito por Scali et al
 10 2005 (figura 44). O transcrito mostrou que o éxon de fêmea estava em uma posição diferente. Há várias explicações para estas diferenças, mas as mais prováveis são ou algum tipo de diferença de cepa nos *Anopheles*, que os requerentes usaram para obter dados, ou a sequência publicada não é de *Anopheles gambiae*, ou se tem mais de uma isoforma de fêmea, como
 15 mostrado para *Stegomyia aegypti* no Exemplo 20.

Os requerentes também usaram com sucesso iniciadores, projetados em torno de sua versão da emenda de *dsx* de *Anopheles gambiae*, que são capazes de distinguir entre machos e fêmeas de *Anopheles gambiae* (figura 45). Isto fornece uma boa evidência de que o sistema será funcional
 20 como um mecanismo de emenda específica de sexo quando fusionado a uma proteína de interesse, como tTAV ou uma matadora.

O gene *dsx* de *Anopheles gambiae* que os requerentes isolaram do DNA genômico, que tem várias mudanças na sequência nucleotídica comparada com a sequência genômica relatada, foi clonado em LA3359 (SEQ
 25 ID NO. 47) e LA3433 (SEQ ID NO. 48), cujos esquemas podem ser encontrados nas figuras 23 e figura 24, respectivamente.

Exemplo 8: *dsx* em *Stegomyia aegypti*

A emenda do gene parece ser similar a *dsx* *Anopheles gambiae* (Scali et al 2005). O gene *dsx* de *Stegomyia aegypti* está ilustrado

diagramaticamente na figura 47 ou 48. O transcrito específico de macho (M1) é produzido que não inclui éxons 5a ou 5b. Duas variantes de emenda específica de fêmea (F1 e F2) têm a seguinte estrutura; F1 compreende éxons 1-4, 5a, 6 e 7 mas não 5b, F2 compreende éxons 1-4 e 5b (figura 46). Além disso, mais um transcrito (C1) está presente em ambos, machos e fêmeas; isto compreende éxons 1-4 e 7, mas não éxons 5a, 5b ou 6.

A emenda do gene parece ser similar a *dsx Anopheles gambiae* (Scali et al 2005). O gene *dsx* de *Stegomyia aegypti* é ilustrado diagramaticamente na figura 47 ou 48.

Actin 4

Exemplo 9: Gene Actin-4 de *Stegomyia aegypti*

Uma forma de obter expressão específica de sexo, de tecido e de estágio de um gene de interesse é ligá-lo ao gene Actin-4 de *Stegomyia aegypti* (AeAct-4). Esse gene é apenas expressado no desenvolvimento de músculos aptos a voar de *Stegomyia aegypti* fêmea (Munoz et al 2004). Eles foram usados em hibridização in-situ para um RNA detectar o perfil de expressão de AeAct-4. Tomou-se um fragmento do gene Actin-4 de *Stegomyia aegypti*, compreendendo uma região de promotor putativa, um íntron emendado alternativamente, e uma seção de região não traduzida 5' (UTR) e colocou-se na frente da seqüência codificando para tTAV (figura 49) para testar a função da emenda específica de sexo quando fusionada a tTAV.

Os requerentes integraram LA1172 no genoma de *Stegomyia aegypti* usando piggyBac. Duas linhagens independentes foram geradas (linhagens 2 e 8). Ambas as linhagens mostraram a emenda correta do gene Actin-4-tTAV (figuras 50 e 51). O promotor de Actin-4 e íntron emendado alternativamente podem, por esta razão, ser usados com sucesso para prover a emenda específica de sexo, de tecido e de estágio de um gene de interesse em *Stegomyia aegypti*.

Descrição das figuras e das listagens seqüência de exemplos 1-9

Figura 19: Um uso do elemento P na geração de expressão específica de linhagem germinal de um gene de interesse (Gene E).

Inserção do elemento P IVS3 e seqüências de flaqueamento exônicas a montante de uma fusão de ubiquitina - gene E que permite a expressão específica de linhagem germinal de gene E sob um promotor ativo de linhagem germinal. A - Promotor ativo de linhagem germinal; B - Matriz de leitura aberta de elemento P; C - P íntron 'IVS3'; D - Ubiquitina; E - Região de codificação de proteína de interesse por exemplo tTAV.

Figura 20: Expressão específica de sexo usando dsx.

A: Íntron usado como íntron Ccra acima, mas dando expressão específica de macho. Um fragmento de dsx (aqui a versão de *Anopheles*) é inserido em uma região de codificação heteróloga (caixas sombreadas). O íntron está completamente removido em machos, mas em fêmeas a região de codificação é prematuramente terminada.

B: Uma abordagem alternativa a expressão específica de macho, no qual uma região de codificação heteróloga é fusionada a um fragmento de dsx.

C: Expressão específica de fêmea: a região de codificação heteróloga está inserida no éxon específico de fêmea, ou como uma fusão na matriz para um fragmento de Dsx, ou com seus próprios códons de partida e parada.

D: Expressão diferencial: projetos B e C podem ser combinados para resultar na expressão de gene a em fêmeas e b em machos.

Figura 21: Emenda alternativa específica de sexo de Ccra

A: Ccra é emendado em fêmeas para produzir três transcritos: F1, que codifica proteína Tra, e M1 e M2, que não codificam, porque eles incluem éxons adicionais com códons de parada (re-escrito de Pane et al. 2002). Machos produzem apenas transcritos M1 e M2 e assim não produzem proteína Tra funcional de todo.

B: Se esse íntron for para funcionar similarmente em uma região de codificação heteróloga, isto similarmente permitiria às fêmeas, mas não aos machos, produzir proteína funcional X.

Figura 22: Representação diagramática de construção pLA3435/ plasmídeo (SEQ ID NO.46).

Figura 23: Mapa de plasmídeo de pLA3359 do gene *dsx* de *Anopheles gambiae* colocado sob o controle de um promotor Hr5-IE1 para avaliar a emenda via expressão transiente .

Figura 24: pLA3433 - gene *dsx* de *Anopheles gambiae* colocado sob o controle de promotor Hr5-IE1, com adição de éxon 2, para avaliar a emenda através de expressão transiente .

Figura 25: Representação esquemática de construção pLA1188.

Figura 26: Representação esquemática de construção pLA3077.

Figura 27: Representação esquemática de construção pLA3097.

Figura 28: Representação esquemática de construção pLA3233.

Figura 29: Representação esquemática de construção pLA3014.

Figura 30: Representação esquemática de construção pLA3166.

Figura 31: Representação esquemática de construção pLA3376.

Figura 32: Representação esquemática de construção pLA3242.

Figura 33: Sequência de flanqueamento de Cctra

Emenda do íntron Cctra em LA3077 e LA3097 é exatamente

como seria vista no íntron CcTra nativo. Emenda em LA1118 resulta na remoção de 4 nucleotídeos adicionais. Em todos os casos os íntrons são flanqueados por éxon TG exônico 5' e 3' GT.

Figura 34: Gel mostrando emenda específica de sexo correta de íntron(s) derivada de CcTra (banda de 776 bp em fêmeas) em *Ceratitis capitata* transformada com LA3077. Trilha 1: Marcador (SmartLadder™ de Eurogentec, bandas de aproximadamente 0,8, 1,0 e 1,5kb são indicadas); Trilhas 2 e 3: *Ceratitis capitata* LA3077 /+ machos; Trilhas 4 e 5: *Ceratitis capitata* LA3077 /+ fêmeas.

Figura 35: Dados fenotípicos de construções específicas de fêmea transformadas em *Ceratitis capitata*.

Coluna 1: Designação de construção LA#, por exemplo LA3077, LA3097, LA3233, etc, é indicada pelo número, com linhas de inserção independentes referidas por letra; Colunas 2 e 3: Resultados sem tetraciclina (NT) resulta de cada linhagem transformada obtida nos machos totais (2) e fêmeas totais (3). Colunas 4 e 5: Tetraciclina (TET) resulta de cada linhagem transformada dada em machos totais (4) e fêmeas totais (5).

Figura 36: transcritos de construções de íntron CcTra em *Drosophila* e *Ceratitis capitata*.

A linha de topo representa o DNA de construção contendo íntron tra flanqueado por gene desejado (a caixa aberta). A caixa vermelha representa os éxons específicos de macho. Íntrons são representados por linhas sólidas. A seta acima da primeira linha representa as posições dos oligonucleotídeos usados nos experimentos de RT-PCR. A barra indica a escala da figura.

Figura 37: Gel mostrando emenda específica de fêmea correta de seqüência derivada de CcTRA (banda de 508bp) em *Ceratitis capitata* de fêmea transformada com LA3014. Trilha 1: Marcador (SmartLadder™ de Eurogentec, bandas de aproximadamente 0,4 e 1,0kb são

indicadas); Trilha 2: *Ceratitis capitata* LA3014 /+ macho; Trilha 4: *Ceratitis capitata* LA3014 /+ fêmea; Trilhas 3 e 5: sem controles negativos de transcriptase reversa (banda de fundo, provavelmente de DNA genômico, pode ser vista nas trilhas 2 e 4).

5 **Figura 38: Dados fenotípicos de *Anastrepha ludens* transgênico transformado com LA3097 ou LA3233.** Coluna 1: Construção La# (LA3097 ou LA3233) indicou, com linhas de inserção independentes referidas por letra; Colunas 2 e 3: não tetraciclina (NT) resulta de cada linhagem transformada obtida em machos totais (2) e fêmeas totais (3).
10 Colunas 4 e 5: Tetraciclina (TET) resulta para cada linhagem transformada dada em machos totais (4) e fêmeas totais (5).

15 **Figura 39: Gel mostrando emenda específica de sexo correta de emenda CcTRA (banda de 348bp em fêmeas) em *Anastrepha ludens* transformada com LA3097.** Trilha 1 : Marcador (SmartLadder™ de Eurogentec, bandas de aproximadamente 0,4 e 1,0kb são indicadas); Trilhas 2, 3 e 4: *A. ludens* LA3097 /+ machos; Trilhas 5, 6 e 7: *A. ludens* LA3097 /+ fêmeas.

20 **Figura 40: Gel mostrando emenda específica de sexo correta de BzTRA em ReaperKR (banda de 200 bp em fêmeas) e tTAV3 (banda de 670 bp em fêmeas) regiões de LA3376, em *Ceratitis capitata* transformada com LA3376.** Trilha 1: Marcador (SmartLadder™ de Eurogentec, bandas de aproximadamente 0,2, 0,6 e 1,0kb são indicadas); Trilhas 2 e 3: *C. capitata* LA3376 /+ machos testadas para emenda em ReaperKR; Trilhas 4 e 5: *C. capitata* LA3376 /+ fêmeas testadas para emenda
25 em ReaperKR; Trilha 6: SmartLadder™; Trilhas 7 e 8: *C. capitata* LA3376 /+ machos testada para emendas em tTAV; Trilhas 9 e 10: *C. capitata* LA3376 /+ fêmeas testadas para emenda e, tTAV; Trilha 11: SmartLadder™.

Figura 41: Gel mostrando emenda de CrTRA específica de sexo correta em CrTRA - ReaperKR (banda de 200 bp em fêmeas) em

***Ceratitis capitata* injetada com LA3242.** Trilha 1: Marcador (SmartLadder™ de Eurogentec, bandas de aproximadamente 0,2, 0,6 e 1,0kb são indicadas); Trilhas 2-7: *C. capitata* machos de tipo selvagem injetados com LA3242; Trilha 8: SmartLadder™; Trilhas 9-14: *C. capitata* fêmeas de tipo selvagem injetadas com LA3242; Trilha 15: SmartLadder™.

Figura 42: Representação esquemática de construção de minigene Bmdsx.

Duas construções de minigenes derivadas do gene dsx de *Bombyx mori* são ilustradas diagramaticamente, junto com a emenda alternativa prevista destas construções (padrão de fêmea mostrado acima da construção, padrão de macho abaixo). (A) é a construção de minigene dsx de *Bombyx mori* usada em Funaguma et al., 2005) (B) é pLA3435. A e B diferem entre si de várias formas: (i) Éxon 1 é excluído de pLA3435, (ii) o íntron entre éxons específicos de fêmea 3 e 4 foi removido e uma sequência heteróloga pequena foi inserida em pLA3435 (iii) Funaguma et al., usam o promotor iel do baculovírus BmNPV e um BmA 33' UTR comparado com pLA3435 que usa melhorador/promoter hr5-IE1 do baculovírus AcNPV de e um 3' SV40 3' UTR. (iv) pLA3435 usa as sequências de íntron ligeiramente mais longas quando comparadas com (A) (ver figura 15 para sequência). Duas construções de minigenes derivadas do gene dsx de *Bombyx mori* são ilustradas diagramaticamente, junto com a emenda alternativa prevista destas construções (padrão de fêmea mostrado acima da construção, padrão de macho abaixo).

Figura 43: Emenda específica de sexo de construção de minigene BMdsx em PBW.

Análise de expressão transiente de pLA3435 usando RT-PCR mostra a presença de um fragmento de 442 bp (Trilhas 1,2,3 e 4) em machos e um fragmento de 612 bp em fêmeas (Trilha 5), mostrando que o minigene BMdsx com um fragmento heterólogo inserido entre éxon 3 e 4 é capaz de

emendar corretamente na traça divergente, PBW. Marcadores são SmartLadder™ de Eurogentec; bandas de aproximadamente 0,2, 0,4 e 0,6 kb são indicadas

Figura 44: Emenda específica de sexo de *dsx* *Anopheles gambiae*.

Anófelos (A) mostram a emenda que foi relatada por Scali et al 2005. No entanto, quando RT-PCR foi realizado usando os iniciadores dos requerentes (*spl-agdsx-e3* (SEQ ID NO. 60) e *spl-agdsx-m* (SEQ ID NO. 61)) um padrão de emenda diferente para fêmeas foi revelado, representado por Anófelos (B).

Figura 45: Identificação de *Anopheles gambiae* macho e fêmea usando iniciadores *dsx*.

RNA foi extraído da *Anopheles gambiae* macho e fêmea e os transcritos de *dsx* foram amplificados por RT-PCR usando os iniciadores *spl-agdsx-e3* (SEQ ID NO. 62) e *spl-agdsx-m* (SEQ ID NO. 63); o padrão de banda resultante é mostrado no gel acima. As bandas esperadas para os transcritos de macho e fêmea. são indicadas por setas brancas, as bandas foram clonadas e seqüenciadas e são idênticas à seqüência prevista da versão dos requerentes de transcrito de *dsx* (ver SEQ ID NO. 47 (LA3359) e SEQ ID NO. 48 (LA3433)). Os marcadores de peso molecular são mostrados em kb (SmartLadder™ de Eurogentec; tamanhos são aproximados).

Figura 46; Identificação de *Stegomyia aegypti* macho e fêmea usando iniciadores *dsx*.

Os iniciadores para o RT-PCR de *Stegomyia aegypti* de A e B foram *AedesxF1* (SEQ ID NO. 64) e *AedesxR5* (SEQ ID NO. 65) e foram testados em pupas, um estágio de vida de *Stegomyia aegypti* que pode ser sexuada convenientemente e corretamente; a amplificação de RT-PCR resultante é mostrada na imagem em gel (A). As pupas, macho e fêmea, mostram uma banda específica de sexo distintiva. Então os iniciadores foram

testados em extrações de RNA de larvas, que não podem ser prontamente sexuada por sua morfologia e a amplificação de RT-PCR resultante mostrada em imagem em gel (B). As larvas mostram um padrão de banda claro que distingue machos de fêmeas sem ambigüidade. Imagem em gel (C) mostra uma banda de aproximadamente 600 bp de RT-PCR usando os iniciadores *AedessxF1* e *AedesxR2* (SEQ ID NO. 66) de pupas de indivíduos machos e fêmeas. O sequenciamento desta banda mostrou uma variante de emenda específica de fêmea a qual parece não possuir o éxon compartilhado de macho para o qual *AedesxR5* é previsto se recombinar (éxon 7, ver figura 56). Os marcadores de peso molecular são mostrados em kb (SmartLadder™ de Eurogentec; tamanhos são aproximados).

Figura 47: Representação diagramática de parte do gene dsx de *Stegomyia aegypti* (fora de escala).

Um fragmento do gene dsx de *Stegomyia aegypti* é representado acima. Éxons 5a e 5b são específicos de fêmea e éxon 6 é um éxon específico de macho. Duas variantes de emendas específicas de fêmea foram encontradas (F1 e F2) as quais compreendem éxons 1-4,5b,6 e 7 (F1) ou 1-4,5a (F2); transcritos em machos (M1) compreendem éxons 1-4,6 e 7 mas não éxon 5a ou 5b e um transcrito (C1) de 1-4 e 7 mas não éxons 5a, 5b ou 6 é mostrado em machos e fêmeas. O número para cada um dos éxons após # refere-se a contig 1.370

(http://www.broad.mit.edu/annotation/disease_vetor/Aedes_aegypti/), que se lêem orientação oposta, e após * refere-se à seqüência nucleotídica mostrada em SEQ ID NO. 43.

Figura 48: Representação diagramática do gene dsx de *Stegomyia aegypti*.

O gene dsx completo de *Stegomyia aegypti* é representado acima. Éxon 5 é o éxon específico de fêmea e éxon 6 é um éxon putativo específico de macho. Em princípio, transcritos em fêmeas compreendem

éxons 1,2,3,4,5 e 7, e machos compreendem éxons 1,2,3,4,6 e 7. O número de cada um dos éxons após # refere-se ao contig 1.370

([http://www.broad.mit.edu/annotation/disease_vector/Aedes aegypti/](http://www.broad.mit.edu/annotation/disease_vector/Aedes_aegypti/)) que se lê na orientação oposto, e após * refere-se à figura 12.

5 **Figura 49: Mapa de Plasmídeo de pLA 1172.**

A região de codificação para tTAV foi colocada sob o controle de um fragmento do gene Actin-4 de *Stegomyia aegypti* (Munoz et al 2005) que inclui o 5' UTR, primeiro íntron, e seqüências a montante (promotor putativo). A construção também contém seqüência tetO Nipper. A construção
10 tem extremidades de tipo piggyBac e um marcador de DsRed2 para integração estável em um genoma.

Figura 50: Emenda específica de sexo de tTAV em transformantes LA1172.

Imagem em gel de RT-PCR de RNA extraído de pupas
15 machos e fêmeas de LA 1172 linhagem 2 . Os iniciadores usados foram Agexon1 (SEQ ID NO. 67) e Tra (tTAV) seq+ (SEQ ID NO. 68). Sequenciamento das bandas de RT-PCR mostrou a emenda esperada ocorrendo em machos e fêmeas. Os dados mostrados no diagrama acima são de LA1172 linhagem 2, a linhagem 8 mostrou exatamente os mesmos
20 resultados (dados não mostrados). Marcadores são SmartLadder™ de Eurogentec; tamanhos aproximados são indicados, em kb).

Figura 51: RT-PCR de amostras de tipo selvagem, mostrando variante de emendas específicas de sexo do gene Actin-4 de *Stegomyia aegypti*.

25 Imagem em gel de RT-PCR de RNA extraída de diferentes estágios de desenvolvimento, e dissecações de adultos, de LA1172 linhagem 8 . Os iniciadores usados foram Agexon1 (SEQ ID NO. 69) e Éxon 3 (SEQ ID NO. 70). A imagem em gel mostra que a expressão forte do gene Actin-4 apenas ocorre no estágio de pupa, e que a expressão de adulto é geralmente

limitada para o tórax de fêmea onde os músculos de vôo são encontrados.

Tabela 17, abaixo mostra os conteúdos de cada trilha.

Tabela 1.

E = conjunto de ~ 100 embriões	MH = cabeça de macho adulto
L4 = larva no quarto instar	MT = tórax de macho adulto
ME = pupa de macho precoce (<4 horas de vida)	MA = abdome de adulto macho
FE = pupa de fêmea precoce (<4)	FH = cabeça de fêmea adulta
MP = pupa de macho	FT = tórax de fêmea adulta
FP = pupas de fêmea	FA = abdome de fêmea adulta
	- vê = controle de água

OUTROS EXEMPLOS

5 Exemplo 10: Traças.

Os requerentes realizaram recentemente construções baseadas em seus dados de expressão transiente usando uma construção de minigene recombinante derivada de *Bombyx mori*. Isto ainda é discutido abaixo na seção intitulada " alinhamento de seqüência dsx de traça e motivos conservados"

10 Exemplo 11: Uso de Bztra

Os requerentes realizaram recentemente duas construções baseadas em Bztra, expressadas na mosca mexicana das frutas (LA3376). LA3376 concede letalidade específica de fêmea repressível. Os requerentes
15 previamente mostraram que LA3376 funciona e emenda corretamente em mosca-do-mediterrâneo. Transformantes em mosca mexicana das frutas (*Anastrepha ludens*) também foram gerados com LA3376. Estes foram analisados para emenda correta do íntron Bztra a fim de demonstrar a faixa filogenética do íntron Bztra por RT-PCR usando iniciadores SRY e AV3F
20 (figura 15 e "géis de RT-PCR de mosca-do-mediterrâneo " seção acima). Isto mostra uma emenda correta do íntron Bztra em mosca mexicana das frutas.

Exemplo 12: Dmdsx em mosca-do-mediterrâneo (DmDsx em exemplo de mosca-do-mediterrâneo transgênica: fusão nipper em #797)

Os requerentes também obtiveram recentemente dados em

uma construção de Dmdsx em mosca-do-mediterrâneo. A construção usou um fragmento do gene *doublesex* em *Drosophila melanogaster* para dar expressão específica de sexo de um fragmento de gene Nipp1Dm de *Drosophila melanogaster* (os requerentes chamam este fragmento "nipper").

- 5 Os requerentes não notaram uma emenda específica de sexo clara. No entanto, os dados fenotípicos mostram alguma especificidade de sexo; os requerentes notaram a letalidade aumentada de fêmeas, a cerca de 75% de penetração. Como evidente, esta penetrância incompleta poderia ser devido a nível de expressão, falta de toxicidade de nipper em mosca-do-mediterrâneo, etc. Os requerentes também notaram uma significativa redução no número de machos, mas a fonte de tTA, LA670, usada neste experimento poderia ela própria matar alguns dos machos.

- Os requerentes testaram três linhagens transgênicas de mosca-do-mediterrâneo independentes que transportam uma fusão de nipper para seqüência DmDsx que se destinava a ser expressada especificamente em fêmeas. Esta construção pode não ter funcionado perfeitamente possivelmente devido à seqüência essencial de emenda alternativa correta e/ou os sítios de ligação Sxl requeridos por DmDsx, e uma vez que a mosca-do-mediterrâneo não usa Sxl na via de determinação de sexo, DmDsx pode estar incapaz de emendar completamente esta fusão na forma correta em mosca-do-mediterrâneo. No entanto, os requerentes alcançaram sucesso em causar de modo reprodutível uma letalidade aumentada em fêmeas comparada aos machos através de três linhagens em uma eficiência muito similar (aproximadamente 75% mais letalidade observada em fêmeas do que em machos). Isto demonstra que o sistema dsx pode funcionar através de espécies de parentesco bastante distante (separação evolucionária é em torno de 120-150 milhões de anos), e se a seqüência Ccdsx foi usada, pode ter bem funcionado devido à exigência de Sxl de Dmdsx .

Os resultados de 797 são mostrados abaixo, usando um

sistema nipper de emenda Tet014 dsx (Pub EGFP). Eles mostram que este sistema é letal em um estágio larval (~50%), e é provável ser atuante com maior sucesso em fêmeas (~75%). 797 é marcado com verde (G), 670 com vermelho (R). 670 é uma fonte de tTAV, assim, espera-se ver um fenótipo nas moscas R+G; G (e R) são apenas controles. NF não fluorescente (isto é tipo selvagem) é também um controle onde incluído. Toda a progênie criada em meio livre de tet.

Todas as três linhagens independentes vistas como atuando de maneira similar.

797A/797A M2 x 670A/+:

	Pupas	Adultos	Machos : Fêmeas
G	184	176	85:91
R+G	74	57	44:13

797C/797C M1 x 670A/+:

	Pupas	Adultos	Machos : Fêmeas
G	169	157	89:68
R+G	94	67	54:13

797C/797C M2 x 670A/+:

	Pupas	Adultos	Machos : Fêmeas
G	406	377	179:198
R+G	171	147	121:26

670A/+ x 797C/+ M2:

	Pupas	Adultos	Machos : Fêmeas
NF	198	192	92:100
G	162	147	67:80
R	149	72	43:29
R+G	45	22	20:2

Média de todas as 3 linhagens: número de fêmeas R+G = 21% do número de machos R+G, assim, a mortalidade em excesso substancial em fêmeas R+G com relação a machos. Esse efeito não é visto em fêmeas de controle R apenas ou G apenas, nem em tipo selvagem.

Exemplos 13-15:

Os requerentes demonstraram recentemente :

(5) emenda específica de sexo em construções de minigene

baseadas em Aadsx recombinante ;

(6) fenótipo específico de sexo de uma construção baseada em Cctra; e

(7) emenda específica de sexo em construções baseadas em *Aedes* - Actin4.

Pelo menos alguns de cada um destes exemplos não apenas mostram minigenes, mas realmente mostram emendas para gerar tTAV/tTAV2 ou ubi tTAV2

Exemplo 13: Minigenes *doublesex* (dsx) de *Aedes*

Ver também seção intitulada Sítios de ligação dsx Tra2 de *Aedes*. Os requerentes isolaram o gene dsx de *Aedes aegypti* (Aadsx) e identificaram 6 transcritos desta região (figura 1). Estes são: 2 transcritos específicos de machos (M1 e M2), 3 transcritos específicos de fêmeas (F1, F2 e F3) e um transcrito encontrado em ambos, machos e fêmeas (MF). Os requerentes realizaram duas construções de minigenes. Nestas construções, a grande maioria da seqüência intrônica foi deletada. Por exemplo, minigene 1 de DSX é aproximadamente 4,4kb em comprimento, considerando que suas seqüências terminais são separadas por aproximadamente 26kb em seu contexto natural, isto é, no DNA genômico de *Aedes aegypti*.

A emenda em minigene 2 da figura 1 é ilustrativa como emenda ocorre na forma de "fêmea" em ambos, machos e fêmeas. Isto pode significar que esse sistema depende de uso de aceitador de emenda alternativa. Nesse modelo, ocorre competição entre aceitadores de emenda alternativas, com algum fator específico de sexo influenciando isso, o fator específico de sexo provavelmente sendo Tra. Mas a deleção dos aceitadores de emenda 3' de M1 e M2 força a emenda nas formas F, por remoção da alternativa.

Assim, prefere-se que um ou mais dos aceitadores de emenda 3' específicos de fêmea (F1 e/ou F2) sejam providos junto com um aceitador

de emenda 3' adicional. O mais preferível, referido aceitador de emenda adicional é o aceitador de emenda 3' de variante de emenda de M1 ou M2 (ou ambos), embora seja previsto que isto não é essencial pois outros aceitadores de emenda 3' conhecidos irão provavelmente funcionar.

5 Figura 1 ilustra os vários transcritos produzidos por emenda alternativa do gene *doublesex* de *Aedes aegypti* (Aadsx). Será notado que *Aedes aegypti* é também conhecido como *Stegomyia aegypti*. A figura mostra o gene Aadsx do quarto éxon, que não é emendado alternativamente, isto é, está presente em todos os transcritos discutidos aqui. Numeração é do primeiro nucleotídeo do quarto éxon (acgacgaact...). Note-se que o diagrama
10 está fora de escala — os íntrons são muito mais longos que os éxons. A região alternativamente emendada total compreende acima de 43kb.

 Este fragmento de minigene foi incluído em uma construção de expressão (LA3515). *Aedes aegypti* transgênicos foram gerados por
15 recombinação específica de sítio no sítio de attP, usando o método de Nimmo et al (2006 : Nimmo, D.D., Alphey, L. Meredith, J.M. e Eggleston, P (2006). High efficiency site-specific genetic engineering of the mosquito genome. Insecto Molecular Biology, 15: 129-136)

 Um segundo minigene, menor, foi construído similarmente
20 (minigene 2 de DSX) e uma construção de expressão para este foi inserida no mesmo sítio de attP como minigene 1 de DSX, para permitir comparação direta (LA3534). Minigene 2 de DSX não mostrou emenda específica de sexo. Isto indica que as seqüências presentes em minigene 1 de DSX mas não em minigene 2 de DSX (aproximadamente 2029bp, ver Fig 1 e SEQ ID NO.
25 150, onde éxons são encontrados nas posições 29-163 e 1535-2572) são essenciais para a emenda alternativa correta, mesmo se o primeiro íntron emendado alternativamente, e a seqüência exônica imediatamente flanqueando o mesmo, estiverem presentes em ambas construções.

 Os requerentes produziram duas linhagens transgênicas

(LA3491 e LA3534) usando construções de minigenes de gene *dsx* de *Aedes aegypti*. LA3491 é uma fusão de éxon 4 compartilhado, éxons de cassete específicos de fêmea, e parte do primeiro éxon 3' compartilhado (éxon 5 em M1 transcrito).

5 Transcritos da região de minigene LA3491 foram analisados por PCR de transcriptase reversa (RT-PCR) e sequenciamento. Transcritos correspondendo à emenda alternativa na forma de F2 foram encontrados em fêmeas mas não em machos (Fig 2 e 3) e na forma F1 ocorreu alguma expressão de macho mas ela foi muito baixa (Fig 4). Enquanto transcritos
10 correspondendo à forma de M1 foram detectados em machos mas não em fêmeas (Fig 2). Uma vez que o minigene não continha o aceitador de emenda de 3' da variante M2, este transcrito não foi possível a partir desta construção. Esse minigene não contém qualquer seqüência exógena, embora o mesmo demonstre claramente emenda específica de sexo de um fragmento de *Aadsx*,
15 de fato um fragmento de "minigene" altamente deletado.

Será evidente que algumas seqüências são importantes para controlar emenda e devem ser, assim, conservadas, como discutido em outra parte. Isto pode ser facilmente estabelecido por deleção de algumas porções e teste para emenda alternativa por RT-PCR por exemplo.

20 Figura 2 mostra RT-PCR de machos e fêmeas de linhagem transgênica de *Aedes aegypti* LA3491 usando os iniciadores 688 - iel-transcr (SEQ ID NO. 4) e 790 - *Aedsx-m-r2* (SEQ ID NO. 5). Usando estes iniciadores, emenda no padrão de F2 iria dar uma banda de aproximadamente 985bp enquanto emenda no padrão de M1 iria dar uma banda de
25 aproximadamente 516bp. Uma banda de aproximadamente 985bp (F2) apareceu apenas em trilhas representando fêmeas e uma banda de aproximadamente 516bp de transcrito 1 específico de macho (M1) apareceu apenas em machos. Estes bandas foram seqüenciadas e mostraram que a emenda correta tinha ocorrido, isto é tipo F2 e tipo M1 respectivamente. A

ausência de bandas nos controles de RT (-RT CON) mostra que não houve contaminação de DNA genômico nas amostras. Trilhas 1 e 11 são marcadores (SmartLadder™ de Eurogentec, bandas de 1,5kb a 0,2kb são indicadas). Trilhas 2 e 3 são controles negativos (sem transcriptase reversa) e trilhas 2-9 representam reações realizadas em extratos dos machos e fêmeas, como marcado.

Figura 3 mostra RT-PCR dos machos e fêmeas de linhagens transgênicas de *Aedes aegypti* LA3491 usando os iniciadores 688 - iel-transcr (SEQ ID NO. 4) e 761 - Aedsx-fem-r (SEQ ID NO. 6). Usando estes iniciadores, emenda no padrão de F2 iria dar uma banda de aproximadamente 525bp. Uma banda de aproximadamente 525bp estava presente nas reações em extratos de fêmeas, mas não de reações correspondentes em extratos de machos. Sequenciamento desta banda de 525bp confirmou estar correto, isto é a emenda de tipo F2 ocorreu. Marcador (SmartLadder™ de Eurogentec, bandas de 1,5kb a 0,2kb são indicadas).

Figura 4 mostra RT-PCR dos machos e fêmeas de linhagens transgênicas de *Aedes aegypti* de LA3491 usando os iniciadores 688 - iel-transcr (SEQ ID NO. 4) e AedsxRl (SEQ ID NO. 4). Usando estes iniciadores, emenda no padrão F1 iria dar uma banda de 283bp. A banda de aproximadamente 283bp está predominantemente presente em fêmeas, embora haja evidência de uma quantidade pequena de emenda em machos. Sequenciamento confirmou que esta banda de fato corresponde à emenda no padrão F1. Marcador (SmartLadder™ de Eurogentec, bandas de 1,5kb a 0,2kb são indicadas).

LA3534 é idêntico a LA3491 exceto para uma deleção 3' de aproximadamente 2kb. Esta construção não mostrou nenhuma emenda diferencial entre machos e fêmeas (Fig 1, minigene 2). Géis de RT-PCR não foram mostrados para este caso. Com base nestes resultados, várias construções foram projetadas para incorporar a emenda específica de sexo de

LA3491 (Fig 1, minigene 1) em um sistema de retro-alimentação positivo. LA3612 (Fig 5), que incorpora uma fusão de ubiquitina e tTAV2 na região de codificação dsx, é engenheirado de modo que quando o transcrito de fêmea F2 é produzido, a ubiquitina é clivada e o tTAV2 é liberado para iniciar e sustentar o sistema de retro-alimentação positiva. LA3619 (Fig 5) tem tTAV2 sem ubiquitina e usando seu próprio códon de partida de tradução. LA3646 (Fig 5) é idêntico a LA3619 exceto os códons de partida para o gene de dsx foi mudado; isto deve melhorar a quantidade de tTAV2 produzida removendo a tradução não específica.

Figura 5 é uma representação diagramática de plasmídeos com base em emenda em minigene dsx de *Aedes aegypti*. Para maior certeza, será entendido que o primeiro íntron de fêmea representa qualquer emenda F1, F2 ou F3, e tTAV no diagrama refere-se a tTAV2 (será notado que outras proteínas ou outras versões de tTA ou tTAV poderiam alternativamente ser usadas). Em cada um destes plasmídeos, além de LA3491, a sequência heteróloga foi adicionada ao éxon F2. "ATG putativo" representa qualquer sequência de triplete ATG em sequência exônica localizada 5' com relação ao DNA heterólogo. Em LA3646, estes códons de partida de tradução putativos ("ATG putativo") foram removidos ou modificados. No caso de construção LA3612, tradução de um ATG a montante (5') que está em matriz com a região de codificação de ubi-tTAV irá ainda (assumindo nenhuma intervenção de códon de parada) produzir tTAV funcional, seguindo separação das porções de ubiquitina e tTAV por ação de protease. Os vários cassetes de emenda alternativa estão ligados operativamente a um promotor apropriado, terminador transcripcional e outras sequências regulatórias.

Esse exemplo mostra emenda específica de sexo de um fragmento de "minigene" altamente comprimido em um contexto heterólogo (isto é promotor heterólogo, 5' UTR e 3' UTR). Embora não mostre expressão

diferencial de uma sequência não *Aedes*, como os éxons emendados alternativamente são derivados do gene *Aadxs* e não contêm material adicional, isto ilustra claramente a possibilidade desta abordagem. Em qualquer caso, os promotores, 5' UTR e 3' UTR, são heterólogos. Os requerentes realizaram construções adicionais que ilustram vários métodos diferentes para obter expressão diferencial (específica de sexo) de uma proteína heteróloga por este *dsx*.

Alinhamento de sequência TRA

Pane et al. (2002) sugeriram que determinadas sequências relacionadas a sítios de ligação conhecidos do complexo TRA/TRA2 em *Drosophila* poderiam ser importantes na regulação da emenda de *Cctra*, e isto também era conhecido para *dsx* de *Drosophila* e foi também sugerido para *dsx* de *Anopheles gambiae* (Scali et al 2005). A sequência de consenso é diferentemente descrita como

UC(U/A)(U/A)C(A/G)AUCAACA (Pane et al), SEQ ID NO. 8, ou UC(U/A)(U/A)CAAUCAACA (Scali et al 2005), SEQ ID NO. 9.

É notável que estas definições sejam extremamente similares. Pane et al identificam 8 alinhamentos parciais para esse consenso na sequência de *Cctra* (7 ou mais nucleotídeos alinhando as 13 sequências nucleotídicas de consenso). Scali et al identificam 6 alinhamentos em *Agdsx* (9/13 ou melhor). Tais sequências são também conhecidas para regular a emenda alternativa do gene *fruitless* de *Drosophila*; Scali et al reviram 3 alinhamentos naquela sequência (12/13 ou melhor). Emenda correta de *dsx* pode também requerer um região rica em purina, como discutido por Scali et al.

Como pode ser visto da Tabela 2 e figura 7, os requerentes identificaram o que se pensa ser agrupamentos significantes de sítios de ligação para Tra/Tra2 em minigene 1 *dsx* de *Aedes aegypti* dos requerentes.

Alinhamento de sequência dsx de traça e motivos conservados

Figura 6 mostra um alinhamento dos segundos éxons específicos de fêmea e seqüências de flaqueamento de gene dsxs da lagarta rosada (*Pectinophora gossypiella*, PBW - dsx, SEQ ID NO. 146), bicho-da-seda (*Bombyx mori*, bombyx-dsx, SEQ DD NO. 147) e traça das maçãs e pêras (*Cydia pomonella*, traça das maçãs e pêras- dsx, SEQ ED NO. 148). O segundo éxon específico de fêmea é mostrado em negrito. Os requerentes identificaram múltiplas cópias de uma seqüência nucleotídica curta, repetida, conservada em seqüência e localização aproximada entre estas traças de parentesco relativamente distante; estas estão localizadas apenas 5' para o éxon específico de fêmea. As repetições AGTGAC/T conservadas estão sublinhadas. Asteriscos (*) representam nucleotídeos idênticos, traços (-) representam intervalos para melhor alinhamento. Os éxons estão representados na SEQ ID NOS. pela seguinte numeração de nucleotídeos: SEQ ID NO. 146289-439; SEQ ID NO. 147 339-492; e SEQ ID NO. 148 285-439.

Sítios de ligação Tra2 *Aedes dsx*.

Em fêmeas de *Drosophila melanogaster*, Tra e um produto do gene tra2 constitutivamente ativo, atuam como reguladores de emenda por ligação dos sítios de melhorador de emenda no pré-mRNA de dsx, que ativa o sítio do aceitador 3' fraco do éxon específico de fêmea (Scali et al). Em machos não há expressão de TRA e sítio do aceitador 3' fraco não é reconhecido e a emenda ocorre no éxon de macho. Para buscar por sítios de ligação Tra/Tra2 putativos, os requerentes usaram a seqüência de consenso destes sítios de ligação deduzidas para *Drosophila* Tra/Tra2 e procuraram a distribuição destes na seqüência de gene dsx de *Aedes aegypti*. Isto é mostrado na Tabela 2, abaixo.

Nome	Seqüência w = T ou A r = A ou G	Presente em minigene 1	Posição	Identidade com consenso	Identidade com wwcrat	SEQ ID NO.
consenso	tcwwcratcaaca	/	/	/13	/6	138

1	tcaacaagcaaca	Y	14917	12	5	10
2	ttatcaacaaca	Y	364	11	5	11
3	tcataattaaaa		1015	11	6	12
4	tcataatcaaac		6502	11	6	13
5	tcttcaaccaacc	Y	14958	11	5	14
6	cctacaatctaca	Y	14973	11	6	15
7	tcttagatcaaaa		16553	11	5	16
8	tcttcgatcatta		17386	11	6	17
9	ccaacaatctaca		28802	11	6	18
10	tcaaagatcacca		42096	11	5	19
11	tcttcggtcgacg	Y	256	11	5	20
12	tcgacaaacaaaa		1277	11	<5	21
13	tattcaaacacg		4061	11	5	22
14	ttttcgataaaaa		4380	10	6	23
15	tcttcagtctgca		5399	10	5	24
16	gattcaatcatca		7723	10	6	25
17	ttatcgagcaaaa		8137	10	5	26
18	tcataactcaaga		9062	10	<5	27
19	tcagaaatcaaaa		9126	10	<5	28
20	tctttaatttaca		10639	10	5	29
21	tttacaatcctca		10646	10	6	30
22	tcatagatcagga		11214	10	5	31
23	acctcaaacaca		11989	10	<5	32
24	tcategaacaccc		12020	10	5	33
25	tcaataatcgta		12199	10	5	107
26	tcataaacgtca		13287	10	5	108
27	ttatcggttaaaca	Y	13439	10	5	109
28	taaacagtcaata	Y	13446	10	5	110
29	tacacgatcagca	Y	14096	10	5	111
30	aatacaaacaca	Y	14637	10	5	112
31	tcataacaagca	Y	14914	10	5	113
32	tctacaaccaga	Y	14980	10	5	114
33	acatcgattcaca		16085	10	6	115
34	cgctcaatcaaca		16175	10	5	116
35	tctaccataaaaa		16511	10	5	117
36	aatgaatcaaca		20044	10	5	118
37	acatcggtcaacg		21374	10	5	119
38	tcttgattcacca		21580	10	<5	120
39	tctgcagacaaca		22408	10	<5	121
40	tcttcggtaatca		23285	10	5	122
41	tctataaacaata	Y	25436	10	<5	123
42	taacaataaata	Y	25440	10	6	124
43	taacaagcaaaa		28242	10	5	125
44	tcaacgatcggcg		30309	10	6	126
45	tgatccatcatca		30910	10	5	127
46	tcaacatgcaaga		32295	10	<5	128
47	tcttaataaaga		32862	10	5	129
48	tcaaagatctata		40551	10	5	130
49	taatgaattaaca		40847	10	5	131
50	ttaccatcaact		41712	10	5	132
51	taatgaaacaaca		43380	10	<5	133
52*	gtttcaattaaaa	Y	13500	9	6	134
53*	tattcaattataa	Y	13602	9	6	135
54*	tcttcaatcgttt	Y	15002	9	6	136
55*	tcaacgatccttt	Y	15533	9	6	137

Tabela 2: * = em 3491, apenas 9/13 mas 6/6 no núcleo. Esta tabela não inclui 9/13 de identidades além daquelas em 3491 com identidade de 6/6 com sequência de núcleo de wwcrat. Esta sequência de núcleo de consenso (WWCRAT) é particularmente preferida.

Figura 7 é uma representação diagramática de sítios de ligação Tra/Tra2 putativos na região de codificação dsx de plasmídeo LA3491. Esse diagrama está aproximadamente em escala e representa uma sequência de aproximadamente 4kb. Os requerentes puderam calcular a possibilidade de um alinhamento aleatório para a sequência de consenso Tra/Tra2. Considerando que todos os 4 nucleotídeos ocorrem em uma frequência igual, as chances de qualquer dado nucleotídeo em uma sequência aleatória ser o primeiro nucleotídeo de um alinhamento 10/13 ou melhor para o consenso são de aproximadamente 7×10^{-4} . Assim, seria de se esperar um pouco menos do que um tal alinhamento por 1000 nucleotídeos de tal sequência aleatória. O cálculo para isto é apresentado abaixo:

Emenda específica de sexo: probabilidades

Perguntas

Um sequência de consenso de sítio de ligação consiste de 13 bases. Dez daquelas posições (fixas) (chamado conjunto X) deve ser cada uma base específica. As outras três (chamado de conjunto Y) pode ser cada de duas bases específicas. Considerando que cada base possível A, G, C e T é igualmente provável e que a base de cada posição é independente das bases em outras posições, qual é a probabilidade de uma sequência de 13 bases selecionada aleatoriamente exatamente alinhar esta sequência? Quais são as probabilidade de tal sequência ser um desalinhamento próximo (permitindo até uma, duas, três ou quatro diferenças)? As respostas são dadas na Tabela 2 abaixo e os aspectos operacionais são mostrados a seguir.

Respostas

Tabela 3

Número de posições desalinhadas	Probabilidade (fração)	Probabilidade (para 3 d.p.)
sem, isto é, alinhadas exatas	$1/2^{23}$	$1,192 \times 10^{-7}$
até 1, isto é, pelo menos 12 posições alinhadas	$17/2^{22}$	$4,053 \times 10^{-6}$
até 2, isto é, pelo menos 11 posições alinhadas	$133/2^{21}$	$6,342 \times 10^{-5}$
até 3, isto é, pelo menos 10 posições alinhadas	$23/2^{15}$	$7,019 \times 10^{-4}$
até 4, isto é, pelo menos 9 posições alinhadas	$33863/2^{23}$	$4,037 \times 10^{-3}$

Aspectos operacionais

$$P(\text{alinhamento exato}) = P_0 = \left(\frac{1}{4}\right)^{10} \left(\frac{1}{2}\right)^3 = \frac{1}{4^{10} \times 2^3} = \frac{1}{2^{23}} = 1.192 \times 10^{-7} \text{ to 3 d.p. (3 d.p todos abaixo)}$$

$P(\text{desalinhamento em exatamente 1 posição}) =$

$P(\text{desalinhamentos em uma das 10 posições de X ou desalinhamento em uma das 3 posições de Y})$

$$= P_1 = 10 \left(\frac{1}{4}\right)^9 \left(\frac{3}{4}\right) \left(\frac{1}{2}\right)^3 + 3 \left(\frac{1}{4}\right)^{10} \left(\frac{1}{2}\right)^2 = \frac{(10 \times 3) + 3}{4^{10} \times 2^3} = \frac{33}{2^{23}} = 3.934 \times 10^{-6}$$

$P(\text{desalinhamento em exatamente 2 posições}) =$

$P(\text{desalinhamentos em 2 dos 10 X ou desalinhamento em 1 dos 10 X e 1 dos 3 Y ou desalinhamento em 2 dos 3 Y})$

$$= P_2 = \frac{10!}{2!8!} \left(\frac{1}{4}\right)^8 \left(\frac{3}{4}\right)^2 \left(\frac{1}{2}\right)^3 + 10 \times 3 \left(\frac{1}{4}\right)^9 \left(\frac{3}{4}\right) \left(\frac{1}{2}\right)^3 + 3 \left(\frac{1}{4}\right)^{10} \left(\frac{1}{2}\right)^2$$

$$= \frac{((45 \times 3^2) + (30 \times 3) + 3)}{2^{23}} = \frac{498}{2^{23}} = \frac{249}{2^{22}} = 5.937 \times 10^{-5}$$

$P(\text{desalinhamento em exatamente 3 posições}) =$

$P(\text{desalinhamentos em 3 dos 10 X ou desalinhamentos em 2 dos 10 X e 1 dos 3 Y ou desalinhamentos em 1 dos 10 X e 2 dos 3 Y ou desalinhamentos em 3 dos 3 Y})$

$$= P_3 = \frac{10!}{3!7!} \left(\frac{1}{4}\right)^7 \left(\frac{3}{4}\right)^3 \left(\frac{1}{2}\right)^3 + \frac{10!}{2!8!} 3 \left(\frac{1}{4}\right)^8 \left(\frac{3}{4}\right)^2 \left(\frac{1}{2}\right)^3 + 10 \times 3 \left(\frac{1}{4}\right)^9 \left(\frac{3}{4}\right) \left(\frac{1}{2}\right)^3 + \left(\frac{1}{4}\right)^{10} \left(\frac{1}{2}\right)^2$$

$$= \frac{((120 \times 3^3) + (45 \times 3^3) + (30 \times 3) + 1)}{2^{23}} = \frac{5356}{2^{23}} = \frac{1339}{2^{21}} = 6.385 \times 10^{-4}$$

$P(\text{desalinhamento em exatamente 4 posições}) =$

$P(\text{desalinhamentos em 4 dos 10 X ou desalinhamentos em 3 dos 10 X e 1 dos 3 Y ou desalinhamentos em 2 dos 10 X e 2 dos 3 Y ou desalinhamentos em 1 dos 10 X e 3 dos 3 Y})$

$$= P_4 = \frac{10!}{4!6!} \left(\frac{1}{4}\right)^6 \left(\frac{3}{4}\right)^4 \left(\frac{1}{2}\right)^3 + \frac{10!}{3!7!} 3 \left(\frac{1}{4}\right)^7 \left(\frac{3}{4}\right)^3 \left(\frac{1}{2}\right)^3 + \frac{10!}{2!8!} 3 \left(\frac{1}{4}\right)^8 \left(\frac{3}{4}\right)^2 \left(\frac{1}{2}\right)^3 + 10 \left(\frac{1}{4}\right)^9 \left(\frac{3}{4}\right) \left(\frac{1}{2}\right)^3$$

$$= \frac{((210 \times 3^4) + (120 \times 3^4) + (45 \times 3^3) + (10 \times 3))}{2^{23}} = \frac{27975}{2^{23}} = 3.335 \times 10^{-3}$$

$$P(\text{desalinhamento em até 1 posição}) = P_0 + P_1 = \frac{1+33}{2^{23}} = \frac{17}{2^{22}} = 4.053 \times 10^{-6}$$

$$P(\text{desalinhamento em até 2 posições}) = P_0 + P_1 + P_2 = \frac{1+33+498}{2^{23}} = \frac{532}{2^{23}} = \frac{133}{2^{21}} = 6.342 \times 10^{-5}$$

$P(\text{desalinhamento em até 3 posições})$

$$= P_0 + P_1 + P_2 + P_3 = \frac{1+33+498+5356}{2^{23}} = \frac{5888}{2^{23}}$$

$$= \frac{23}{2^{15}} = 7.019 \times 10^{-4}$$

$P(\text{desalinhamento em até 4 posições})$

$$= P_0 + P_1 + P_2 + P_3 + P_4 = \frac{1+33+498+5356+27975}{2^{23}} = \frac{33863}{2^{23}}$$

$$= 4.037 \times 10^{-3}$$

Experimento 14: Cctra

Os requerentes possuem uma linhagem de LA3097 (LA3097A) que mostra expressão muito boa de seu marcador fluorescente; é desconhecido se esta linhagem é um evento de integração simples. Esta linhagem mostra evidência de emenda específica de sexo, quando criadas sem tetraciclina todas as fêmeas morreram como embriões, e quando administrados com 30µg/ml de tetraciclina ambos, machos e fêmeas, sobrevivem.

Esse exemplo é importante. Ele mostra que Cctra provê emenda alternativa específica de sexo em *Aedes*, e que isto pode ser usado para resultar em letalidade específica de sexo. Assim, por esta razão, ele fornece uma evidência da faixa filogenética de emenda CcTRA. Assim, é inteiramente plausível que a presente invenção possa ser aplicada para todos dípteros, como os requerentes demonstraram que Cctra trabalha em *Drosophila*, tefritídas e mosquitos, que essencialmente estendem-se sobre toda ordem de Dípteros.

É surpreendente que Cctra funcione em *Aedes*, dado a rápida evolução de seqüência de tra.

Os requerentes transformaram *Aedes aegypti* com construção LA3097. Machos heterozigóticos da linhagem transgênica resultante foram cruzados com tipo selvagem e a progênie criada em meio de cultura aquoso suplementado com tetraciclina para uma concentração final de 30 µg/ml. Adultos foram recuperados como segue: 14 machos e uma fêmea, deste modo mostrando letalidade específica de fêmea resultante.

Esta espécie e cepa têm normalmente uma relação de sexo de aproximadamente 1:1, por esta razão esta construção resultou em letalidade específica de fêmea em *Aedes aegypti*. Construções equivalentes que não continham a seqüência intrônica de Cctra resultaram em letalidade não específica de sexo. Por esta razão, o íntron Cctra pode ser usado para prover regulação (isto é, específica de sexo) de expressão de gene diferencial em

mosquitos, e isto pode ainda ser usado para prover letalidade específica de sexo e um método para eliminação seletiva de fêmeas de uma população.

Em maiores detalhes: em 0 µg/ml de tetraciclina, machos sobrevivem apenas para pupas, isto é não se tornam adultos. Fêmeas morrem
5 tão prematuramente que não foram vistas, provavelmente como embriões, assim, ainda há um efeito diferencial entre os sexos. No entanto, a letalidade pupal em machos sugere que o sistema não é completamente desligado em machos. A linhagem de inserção única que os requerentes recuperaram é incomum, em que ela mostra expressão extremamente forte do marcador;
10 outras inserções com níveis de expressão mais típicos poderiam bem não mostrar letalidade para machos.

Emenda em LA3097A

Análise de emenda de LA3097 de mosquitos transgênicos de LA3097A por RT-PCR mostrou que machos e fêmeas compartilharam dois
15 transcritos, uma banda de aproximadamente 950 bp e uma banda fraca de aproximadamente 800 bp (figura 59). Sequenciamento destas bandas mostraram que banda de 900 bp corresponde a uma variante de emenda não específica de sexo (AeM2, -920 bp), e a banda fraca foi uma mistura de uma variante de emenda não específica de sexo (AeM1, ~804bp) e a forma de
20 fêmea (AeF1, ~765bp), ver figura 60. A emenda do transcrito de AeF1 foi idêntica à mostrada para esta construção em mosca-do-mediterrâneo (figura 33). A emenda dos transcritos de M diferem um pouco da vista no contexto nativo (Emenda de CcTRA em mosca-do-mediterrâneo, que o gene nativo ou como observado pelos requerentes de LA3097 em mosca-do-mediterrâneo
25 transgênica); em AeM1, o segundo éxon emendado alternativamente (ME1b) não está incluído no transcrito de AeM1 maduro e, em AeM2, o segundo éxon de emenda alternativa (ME2b) está similarmente não incluído no transcrito AeM2 maduro. Em outras palavras, para cada um destes transcritos o primeiro, mas não o segundo, éxon de cassete está presente, relativo ao

protótipo de mosca-do-mediterrâneo. Note-se que, como uma consequência da ausência do segundo éxon de cassete em AeM1, e a matriz de leitura de tTAV2 relativa ao primeiro éxon de cassete nesta construção, a emenda no padrão de AeM1 não conduz a interrupção da matriz de leitura aberta de tTAV2, mas sim à adição de 39 nucleotídeos (correspondendo a 13 aminoácidos) entre a ATG e o resto da matriz de leitura aberta de tTAV2. É provável que esta variante de tTAV2 possa reter alguma atividade, relativa ao tTAV2 normal ou prototípico (como codificado pela variante de emenda de F1). Na ausência de tetraciclina, um efeito fenotípico foi observado em machos bem como em fêmeas, embora mais fraco em machos do que em fêmeas. Produção de uma variante parcialmente ativa de tTAV2 a partir do transcrito de AeM1 em machos (e fêmeas) pode explicar isso.

Figura 59 mostra RT-PCR dos machos e fêmeas de linhagem transgênica de *Aedes aegypti* LA3097A usando os iniciadores HSP (SEQ ID NO. 139) e VP16 (SEQ ID NO. 140). Usando estes iniciadores, emenda no padrão de CcF1 (isto é correspondendo à variante de F1 de *Ceratitis capitata*) resultaria em uma banda de aproximadamente 765bp e emenda na CcM1 1005bp e CcM2 1094bp. Em ambos, machos e fêmeas, uma banda forte de aproximadamente 950bp (1) foi observada junto com uma banda fraca de aproximadamente 800bp (2). Marcador (SmartLadder™ de Eurogentec, bandas de 1,5kb a 0,4kb são indicadas).

Análise de sequência de vários clones de banda 2 (isto é variantes de emenda de AeM1/Ae F1) de machos e fêmeas mostraram que um dentre cinco clones de fêmeas mostraram emenda de AeM2 (20%), enquanto em machos, três dos quatro clones mostraram emenda de AeM2 (75%); todos os outros clones mostraram emenda de AeF1. Isto indica que há mais transcrito de AeF1 presente em fêmeas que em machos e isto explicaria o efeito de morte diferencial visto entre os mesmos.

Figura 60 ilustra os vários transcritos produzidos por emenda

alternativa de Cctra de linhagem transgênica de *Aedes aegypti* LA3Q91A. 3097 representa a sequência de DNA de Cctra e os números referem-se à figura descrita em outra parte. Sombreamento e caixas também referem-se à figura 33. Nota-se que o diagrama está fora de escala.

5 **Exemplo 15: Actin-4 de *Aedes***

Os requerentes obtiveram onze linhagens de LA3545, que usam o gene actin-4 de *Aedes* (AeAct-4 ou AaAct4) para dirigir a expressão de tTAV2. Em construção LA3545, uma sequência codificando tTAV2 foi inserida no segundo éxon de AaAct4 (fig 10). Para transcritos emendados no

10 padrão característico de emenda AaAct4 em fêmeas, o ATG da região de codificação de tTAV2 será o primeiro ATG (o mais a 5') do transcrito. Emenda no padrão característico de emenda AaAct4 em machos introduz um arranjo de códons de partida e de parada antes da sequência de tTAV2 que

15 de tTAV2. Estas linhagens devem apenas expressar tTAV2 em pupas de fêmeas. A emenda é mostrada na figura 8, abaixo.

Figura 8 mostra RT-PCR de adultos machos e fêmeas de linhagem transgênica de *Aedes aegypti* LA3545AeC usando os iniciadores Agéxon1F (SEQ ID NO. 141) e TETRRI (SEQ ID NO. 142). Usando estes

20 iniciadores, emenda em um padrão equivalente àquele do gene AaAct4 nativo iria dar bandas de aproximadamente 347bp para a variante de emenda de tipo fêmea e de aproximadamente 595bp para a variante de emenda de tipo macho. Uma banda de aproximadamente 347 bp (F) foi encontrada apenas em reações em extratos de fêmeas; uma banda de aproximadamente 595bp (M) foi

25 encontrada em ambos, machos e fêmeas. Sequenciamento confirmou que a emenda correta ocorreu tanto em machos como em fêmeas. Marcador (SmartLadder™ de Eurogentec, bandas de 1,5kb a 0,2kb são indicadas).

Os requerentes também obtiveram *Aedes aegypti* transgênico portando construção LA3604, que é similar a LA3545 exceto que este tem um

códon de partida engenheirado na porção de éxon 1 que está presente em ambos, transcritos de tipo macho e de tipo fêmea (Fig 10). Isto está disposto para ser o primeiro ATG em qualquer tipo de transcrito. LA3604 codifica tTAV2 fusionado a ubiquitina (LA3545 codifica tTAV, enquanto LA3604
 5 codifica ubi-tTAV2). Esta construção deve produzir proteína de tTAV2 completamente funcional em fêmeas apenas, mesmo se a forma de macho for expressada em fêmeas o éxon de macho extra contém vários códons de partida e de parada que poderiam prevenir a tradução da proteína de fusão ubi-tTAV2.

10 A emenda alternativa de AaAct4 ocorre no 5' UTR (do gene nativo). Ela pode ou não ter uma parte regulatória no gene nativo. Uma possibilidade é como a seguir: na variante de emenda específica de fêmea, o códon de partida da região de codificação de AaAct4 é o primeiro ATG do transcrito. No entanto, na variante de emenda específica de macho, existem
 15 várias seqüências de ATG adicionais a 5' para o códon de partida da região de codificação de AaAct4; a maior destes tem códons de parada na matriz em uma distância pequena 3'. Esta arrumação de seqüência pode interferir com a tradução eficiente da proteína AaAct4 e assim reduzir expressão da proteína em machos como comparado com fêmeas. Esta é a arrumação em LA3545.

20 No entanto, um efeito diferencial maior entre machos e fêmeas seria esperado se o íntron fosse incluído na região de codificação (em vez de 5' UTR), isto é, inserido entre os códons de partida e parada do polinucleotídeo para expressão no organismo. Nesse caso, o éxon de cassete específico de macho mudaria o potencial de codificação do transcrito, em vez
 25 de simplesmente interferir com tradução.

Isso é obtido na construção LA3604. Os requerentes modificaram o primeiro éxon compartilhado para incluir uma seqüência de ATG em um contexto de seqüência apropriado para uma iniciação translacional. Nesta seqüência modificada, esse é o primeiro ATG em ou as

variantes de emenda de tipo macho (M) ou de tipo fêmea (F). Seguindo a emenda na forma de F, esse (projetado) ATG de 5' está na matriz com a região de codificação de ubi-tTAV. Transcritos de tipo F codificariam, por esta razão, uma proteína de fusão, compreendendo seções codificadas por (i) 5 parte de qual é normalmente 5' UTR de Act4 (mas aqui obviamente traduzida, e assim, não UTR em todas), (ii) região de codificação de ubiquitina e (iii) região de codificação de tTAV2.

Atividade de ubiquitina proteases celulares irão liberar a proteína de tTAV2. Tradução do ATG 5' engenheirado seria terminada por 10 códons de parada em matriz na seqüência adicional (éxon de cassete) presente em transcritos emendados na forma M. Isto poderia, assim, evitar a expressão de tTAV2 funcional em machos, assim dando expressão específica de sexo de tTAV2. Obviamente, isto proporcionaria um método geral de expressão específica de sexo de uma proteína, substituindo-se o segmento tTAV2 com 15 outra proteína ou seqüência de interesse. Usando esta estratégia os requerentes proveram transgênicos e mostraram a emenda específica de sexo (Fig 9).

Figura 9 mostra RT-PCR dos machos e fêmeas de linhagem transgênica de *Aedes aegypti* LA3604AeA usando os iniciadores AgexonlF 20 (SEQ ID NO. 141) e TETRRI (SEQ ID NO. 142). Usando estes iniciadores, emenda na forma de fêmea iria dar uma banda de aproximadamente 575 bp, enquanto inclusão do éxon de cassete específico de macho aumentaria o mesmo a aproximadamente 823bp. Uma banda de aproximadamente 575bp foi vista em cada fêmea analisada, enquanto uma banda de aproximadamente 25 823bp foi vista em cada macho analisado. Estas bandas parecem ser substancialmente específicas para os respectivos sexos. Sequenciamento destas bandas mostrou que a emenda correta ocorreu em machos e fêmeas. Marcador: SmartLadder™ de Eurogentec, bandas de 1,5 kb a 0,2 kb são indicadas.

Figura 10, abaixo, é uma representação diagramática de LA3545 e LA3604 de plasmídeos. S1: éxon 1 compartilhado; M1: seqüência adicional incluída em éxon 1 específico de macho; S2: éxon 2 compartilhado (extremidade de 5' apenas); ubi: seqüência codificando ubiquitina; tTAV2: seqüência codificando tTAV2.

Em várias das linhagens transgênicas LA3545, um efeito específico de sexo e tecido foi observado: fêmeas não são aptas a voar. Duas das linhagens mostraram 90-100% de um fenótipo de inaptidão ao vôo de fêmeas, uma linhagem mostra 70% sem a aptidão ao vôo e outros 50%. Este fenótipo é provavelmente devido à expressão específica de fêmea de tTAV2 no desenvolvimento de músculos aptos a voar. A diferença nos fenótipos entre as linhagens é devido aos efeitos posicionais na expressão do promotor AaAct4. Devido a uma posição de genes na expressão de genoma poder ser influenciada pelo número de fatores (regiões de heterocromatina ou eucromatina, elementos melhoradores e supressores, proximidade de outros genes) que podem ser vistos prontamente nos marcadores fluorescentes usados para identificar transgênicos. Todas as onze linhagens de LA3545 foram identificadas porque elas têm perfis fluorescentes diferentes, mesmo se elas tem os mesmos promotores e marcador. Esta variação é devido aos efeitos posicionais. Isto poderia então significar que os requerentes esperaram que algumas linhagens de LA3545 expressassem mais tTAV2 do que outras, devido aos efeitos posicionais, e as linhagens que expressam mais irão dar um fenótipo de inaptidão ao vôo específico de fêmea.

Para testar esta hipótese, os requerentes desenvolveram uma linhagem separada de *Aedes aegypti* com um gene repórter tetO-DsRed2 (LA3576 ver Fig 17 e SEQ ID NO. 143), quando cruzados com linhagens LA3545 diferentes isto iria permitir a visualização de onde e quando Actin4-tTAV2 estava sendo expressado. Dentre as 8 linhagens de LA3545 cruzadas para LA3576 todas mostraram fluorescência de músculo de vôo indireto

específico de fêmeas em larvas L4 tardias, pupas e adultos. m quatro das linhagens, a expressão de DsRed2 pareceu ser específica (isto é exclusiva) para os músculos aptos ao vôo indiretos de fêmeas; em outros, quatro tecidos adicionais mostraram expressão de DsRed2. Este fenômeno, onde expressão
5 de um transgene depende, em parte, da região ou ponto no genoma em que está inserido, é chamado efeito de posição, e será bem conhecido e entendido pela pessoa versada na técnica.

Usando LA3576, provou-se que a expressão de tTAV2 em LA3604 foi específica de fêmea, ocorre principalmente nos músculos de vôo
10 indiretos e é específica de estágio. Várias construções de efetuator tetO foram então construídas para analisar seus efeitos. Os transgênicos tetO- MichelobX (LA3582, ver Fig 15 e SEQ ID NO. 144) quando cruzados para LA3545 demonstraram todos fenótipos de inaptidão ao vôo específicos de fêmea que poderiam ser reprimidos por tetraciclina. Isto prova que Actin4 pode ser usada
15 para dirigir um gene efetuator em um modo específico de estágio, tecido e sexo.

Porque algumas linhagens de LA3545 têm um fenótipo de inaptidão ao vôo específico de fêmea sem a presença de um gene efetuator induzido, isto mostrou que tTAV2 poderia atuar como uma molécula de
20 efetuator. tTAV2 é composto de um tTA, um domínio de ligação de tetO e VP16, uma proteína de vírus de herpes simplex. VP16 ativa a transcrição de genes virais prematuros imediatos usando suas seqüências amino-terminais para fixar a uma ou mais proteínas codificadas de hospedeiro que reconhecem seqüências de DNA em seus promotores. Em LA3604, um efetuator de tetO -
25 VP16 foi adicionado para melhorar o efeito de tTAV2. Em três linhagens transgênicas de LA3604 isto levou a um fenótipo de inaptidão ao vôo específico de fêmeas de 100% quando criadas sem tetraciclina, mostrando que VP16 é uma molécula efetuidora efetiva. Nota-se que LA3604 tem um códon de partida potencial (ATG) engenheirado a 5' de para o íntron emendado

alternativamente. Assim, nesta construção, o éxon específico de macho é esperado para interromper a matriz de leitura aberta codificando tTAV (ubi-tTAV); uma vez que a seqüência específica de macho contém vários códons de parada, isto tenderá a reduzir ou eliminar a produção de tTAV funcional em machos. Por meio de comparação, o éxon específico de macho 5' para o códon de partida de tTAV em LA3545. No entanto, inserindo um número de códons de partida 5' para o códon de partida de tTAV (que é o primeiro ATG do transcrito de fêmea mas não do transcrito de macho), nenhum destes códons de partida adicionais sendo apropriados para a produção eficiente de tTAV funcional devido a estar fora da matriz ou ter códons de parada intervindo, esta arrumação também irá tender a reduzir ou eliminar a produção de tTAV funcional em machos, consistente com dados fenotípicos acima.

Exemplo 16: Uso de ubiquitina e posicionamento de íntrons

Os requerentes realizaram recentemente construções baseadas em Ccra com o íntron de cassete de Ccra em uma variedade de contextos diferentes, isto é flanqueado por seqüências diferentes. Várias linhagens de mosca-do-mediterrâneo transgênicas transportando estes foram construídas. Isto mostra que o sistema é geral e robusto, isto é, que vai funcionar para uma extensa faixa de seqüências de interesse heterólogas.

Os requerentes também obtiveram pelo menos um exemplo feito recentemente de uma fusão Ccra -ubi-tTAV dando uma emenda correta (DsRed- Ccra - ubi-tTAV).

Exemplos preferidos da proteína funcional colocam a seqüência de codificação para ou ubiquitina ou tTA, ou seus mutantes funcionais e ou variantes tais como tTAV, tTAV2 ou tTAV3, 3' para o íntron. Estes estão dispostos de modo que estes elementos estão substancialmente adjacentes à extremidade 3' do íntron, mais preferivelmente de modo que a região de codificação inicia dentro de 20 nucleotídeos ou menos do limite do

íntron 3'), e, o mais preferivelmente, imediatamente adjacente à extremidade 3' do íntron, embora isto seja menos relevante se o sistema de ubiquitina for usado.

Exemplos preferidos de construções de acordo com a presente invenção estão listados na Tabela 4, abaixo. Será notado que LA1118 não está dentro do escopo da presente invenção, como ele não codifica uma proteína funcional, isto é não funciona propriamente. Pensa-se que isto é devido ao uso inesperado de um doador de emenda 5' de 4bp para a junção com sequência de íntron Cctra, conduzindo a uma armação que é induzida em todas as emendas. Assim, isto é incluído apenas com a finalidade de informação.

Tabela 4

Número construção (Figuras #.)	Espécie íntron tra é de	Posição de ATG (bp)	Íntron tra é fusionado para
LA1188 (80)	mosca-do-mediterrâneo	+132	tTAV
LA3014 (29)	mosca-do-mediterrâneo	+22	ubiquitina
LA3166 (30)	mosca-do-mediterrâneo	+136	ubiquitina
LA3097 (27)	mosca-do-mediterrâneo	+0	tTAV
LA3077 (26)	mosca-do-mediterrâneo	+61	tTAV
LA3233 (28)	mosca-do-mediterrâneo	+0	tTAV2
LA3376 (31)	mosca-do-mediterrâneo	+0	tTAV2
LA3376 (31)	<i>B. zonata</i>	+3	reaper KR
LA3376 (31)	<i>B. zonata</i>	+0	tTAV3
LA3242 (32)	<i>C. rosa</i>	+3	reaper KR
LA1038 (14)	mosca-do-mediterrâneo	+21	Nipp1 (nipper)
LA3054 (61)	mosca-do-mediterrâneo	+811	DsRed-ubiquitina
LA3056 (62)	mosca-do-mediterrâneo	+811	DsRed-ubiquitina
LA3488 (63)	mosca-do-mediterrâneo	+949	Ubiquitina
LA3596 (67)	mosca-do-mediterrâneo	+949	Ubiquitina

Tabela 4 mostra construções que contêm uma sequência de controle de emenda que é derivada de um íntron tra. Os íntrons foram derivados de *C. capitata* (mosca-do-mediterrâneo), *B. zonata* ou *C. rosa* (ver coluna 2). Referido íntron foi inserido dentro da região de codificação tal que a distância entre ATG de iniciador putativo e o último nucleotídeo do éxon imediatamente precedendo o íntron tra foi como deve estar indicado na coluna 3. Íntron é inserido em ou adjacente à região de codificação para ubiquitina, tTAV, ReaperKR, nipper ou ubiquitina- DsRed como mostrado na coluna 4. Estes foram gerados e mostrados como emendando com sucesso, por RT-PCR ou fenotipicamente em mosca-do-mediterrâneo em, em alguns casos, também

quer em *Drosophila melanogaster* (LA3077) ou *Anastrepha ludens* (LA3097, LA3233, LA3376). Além disso, a distância entre o ATG e a extremidade do éxon imediatamente precedendo o íntron tra (assumindo emenda em forma de tipo F1) pode variar de 0bp a pelo menos +949bp sem conseqüências adversas para emenda (ver Tabela 4, coluna 3). Assim, é razoável assumir que esta distância pode ser até pelo menos 900 e preferivelmente até pelo menos 949 bp.

Mais informação nestes exemplos está sumarizada na Tabela 5. A opção preferida é usar nenhuma seqüência endógena para obter o controle de emenda alternativa correto de expressão (+0bp na tabela 4). Os requerentes preferiram inserir o íntron tra entre os dinucleotídeos de flaqueamento TG... GT na região de codificação da proteína de interesse para ser emendados alternativamente de modo a assegurar emenda correta como isto pode ser importante, No entanto os requerentes não irão se limitar a isto, se necessário, como outros nucleotídeos de flaqueamento podem funcionar corretamente também. Exemplos LA1038, LA3054 e LA3056 inclui alguma seqüência exônica de flaqueamento endógena do gene CCtra natural. Na tabela 5, se 6 nucleotídeos ou menos (incluindo o códon de partida ATG) estão incluídos de fusões particulares para 3' ou 5' da junção de emenda, para os propósitos de sumário desta tabela estes não serão considerados como fazendo parte da fusão. Tabela 4 pode estar correlacionada com a tabela 3 para encontrar qual íntron tra (Cctra, Bztra ou Crtra) é usado em cada exemplo. Novamente, LA1118 está apenas incluído para os fins de informação e está fora do escopo da presente invenção.

Tabela 5

Número construção (Figuras #.)	Íntron tra é fusionado a 5'	Íntron tra é fusionado a 3'	Seqüência tra exônica fusionada a 5' (bp)	Seqüência tra exônica fusionada a 3' (bp)
LA1188 (80)	Hsp70-tTAV	tTAV	+0bp	+0bp
LA3014 (29)	Hsp70-ubiquitina	ubiquitina- reaperKR-sv40	+0bp	+0bp
LA3166 (30)	Hsp70-ubiquitina	ubiquitina- reaperKR-sv40	+0bp	+0bp
LA3097 (27)	Hsp70	tTAV-K10	+0bp	+0bp

LA3077 (26)	Hsp70-tTAV	tTAV-K10	+0bp	+0bp
LA3233 (28)	Hsp70	tTAV2-K10	+0bp	+0bp
LA3376 (31)	Hsp70	tTAV2-K10	+0bp	+0bp
LA3376 (31)	Sry-a	tTAV3-sv40	+0bp	+0bp
LA3376 (31)	HB	reaperKR-sv40	+0bp	+0bp
LA3242 (32)	HB	reaperKR-sv40	+0bp	+0bp
LA1038 (14)	Hsp70-tra	Tra-Nipp1 (nipper)-sv40	+22bp	+20bp
LA3054 (61)	Opie2-nls- DsRed-tra	tra-ubiquitina- tTAV-sv40	+22bp	+20bp
LA3056 (62)	Opie2-nls- DsRed-tra	tra-ubiquitina- tTAV-sv40	+22bp	+242bp
LA3488 (63)	Iel-nls- TurboGreen-nls- ubiquitina	ubiquitina-nls- DsRed-nls- sv40	+0bp	+0bp
LA3596 (67)	Iel-nls- TurboGreen-nls- ubiquitina	ubiquitina-nls- DsRed-nls- sv40	+0bp	+0bp

Como mencionado acima quando um íntron é colocado 5' para uma região de codificação de proteína (ORF-X), prefere-se posicionar ou usar ubiquitina 3' para o íntron, 5' para ORF-X, assim, e provendo regulação específica de fêmea de ORF-X, enquanto introduzindo separação física entre esta sequência e o íntron tra, assim reduzindo a possibilidade de que sequências dentro de ORF-X irão interferir com a emenda do íntron tra.

Construções e sequências compostas também são previstas, por exemplo da forma:

X-ubi-Y

com o íntron alternativamente emendado inserido entre a região de codificação X e a região codificando ubiquitina (ubi), ou dentro da região de codificação de ubiquitina, ou entre a região codificando ubiquitina e região de codificação Y. Assim X será expressado sem restrição de emenda do íntron, enquanto Y será apenas expressado quando o íntron é emendado em uma forma apropriada. Mais configurações e arrumações deste tipo geral serão evidentes para o versado na técnica. Alguns exemplos dos mesmos são LA3014, LA3054, LA3056, LA3166, LA3488 e LA3596, sendo que todos usam fusões de ubiquitina, desta forma demonstrando a capacidade desta idéia ser aplicada com sucesso em mosca-do-mediterrâneo transgênica. Exemplos alternativos em mosquitos transgênicos incluem LA3604 e

LA3612, mostrando a extensa aplicabilidade filogenética deste sistema em não apenas espécies diferentes (mosquitos e mosca-do-mediterrâneo), mas também em diferentes contextos incluindo AaActin4, Aadxs e Cctra.

LA3596 (ver Fig 67 e SEQ ID NO. 145) é de projeto similar a LA3488, planejado para gerar fluorescência verde (por expressão de proteína fluorescente TurboGreen localizada nuclear) em ambos sexos, mas fluorescência vermelha apenas em fêmeas (por expressão de proteína fluorescente DsRed2 localizada nuclear). Isto é obtido pela fusão destas duas proteínas, dirigida por cassete de melhorador/promotor de Hr5-Iel, ligado junto com um ligador de 11 aminoácidos curto (ligador de SG4) e a região de codificação compreendendo ubiquitina (com uma mutação em pontos planejada para estabilizar a proteína resultante reduzindo sua propensão à degradação mediada por ubiquitina) e o íntron Cctra para limitar expressão de DsRed2 para fêmeas. Moscas-do-mediterrâneo transgênicas foram geradas com esta construção. Fluorescência vermelha foi limitada a fêmeas nesta linhagem como esperado, enquanto fluorescência verde foi observada em ambos, machos e fêmeas. Isto poderia ser usado para separação de sexos por triagem de fluorescência para uma particular proteína fluorescente, nesse caso fluorescência vermelha representando expressão de DsRed2.

Exemplo 17: Outra exemplificação de Cctra

Referência também é feita a LA3014 e LA3166 e dados fenotípicos dos mesmos em outros exemplos.

Os requerentes previamente realizaram e obtiveram transgênicos com o íntron Cctra em uma proteína funcional diferente de tTAV, ver LA3014 e LA3166. LA3014 contém uma fusão de ubiquitina - reaper^{KR} a jusante de um íntron Cctra. Dados fenotípicos mostram que mosca-do-mediterrâneo transgênica LA3014 concedeu letalidade específica de fêmea repressível. Análise de RT-PCR em RNA extraído de machos e fêmeas adultos criados sem tetraciclina, usando iniciadores e ReaperKR, demonstram que emenda correta estava ocorrendo em fêmeas (banda de 508 bp) e

nenhuma desta banda foi encontrada em machos (figura 37). LA3166 é outra construção com o íntron Cctra colocado dentro da região de codificação de ubiquitina fusionado a reaper^{KR}, mas colocado em uma posição diferente em ubiquitina. LA3166 também produz um efeito letal específico de fêmea repressível dominante em mosca-do-mediterrâneo.

LA1038 é um novo exemplo do uso do íntron Cctra em um contexto de seqüência diferente, aqui colocado em um fragmento de Nipp1Dm chamado 'nipper' que também emenda-se corretamente em mosca-do-mediterrâneo transgênica quando analisada por RT-PCR (figura 12). LA670 foi requerido como uma fonte de tTAV para dirigir a expressão de nipper alternativamente emendado.

Os requerentes também realizaram recentemente e obtiveram transgênicos com, construções baseadas em Cctra "somente íntron" com o íntron em um gene diferente (muitos dos exemplos acima, salvo especificado em contrário, estão em tTAV ou uma de suas variantes, isto é tTAV2 ou tTAV3). Estas construções funcionam como previsto. Isto é um resultado importante, deste modo mostrando que existem seqüências exônicas não essenciais em Cctra que os requerentes simplesmente duplicaram (em função, se não necessariamente em seqüência) por acaso, em tTAV. Também os requerentes realizaram construções de ubi - rpr^{KR} deste tipo (LA3014 e LA3166), que também validam o método de fusão de ubiquitina descrito acima. O método de fusão de ubiquitina é ainda exemplificado por análise RT-PCR de LA3054, LA3056 e LA3488 (figuras 11, 13, 14), como descrito no Exemplo 16, acima.

Exemplo 17: Outra exemplificação de Cctra

Referência também é feita a LA3014 e LA3166 e dados fenotípicos a partir dos mesmos em outros exemplos.

Os requerentes previamente realizaram e obtiveram transgênicos com, o íntron Cctra em uma proteína funcional diferente de tTAV, ver LA3014 e LA3166. LA3014 contém uma fusão de ubiquitina -

Reaper^{KR} a jusante de um íntron Ccra. Dados fenotípicos mostram que mosca-do-mediterrâneo transgênica de LA3014 concedeu letalidade específica de fêmea repressível. Análise de RT-PCR em RNA extraído de adultos machos e fêmeas criados sem tetraciclina, usando iniciadores e

5 ReaperKR, demonstram que a emenda correta estava ocorrendo em fêmeas (banda de 508bp) e nenhuma desta banda foi encontrada em machos (figura 37). LA3166 é outra construção com o íntron Ccra substituído dentro da região de codificação de ubiquitina fusionada ao Reaper^{KR}, mas substituído em uma posição diferente em ubiquitina. LA3166 também produz um efeito

10 letal específico de fêmea repressível dominante em mosca-do-mediterrâneo.

LA1038 é um novo exemplo do uso do íntron Ccra em um contexto de seqüência diferente, aqui substituído em um fragmento de Nipp1Dm chamado 'nipper' que também emenda corretamente e, mosca-do-mediterrâneo transgênica quando analisada por RT-PCR (figura 12). LA670 foi requerida como

15 uma fonte de tTAV para conduzir expressão nipper alternativamente emendada.

Os requerentes realizaram recentemente e obtiveram transgênicos com, construções baseadas em Ccra apenas de íntron com o íntron em um gene diferente (muitos dos exemplos acima, salvo aparente, ocorrem em tTAV ou uma das suas variantes, isto é tTAV2 ou tTAV3). Estas

20 construções funcionam como previsto. Isto é um resultado importante, deste modo mostrando que existem seqüências exônicas não essenciais em Ccra que os requerentes simplesmente duplicaram (em função, se não necessariamente em seqüência) aleatoriamente, em tTAV. Também os requerentes notaram construções de ubi - rpr^{KR} deste tipo (LA3014 e LA3166), que também validam

25 o método de fusão de ubiquitina descrito acima. O método de fusão de ubiquitina é ainda exemplificado por análise de RT-PCR de LA3054, LA3056 e LA3488 (figuras 11, 13, 14), e como descrito no Exemplo 16, acima.

Figura 11: Gel mostrando emenda específica de sexo de íntron(s) derivados de Ccra (banda de 780bp em fêmeas) em *Ceratitis*

capitata transformadas com LA3488. Emenda na forma de F1 poderia dar um produto de aproximadamente 780bp. Uma banda deste tamanho é claramente visível de fêmeas (trilha 4), mas não de machos, nem em trilhas com reações das quais a enzima de transcriptase reversa foi omitida ("sem RT"). Por esta razão, o íntron derivado de Cctra é capaz de emenda alternativa específica de sexo nesse contexto de seqüência novo. Trilha 1: Marcador (SmartLadder™ de Eurogentec, bandas de aproximadamente 0,8, 1,0 e 1,5kb são indicadas); Trilhas 2 e 3: LA3488 de *Ceratitidis capitata* /+ machos (controle de RT e sem controle RT, respectivamente); Trilhas 4 e 5: LA3488 de *Ceratitidis capitata* /+ fêmeas (controle de RT e sem controle de RT, respectivamente).

Figura 12: Gel mostrando emenda específica de sexo de íntron(s) derivados de Cctra em *Ceratitidis capitata* transformada com LA1038. Emenda na forma de F1 poderia dar um produto de aproximadamente 230bp. Uma banda deste tamanho é claramente visível de fêmeas (trilhas 1, 2, 7, 8, 9 e 10), mas não de machos. Por esta razão, o íntron derivado de Cctra é capaz de emenda alternativa específica de sexo nesse contexto de seqüência novo. Trilha 15: Marcador (SmartLadder™ de Eurogentec, bandas de aproximadamente 0,2, 0,4 e 0,6kb são indicadas); Trilhas 1, 2, 7, 8, 9 e 10: LA670 *Ceratitidis capitata*; fêmeas de LA1038; Trilhas 3, 4, 5, 6, 11, 12, 13 e 14: LA670 de *Ceratitidis capitata*; machos de LA1038.

Figura 13: Gel mostrando emenda específica de sexo de íntron(s) derivados de Cctra em *Ceratitidis capitata* transformada com LA3054. Emenda na forma de F1 poderia dar um produto de aproximadamente 340 bp. Uma banda deste tamanho é claramente visível na trilha 7, mas não de machos. Por esta razão, o íntron derivado de Cctra é capaz de emenda alternativa específica de sexo nesse contexto de seqüência novo. Trilha 1 : Marcador (SmartLadder™ de Eurogentec, bandas de aproximadamente 0,4, 0,6, 0,8 e 1,0kb são indicadas); Trilhas 2-5: machos de LA3054 de *Ceratitidis capitata*; Trilha 7: fêmea de LA3054 *Ceratitidis capitata*.

Figura 14: Gel mostrando emenda específica de sexo de íntron(s) derivados de Cctra e, *Ceratitis capitata* transformada com LA3056. Emenda na forma de F1 poderia dar um produto de aproximadamente 200 bp. Uma banda deste tamanho é claramente visível de uma fêmea (trilha 6), mas não de machos (trilhas 2-4). Por esta razão, o íntron derivado de Cctra é capaz de emenda alternativa específica de sexo nesse contexto de seqüência novo. Trilha 1: Marcador (SmartLadder™ de Eurogentec, bandas de aproximadamente 0,2, 0,4, 0,6 e 0,8kb são indicadas); Trilhas 2-5: LA3056 de *Ceratitis capitata* /+ machos; Trilhas 6-7: LA3056 *Ceratitis capitata* /+ fêmeas.

Figura 15: Gel mostrando emenda específica de sexo de íntron(s) derivados de BzTra em *Anastrepha ludens* transformada com LA3376. Emenda na forma F1 poderia dar um produto de aproximadamente 672 bp. Uma banda deste tamanho é claramente visível de fêmeas (trilha 4), mas não de machos, nem em trilhas com reações de que enzima de transcriptase reversa foi omitida ("sem RT"), iniciadores usados foram SRY e AV3F. Por esta razão, o íntron derivados de Bztra é capaz de emenda alternativa específica de sexo nesse contexto de seqüência novo e espécie. Trilha 1: Marcador (SmartLadder™ de Eurogentec, bandas de aproximadamente 0,6, 0,8, e 1,0kb são indicadas); Trilhas 2 e 3: LA3376 *Anastrepha ludens* /+ machos (controle de RT e sem controle RT, respectivamente); Trilhas 4 e 5: LA3376 de *Anastrepha ludens* /+ fêmeas (controle de RT e sem controle de RT, respectivamente).

Figura 18 e SE ID NOs 149 e 150 mostra seqüências de minigene1 de DSX, minigene2 de DSX e mapa de plasmídeo de LA3619.

Figuras 19-51 são como por Exemplos 1-9 acima. Figuras 52-58, 68 e 69 mostram vários diagramas e seqüências de plasmídeo. Figuras 59-60 estão descritas acima e Figuras 61-66 mostram vários diagramas e seqüências de plasmídeo. Fig 67 é pLA3596, como discutido em outra parte.

REFERÊNCIAS

- Allen ML, Christensen BM. Related 2004 Flight muscle-specific expression of act88F: GFP in transgenic *Culex quinquefasciatus* Say (Diptera: Culicidae). *Parasitol Int.* 55 (4): 307- 14.
- 5 Bennett D5 Szoor B, Gross S, VereshchaginaN, Alphey L. 2003 Ectopic expression of inhibitors of protein phosphatase type 1 (PPI) can be used to analyze roles of PPI in *Drosophila* development. *Genetics.* 164(1) :235-45.
- Black, D. (2003). Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem* 72, 291-336.
- 10 Burset, M., Seledtsov, L, and Solovyev, V. (2001). SpliceDB: database of canonical and non- canonical de emenda sites in mammalian genomes. *Nucleic Acids Research* 2P₅ 255-259.
- Caceres JF, Kornblihtt AR. 2002 Alternative splicing: multiple control mechanisms and involvement in human disease. *Trends Genet.* 18(4): 15 186-93.
- Cande C, Cecconi F, Dessen P, Kroemer G. 2002 Apoptosis-inducing factor (AIF): key to the conserved caspase-independent pathways of cell death? *J Cell Sci.* 115(24):4727-34.
- Cartegni, L., Chew, S., and Krainer, A. (2002). Listening to 20 silence and understanding nonsense: éxonic mutations that affect splicing. *Nature Reviews Genetics* 3, 285-298.
- Clark, F., and Thanaraj, T. (2002). Categorization and characterization of transcript-confirmed constitutively and alternatively spliced introns and exóns from human. *Human Molecular Genetics* 11, 451-464.
- 25 Funaguma, S., Suzuki, M., Tamura, T., and Shimada, T. (2005). The *Bmdsx* transgene including trimmed introns is sex-specifically spliced in tissues of the silkworm, *Bombyx mori*. *J Insect Sci* 5, 17.
- George, EX., Ober, M.B. and Emerson Jr, CP. (1989). Functional domains of the *Drosophila melanogaster* muscle myosin heavy-

chain gene are encoded by alternatively spliced exons. *Mol Cell Biol.* P:2957-2974.

Graveley BR. 2001 Alternative splicing: increasing diversity in the proteomic world. *Trends Genet.* 17(2): 100-7.

5 Hammes, A., Guo, J.K., Lutsch, G., Leheste, J.R., Landrock, D., Zeigler, U., Gubler, M.C. and Schedl, A. (2001). Two de emenda variants of the Wilms' Tumour 1 gene have distinct functions during sex determination and nephron formation. *Cell* 106:319-329.

10 Hastings, G.A. and Emerson Jr, CP (1991). Myosin functional domains encoded by alternative exons are expressed in specific thoracic muscles of *Drosophila*. *J. Cell Biol.* 114: 263-276.

Hedley, MX. and Maniatis (1991). Sex-specific splicing and polyadenylation of *dsx* pre-mRNA requires a sequence that binds specifically to a *tra-2* protein in vivo. *Cell* 65:579-586.

15 Heinrich J. C. and Scott MJ. 2000 A repressible female-specific lethal genetic system for making transgenic insect strains suitable for a sterile-release program *PNAS* 97 (15): 8229-8232

Horn C, Wimmer EA. 2003 A transgene-based, embryo-specific lethality system for insect pest management. *Nat Biotechnol.* 21(1): 64-70.

20 Hoshijima, K.K, Inoue, L., Higuchi, L, Sakamoto, H. and Shimura, Y. (1991). Control of *doublesex* alternative splicing by *transformer* and *transformer-2* in *Drosophila*. *Science* 252:833-836.

25 Huang, Q., Deveraux, Q.L., Maeda, S., Salvesen, G.S., Stennicke, H.R., Hammock, B.D. e Reed, J.C. (2002). Evolutionary conservation of apoptosis mechanisms: Lepidopteran and baculoviral inhibitor of apoptosis proteins are inhibitor of mammalian caspase-9. *Agricultural Sciences* 97(4): 1427-1432.

Ito, Y., Hirochicka, H. and Kurata, N. (2002). Organ-specific alternative transcripts of KNOX family class 2 homeobox genes of rice. *Gene* 288:41-47.

Johnson JM, Castle J, Garrett-Engele P, Kan Z, Loerch PM, Armour CD, Santos R, Schadt EE, Stoughton R, Shoemaker DD. 2003 Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science*. 302(5653):2141-4.

5 Jurica MS, Moore MJ. 2003 Pre-mRNA splicing: awash in a sea of proteins. *Mol Cell*. 12(1):5- 14.

Kazzaz JA, Rozek CE. 1989 Tissue-specific expression of the alternately processed *Drosophila* myosin heavy-chain messenger RNAs. *Dev Biol*. 133 (2): 550-61.

10 Maniatis, T., e Tasic, B. (2002). Alternative pre-mRNA splicing and proteome expansion in metazoans. *Nature* 418, 236-243.

Munoz, D., Jimenez, A., Marinotti, O, and James, A. (2004). The AeAct-4 gene is expressed in the developing flight muscles of females *Aedes aegypti*. *Insect Molecular Biology* 13, 563-568.

15 Nishiyama, R., Mizuno, H., Okada, S., Yamaguchi, T., Takenaka, M., Fukuzawa, H. e Ohyama, K. (1999). Two mRNA species encoding calcium-dependent protein kinases are differentially expressed in sexual organs of *Marchantia polymorpha* through alternative splicing. *Plant Cell Physiol*. 40(2):205-212.

20 Nishiyama, R., Yamato, K.T., Miura, K., Sakida, M., Okada, S., Kono, K., Takahama, M., Sone, T., Takenaka, M., Fukuzawa, H. e Ohyama, K. (2000). Comparison of expressed sequence tags from male and female sexual organs of *Marchantia polymorpha*. *DNA Res*. 7:165-174.

25 Olson, M.R., Holley, C.L., Ji Yoo, S., Huh, J.R, Hay, B.A. and Kornbluth, S. (2003). Reaper is regulated by IAP-mediated Ubiquitination. *J.Biol.Chem.*, 278(15):4028-4034.

Olson, M.R., Holley, C.L., Gan, E.C., Colon-Ramos, D.A., Kaplan, B. and Kornbluth, S. (2003). A GH3-like domain in reaper is required for mitochondrial localization and induction of IAP degradation. *J. Biol.*

Chem. 218(45). -44758-44768.

Pan, Q., Shai, O., Misquitta, C, Zhang, W., Saltzman, A., Mohammad, N., Babak, T., Siu, H., Hughes, T., Morris, Q., et al. (2004). Revealing global regulatory features of mammalian alternative splicing using a quantitative microarray platform. Mol Cell 16, 929-941.

Pane, A., Salvemini, M., Delli Bovi, P., Polito, C, and Saccone, G. (2002). The *transformer* gene in *Ceratitis capitata* provides a genetic basis for selecting and remembering the sexual fate. Development 129, 3715-3725.

Park, J., Parisky, K., Celotto, A., Reenan, R., and Graveley, B. (2004). Identification of alternative splicing regulators by RNA interference in *Drosophila*. Proc Nat'l Acad Sci (USA) 101, 15974-15979.

Parker L, Gross S, Beullens M, Bollen M, Bennett D, Alphey L. 2002 Functional interaction between nuclear inhibitor of protein phosphatase type 1 (NIPPI) and protein phosphatase type 1 (PPI) in *Drosophila*: consequences of over-expression of NIPPI in flies and suppression by co- expression of PPI. Biochem J. 368 (3) :1 '89-97.

Raphael, K.A., Whyard, S., Shearman, D., An, X. e Frommer, M. (2004). *Bactrocera tyroni* and closely related pest-tephritids-molecular analysis and prospects for transgenic control strategies. Insect Biochem. Mol. Biol. 34:167-176.

Ryner, L. and Baker, B. S. (1991). Regulation of *doublesex* pre-mRNA processing occurs by 3'- de emenda site activation. Genes Dev. 5:2071-2085.

Saccone, G., Pane, A., e Polito, C. (2002). Sex determination in flies, fruit flies and butterflies. Genetica 116, 15-23.

Scali, C, Catteruccia, F., Li, Q., e Crisanti, A. (2005). Identification of sex-specific transcripts of the *Anopheles gambiae doublesex* gene. J Exp Biol 208, 3701-3709.

Scott, M., Heinrich, J., and Li, X. (2004). Progress towards the development of a transgenic strain of the Australian sheep blowfly (*Lucilia cuprina*) suitable for a male-only sterile release program. *Insect Biochem Mol Biol* 34, 185-192.

5 Seo, S-J., Cheon, H-M., Sun, J., Sappington, T. W. and Raikhel, A.S. (2003). Tissue- and stage- specific expression of two lipophorin receptor variants with seven and eight ligand-binding repeats in the adult mosquito. *J. Biol. Chem.* 275(41):41954-41962.

10 Siebel CW, Fresco LD, Rio DC. 1992 The mechanism of somatic inhibition of *Drosophila* P- element pre-mRNA splicing: multiprotein complexes at an exon pseudo-5' splice site control U1 snRNP binding. *Genes Dev.* (5):1386-401.

15 Shivikrupa, Singh., R e Swarup, G. (1999). Identification of a novel splice variant of C3G which shows tissue-specific expression. *DNA Cell Biol.* 18(5): 701-708.

 Smith, C, and Valcarcel, J. (2000). Alternative pre-mRNA splicing: the logic of combinatorial control. *Trends Biochem Sci* 25, 381-388.

20 Stoss, O., Stoilov, P., Hartmann, A. M., Naylor, O., and Stamm, S. (1999). The in vivo minigene approach to analyze tissue-specific splicing. *Brain Research Protocols* 4, 383-394.

 Stoss, O., Olbrich, M, Hartmann, A. M., Konig, H., Memmott, J., Andreadis, A and Stamm, S. (2001). The STAR/GSG family protein rSLM-2 regulates the selection of alternative splice sites. *J. Biol. Chem.* 276(12):8665-8673.

25 Streuli, M. and Saito, H. (1989). Regulation of tissue-specific alternative splicing: exon-specific cis-elements govern the splicing of leukocyte common antigen pre-mRNA. *EMBO J.* 8(3): 787-796.

 Suzuki, M., Ohbayashi, F., Mita, K., and Shimada, T. (2001). The mechanism of sex-specific splicing at the *doublesex* gene is different

between *Drosophila melanogaster* and *Bombyx mori*. Insect Biochem Mol Biol 31, 1201-1211.

Thanaraj, T., and Clark, F. (2001). Human GC-AG alternative intron isoforms with weak donor sites show enhanced consensus at acceptor
5 exon positions. Nucleic Acids Research 29, 2581-2593.

Thanaraj, T., Stamm, S., Clark, F., Reithoven, J., Le Texier, V., and Muilu, J. (2004). ASD: the Alternative Splicing Database. Nucleic Acids Research 32, D64-D69.

Varshavsky, A. (2000). Ubiquitin fusion technique and its
10 descendants. Meth Enz 327.

Venables, J. (2002). Alternative splicing in the testes. Curr Opin Genet Dev 12, 615-619.

Venables JP. 2004 Aberrant and alternative splicing in cancer. Cancer Res. 64(21): 7647-54.

15 Vernooy, S.Y., Copeland, J., Ghaboosi, N., Griffin, E.E., Yoo, S.J. and Hay, B.A. (2000). J. Cell Biol. 150(2):F69-¥75.

White, K., Tahoaglu, E. and Steller, H. (1996). Cell killing by the *Drosophila* gene reaper. Science 271 (5250): 805-807.

Wing, J.P., Zhou, L., Schwartz, L.M. e Nambu, J.R. (2001)
20 Distinct cell killing properties of the *Drosophila* reaper, head involution defective, and grim genes. Cell Death Diffn 5(11): 930-939

YaLi Chiu A., e Pin Ouyang, A.B.,(2006).Loss of Pnn expression attenuates expression levels of SR family splicing factors and modulates alternative pre-mRNA splicing in vivo. Bioch. Biophys.Res.
25 Comm. W:663-671.

Yoshimura, K., Yabuta, Y., Ishikawa, T. e Shigeoka, S. (2002). Identification of a cis element for tissue-specific alternative splicing of chloroplast Ascorbate Peroxidase pre-mRNA in higher plants. J. Biol.Chem 277 (43,1:40623-40632.

ANOTAÇÕES DE SEQUÊNCIA

O seguinte refere-se aos vários plasmídeos do presente e enfatiza a posição de certos elementos aqui preferidos.

<223> Sequência de pLA3359 (SEQ ID NO. 47).

5 <***> Aspectos chaves incluem:

1. Mini-gene de *Anopheles gambiae* dsx (Agdsx), [um mini-gene é uma sequência recombinante derivado de um gene particular (o gene Agdsx neste exemplo) por ligação junta de segmentos não contíguos enquanto retendo a ordem 5'-3' original; isto é equivalente para deleção de alguns
10 segmentos internos de um fragmento mais longo de sequência genômica derivado do gene], (1-3135): incluindo parte de Agdsx de éxon3, éxon 4a (fêmea), éxon 4b (fêmea) e parte do éxon5 (macho e fêmea).

<***> éxons derivados de Agdsx de posições 426 a 560 (parte do éxon 3); 1068 a 2755 (incluindo parte do éxon 4, encontrado em fêmeas);
15 1809 a 2755 (incluindo parte do éxon 4, encontrado em fêmeas); e 2914 a 3135 (incluindo parte do éxon 5, encontrado em machos).

<***> Alternativamente transcrito emendado começa no segmento derivado de baculovírus AcMNPV Ie1 (prematuro imediato 1) em posição ~8031 (fragmento Ie1 é da posição 7431 a 8060).

20 <***> Aspecto incluído:

1. íntron adicional derivado de gene "scraps" de *Drosophila* ("íntron 'scraps'") a montante para sequência Agdsx de posição 8075 a 8137.

<223> Sequência de pLA3433 (SEQ ID NO. 48).

<***> Aspectos chaves incluem:

25 1. Mini-gene de Agdsx (778-4623): incluindo parte de Agdsx do éxon 2, éxon3, éxon 4a (fêmea), éxon 4b (fêmea) e parte do éxon5 (macho e fêmea).

<***> éxons derivados de Agdsx de posição 778 a 908 (parte do éxon 2); 1913 a 2048 (parte do éxon 3); 2556 a 2642 (parte do éxon 4a);

3297 a 4243 (parte do éxon 4b) e 4402 a 4623 (parte do éxon 5).

<***> Alternativamente transcrito emendado começa no segmento derivado de baculovírus AcMNPV Ie1 (premature imediato 1) em posição ~606 (fragmento Ie1 é da posição 6 a 635).

5 <***> Aspecto incluído:

1. íntron adicional derivado de gene "scraps" de *Drosophila*("íntron 'scraps'") a montante para seqüência Agdsx de posição 650 a 712.

<223> Seqüência de pLA3491.

10 <***> Aspectos chaves incluem:

1. Mini-gene de *Aedes aegypti* dsx (Aadsx): incluindo parte de Aadsx éxon 4, éxon 5a (fêmea), éxon 5b (fêmea), e parte do éxon6 (macho e fêmea).

15 <***> éxons derivados de Aadsx de posição 1316 a 1450 (parte do éxon 4); 2626 a 3761 (parte do éxon 5a); 3293 a 3761 (parte do éxon 5b); e 5215 a 5704 (parte do éxon 6).

<***> Parte do transcrito F1 é previsto como compreendendo nucleotídeos ~1174-1450, 2626-3761, 5215~5850.

20 <***> Parte do transcrito F2 é previsto como compreendendo nucleotídeos ~1174-1450, 3293-3761, 5215--5850.

<***> Parte do transcrito F3 é previsto como compreendendo nucleotídeos ~1174-1450, 2626-3083, 3293-3761, 5215--5850.

<***> Parte do transcrito M1 é previsto como compreendendo nucleotídeos ~1174-1450, 5215--5850.

25 <***> O transcrito alternativamente emendado começa em segmento derivado de baculovírus AcMNPV Ie1 (premature imediato 1) em posição ~1174 (fragmento Ie1 é de posição 574 a 1203).

<***> Aspecto incluído:

1. íntron adicional derivado de gene "scraps" de *Drosophila*

(“íntron 'scraps’”) a montante para sequência Aadsx de posição 1218 a 1280.

<223> Sequência de pLA3646.

<***> Aspectos chaves incluem:

1. Mini-gene de Aadsx (17218-11707): incluindo parte de
5 Aadsx éxon 4 de posição 17113 a 16979, éxon 5a de posição 15803 a 15025 +
14010 a 13650, éxon 5b de posição 15136 a 15025 + 14010 a 13650 e éxon 6
de posição 12196 a 11707 (nota: orientação reversa).

<***> parte do éxon 4 contém mutações de 4 pontos com
relação ao tipo selvagem em posições 17087 (ATG-ACG), 17053 (ATG-
10 ACG), 17050 (ATG-ACG) e 17041 (ATG-ACG) (nota: orientação reversa);
parte do éxon 5a e 5b contém mutações de 3 pontos com relação ao tipo
selvagem em posições 15129 (ATG-ATA), 15116 (ATG-ATA) e 15113
(ATG-ATA) (nota: orientação reversa). Todas destas mutações são para
eliminar sequências de ATG.

15 <***> tTAV2 é inserido nos éxons de sobreposição 5a e 5b da
posição 15024 a 14011 (nota: orientação reversa).

<***> O transcrito alternativamente emendado começa em
fragmento derivado de hsp70 de posição ~17312 (fragmento hsp70 é de
posição 17354 a 17225); (nota: orientação reversa).

20 <***> Aspecto incluído:

1. íntron adicional derivado de gene "scraps" de *Drosophila*
(“íntron 'scraps’”) a montante para sequência Aadsx de posição 1107 a 1045
(nota: orientação reversa)

Sequência de pLA3435 (SEQ ID NO. 46).

25 <223> Aspectos chaves incluem:

1. Mini-gene de dsx *Bombyx mori* (Bmdsx) (1411-3161) com
um ligador exógeno entre éxons 3 e 4 de fêmea fusionados.

<***> Fragmento de éxons dois compartilhados (1411bp-
1554bp)

<***> Parte de éxon três específico de fêmea (2121bp-2202) fusionado para parte de éxon 4 específico de fêmea (2225bp-2290bp) usando um ligador exógeno (2203bp-2224bp)

5 <***> Fragmento de éxon cinco compartilhado (3007bp-3161bp)

<***> Produto de emenda de minigene dsx fêmea é codificado por 1411-1554 + 2121-2290 + 3007-3161.

<***> Produto de emenda de minigene dsx macho é codificado por 1411-1554 + 3007-3161.

10 <***> Transcrição é prevista para começar em aproximadamente posição ~1239 dentro do segmento derivado de baculovírus AcMNPV Ie1 (premature imediato 1) promotor (639bp-1268bp).

<223> Seqüência de pLA3534.

<***> Aspectos chaves incluem:

15 1. Mini-gene de Aadsx (6996-4425): contendo Aadsx éxon 4, parte de éxon 5a (fêmea) e parte de éxon 5b (fêmea), inclusive de fragmentos de íntron Aadsx.

20 <***> éxons derivados de Aadsx de posição 6968 a 6834 (parte de éxon 4), 5462 a 4425 (parte de éxon 5a) e 4795 a 4425 (parte de éxon 5b); (nota: orientação reversa).

<***> Parte do transcrito F1 é prevista como compreendendo nucleotídeos ~7146-6834, 5462-~4300 (nota: orientação reversa).

<***> Parte do transcrito F2 é prevista como compreendendo nucleotídeos ~7146-6834, 4795-~4300 (nota: orientação reversa).

25 <***> Parte do transcrito F3 é prevista como compreendendo nucleotídeos ~7146-6834, 5462-5005, 4795-~4300 (nota: orientação reversa).

<***> O transcrito alternativamente emendado começa em segmento derivado de baculovírus AcMNPV Ie1 (premature imediato 1) em posição ~7146 (fragmento Ie1 é de posição 7746 a 7117, orientação reversa).

<223> Seqüência de pLA3612.

<***> Aspectos chaves incluem:

1. Região de codificação de ubiquitina-tTAV2 inserida em um éxon fêmea de gene Aadsx.

5 <***> Ubiquitina-tTAV2 é de posição 15185-16429 em Aadsx (ubiquitina é de 15185-15412; tTAV2 é de 15413-16429), inclusive de códon de partida e de parada.

 <***> Seqüência derivada de Aadsx: 13150-15184, 16438-18805.

10 <***> Transcrito alternativamente emendado Aadsx-ubiquitina-tTAV2 começa em segmento derivado de hsp70 (fragmento hsp70 é de 13014-13143).

<223> Seqüência de pLA3619.

<***> Aspectos chaves incluem:

15 1. Região de codificação de tTAV2 inserida em um éxon fêmea de gene Aadsx.

 <***> Seqüência derivada de Aadsx: 5635-3641, 2610-243 (nota: orientação reversa).

20 <***> Transcrito alternativamente emendado Aadsx-tTAV2 começa em segmento derivado de hsp70 de 5642-5771 (nota: orientação reversa).

 <***> O transcrito tTAV2 é previsto como sendo traduzido entre 2619-3635, inclusive de códon de partida e de parada (nota: orientação reversa).

25 <223> Seqüência de pLA3545.

<***> Aspectos chaves incluem:

1. Promotor AaActin4 e 5' UTR incluindo primeiro expressão de tTAV íntron regula.

<***> Seqüência derivada de AaActin4 é de posição 923-

4285.

<***> O transcrito alternativamente emendado é previsto para começar em aproximadamente ~2366.

5 <***> O primeiro íntron de AaActin4 (variante de emenda fêmea) é de 2458-4259.

<***> tTAV é previsto como sendo traduzido entre 4300-5316, inclusive de códon de partida e de parada.

<223> Seqüência de pLA3604.

<***> Aspectos chaves incluem:

10 1. Promotor AaActin4 e 5' UTR regula expressão ubiquitina-tTAV2.

<***> Seqüência derivada de AaActin4 é de posição 5795-2407 (nota: orientação reversa).

15 <***> O transcrito alternativamente emendado é previsto para começar em aproximadamente ~4353 (nota: orientação reversa).

<***> O primeiro íntron de AaActin4 (variante de emenda fêmea) é de 2455-4254 (nota: orientação reversa).

20 <***> O transcrito ubiquitina-tTAV2 é previsto como sendo traduzido a partir do códon de partida engenheirado no primeiro éxon do gene AaAct4 a 4299-4297 (ubiquitina é de 2406-2179; tTAV2 é de 2178-1162); (nota: orientação reversa).

<223> Seqüência de pLA3641.

<***> Aspectos chaves incluem:

25 1. Região de codificação de tTAV inserida em um éxon fêmea de gene CodlingDsx.

<***> tTAV é de posição 2731-3747 no gene CodlingDsx.

<***> Transcrito alternativamente emendado Dsx-tTAV começa em segmento derivado de hsp70 (fragmento hsp70 é de 4811-4940).

<***> O transcrito tTAV é previsto como sendo traduzido

entre 2731-3747, inclusive do códon de partida e de parada (nota: orientação reversa).

<223> Seqüência de pLA3570

<***> Aspectos chaves incluem:

- 5 1. Região de codificação de tTAV inserida em um éxon fêmea de gene PBW-Dsx.

<***> Região de codificação de tTAV é de 2336-3352.

<***> Transcrito alternativamente emendado Dsx-tTAV começa em segmento derivado de hsp70 (fragmento hsp70 é de 4683-4812).

- 10 <***> O transcrito tTAV é previsto como sendo traduzido entre 2336-3352, inclusive do códon de partida e de parada (nota: orientação reversa).

<223> Seqüência de pLA1188 (SEQ ID NO. 49)

<***> Aspectos chaves incluem:

- 15 1. Região de codificação de tTAV com íntron Cctra inserido.

<***> Íntron Cctra é de posição 3905-2561 em tTAV (nota: orientação reversa).

<***> Transcrito alternativamente emendado tTAV começa em segmento derivado de hsp70 de posição 4217 (fragmento hsp70 é de 4260-4131); (nota: orientação reversa).

<***> O transcrito tTAV F1 é previsto como sendo traduzido entre 4040-1679 (nota: orientação reversa).

<***> Aspecto incluído:

- 25 1. Íntron Adh dentro do transcrito F1 previsto de posição 4118-4049 (nota: orientação reversa).

<223> Seqüência de pLA3077 (SEQ ID NO. 50).

<***> Aspectos chaves incluem:

1. Região de codificação de tTAV com íntron Cctra inserido.

<***> Íntron Cctra é de posição 3975-2631 em tTAV (nota:

orientação reversa).

<***> Transcrito alternativamente emendado tTAV começa em segmento derivado de hsp70 de posição ~4217 (fragmento hsp70 é de 4260-4131); (nota: orientação reversa).

5 <***> O transcrito tTAV F1 é previsto como sendo traduzido entre 4039-1678, inclusive de códon de partida e de parada (nota: orientação reversa).

<***> Aspecto incluído:

1. Íntron Adh dentro do transcrito F1 previsto de posição
10 4117-4048 (nota: orientação reversa).

<223> Seqüência de pLA3097 (SEQ ID NO. 51).

<***> Aspectos chaves incluem:

1. Região de codificação de tTAV com íntron Cctra inserido.

<***> Íntron Cctra é de posição 3282-1938 em tTAV (nota:
15 orientação reversa).

<***> Transcrito alternativamente emendado tTAV começa em segmento derivado de hsp70 de posição ~3382 (fragmento hsp70 é de 3425-3296); (nota: orientação reversa).

<***> O transcrito tTAV F1 é previsto como sendo traduzido
20 entre 3285-924, inclusive de códon de partida e de parada (nota: orientação reversa).

<223> Seqüência de pLA3233 (SEQ ID NO. 52).

<***> Aspectos chaves incluem:

1. Região de codificação de tTAV2 com íntron Cctra inserido.

<***> Íntron Cctra é de posição 3289-1945 em tTAV2 (nota:
25 orientação reversa).

<***> Transcrito alternativamente emendado tTAV2 começa em segmento derivado de hsp70 de posição ~3389 (fragmento hsp70 é de 3432-3303); (nota: orientação reversa).

<***> O transcrito tTAV2 F1 é previsto como sendo traduzido entre 3292-931, inclusive de códon de partida e de parada (nota: orientação reversa).

<223> Seqüência de pLA3014 (SEQ ID NO. 53).

5 <***> Aspectos chaves incluem:

1. Região de codificação de ubi-reaper[KR] com íntron Cctra inserido.

<***> Íntron Cctra é de posição 3356-4700 em ubi-reaper[KR].

10 <***> Transcrito alternativamente emendado ubi-reaper[KR] começa em segmento derivado de hsp70 de posição ~3234 (fragmento hsp70 é de 3191-3320).

15 <***> O transcrito ubi-reaper[KR] F1 é previsto como sendo traduzido entre 3331-5143, inclusive de códon de partida e de parada (ubiquitina é de 3331-3355, 4701-4948; reaper[KR] é de 4949-5143).

<223> Seqüência de pLA3166 (SEQ ID NO. 54).

<***> Aspectos chaves incluem:

1. Região de codificação de ubi-reaper[KR] com íntron Cctra inserido.

20 <***> Íntron Cctra é de posição 9987-8643 em ubi-reaper[KR] (nota: orientação reversa).

<***> Transcrito alternativamente emendado ubi-reaper[KR] começa em segmento derivado de hsp70 de posição ~10227 (fragmento hsp70 é de 10270-10141); (nota: orientação reversa).

25 <***> O transcrito ubi-reaper[KR] F1 é previsto como sendo traduzido entre 10126-8359, inclusive de códon de partida e de parada (ubiquitina é de 10126-9988, 8642-8554; reaper[KR] é de 8553-8359); (nota: orientação reversa).

<223> Seqüência de pLA3376 (SEQ ID NO. 55).

<***> Aspectos chaves incluem:

1. Região de codificação de tTAV2 com íntron Cctra inserido.
2. Região de codificação de tTAV3 com íntron Bztra inserido.
3. Região de codificação de reaper[KR] com íntron Bztra

5 inserido.

<***> Íntron Cctra é de posição 3289-1945 em tTAV2 (nota: orientação reversa).

<***> Íntron Bztra é de posição 5981-5014 em tTAV3 (nota: orientação reversa).

10 <***> Íntron Bztra é de posição 16391-17358 em reaper[KR].

<***> Transcrito alternativamente emendado tTAV2 começa em segmento derivado de hsp70 de posição ~3389 (fragmento hsp70 é de 3432-3303); (nota: orientação reversa).

15 <***> Transcrito alternativamente emendado tTAV3 começa em segmento derivado de sry-alfa de posição ~6019 (fragmento sry-alfa é de 6243-5999); (nota: orientação reversa).

<***> Transcrito alternativamente emendado reaper[KR] começa em segmento derivado de hunchback de posição ~16339 (fragmento hunchback é de 16289-16372).

20 <***> O transcrito tTAV2 F1 é previsto como sendo traduzido entre 3292-931, inclusive de códon de partida e de parada (nota: orientação reversa).

25 <***> O transcrito tTAV3 F1 é previsto como sendo traduzido entre 5984-4006, inclusive de códon de partida e de parada (nota: orientação reversa).

<***> O transcrito reaper[KR] é previsto como sendo traduzido entre 16385-17550, inclusive de códon de partida e de parada.

<223> Seqüência de pLA3242 (SEQ ID NO. 56).

<***> Aspectos chaves incluem:

1. Região de codificação de tTAV com íntron Cctra inserido.

<***> Região de codificação de reaper[KR] com íntron Crtra inserido.

5 <***> Íntron Cctra é de posição 3282-1938 em tTAV (nota: orientação reversa).

<***> Íntron Crtra é de posição 5488-4180 em reaperKR (nota: orientação reversa).

10 <***> Transcrito alternativamente emendado reaperKR começa em segmento derivado de hunchback de posição ~5540 (fragmento hunchback é de 5590-5507); (nota: orientação reversa).

<***> Transcrito alternativamente emendado tTAV começa em segmento derivado de hsp70 de posição ~3382 (fragmento hsp70 é de 3425-3296); (nota: orientação reversa).

15 <***> O transcrito reaperKR F1 é previsto como sendo principalmente traduzido entre 4088-5494, inclusive de códon de partida e de parada (nota: orientação reversa).

<***> O transcrito tTAV F1 é previsto como sendo principalmente traduzido entre 924-3285, inclusive de códon de partida e de parada (nota: orientação reversa).

20 <223> Seqüência de pLA1172 (SEQ ID NO. 106).

<***> Aspectos chaves incluem:

1. Região de codificação de tTAV entre fragmentos derivados de AaActin4.

25 <***> Fragmentos derivados de AaActin4 são de 7868-11257 e 12366-13100.

<***> O transcrito tTAV é previsto como sendo traduzido entre 11342-12358, inclusive de códon de partida e de parada.

<***> O transcrito AaActin4-tTAV é previstos para começar na posição ~9312.

<***> AaActin4 contém um íntron (variante de emenda tipo fêmea) de posição 9403-11204.

<223> Seqüência de pLA1038 (Fig. 12).

<***> Aspectos chaves incluem:

- 5 1. Fragmento de região de codificação de Nipp1Dm ('nipper') com íntron Cctra inserido com seqüência exônica tra de flanqueamento.

<***> Íntron Cctra é de posição 3365-4709 em nipper.

<***> Íntron Cctra é flanqueado por seqüência exônica Cctra de posições 3343-3364 e 4710-4729.

- 10 <***> Transcrito alternativamente emendado nipper começa em segmento derivado de hsp70 de posição ~3243 (fragmento hsp70 é de 3200-3329).

<***> O transcrito nipper F1 é previsto como sendo traduzido entre 3340-5014, inclusive de códon de partida e de parada.

- 15 <223> Seqüência de pLA3054 (SEQ ID NO. 158).

<***> Aspectos chaves incluem:

1. Região de codificação de DsRed-ubi-tTAV com íntron Cctra inserido com seqüência exônica tra de flanqueamento.

- 20 <***> Íntron Cctra é de posição 3509-2165 em DsRed-ubi-tTAV (nota: orientação reversa).

<***> Íntron Cctra é flanqueado por seqüência exônica Cctra de posições 3531-3510 e 2164-2145 (nota: orientação reversa).

- 25 <***> Transcrito alternativamente emendado DsRed-ubi-tTAV começa ou em segmento derivado de hsp70 de posição ~3243 (fragmento hsp70 é de 4930-4801) ou segmento derivado de Opie2 de posição ~4353 (fragmento Opie2 é de 4795-4255); (nota: orientação reversa).

<***> O transcrito DsRed-ubi-tTAV F1 é previsto como sendo traduzido entre 4320-888, inclusive de códon de partida e de parada (DsRed é de 4212-3538; ubiquitina é de 2135-1908; tTAV é de 1907-888);

(nota: orientação reversa).

<223> Sequência de pLA3056 (SEQ ID NO. 159).

<***> Aspectos chaves incluem:

1. Região de codificação de DsRed-ubi-tTAV com íntron
5 Ccra inserido com sequência exônica tra de flanqueamento.

<***> Íntron Ccra é de posição 3731-2387 em DsRed-ubi-tTAV (nota: orientação reversa).

<***> Íntron Ccra é flanqueado por sequência exônica Ccra de posições 3753-3732 e 2386-2145 (nota: orientação reversa).

- 10 <***> Transcrito alternativamente emendado DsRed-ubi-tTAV começa ou em segmento derivado de hsp70 de posição ~5109 (fragmento hsp70 é de 5152-5023) ou segmento derivado de Opie2 de posição ~4575 (fragmento Opie2 é de 5017-4477); (nota: orientação reversa).

- 15 <***> O transcrito DsRed-ubi-tTAV F1 é previsto como sendo traduzido entre 4542-888, inclusive de códon de partida e de parada (DsRed é de 4434-3760; ubiquitina é de 2135-1908; tTAV é de 1907-888); (nota: orientação reversa).

<***> Aspecto incluído:

- 20 1. Íntron adicional derivado de gene Ccra (segundo íntron de transcrito Ccra F1) dentro do transcrito F1 previsto de posição 2222-2168 (nota: orientação reversa).

<223> Sequência de pLA3488 (SEQ ID NO. 160).

<***> Aspectos chaves incluem:

- 25 1. Região de codificação de TurboGreen-ubi-DsRed com íntron Ccra inserido.

<***> Íntron Ccra é de posição 2263-3607 em TurboGreen-ubi-DsRed.

<***> Transcrito alternativamente emendado TurboGreen-ubi-DsRed começa no segmento derivado de baculovírus de AcMNPV Ie1

(premature imediato 1) em posição ~1180 (fragmento Ie1 é de 580-1209).

<***> O transcrito TurboGreen-ubi-DsRed F1 é previsto como sendo traduzido entre 1311-4467, inclusive de códon de partida e de parada (TurboGreen é de 1311-2093; ligador SG4 é de 2094-2123; ubiquitina é de 2124-3696, inclusive de íntron Ccra; DsRed é de 3697-4467).

<***> Aspecto incluído:

1. Íntron adicional derivado de gene "scraps" de *Drosophila* ("íntron 'scraps'") dentro do transcrito F1 previsto de posição 1224-1286.

<223> Seqüência de pLA3596 (SEQ ID NO. 145).

<***> Aspectos chaves incluem:

1. Região de codificação de TurboGreen-ubi-DsRed2 com íntron Ccra inserido.

<***> Íntron Ccra é de posição 5947-7291 em TurboGreen-ubi-DsRed2.

<***> TurboGreen-ubi-DsRed2 o transcrito alternativamente emendado começa no segmento derivado de baculovírus AcMNPV Ie1 (premature imediato 1) em posição ~4864 (fragmento Ie1 é de 4264-4893).

<***> O transcrito TurboGreen-ubi-DsRed2 F1 é previsto como sendo traduzido entre 4995-8148, inclusive de códon de partida e de parada (TurboGreen é de 4995-5777; ligador SG4 é de 5778-5807; ubiquitina é de 5808-7380, inclusive de íntron Ccra; DsRed2 é de 7381-8151).

<***> Aspecto incluído:

1. Íntron adicional derivado de gene "scraps" de *Drosophila* ("íntron 'scraps'") dentro do transcrito F1 previsto de posição 4908-4970.

2. Mutação de aminoácido destinada para comparar a LA3488 de posição 7294-7296.

REIVINDICAÇÕES

1. Sistema de expressão de polinucleotídeos, caracterizado pelo fato de compreender:

pelo menos uma seqüência polinucleotídica heteróloga codificando uma proteína funcional, definida entre um códon de partida e um códon de parada, e/ou polinucleotídeos para RNA interferência (RNAi), a ser expressado em um organismo;

pelo menos um promotor ligado operativamente ao mesmo; e

pelo menos uma seqüência de controle de emenda que, em cooperação com um espliceossoma, é capaz de (i) mediar a emenda de um transcrito de RNA da seqüência de codificação para dar um primeiro produto de RNA mensageiro (mRNA) emendado, e (ii) mediar pelo menos uma emenda alternativa de referido transcrito de RNA para dar um produto de mRNA emendado alternativo;

em que, quando a pelo menos uma seqüência polinucleotídica heteróloga codifica uma proteína funcional, pelo menos um dentre os produtos de mRNA maduros compreendendo uma matriz de leitura aberta contínua (ORF) se estendendo a partir do referido códon de partida para referido códon de parada, assim definindo uma proteína, que é referida proteína funcional, ou está relacionada com referida proteína funcional por, pelo menos, uma deleção de aminoácido e que é funcional quando traduzida e, opcionalmente, sofreu uma modificação pós-translacional;

a mediação sendo selecionada dentre o grupo consistindo de: mediação específica de sexo, mediação específica de estágio, mediação específica de linhagem germinal, mediação específica de tecido, e combinações das mesmas.

2. Sistema de expressão de polinucleotídeos de acordo com a reivindicação 1, caracterizado pelo fato de que a mediação é específica do sexo.

3. Sistema de expressão de polinucleotídeos de acordo com a reivindicação 1 ou 2, caracterizado pelo fato de que a sequência polinucleotídica a ser expressada compreende dois ou mais éxons de codificação para a proteína funcional.

5 4. Sistema de expressão de polinucleotídeos de acordo com qualquer reivindicação precedente, caracterizado pelo fato de que a proteína é um marcador, ou tem um efeito letal, deletério ou esterilizante.

5. Sistema de expressão de polinucleotídeos de acordo com a reivindicação 4, caracterizado pelo fato de que a proteína tem um efeito letal
10 resultando em esterilização.

6. Sistema de expressão de polinucleotídeos de acordo com a reivindicação 5, caracterizado pelo fato de que o efeito letal da proteína é condicionalmente suprimível.

7. Sistema de expressão de polinucleotídeos de acordo com a reivindicação 4, caracterizado pelo fato de que a proteína é selecionada dentre
15 o grupo consistindo de um fator indutor de apoptose, Hid, Reaper (Rpr) e Nipp1Dm.

8. Sistema de expressão de polinucleotídeos de acordo com qualquer reivindicação precedente, caracterizado pelo fato de que o sistema
20 compreende pelo menos um mecanismo de retro-alimentação positivo, sendo pelo menos uma proteína funcional a ser diferencialmente expressada, via emenda alternativa, e pelo menos um promotor para o mesmo, em que um produto de um gene a ser expressado serve como um fator de controle transcricional positivo para, pelo menos, um promotor, e assim o produto, ou
25 a expressão do produto, é controlável.

9. Sistema de expressão de polinucleotídeos de acordo com a reivindicação 8, caracterizado pelo fato de que um melhorador está associado com o promotor, o produto de gene servindo para melhorar a atividade do promotor *via* o melhorador.

10. Sistema de expressão de polinucleotídeos de acordo com a reivindicação 9, caracterizado pelo fato de que o fator de controle é o produto de gene tTA ou um análogo do mesmo, e em que uma ou mais das unidades de operador tetO é ligada operativamente com o promotor e é o melhorador, tTA ou seu análogo servindo para melhorar a atividade do promotor *via* tetO.

11. Sistema de expressão de polinucleotídeos de acordo com qualquer reivindicação precedente, caracterizado pelo fato de que a proteína funcional é, ela mesma, um transativador transcricional, como o sistema tTAV, compreendendo tTAV, tTAV2, ou tTAV3.

12. Sistema de expressão de polinucleotídeos de acordo com qualquer reivindicação precedente, caracterizado pelo fato de que o promotor é ativado por condições ambientais, por exemplo a presença ou ausência de um fator particular como tetraciclina no sistema *tet* ou por variação da temperatura ambiental.

13. Sistema de expressão de polinucleotídeos de acordo com qualquer uma das reivindicações 1 a 11, caracterizado pelo fato de que o promotor é selecionado dentre o grupo consistindo de promotor específico do embrião *srya*, ou seus homólogos, o gene de *Drosophila* "*slow as molasses*" (*slam*) ou seus homólogos.

14. Sistema de expressão de polinucleotídeos de acordo com qualquer reivindicação precedente, caracterizado pelo fato de ainda compreender um melhorador.

15. Sistema de expressão de polinucleotídeos de acordo com qualquer reivindicação precedente, caracterizado pelo fato de que a mediação de emenda alternativa é específica de sexo e a seqüência de controle de emenda é derivada de um intron *tra*.

16. Sistema de expressão de polinucleotídeos de acordo com a reivindicação 15, caracterizado pelo fato de que a seqüência de controle de emenda é derivada do gene *transformer* de mosca-do-mediterrâneo *Cetra*, ou

de outro ortólogo ou homólogo do gene *transformer* de *Drosophila*.

17. Sistema de expressão de polinucleotídeos de acordo com a reivindicação 16, caracterizado pelo fato de que referido outro ortólogo ou homólogo do gene *transformer* de *Drosophila* é de uma mosca tefrítida da
5 fruta.

18. Sistema de expressão de polinucleotídeos de acordo com a reivindicação 17, caracterizado pelo fato de que de que a mosca tefrítida da fruta é *C. rosa*, ou *B. zonata*.

19. Sistema de expressão de polinucleotídeos de acordo com
10 qualquer uma das reivindicações 1 a 14, caracterizado pelo fato de que a sequência de controle de emenda é derivada de mecanismo de emenda alternativa do gene *Actin-4*.

20. Sistema de expressão de polinucleotídeos de acordo com a reivindicação 19, caracterizado pelo fato de que o gene *Actin-4* é de *Aedes*
15 *spp*.

21. Sistema de expressão de polinucleotídeos de acordo com a reivindicação 19, caracterizado pelo fato de que o gene *Actin-4* é de *Aedes aegypti* *AeActin-4*.

22. Sistema de expressão de polinucleotídeos de acordo com
20 qualquer uma das reivindicações 1 a 14, caracterizado pelo fato de que o mecanismo de emenda compreende pelo menos um fragmento do gene *doublesex (dsx)*, preferivelmente o derivado de *Drosophila*, *B. mori*, lagarta rosada, traça das maçãs e pêras, ou um mosquito, particularmente *Aedes gambiae* ou especialmente *Aedes aegypti*.

23. Sistema de expressão de polinucleotídeos de acordo com a reivindicação 19 a 22, caracterizado pelo fato de que a sequência de controle de emenda e a sequência polinucleotídica heteróloga codificando uma proteína funcional, definida entre um códon de partida e um códon de parada, e/ou polinucleotídeos para RNA interferência (RNAi), a ser expressado em
25

um organismo, são providas na forma de uma construção de minigene ou um éxon de cassete.

24. Sistema de expressão de polinucleotídeos de acordo com a reivindicação 4, caracterizado pelo fato de que o sistema é um plasmídeo ou construção selecionado dentre o grupo consistindo de qualquer uma das figuras 16-18, 22-24, 26-32, 49, 52-55, e 61-69, e/ou SEQ ID NO: 46-48, 50-56, 143-145 e 151-162.

25. Sistema de expressão de polinucleotídeos de acordo com qualquer reivindicação precedente, caracterizado pelo fato de que pelo menos uma seqüência de controle de emenda é intrônica e compreende em sua extremidade 5' nucleotídeo guanina (G), em RNA.

26. Sistema de expressão de polinucleotídeos de acordo com qualquer reivindicação precedente, caracterizado pelo fato de que pelo menos uma seqüência de controle de emenda é intrônica e compreende em sua extremidade 5' nucleotídeos UG e UT em sua extremidade 3', em RNA.

27. Sistema de expressão de polinucleotídeos de acordo com qualquer reivindicação precedente, caracterizado pelo fato de que a mediação é específica de sexo e ainda mediada ou controlada por ligação da proteína TRA ou complexo de proteína TRA/TRA2, ou homólogos das mesmas.

28. Sistema de expressão de polinucleotídeos de acordo com a reivindicação 27, caracterizado pelo fato de que o sistema compreende a seqüência de consenso: TCWWCRATCAACA, em que W = A ou T e R = A ou G.

29. Sistema de expressão de polinucleotídeos de acordo com qualquer reivindicação precedente, caracterizado pelo fato de que o organismo é um mamífero, um peixe, um invertebrado, um artrópode, um inseto ou uma planta.

30. Sistema de expressão de polinucleotídeos de acordo com qualquer reivindicação precedente, caracterizado pelo fato de que o

organismo é um inseto da Ordem Díptera.

31. Sistema de expressão de polinucleotídeos de acordo com a reivindicação 30, caracterizado pelo fato de que o inseto é uma mosca tefrítida das frutas selecionada dentre o grupo consistindo de: mosca-do-mediterrâneo (*Ceratis capitata*), mosca mexicana das frutas (*Anastrepha ludens*), mosca oriental das frutas (*Bactrocera dorsalis*), mosca da azeitona (*Bactrocera oleae*), mosca do melão (*Bactrocera cucurbitae*), mosca das frutas do natal (*Ceratitis rosa*), mosca da cereja (*Rhagoletis cerasi*), mosca-de-queensland (*Bactrocera tyroni*), mosca-do-pêssego (*Bactrocera zonata*),
 5 mosca das frutas do Caribe (*Anastrepha suspensa*) e mosca das frutas da índia ocidental (*Anastrepha obliqua*).

32. Sistema de expressão de polinucleotídeos de acordo com a reivindicação 30, caracterizado pelo fato de que o inseto é um mosquito dos gêneros *Stegomyia*, *Aedes*, *Anopheles* ou *Culex*.

15 33. Sistema de expressão de polinucleotídeos de acordo com a reivindicação 32, caracterizado pelo fato de que o mosquito é selecionado dentre *Aedes aegypti*, *Aedes albopictus*, *Anopheles stephensi*, *Anopheles albimanus* e *Anopheles gambiae*.

20 34. Sistema de expressão de polinucleotídeos de acordo com a reivindicação 30, caracterizado pelo fato de que o inseto é selecionado dentre o grupo consistindo de: verme-parafuso do Novo Mundo (*Cochliomyia hominivorax*), verme-parafuso do Velho Mundo (*Chrysomya bezziana*), e mosca varejeira da ovelha australiana (*Lucilia cuprina*), traça das maçãs e pêras (*Cydia pomonella*), o bicho-da-seda (*Bombyx mori*), a lagarta-rosada
 25 (*Pectinophora gossypiella*), a lagarta da couve (*Plutella xylostella*), a mariposa-cigana (*Lymantria dispar*), a lagarta da laranja-da-baía (*Amyelois transitella*), a lagarta mineira do pessegueiro (*Anarsia lineatella*), e a broca da haste do arroz (*Tryporyza incertulas*), as traças noctuídas, especialmente *Heliothianae*, o besouro japonês (*Popilla japonica*), o besouro branco

(*Graphognathus spp.*), o gorgulho do algodão (*Anthonomous grandis*), a lagarta da raiz do milho (*Diabrotica spp*), e escaravelho da batateira (*Leptinotarsa decemlineata*).

5 35. Sistema de expressão de polinucleotídeos de acordo com a reivindicação 30, caracterizado pelo fato de que o inseto não é uma drosófila.

36. Sistema de expressão de polinucleotídeos de acordo com qualquer reivindicação precedente, caracterizado pelo fato de que a expressão da sequência polinucleotídica heteróloga leva a uma consequência fenotípica no organismo.

10 37. Sistema de expressão de polinucleotídeos de acordo com a reivindicação 1 ou 2, caracterizado pelo fato de que a sequência polinucleotídica a ser expressada compreende um polinucleotídeo para RNA interferência (RNAi).

15 38. Método de controle da população de um organismo em um meio ambiente natural para o mesmo, caracterizado pelo fato de compreender:

i) reproduzir um lote do organismo,

o organismo transportando um sistema de expressão de genes compreendendo um sistema de acordo com qualquer uma das reivindicações 1-36, que é um sistema genético letal dominante,

20 ii) distribuir os referidos animais do lote no meio ambiente em um *locus* para controle da população; e

25 iii) obter um controle da população através da letalidade em estágio prematuro por expressão do sistema letal na descendência que resulta da inter-reprodução dos referidos indivíduos no lote com os indivíduos do sexo oposto da população selvagem.

39. Método de acordo com a reivindicação 38, caracterizado pelo fato de que a letalidade em estágio prematuro é embriônica ou antes da maturidade sexual.

40. Método de acordo com a reivindicação 39, caracterizado

pelo fato de que a letalidade em estágio prematuro ocorre de modo prematuro no desenvolvimento.

41. Método de acordo com a reivindicação 38 ou 39, caracterizado pelo fato de que o efeito letal do sistema letal é condicional e ocorre no referido meio ambiente natural *via* a expressão de um gene letal, a expressão de referido gene letal estando sob o controle de uma proteína de transativador reprimível, a referida reprodução ocorrendo sob condições permissivas na presença de uma substância, a substância estando ausente do referido meio ambiente natural e capaz de reprimir referido transativador.

42. Método de controle biológico, caracterizado pelo fato de compreender:

i) reproduzir um lote de organismos machos e fêmeas transformados com o sistema de acordo com qualquer uma das reivindicações 1-36, sob condições permissivas, permitindo a sobrevivência de machos e fêmeas, para dar um agente de controle biológico de sexo dual;

ii) opcionalmente antes da próxima etapa impor ou permitir condições restritivas para causar a morte de indivíduos de um sexo e, assim, prover um agente de controle biológico de sexo único compreendendo indivíduos do outro sexo transportando o sistema genético letal condicional;

iii) liberar o agente de controle biológico de sexo dual ou de sexo único no meio ambiente em um *locus* para controle biológico; e

iv) obter controle biológico através da expressão do sistema genético na descendência resultante de inter-reprodução dos indivíduos do agente de controle biológico com indivíduos do sexo oposto da população selvagem.

43. Método de separação de sexos, caracterizado pelo fato de compreender:

i) reproduzir um lote de organismos machos e fêmeas transformados com o sistema de expressão como definido em qualquer uma

das reivindicações 1 a 36 sob condições permissivas ou restritivas, permitindo a sobrevivência de machos e fêmeas; e

ii) remover as condições permissivas ou restritivas para induzir o efeito letal do gene letal em um sexo e não no outro por emenda alternativa específica de sexo do gene letal.

44. Método ou controle biológico ou da população, caracterizado pelo fato de compreender:

i) reproduzir um lote de organismos machos e fêmeas transformados com o sistema de expressão de genes como definido em qualquer uma das reivindicações 1 a 36, sob condições permissivas ou restritivas, permitindo a sobrevivência de machos e fêmeas;

ii) remover as condições permissivas ou restritivas para induzir o efeito letal do gene letal em um sexo e não no outro por emenda alternativa específica do sexo do gene letal para obter separação de sexos;

iii) esterilizar ou parcialmente esterilizar os indivíduos separados, e

iv) obter referido controle através da liberação de indivíduos estéreis ou parcialmente estéreis separados no meio ambiente natural do organismo.

**Emenda
Aedes DSX
endógena**

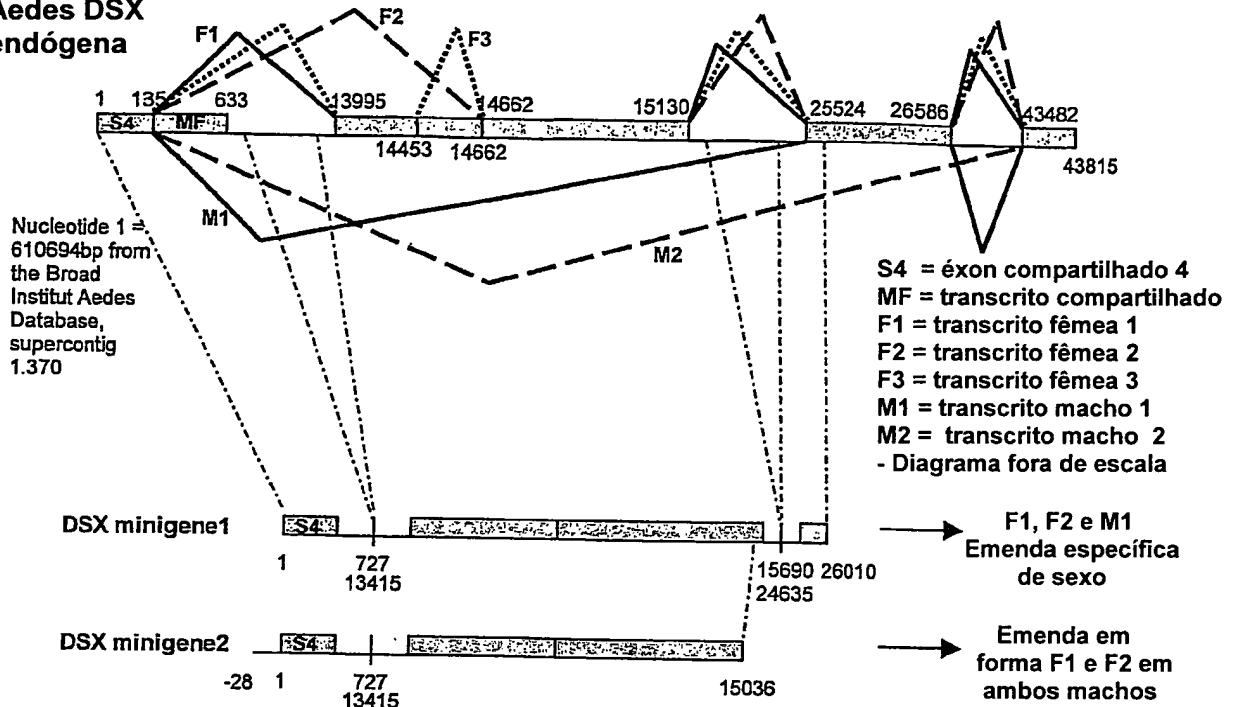


FIGURA 1

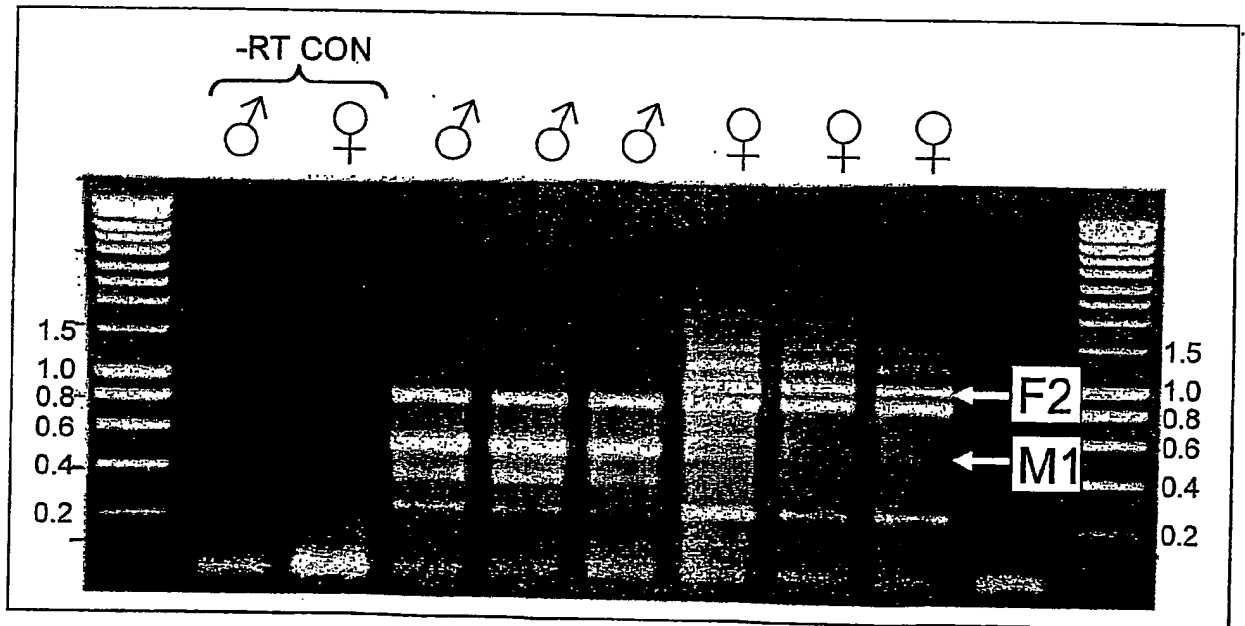


FIGURA 2

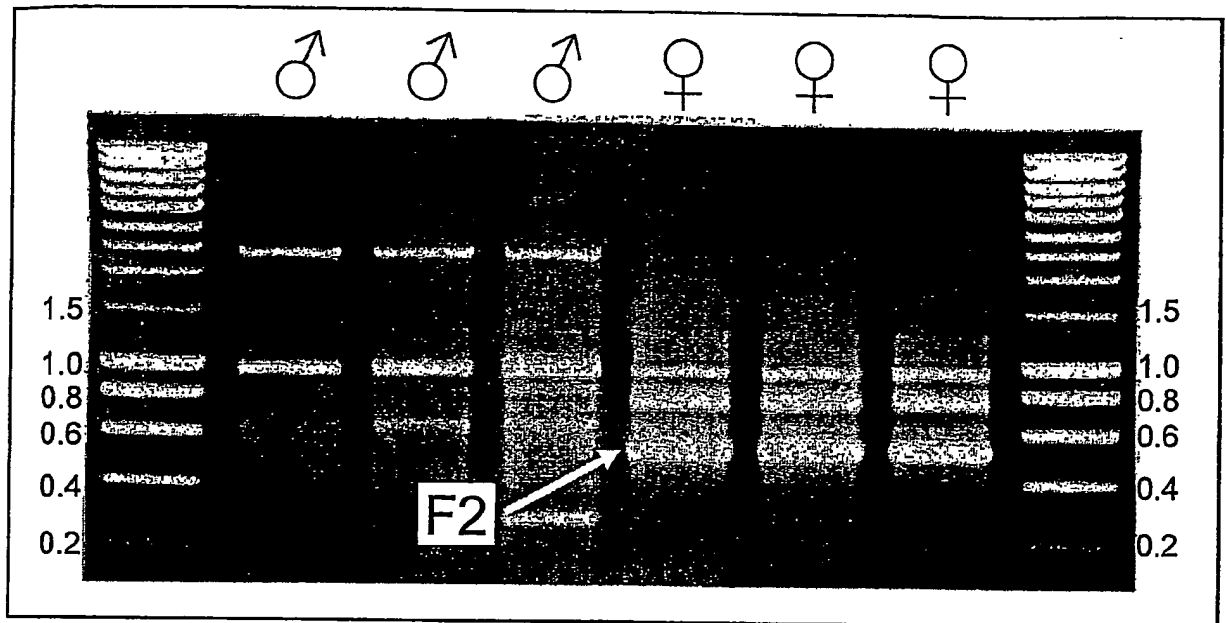


FIGURA 3

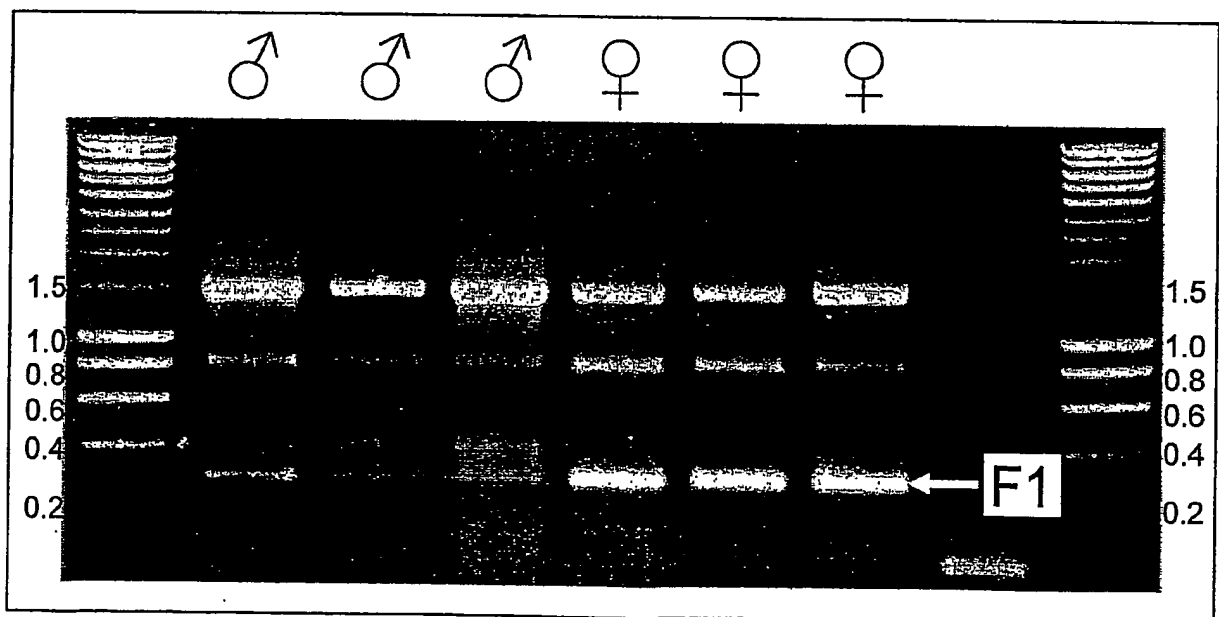
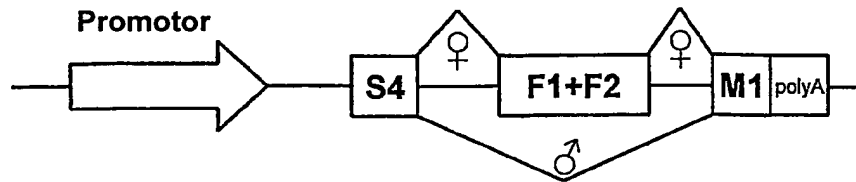
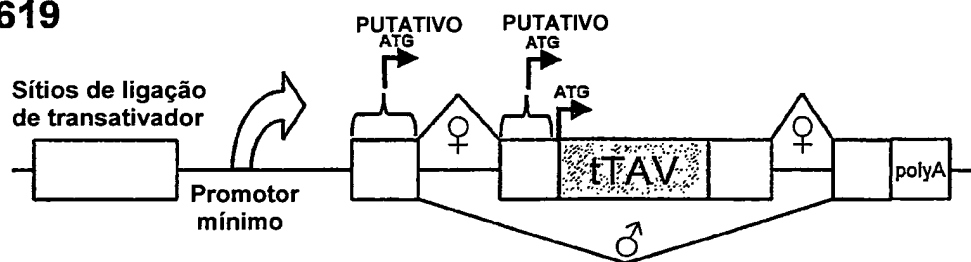
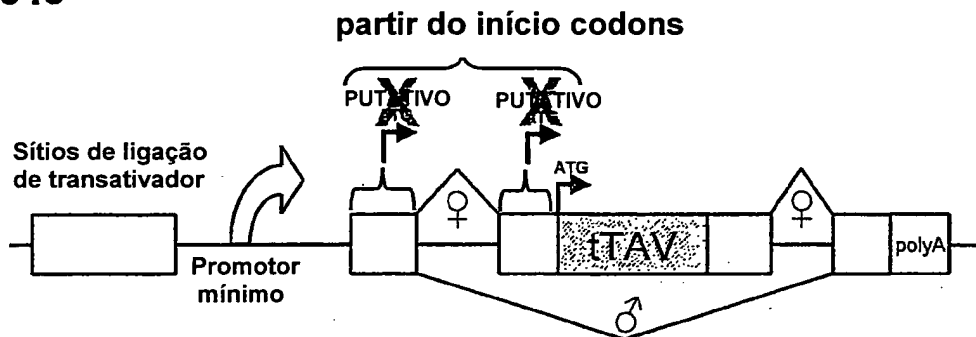
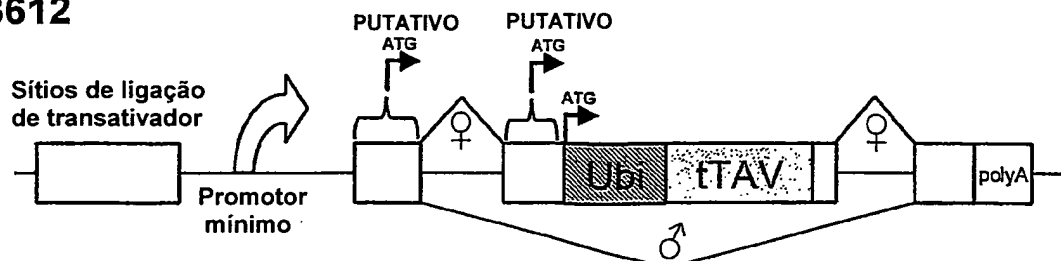


FIGURA 4

LA3491**LA3619****LA3646****LA3612****FIGURA 5**

PBW-dsx	TGTGCGTTGCTGTT--TCCGATGGAAGGACTAT-TGTGTCGTGCGCCACGCTGGACTATTC	4287
bombyx-dsx	TGTGCGATGCTGTG---CGAATTCAACGGAAATATTGTTGTGTCGTAACATTTGGATCTATG	1575
codling-dsx	TGACTGTTCCCTGTAATAAAGCTTCGTTGGACAT-TGTGTC-TCAC-ACATCGGATCTCAT	3428
	** * * * * * * * * * * * * * * * *	
PBW-dsx	GGTGAGTGG-----TAGAATAATA-TTTTATCTA-----TTTCATCGCGGT	4327
bombyx-dsx	GGTAAGTT-----TAGTATAATAACTTTACTCT-----GTTACATTAGT	1615
codling-dsx	GGTAAGTGCTAGTGCTAGCATYRMAACTTAACTCTCTGAGCGAATTCCTTTGACTCTAAA	3488
	*** ** * * * * * * * * * * *	
PBW-dsx	ACAATTGACTTTTATTACTACTGCTATGGAGGAATCTCAGGAACAT-----CGTAA	4383
bombyx-dsx	GAAACATACATTG--TAAAAATTG-TGTTTT--ACTAATGTGAAATTTAT-----TTTG	1666
codling-dsx	GTCACACGRACAGCCATACAATCAA-AGCTACGCTCTAATTTTAAAGATGACAWTCTGTAA	3547
	* ** * * * * * * * * * *	

Figura 6 - O segundo éxon específico de fêmea é marcado por nucleotídeos em negrito. As repetições conservadas AGTGAC/T são sublinhadas.

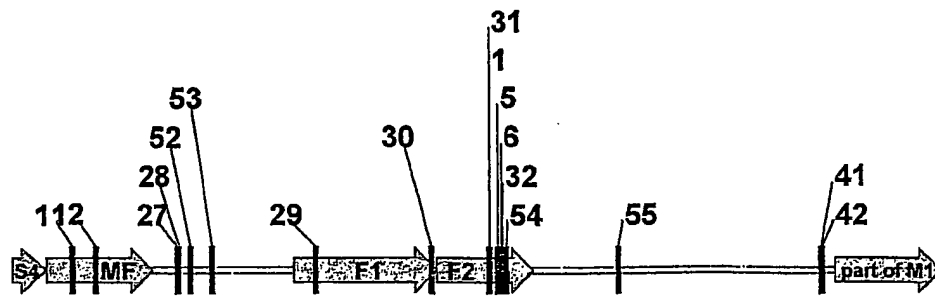


FIGURA 7

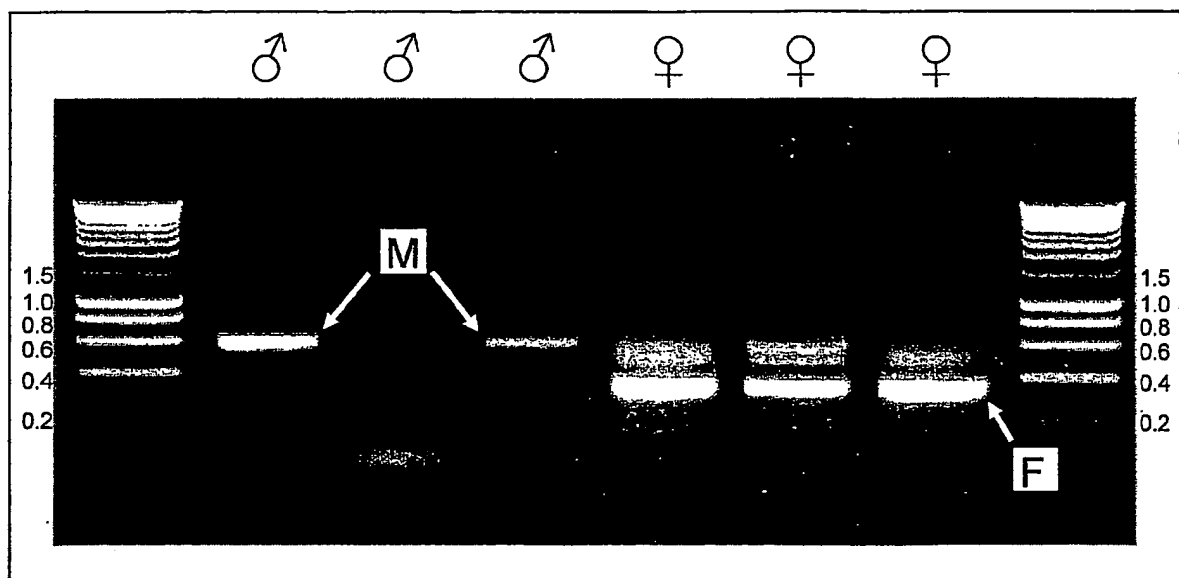


FIGURA 8

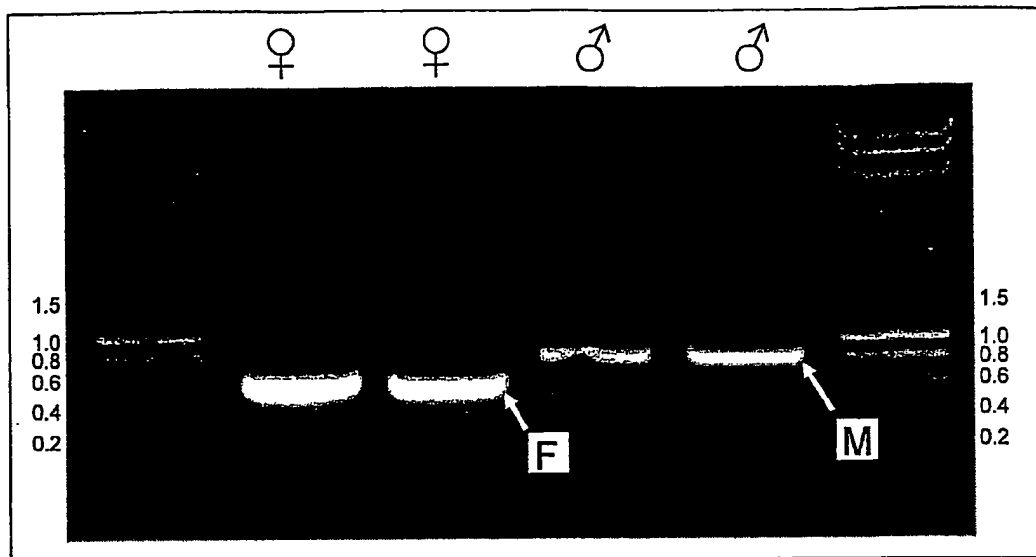


FIGURA 9

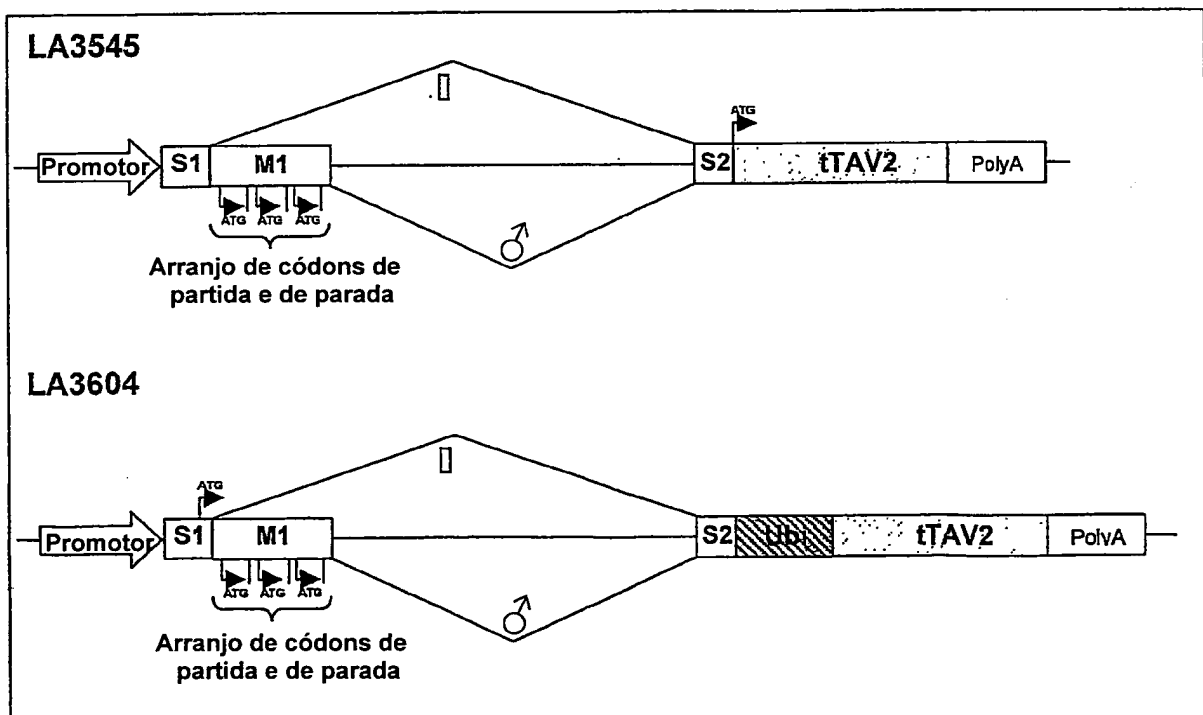


FIGURA 10

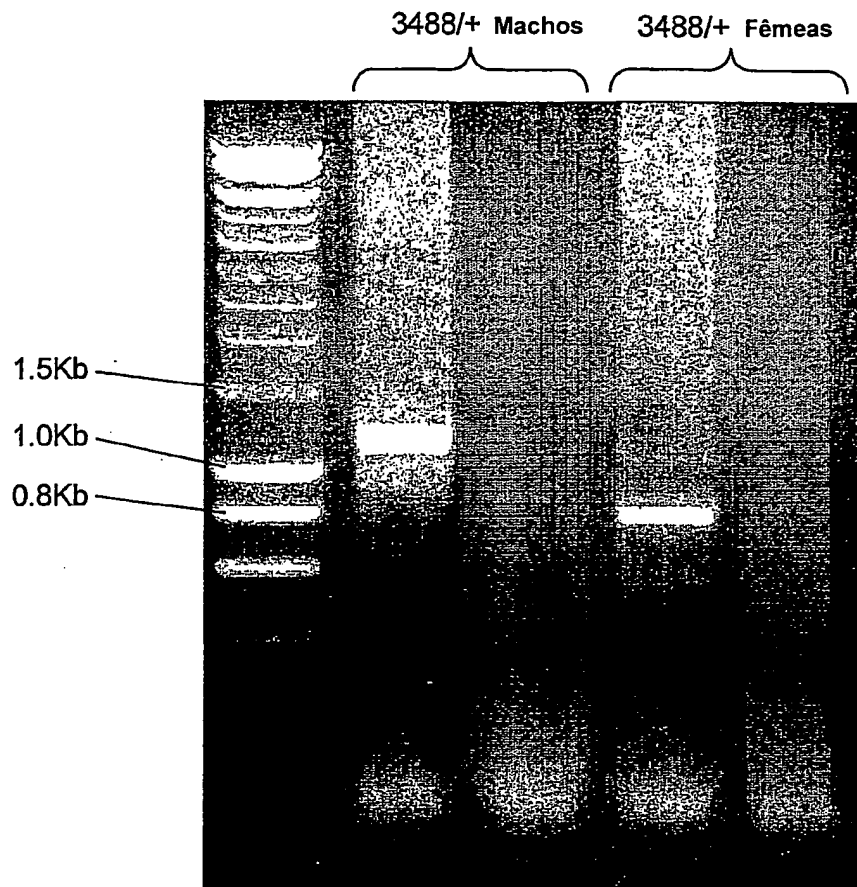


FIGURA 11

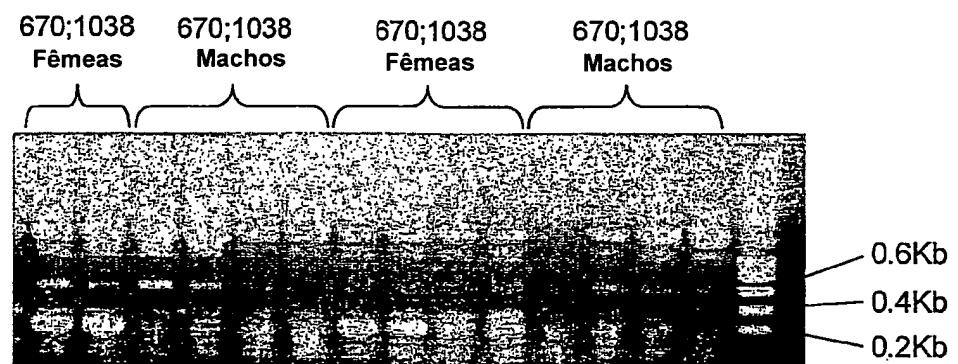


FIGURA 12

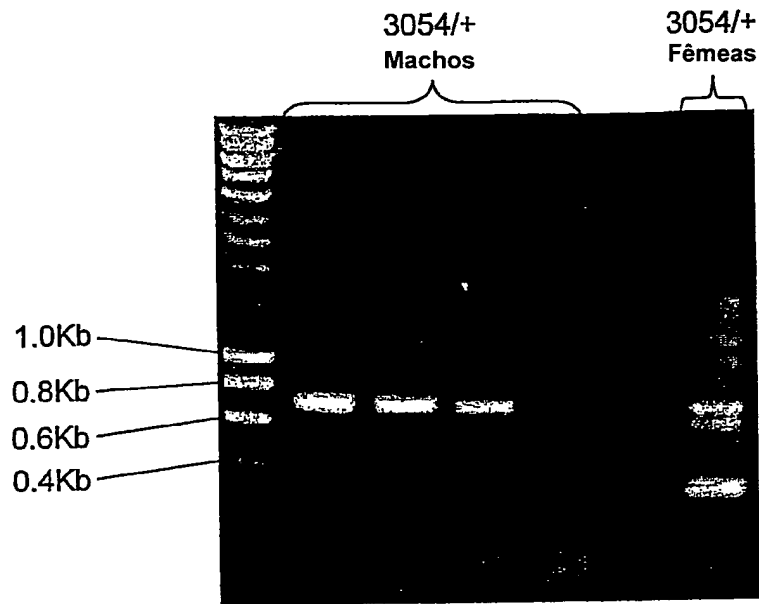


FIGURA 13

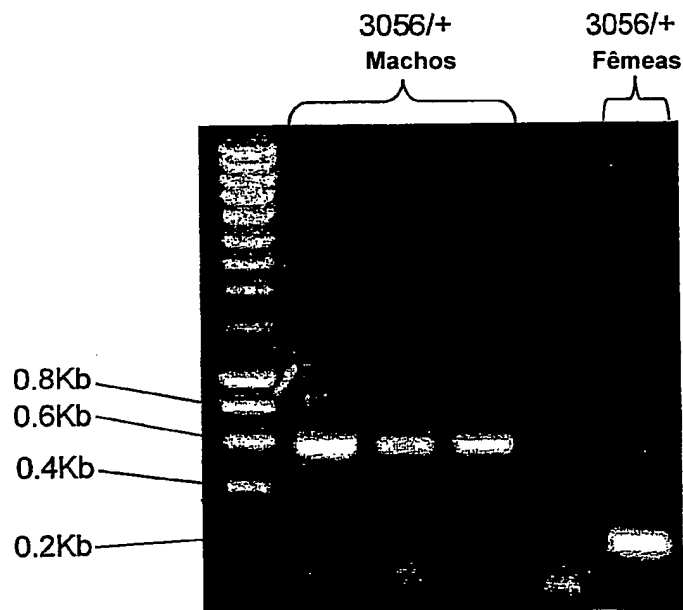


FIGURA 14

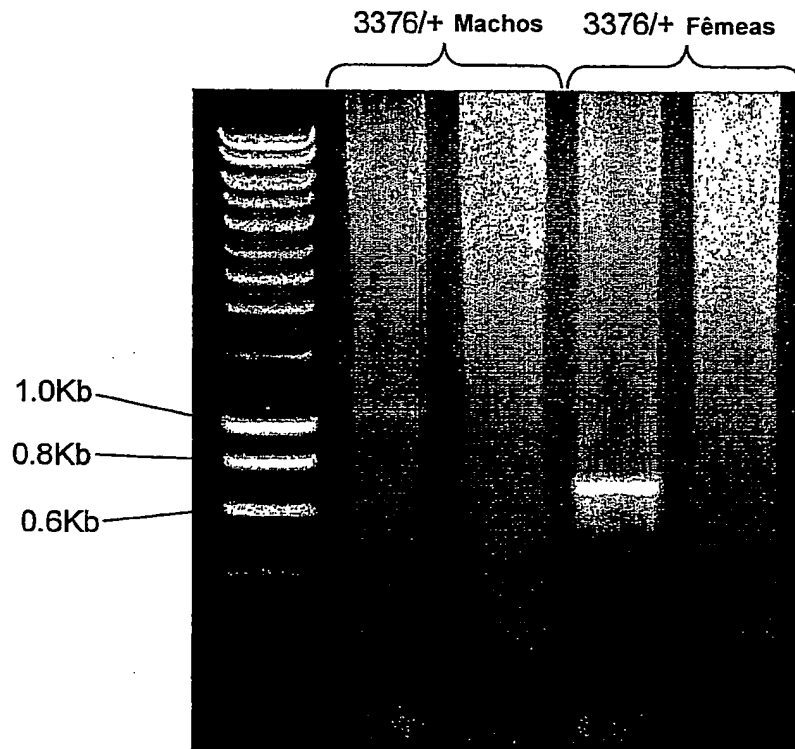
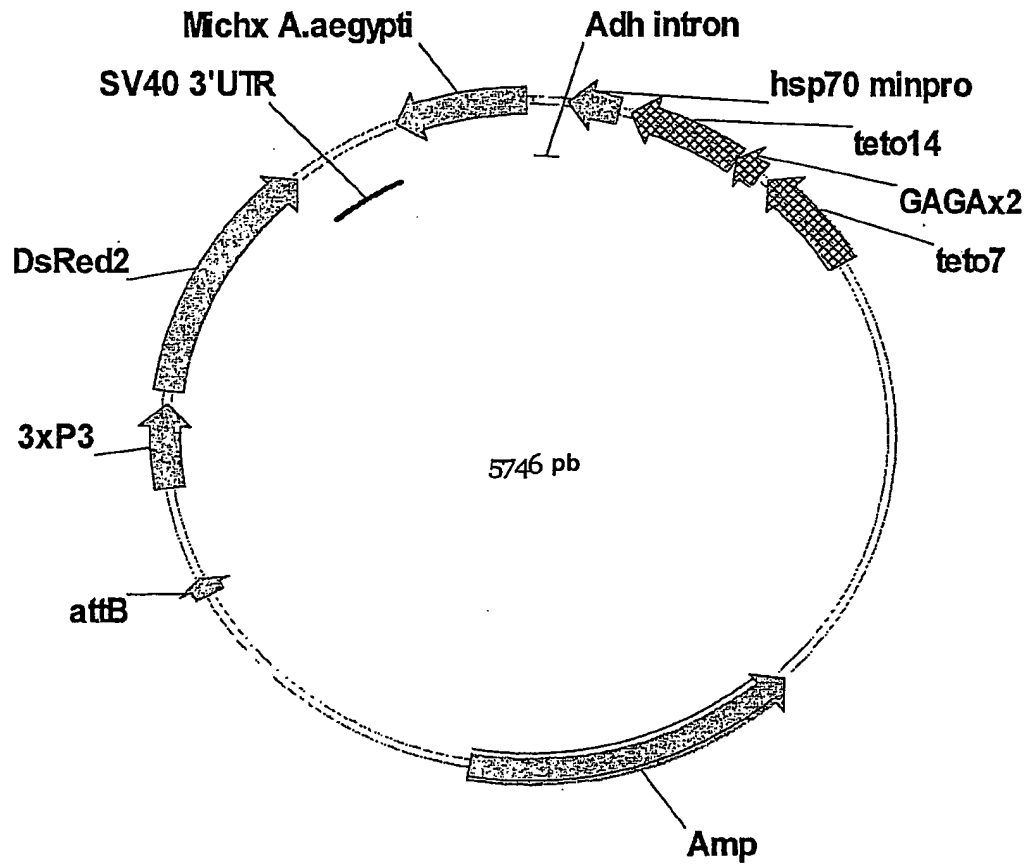


FIGURA 15

FIGURA 16

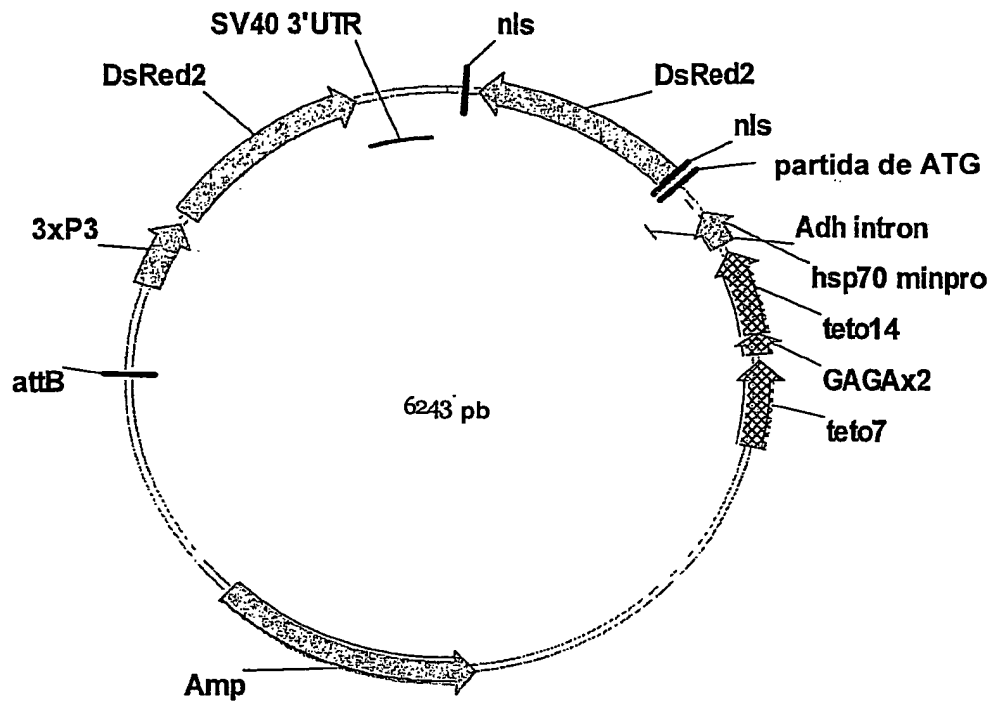
Plasmideo LA3582



#3581,2 AttB-3xP3DsRed2-teto21-hsp-adh-michxc

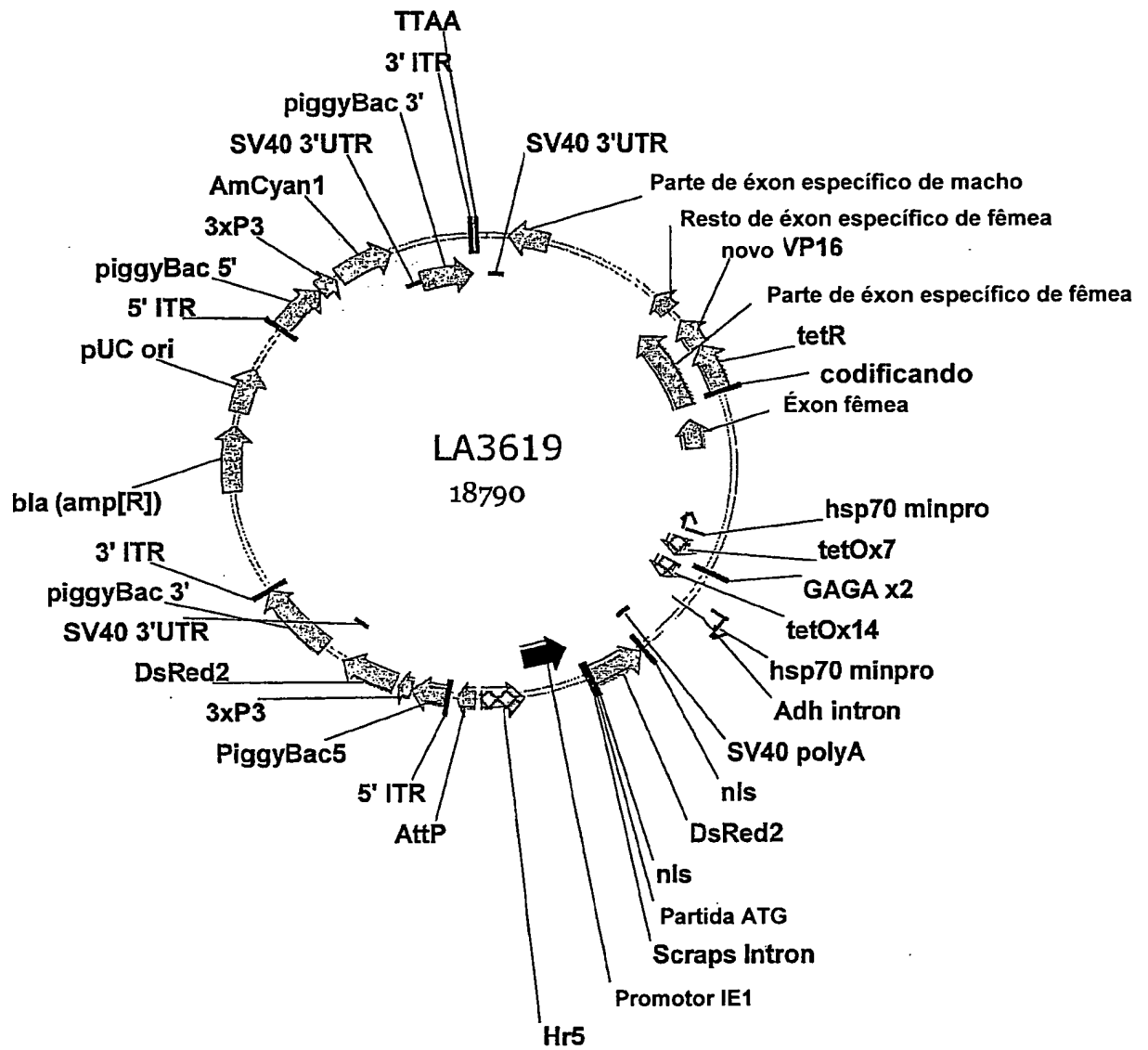
FIGURA 17

Plasmídeo LA3576



#3575,6 AttB-3xP3DsRed2-teto21-hsp-adh-dsred

Figura 18 - Mapa de plasmídeo LA3619



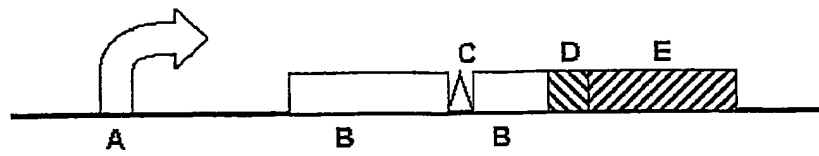
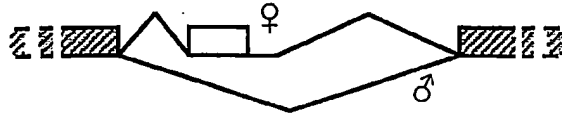
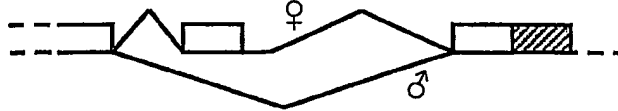


FIGURA 19

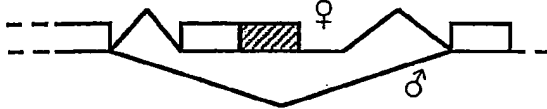
A: Expressão específica de macho



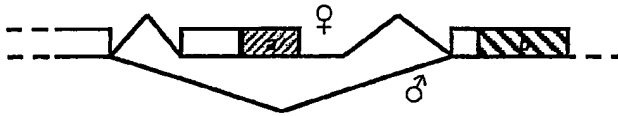
B: Expressão específica de macho



C: Expressão específica de fêmea



D: Expressão diferencial

**FIGURA 20**

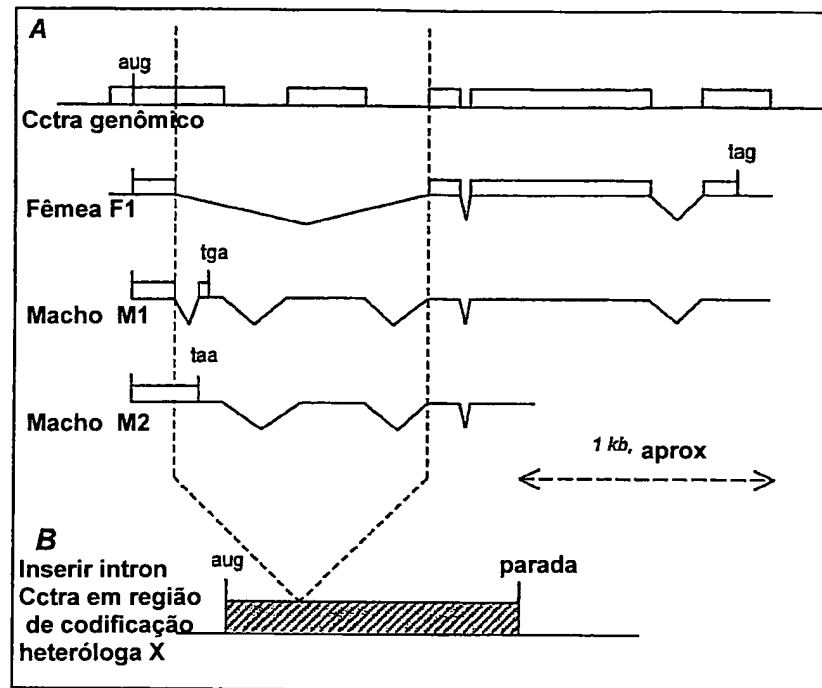


FIGURA 21

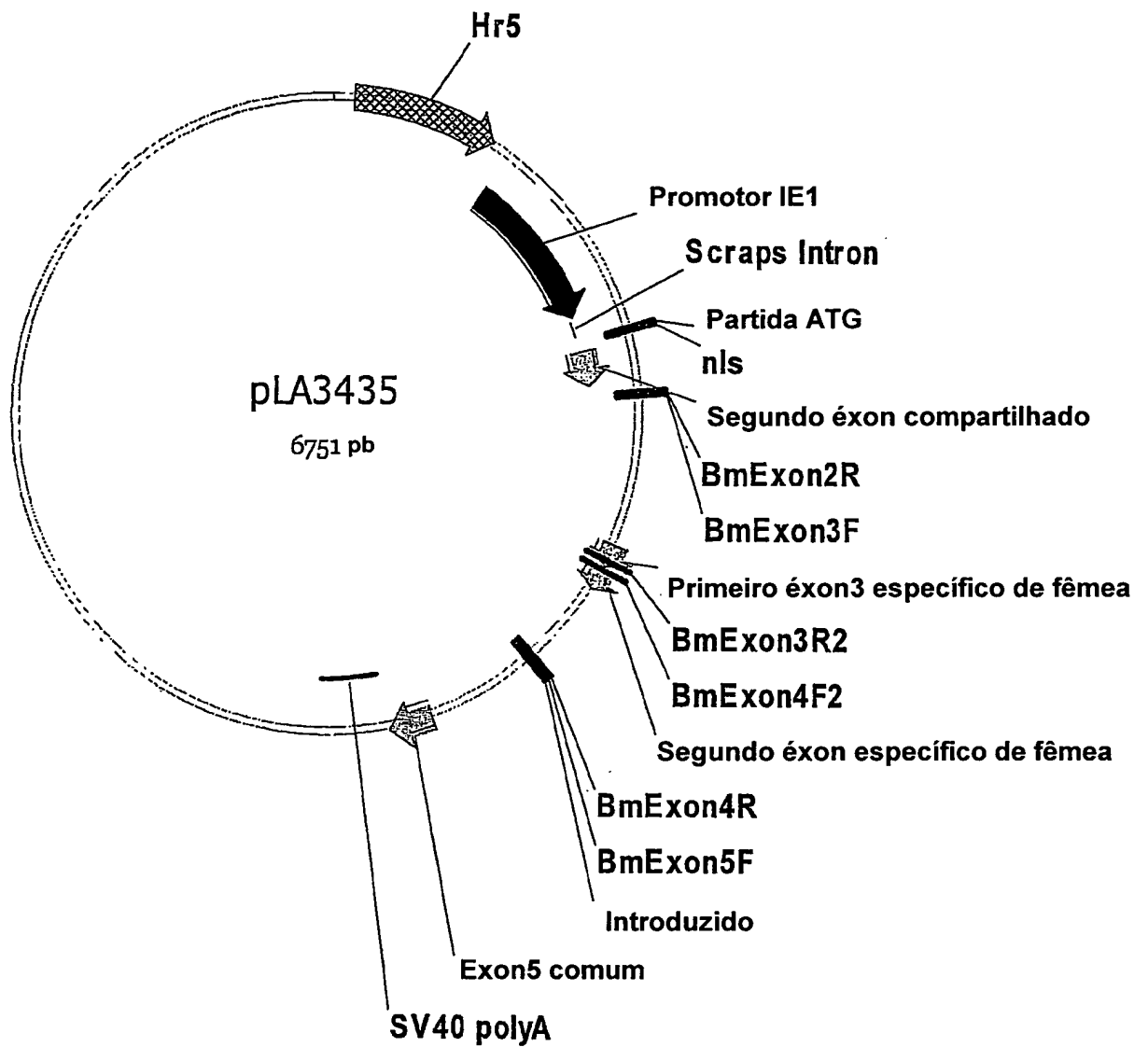


FIGURA 22

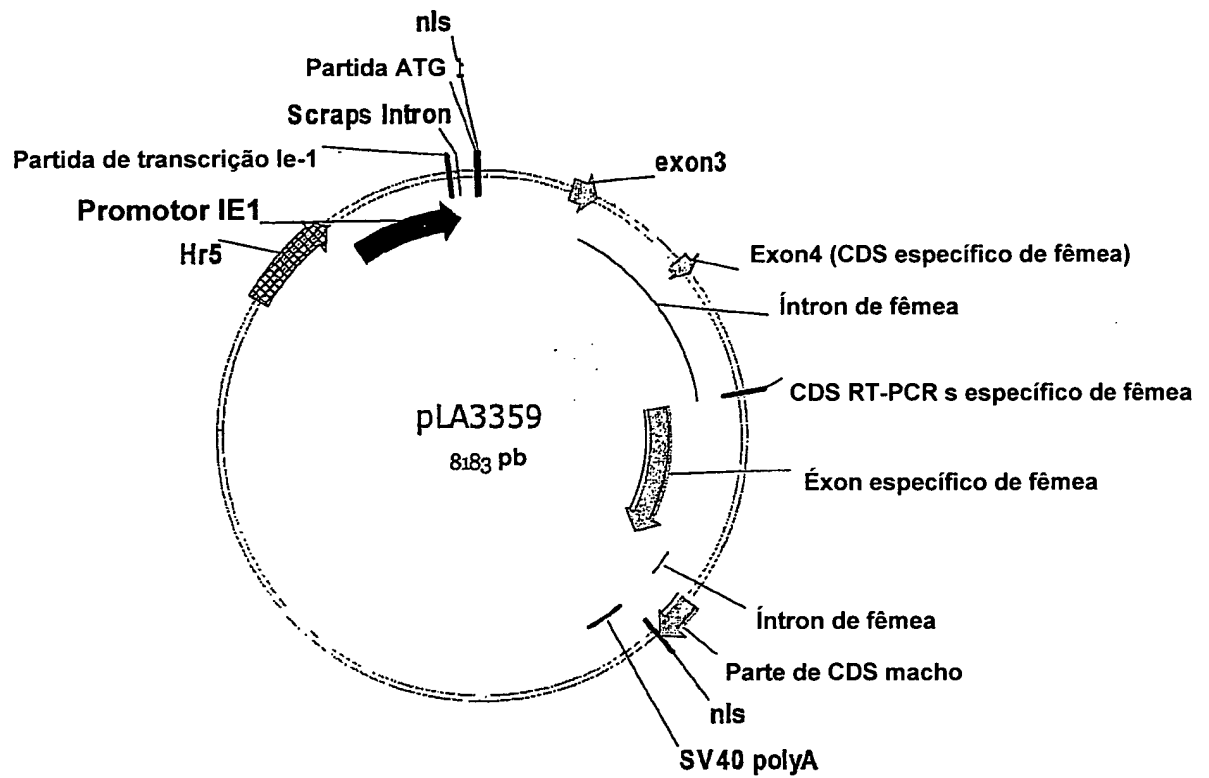


FIGURA 23

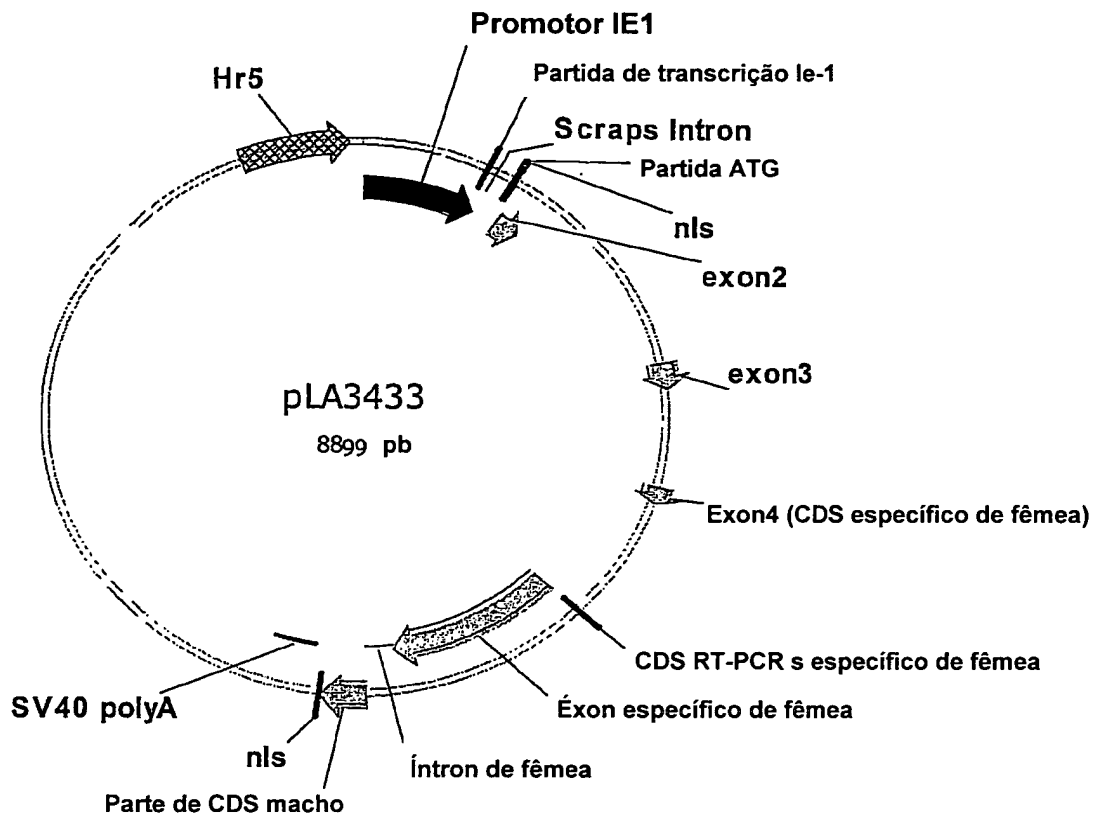


FIGURA 24

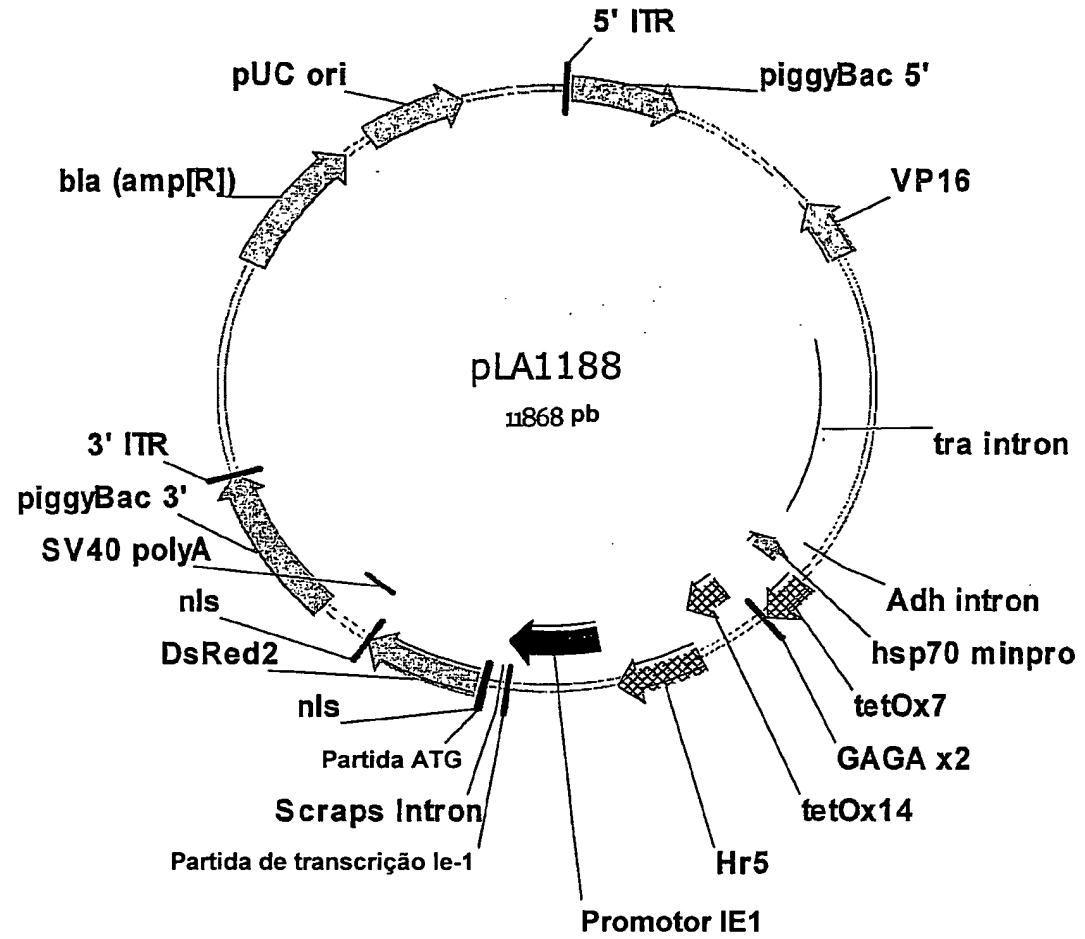


FIGURA 25

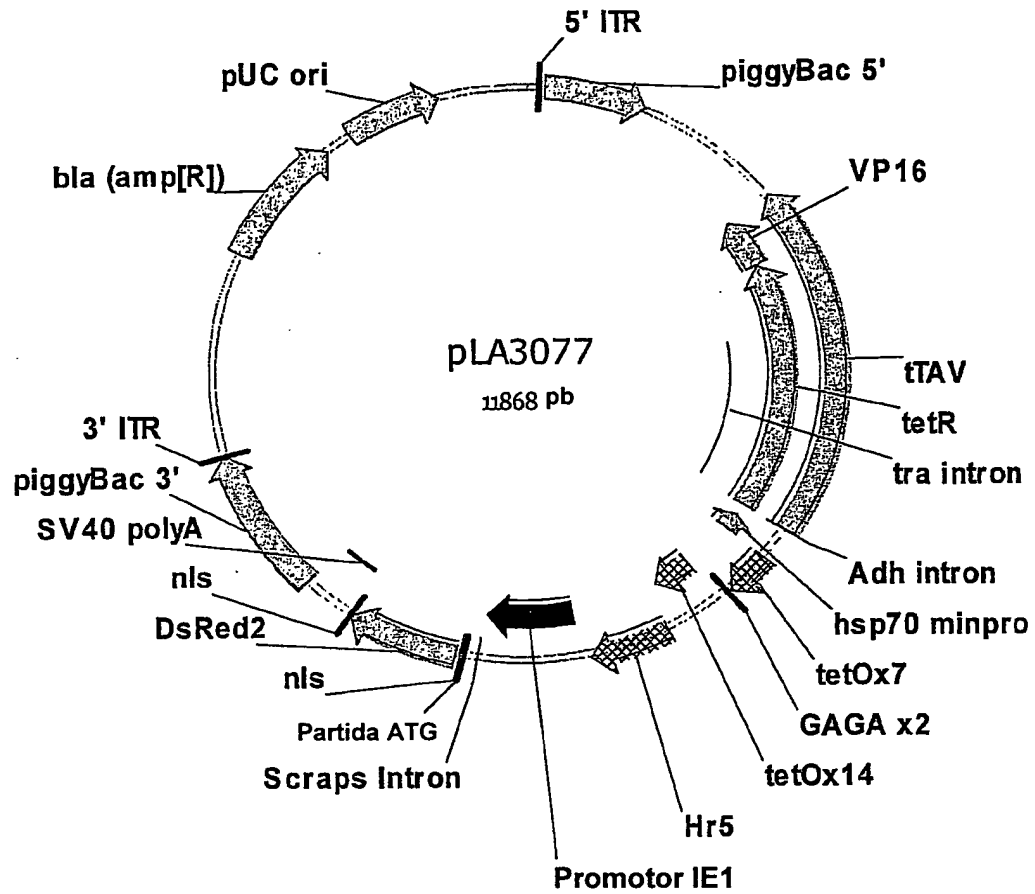


FIGURA 26

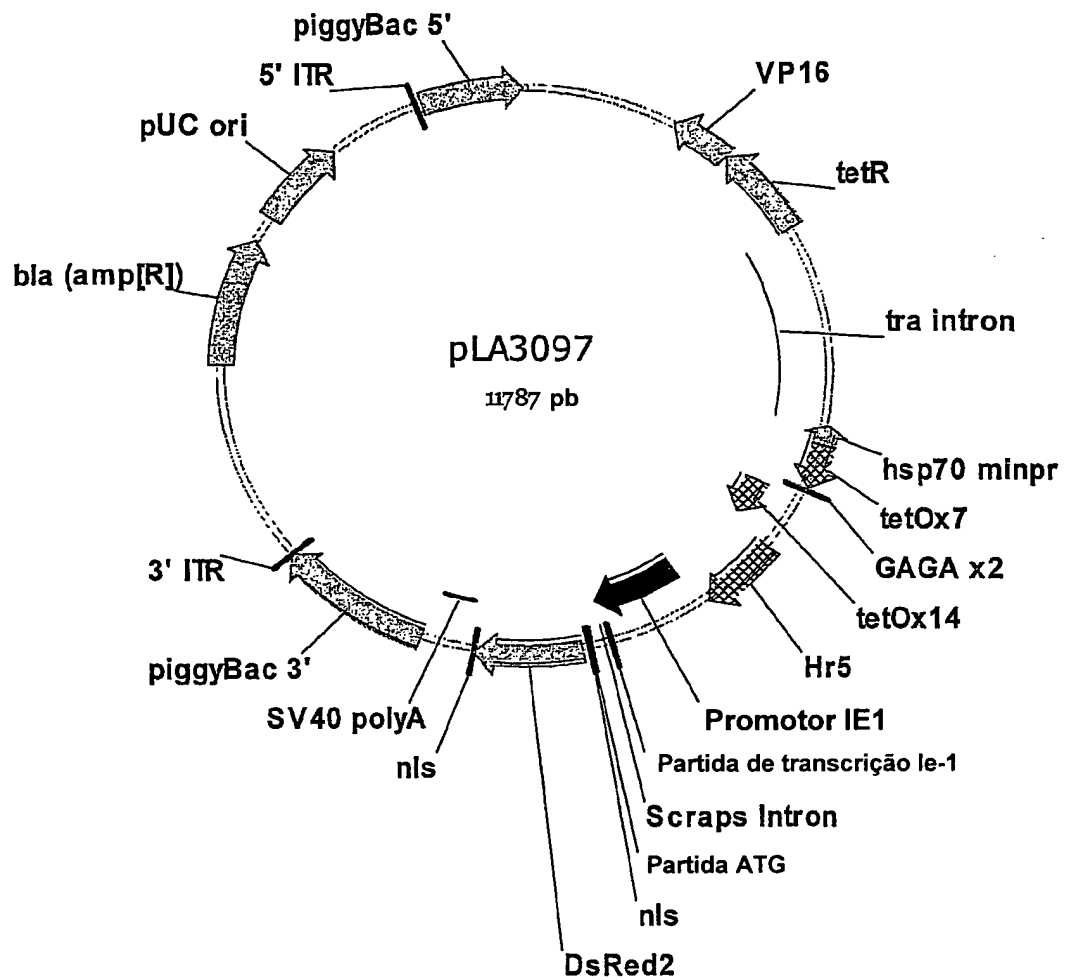


FIGURA 27

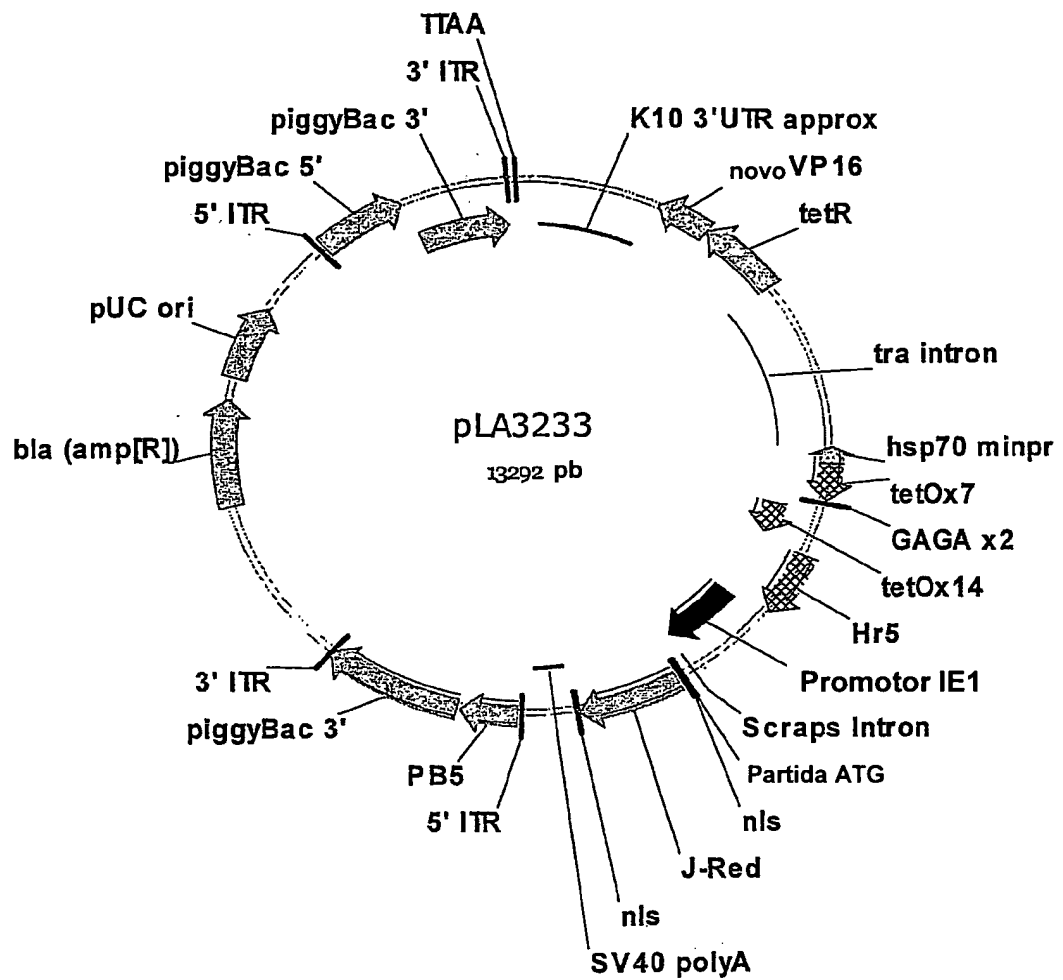


FIGURA 28

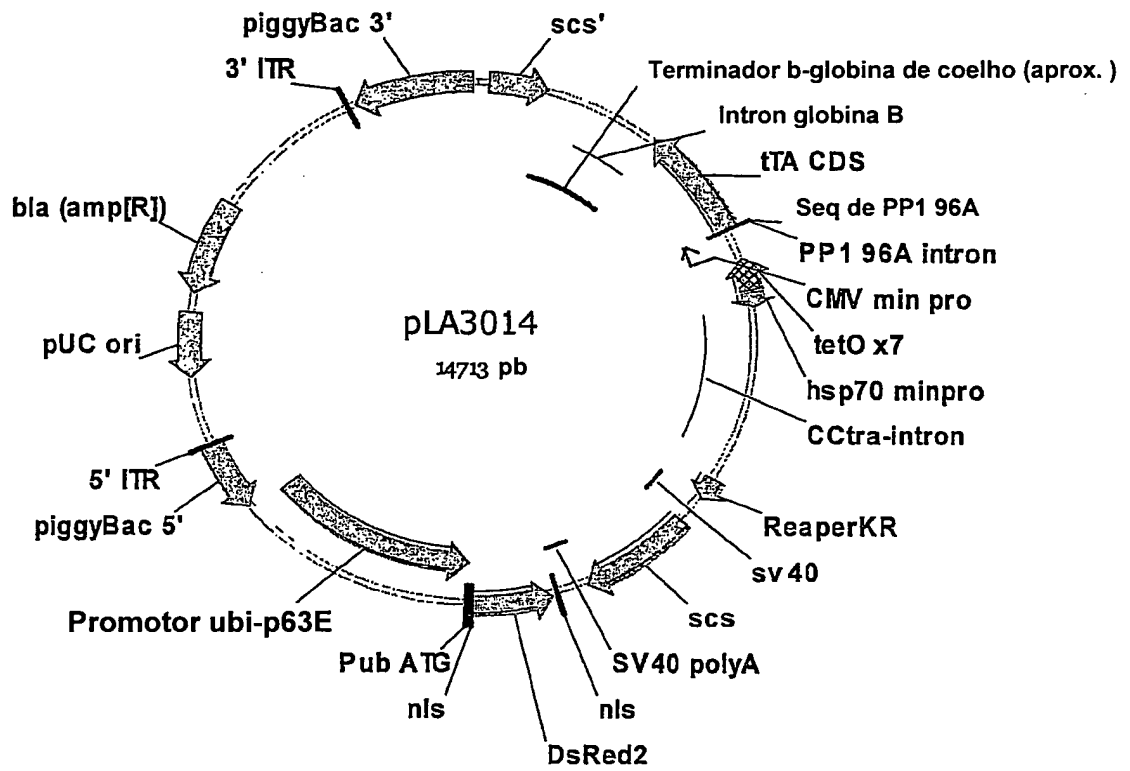


FIGURA 29

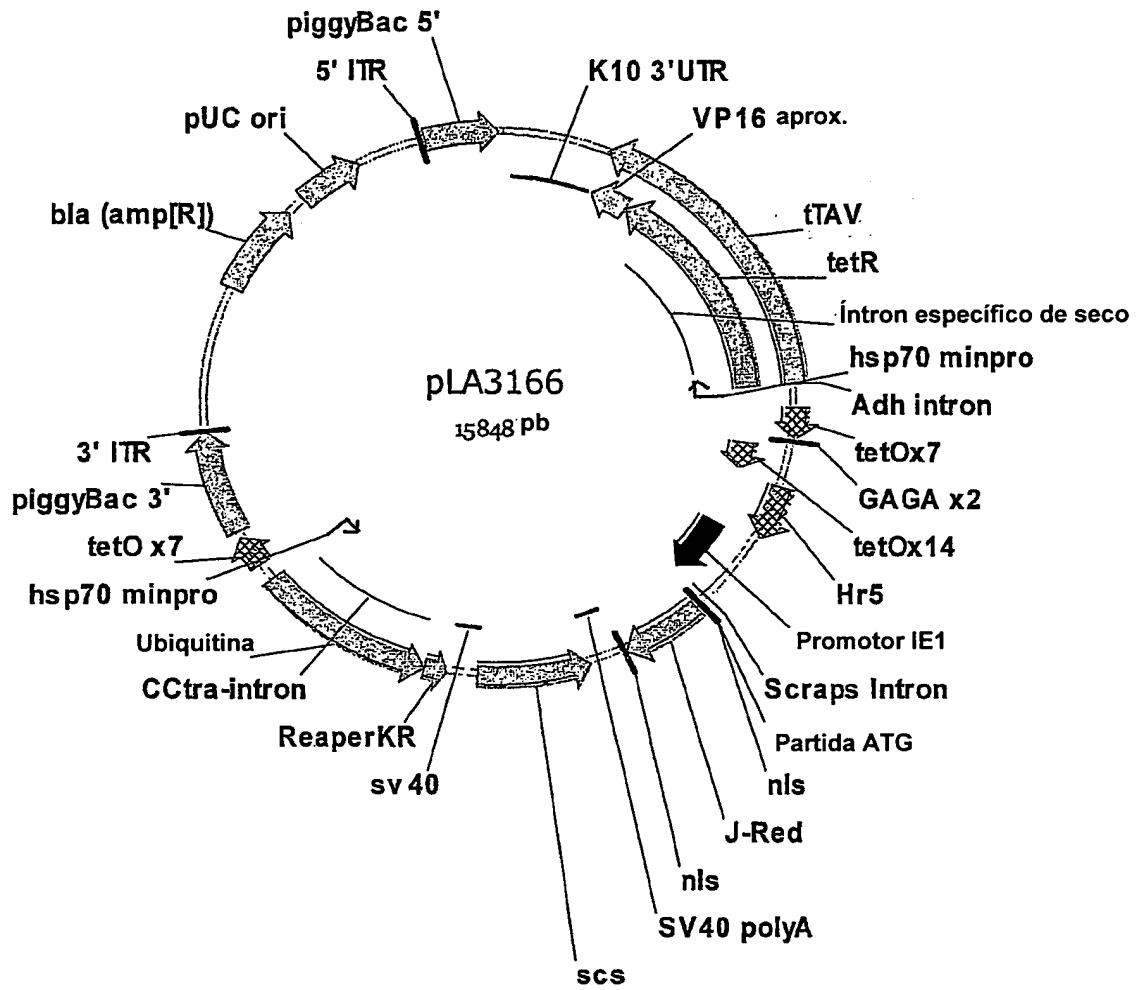


FIGURA 30

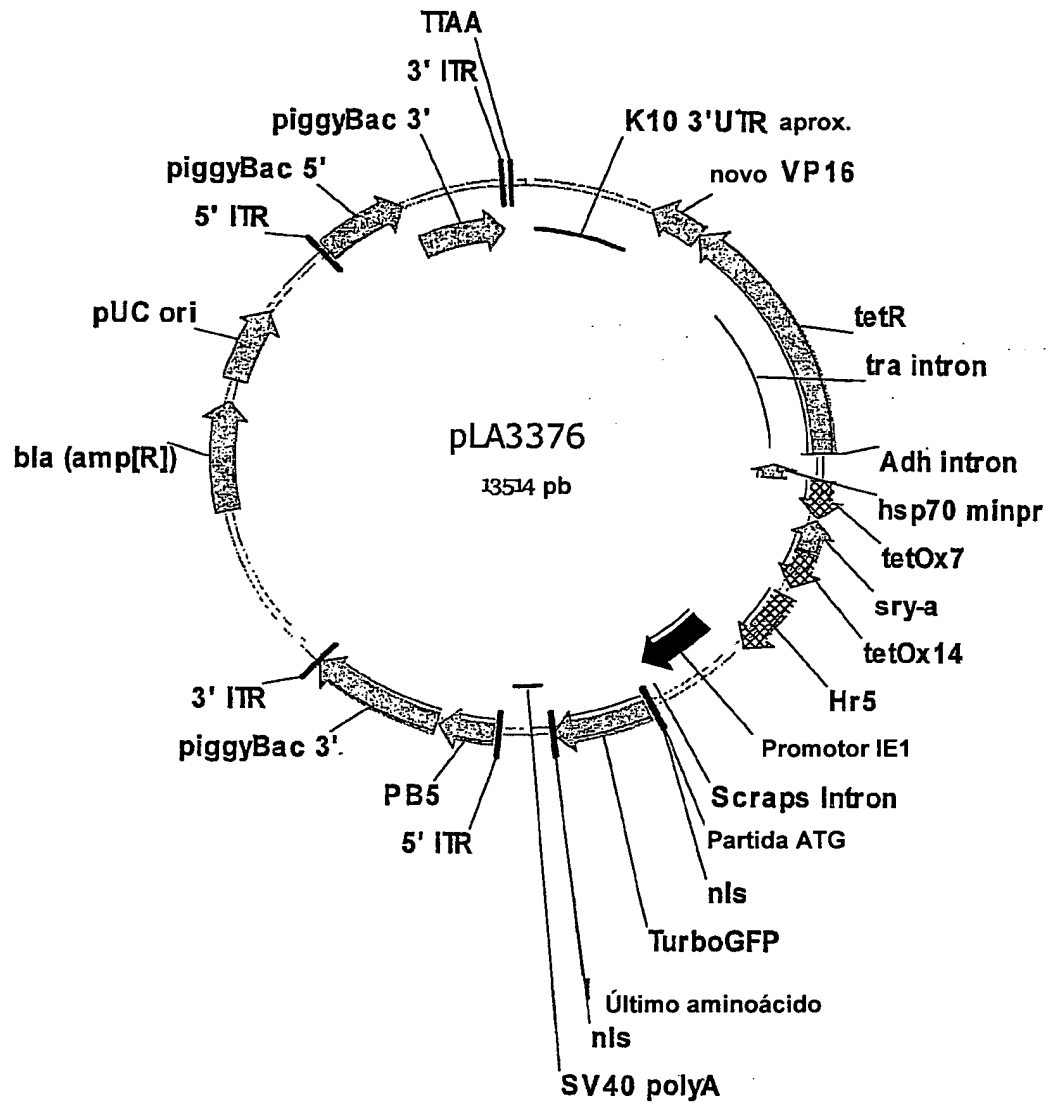


FIGURA 31

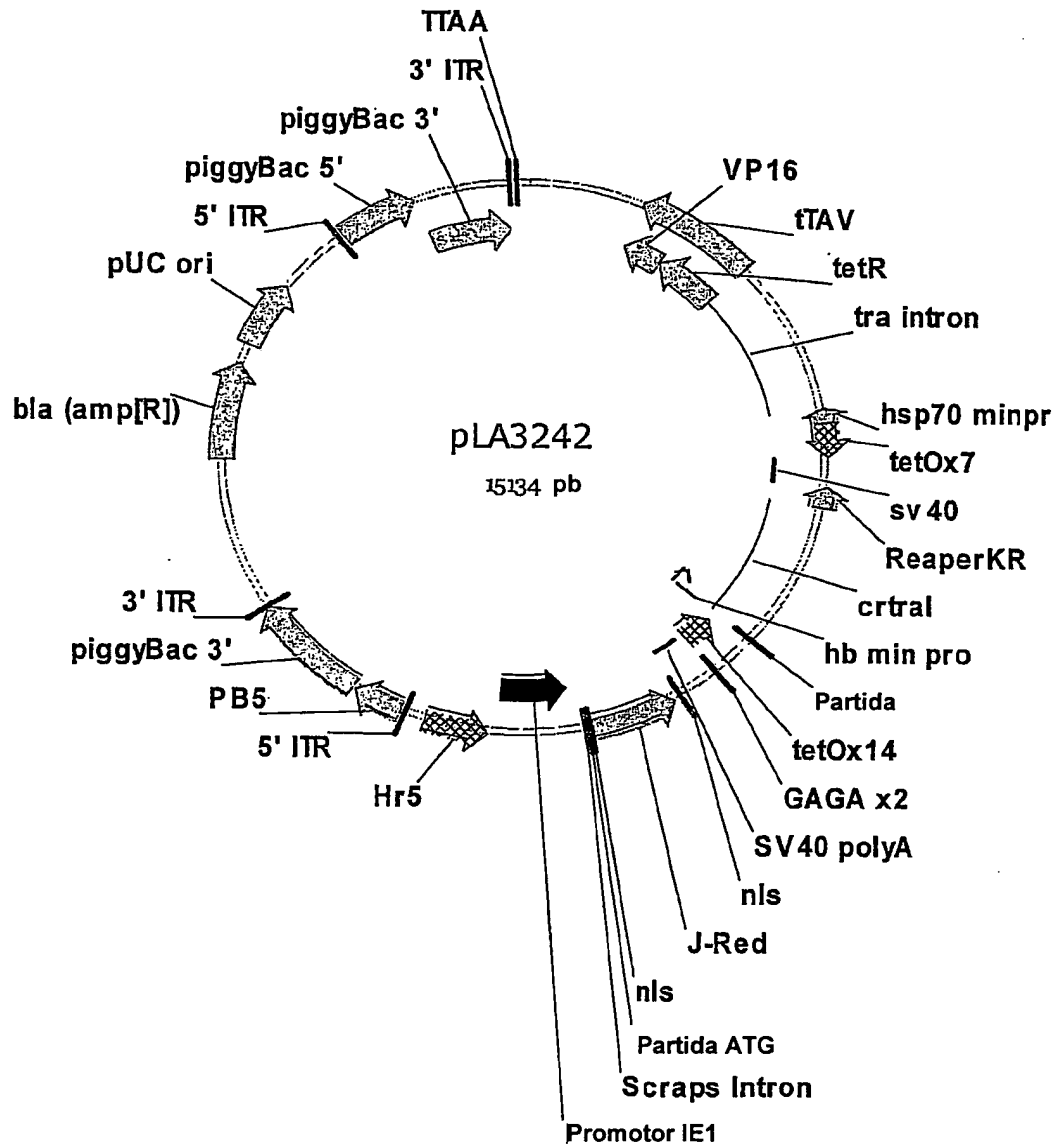


FIGURA 32

Nativo: CGTAGATTG|GT...intron...AG|GTGAAGGCTC
 LA1188: CTA CTG|GCACGT...intron...AG|GTGAAGAATA
 LA3077: AACGAAGTTG|GT...intron...AG|GTATTGAGGG
 LA3097: AGCCACCATG|GT...intron...AG|GTCAGCCGCC

FIGURA 33

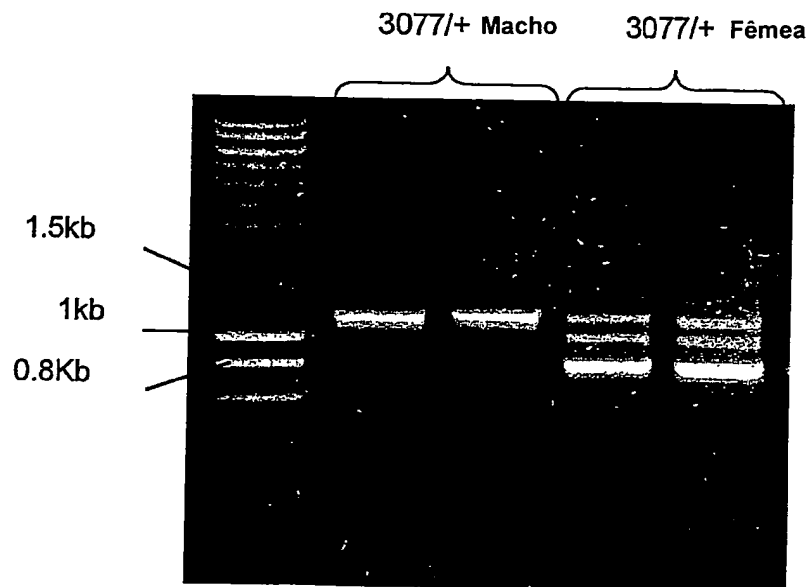


FIGURA 34

3077A	111	32	73	44
3077B	314	157	132	121
3077C	161	116	60	84
3077D	445	85	194	190
3097A	179	5	89	90
3097B	440	0	59	27
3097C	172	0	46	44
3233A	457	1	79	58
3233B	171	0	14	13
3014;1217	136	0	48	10
3166;1217	64	0	5	7

FIGURA 35

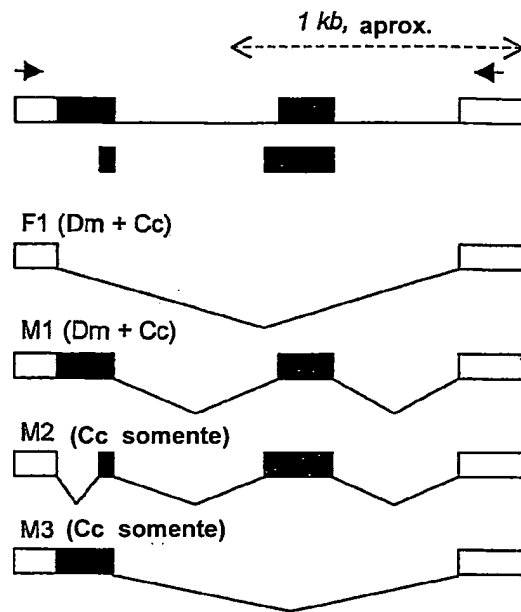


FIGURA 36

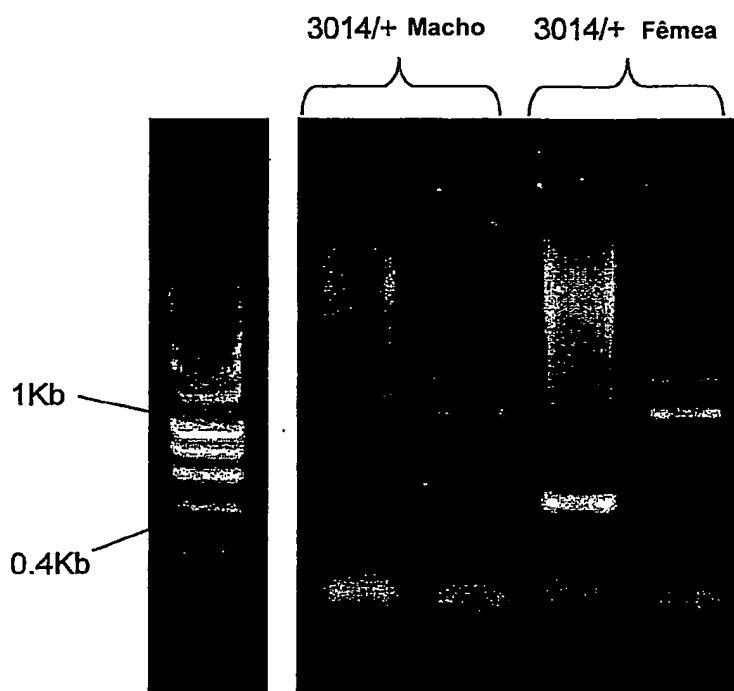


FIGURA 37

	NT Macho	NT Fêmea	TET Macho	TET Fêmea
3097A	136	0	21	19
3097B	295	11	14	11
3097C	96	12	22	21
3097D	103	15	82	67
3233A	78	6	32	5

FIGURA 38

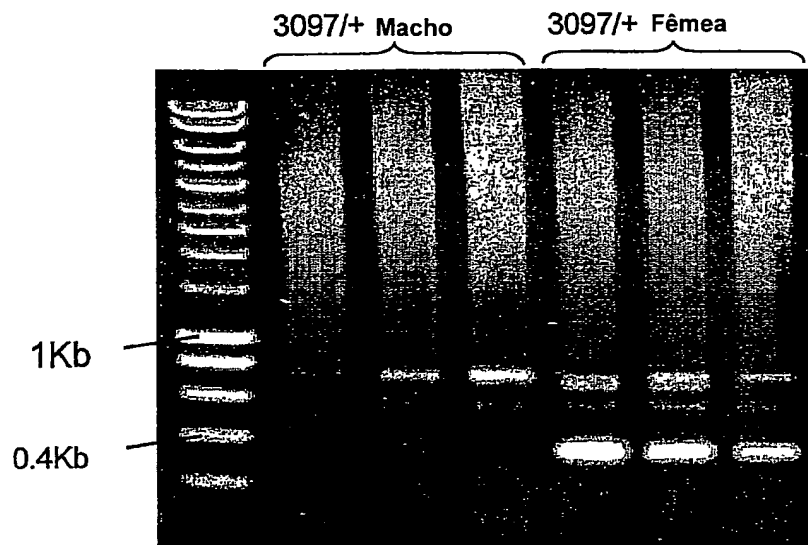


FIGURA 39

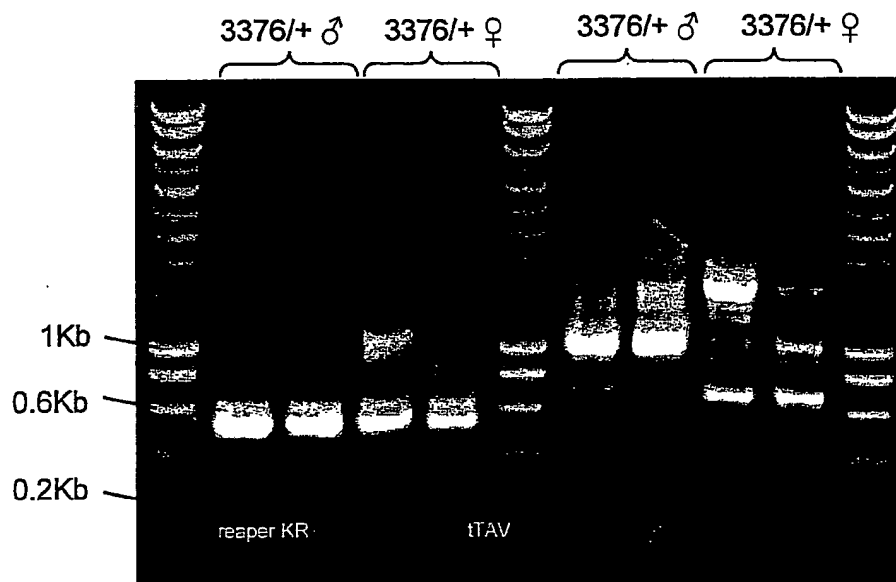


FIGURA 40

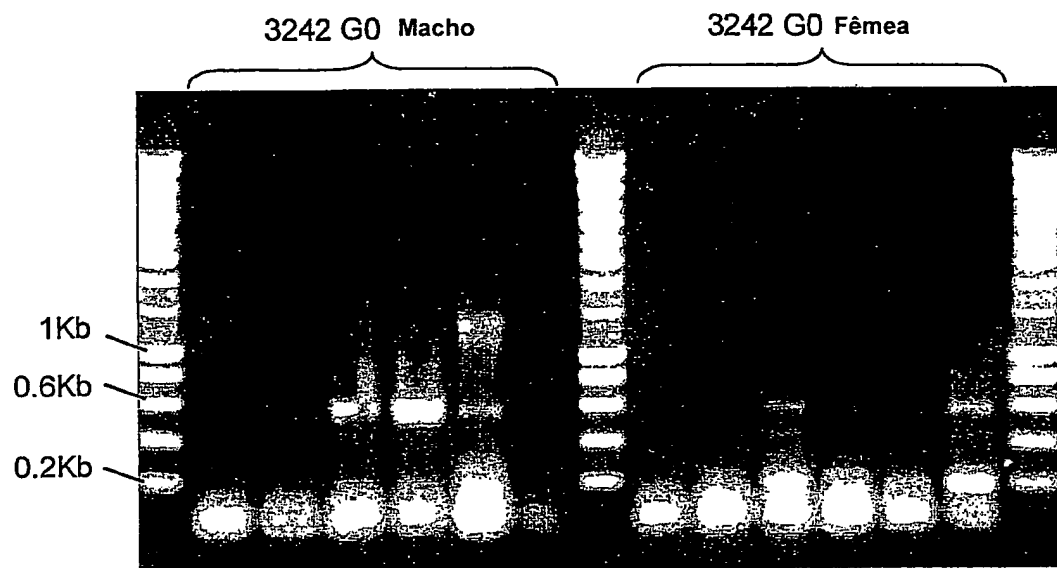


FIGURA 41

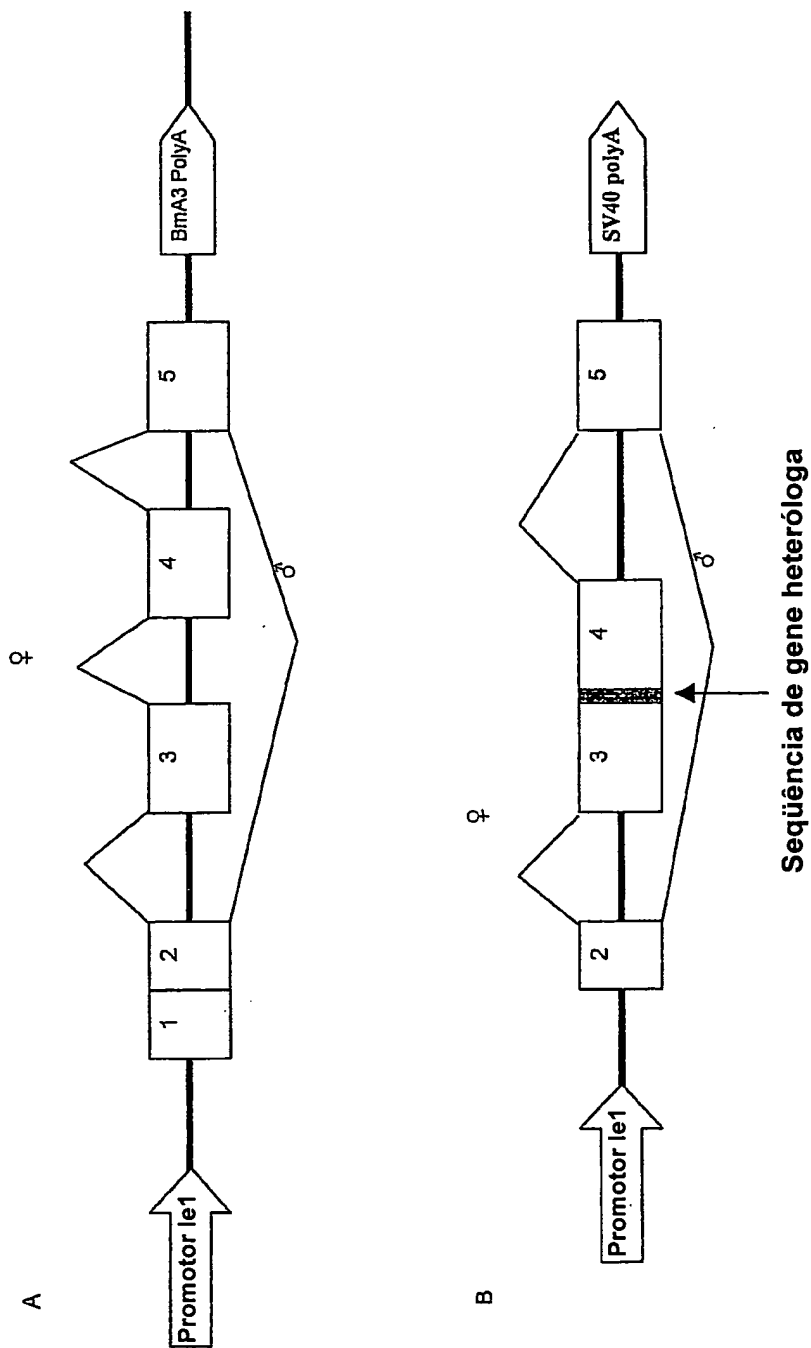


FIGURA 42

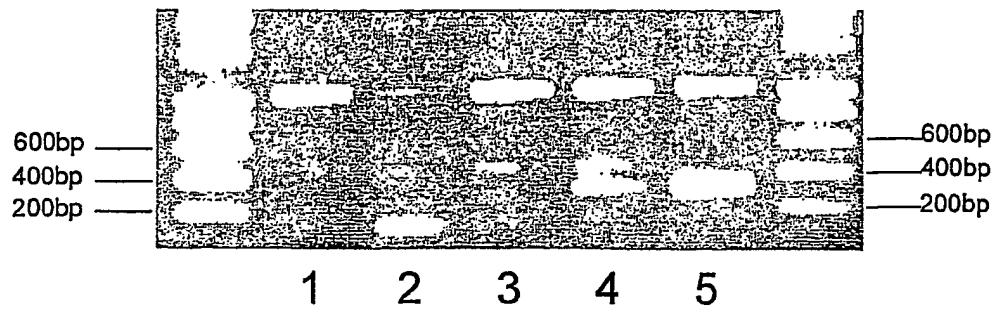


FIGURA 43

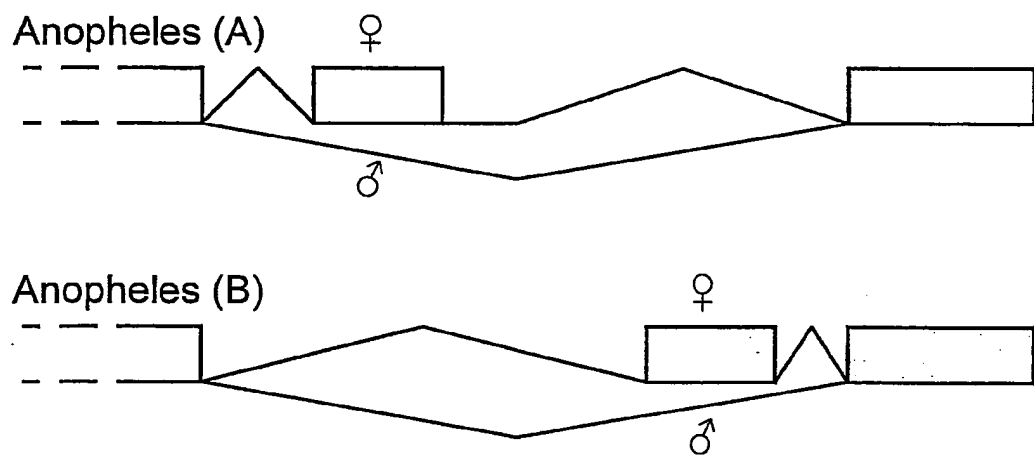


FIGURA 44

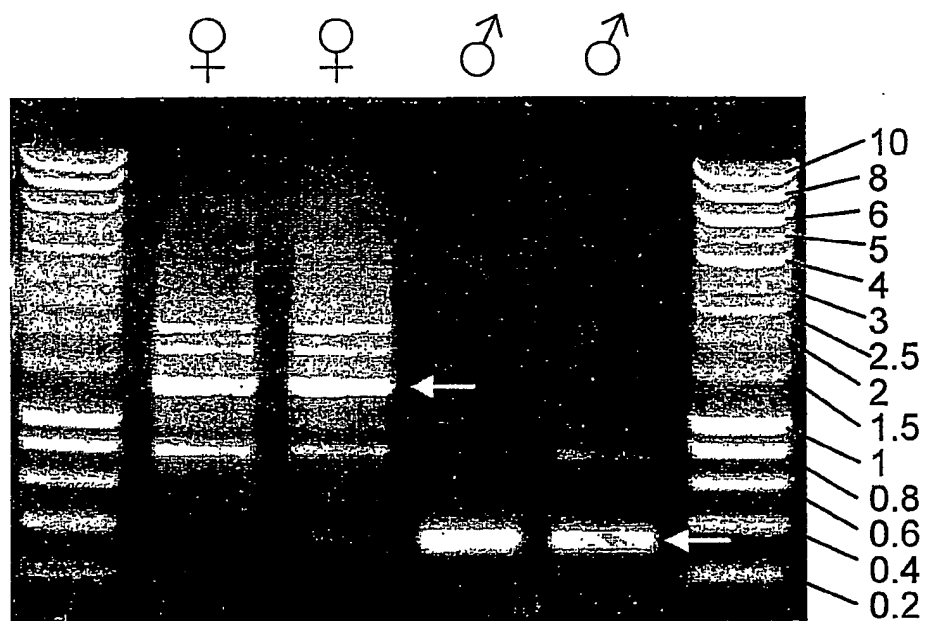


FIGURA 45

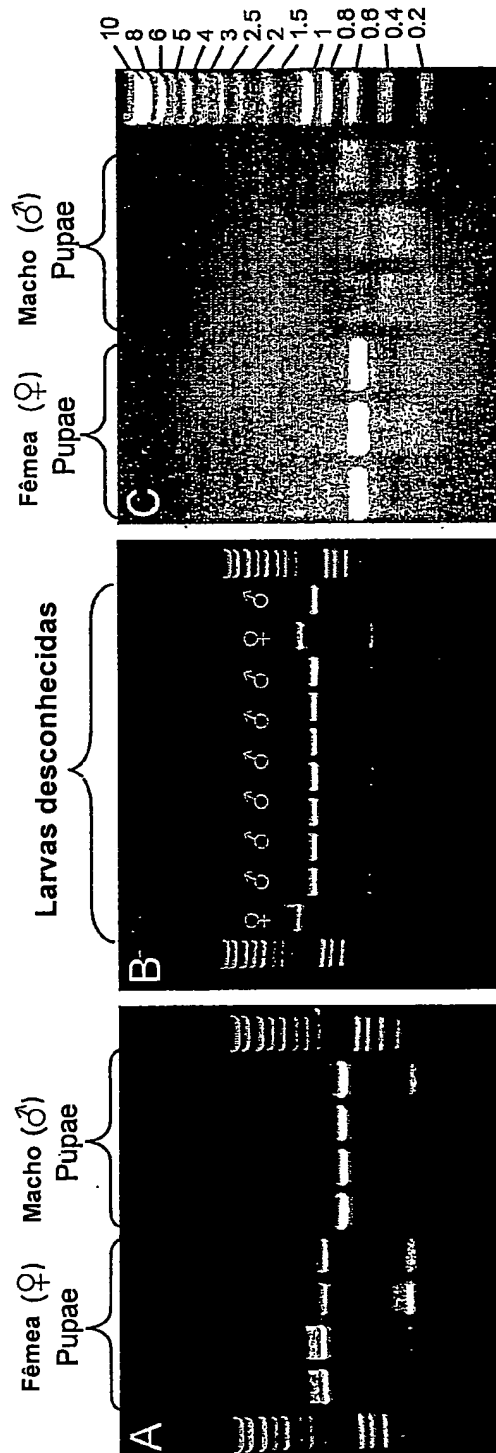


FIGURA 46

Gene *dsx* de *Stegomyia aegypti*

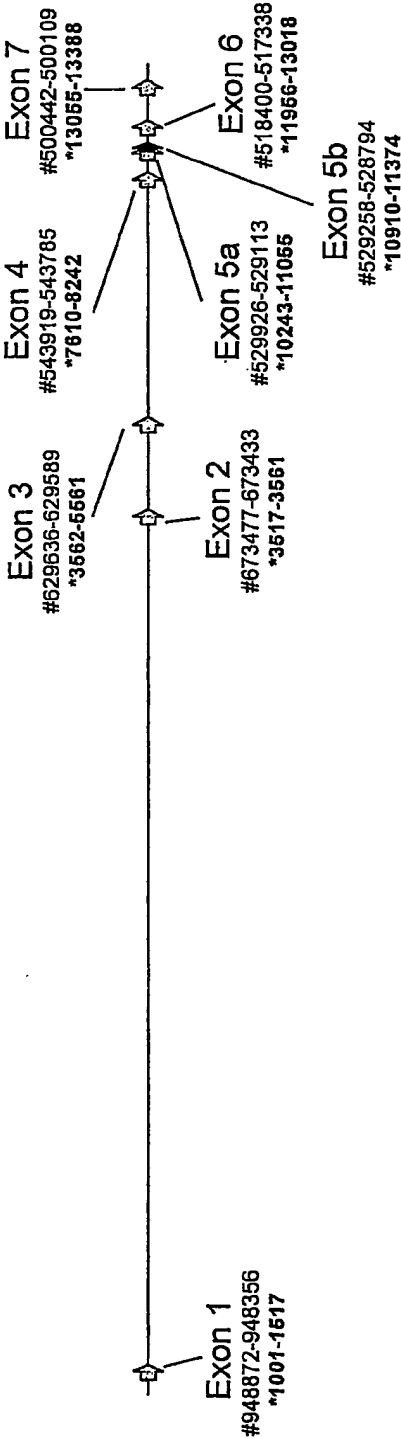


FIGURA 47

Gene *dsx* de *Stegomyia aegypti*

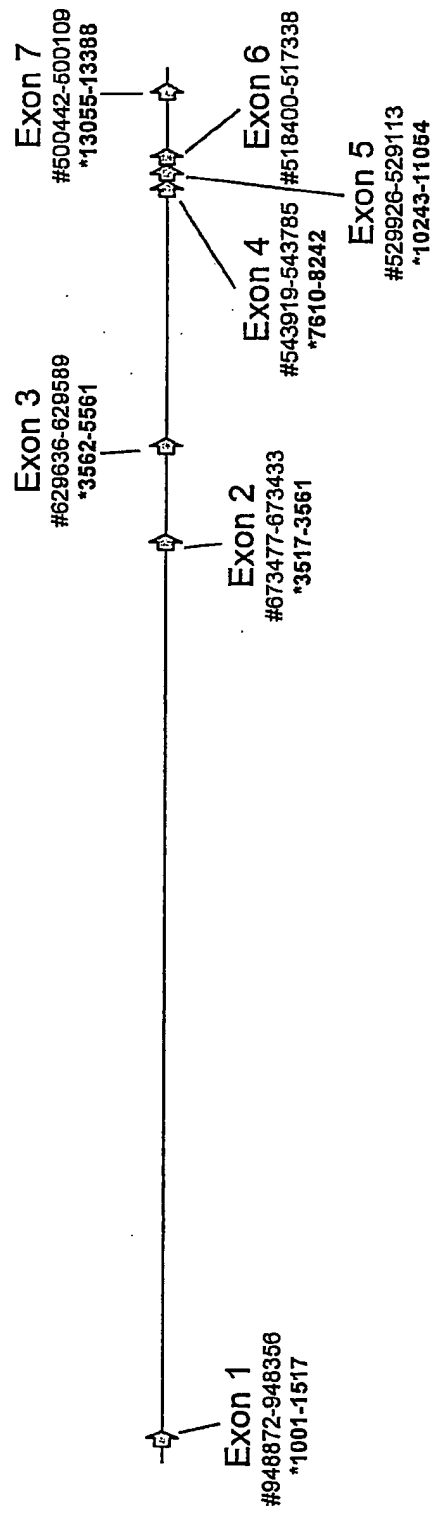


FIGURA 48

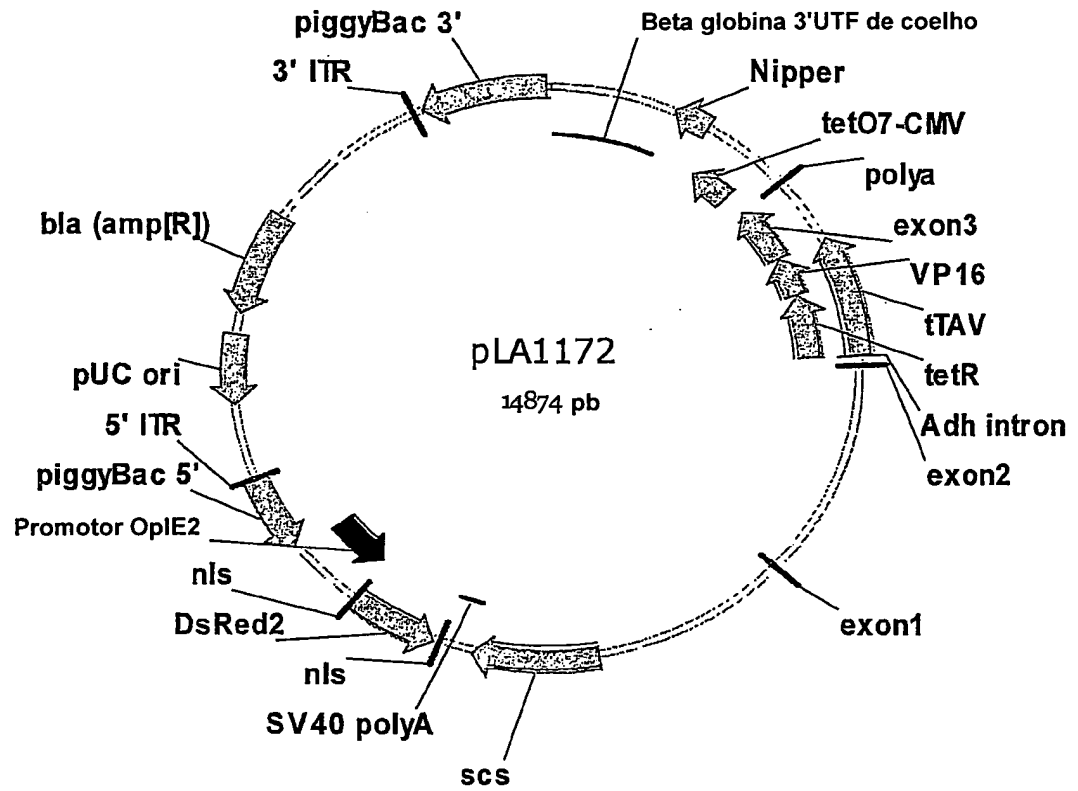


FIGURA 49

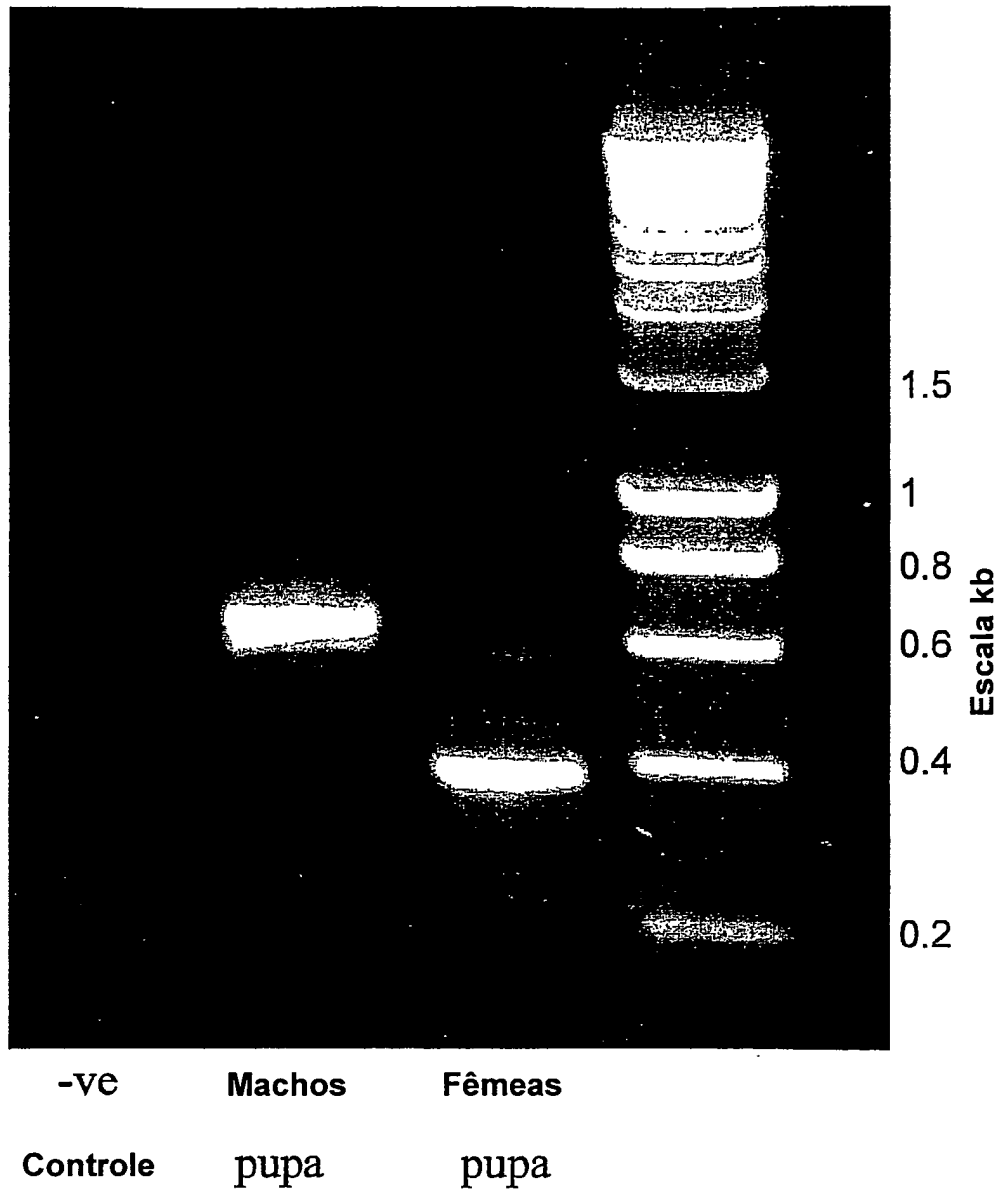


FIGURA 50

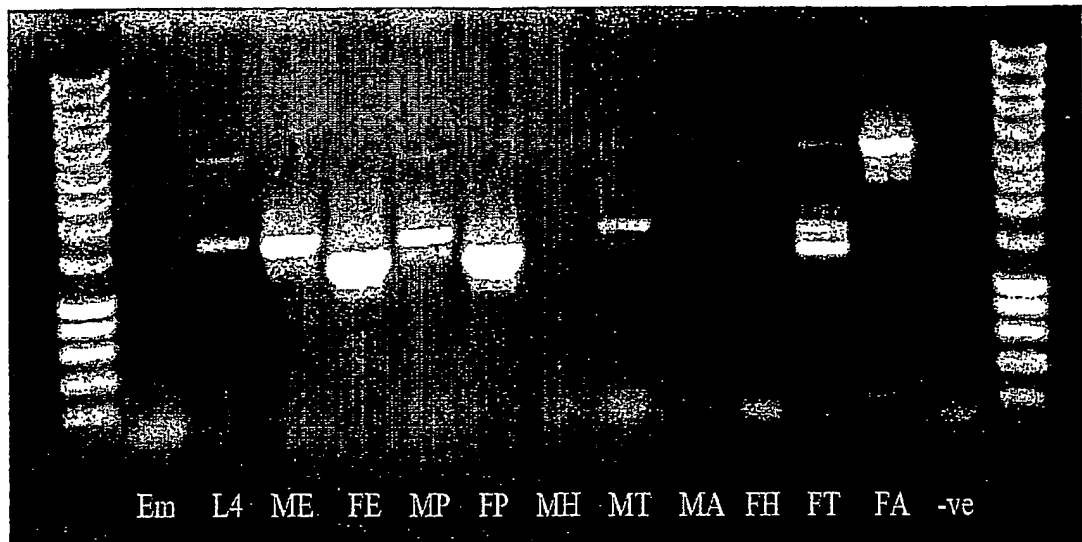


FIGURA 51

FIGURA 52: Mapa Plasmídeo LA3515

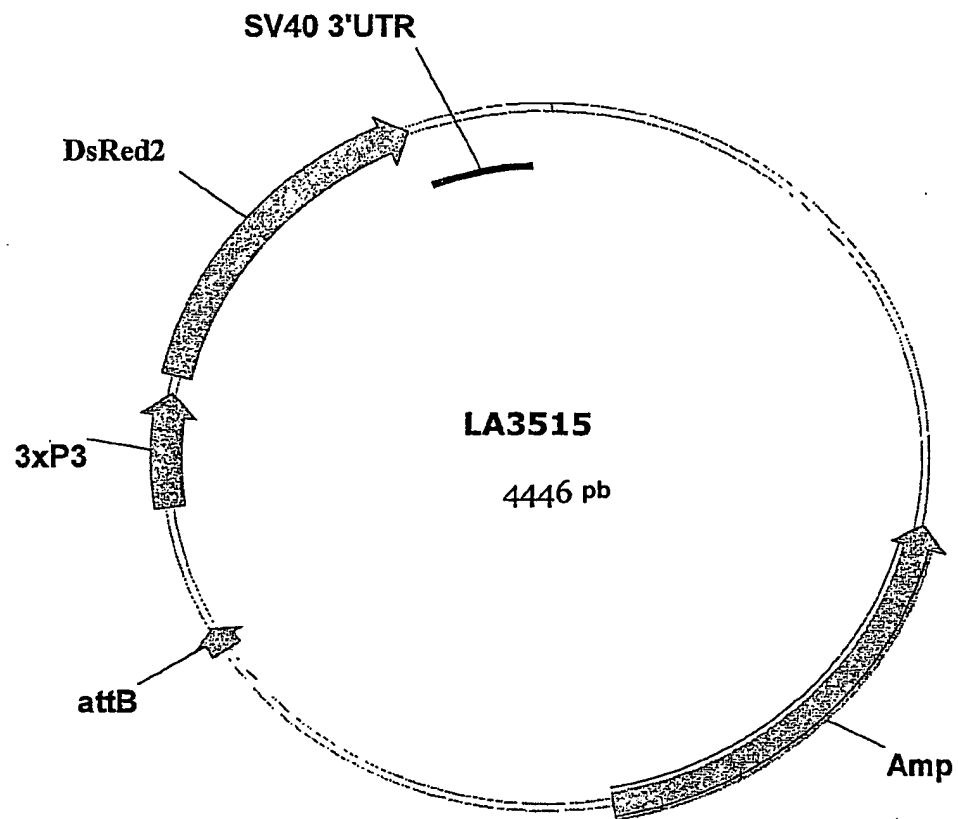


FIGURA 53: Mapa Plasmídeo LA3545

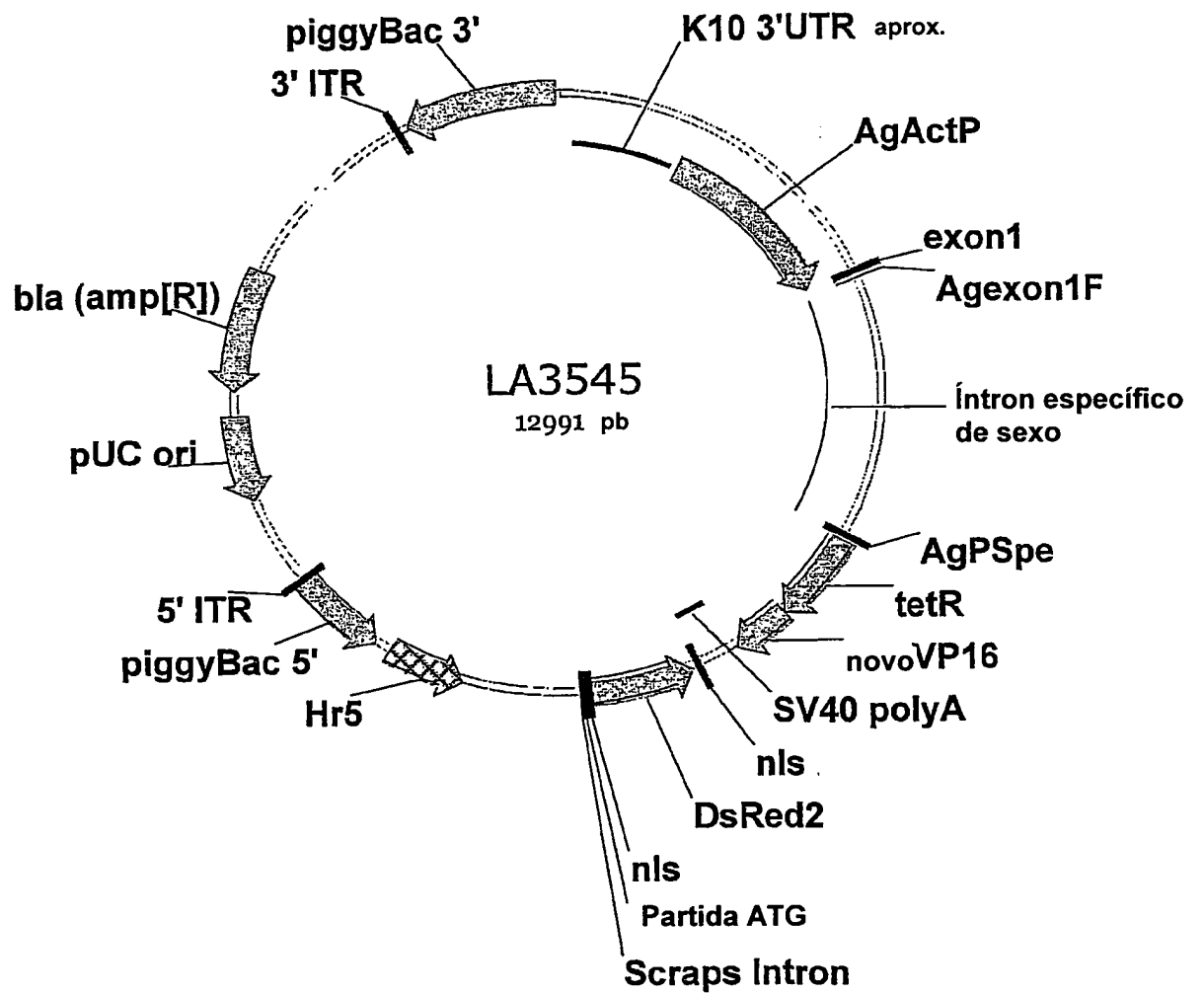


FIGURA 54: Mapa Plasmídeo LA3604

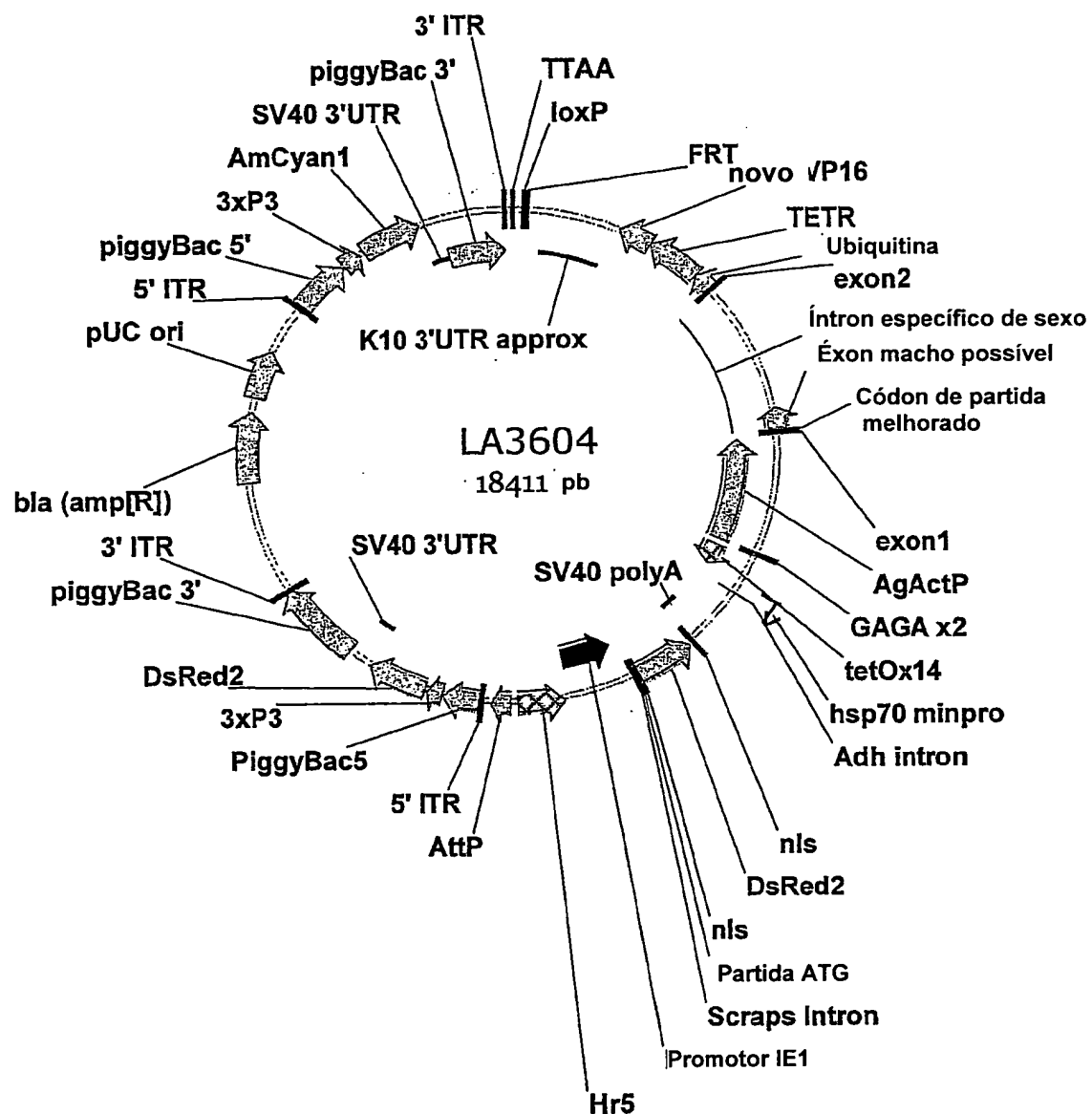


FIGURA 55: Mapa Plasmídeo LA3646

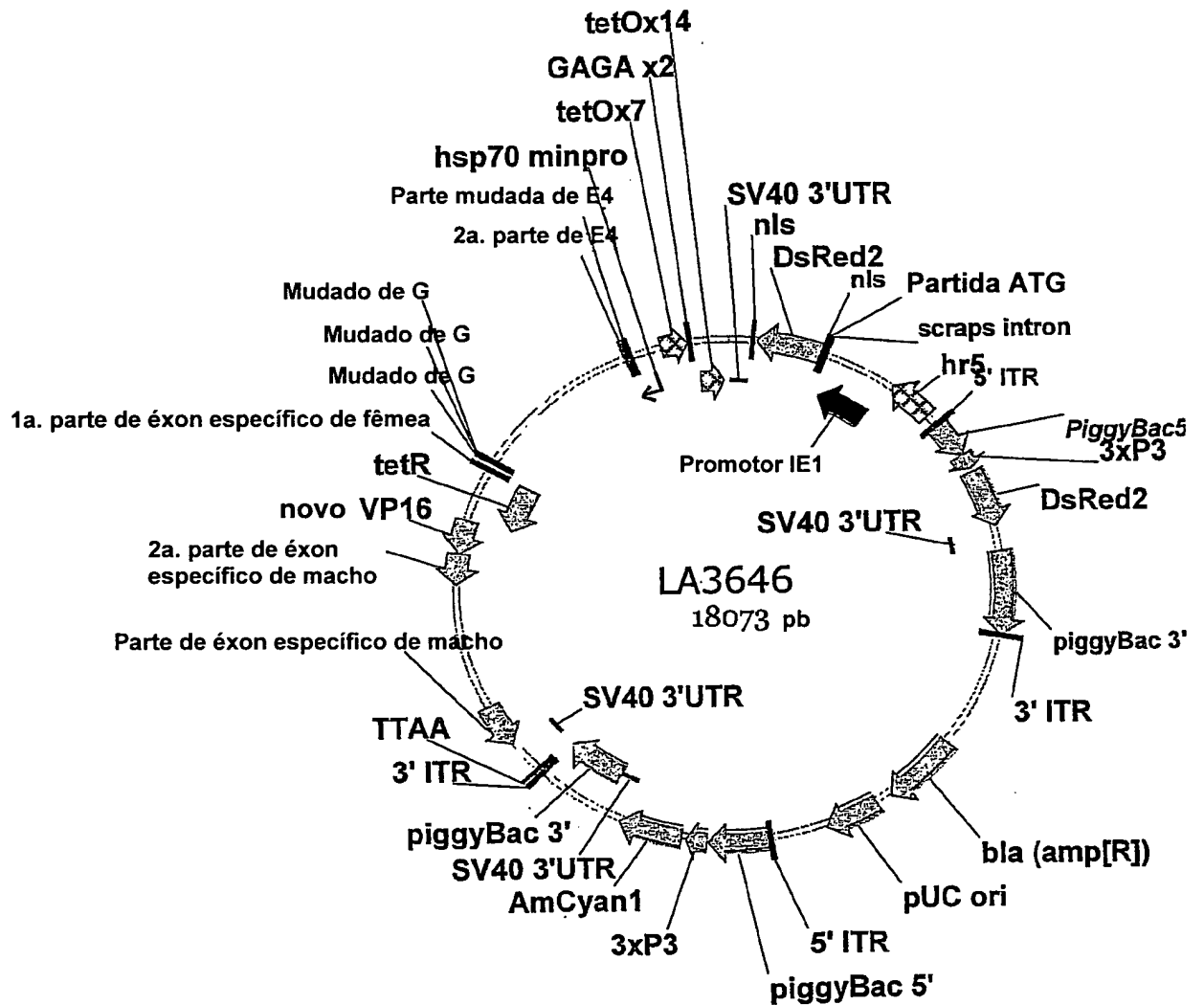


FIGURA 56

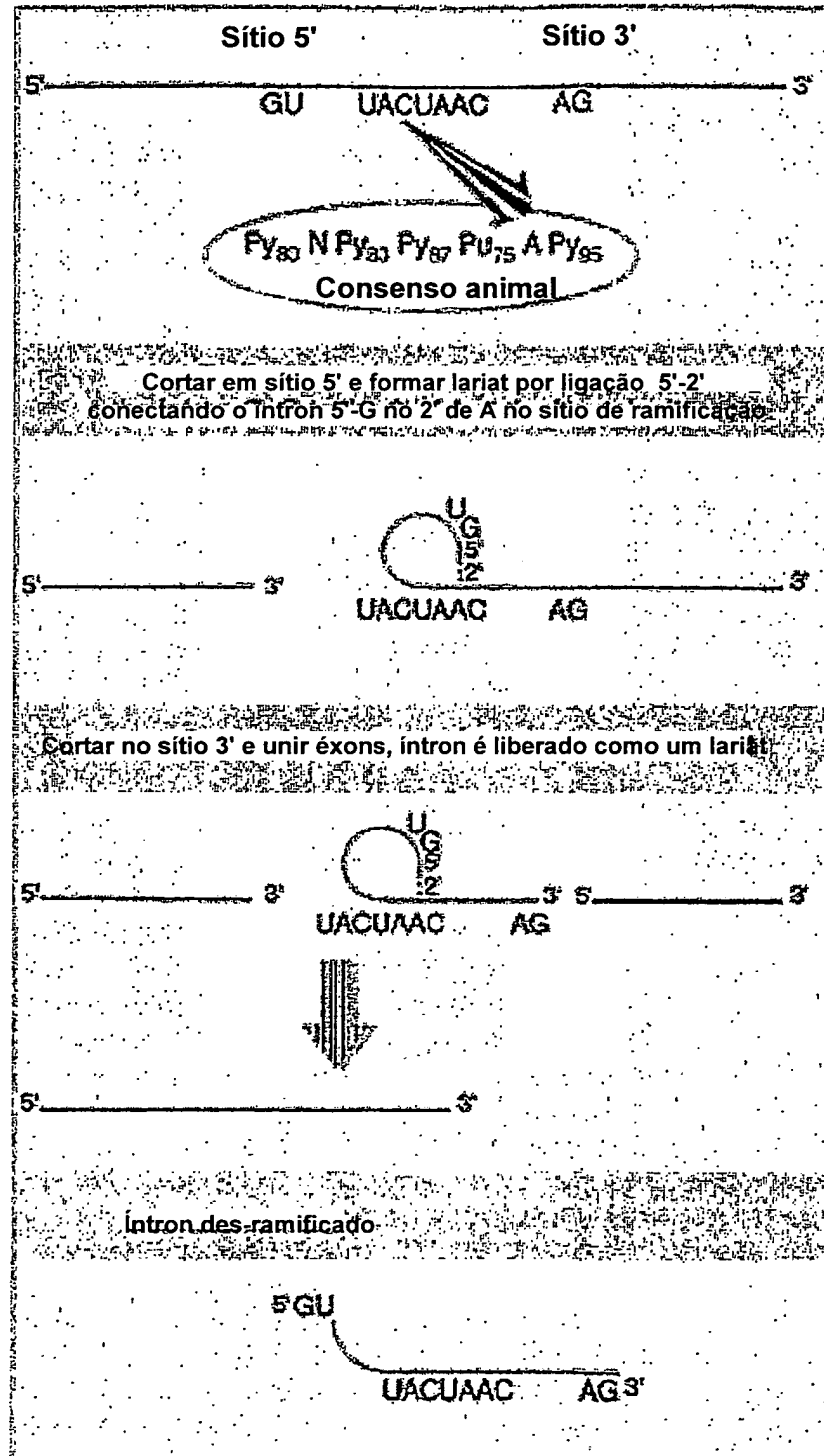
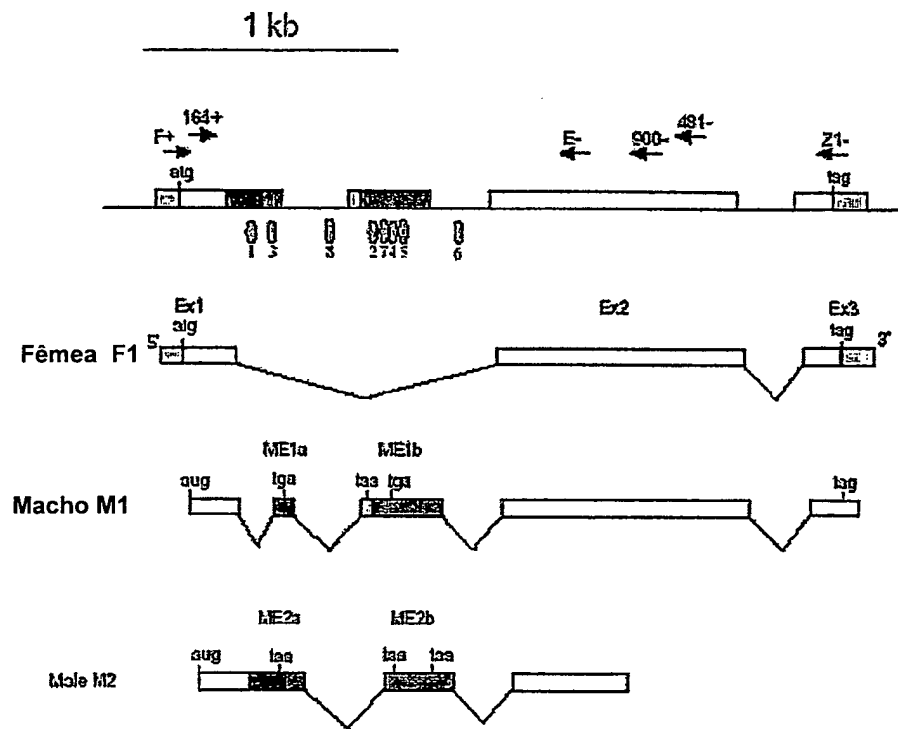


FIGURA 57



A

FIGURA 58

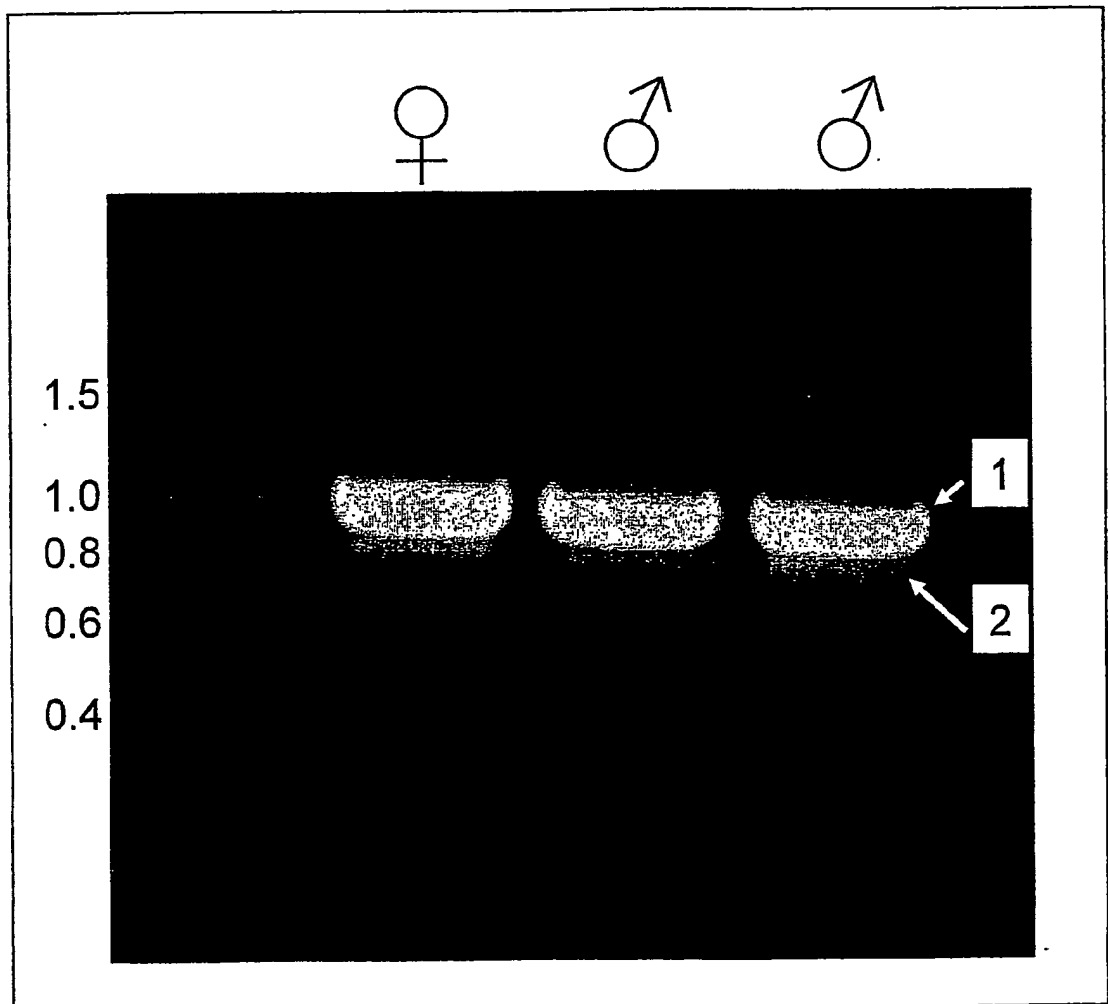
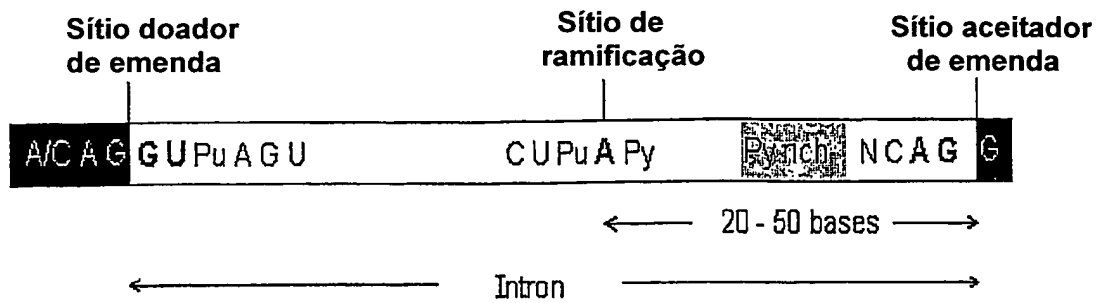


FIGURA 59

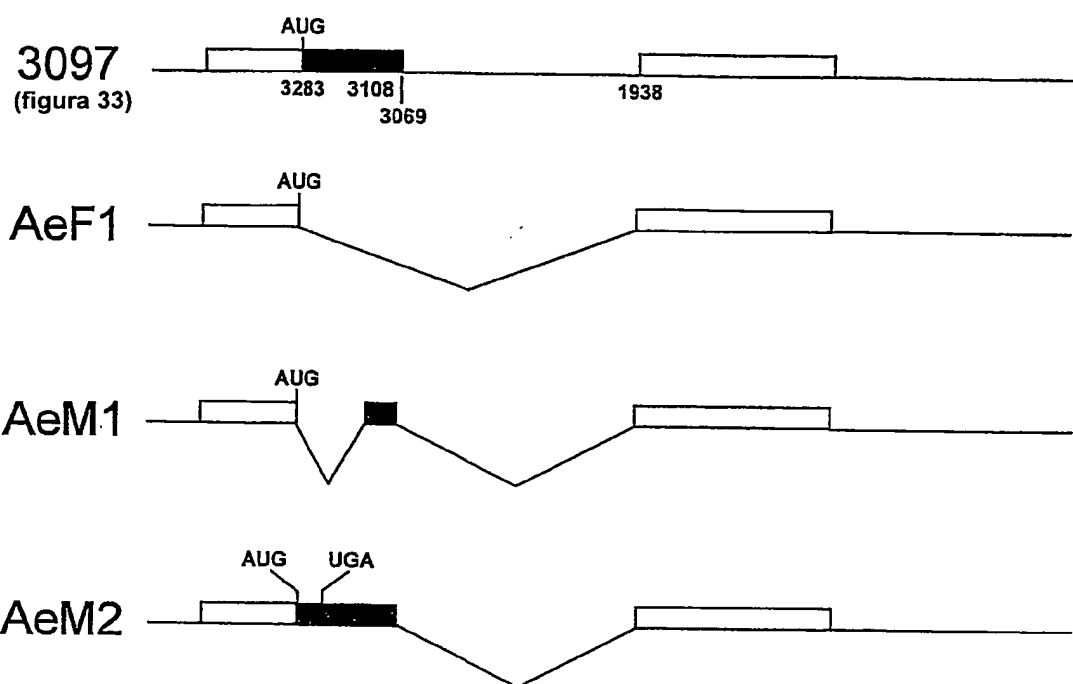


FIGURA 60

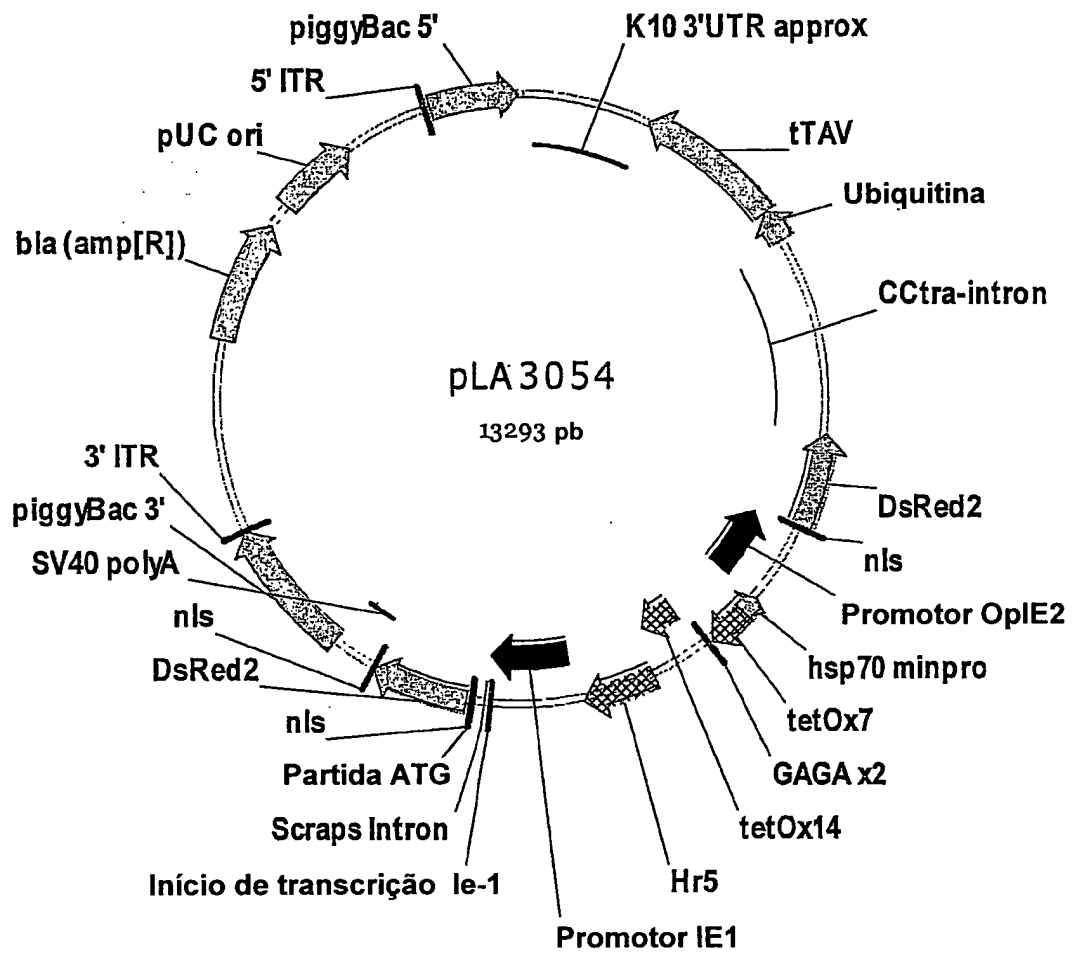


FIGURA 61

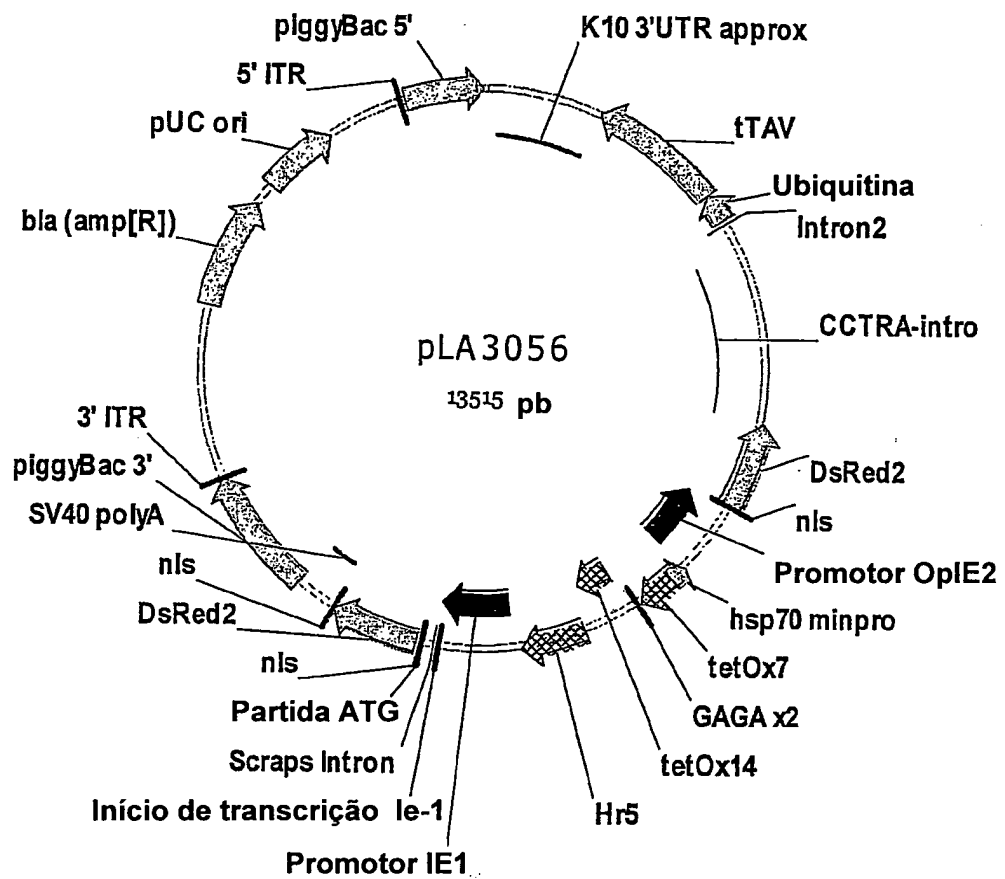


FIGURA 62

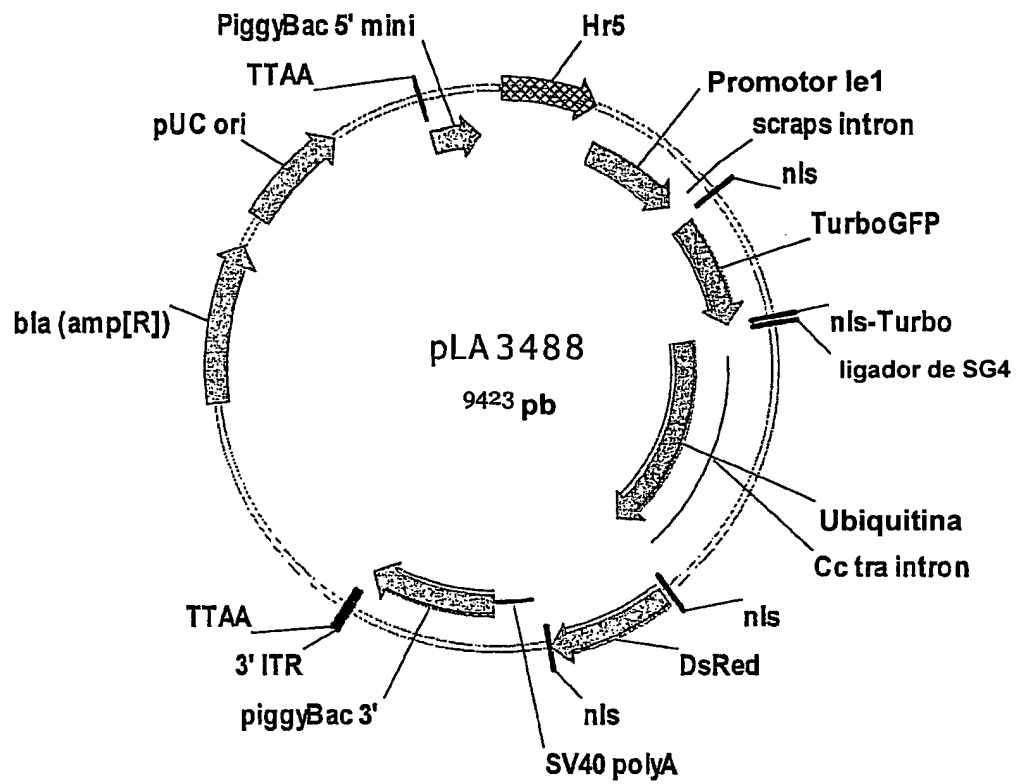


FIGURA 63

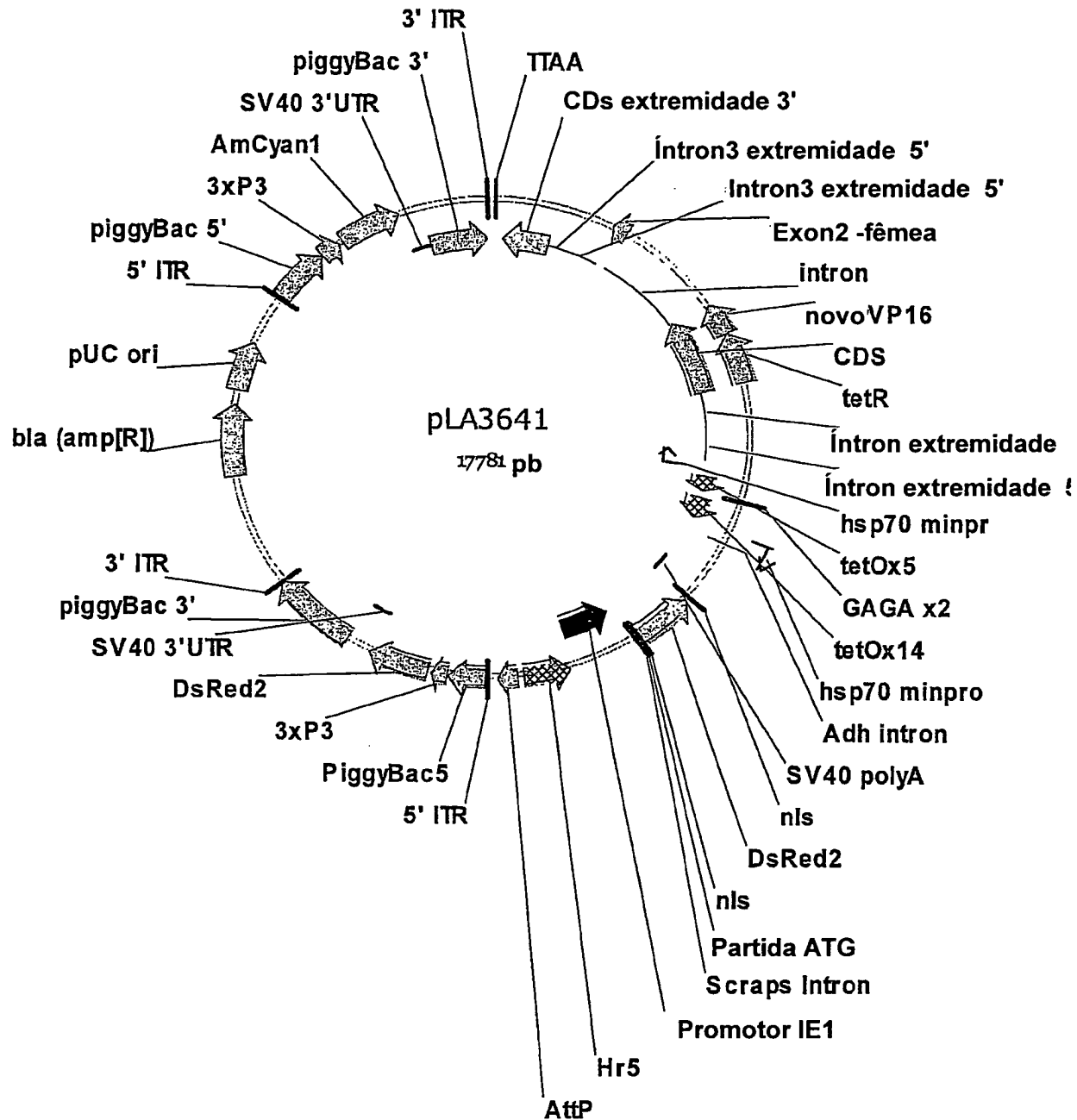


FIGURA 64

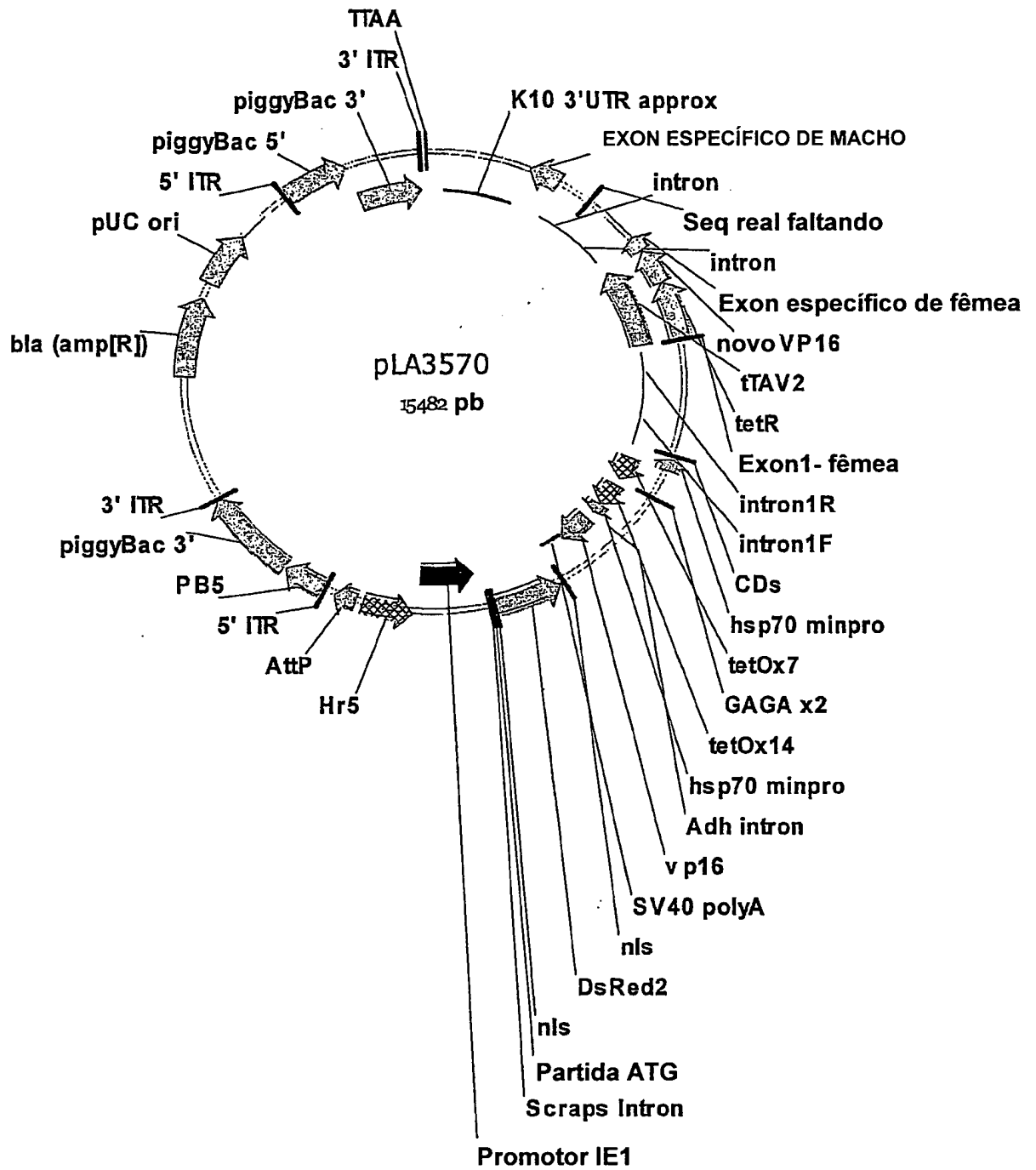


FIGURA 65

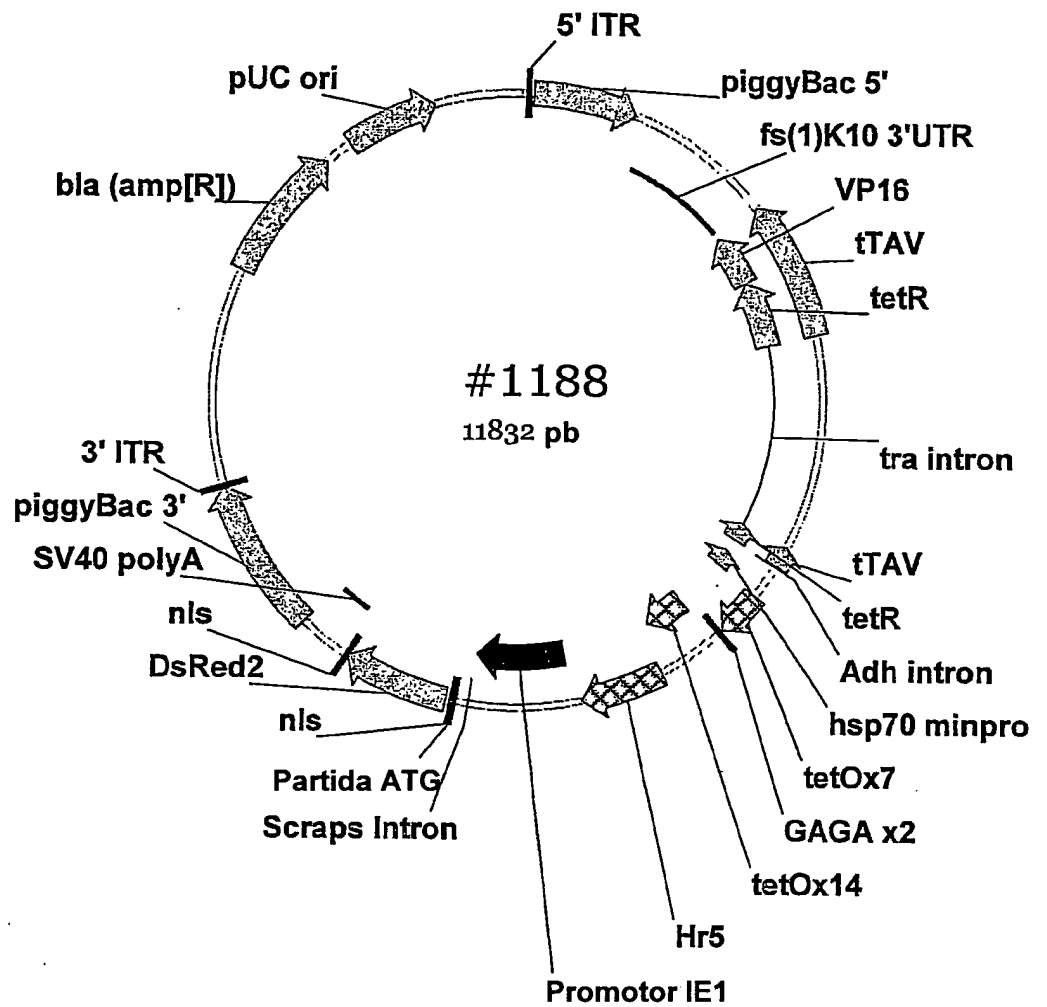


FIGURA 66

**FIGURA 67**

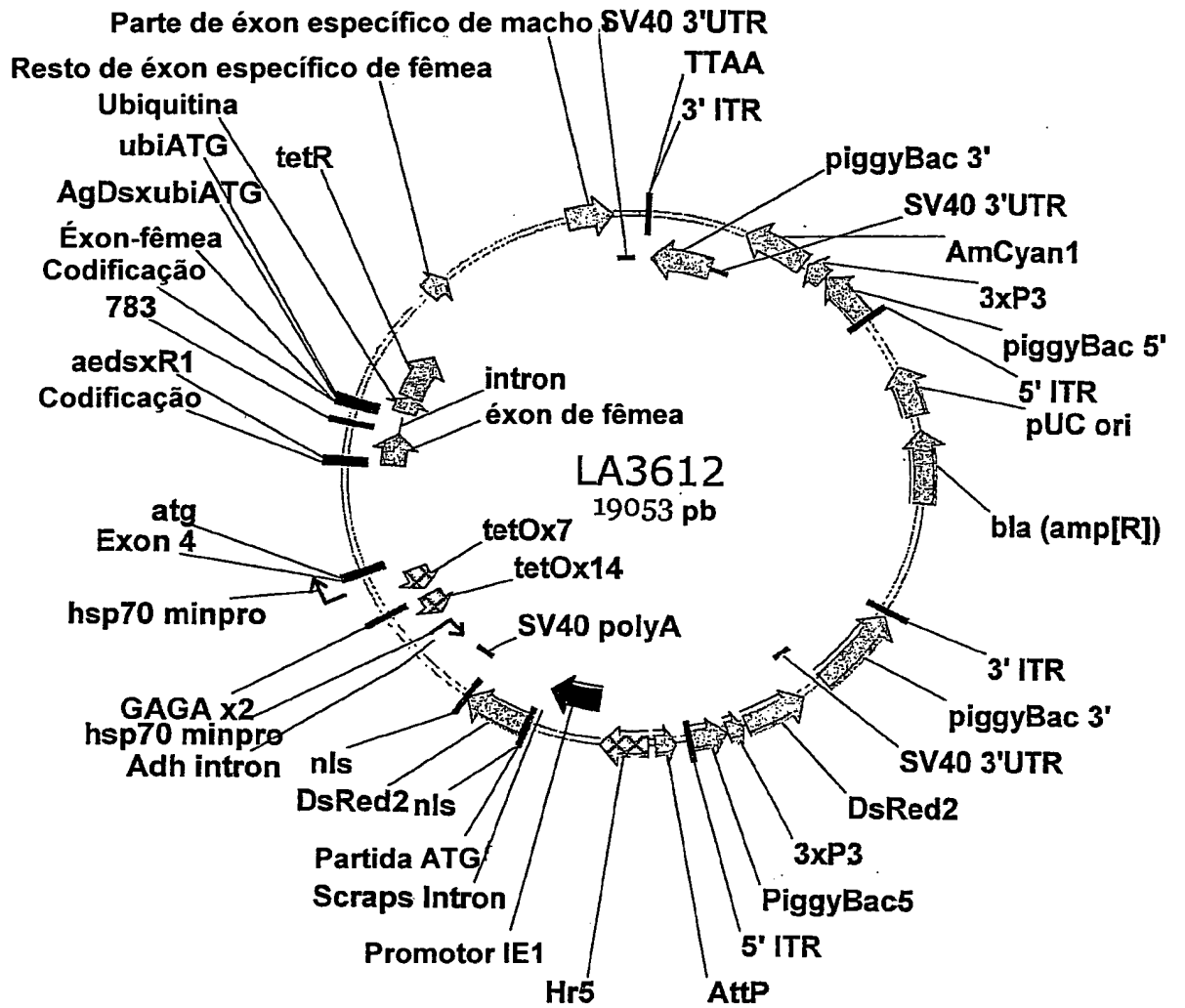


FIGURA 68

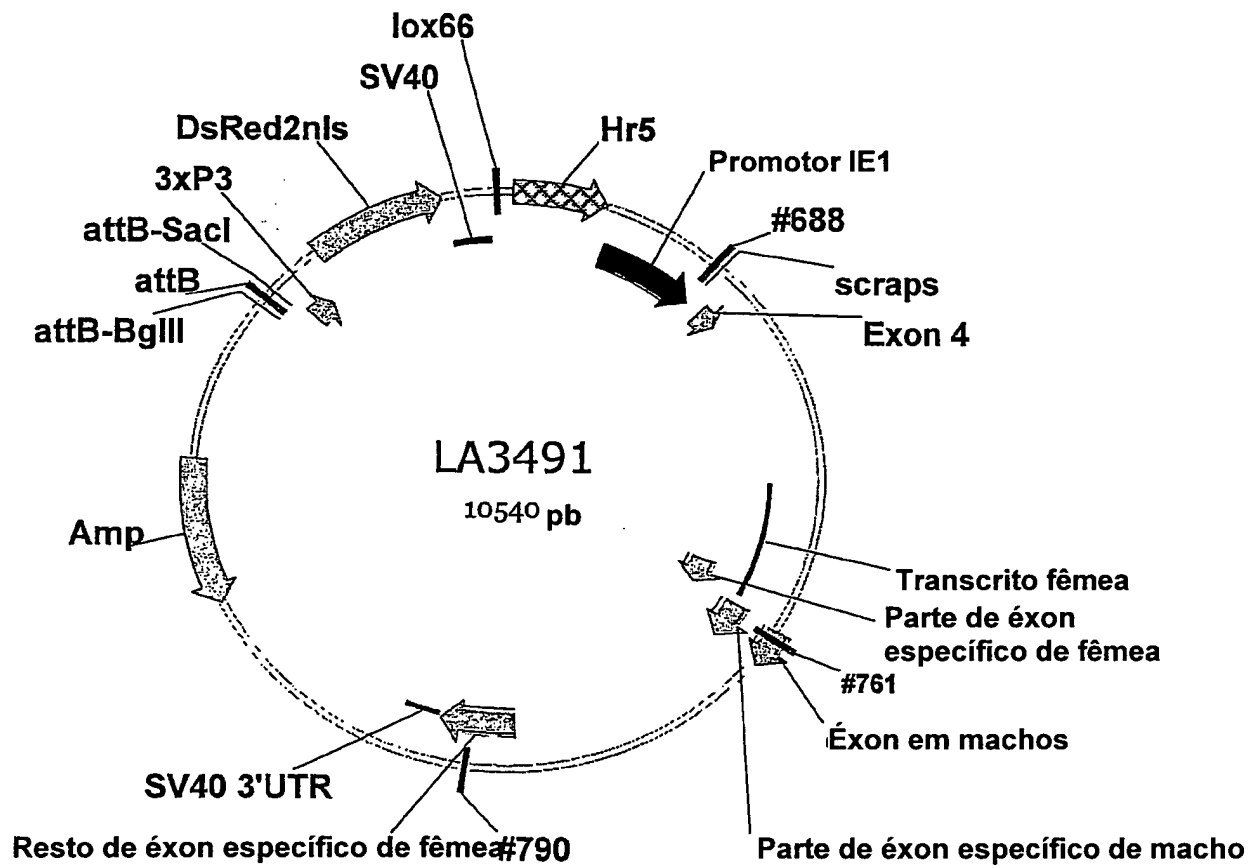


FIGURA 69

RESUMO

“SISTEMA DE EXPRESSÃO DE POLINUCLEOTÍDEOS, MÉTODOS DE
CONTROLE DA POPULAÇÃO DE UM ORGANISMO EM UM MEIO
AMBIENTE NATURAL PARA O MESMO, DE CONTROLE
5 BIOLÓGICO, E DE SEPARAÇÃO DE SEXOS, E, MÉTODO OU
CONTROLE BIOLÓGICO OU DA POPULAÇÃO”

Um sistema de expressão de polinucleotídeos é provido, o qual
é capaz de emendar alternativamente os transcritos de RNA de uma sequência
polinucleotídica a ser expressada em um organismo.