

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

(21) BR 10 2012 015992-9 A2



(22) Data de Depósito: 28/06/2012

(43) Data da Publicação: 23/12/2014
(RPI 2294)

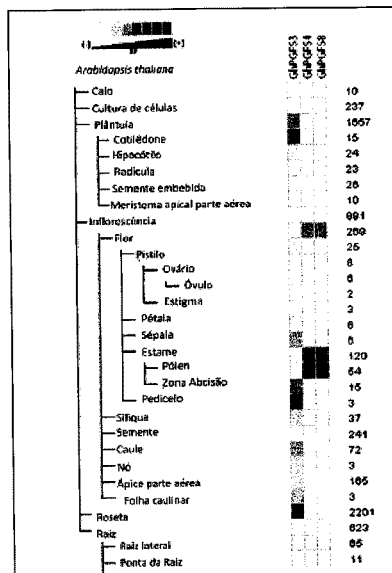
(54) **Título:** COMPOSIÇÕES E MÉTODOS PARA MODIFICAR A EXPRESSÃO DE GENES USANDO PROMOTOR DE GENES ESPECÍFICOS DE FLOR E FRUTO DE PLANTAS

(51) **Int.Cl.:** C12N15/82; C12N15/11; C12N15/29; A01N63/02; A01H5/00

(73) **Titular(es):** Empresa Brasileira de Pesquisa Agropecuária, Universidade Federal do Rio de Janeiro

(72) **Inventor(es):** Marcio Alves Ferreira

(57) **Resumo:** COMPOSIÇÕES E MÉTODOS PARA MODIFICAR A EXPRESSÃO DE GENES USANDO PROMOTOR DE GENES ESPECÍFICOS DE FLOR E FRUTO DE PLANTAS. A presente invenção refere-se a promotores de expressão de genes específicos em plantas. Mais especificamente, a presente invenção proporciona uma sequência de nucleotídeos para a expressão de genes de interesse em flores e frutos de *Gossypium Hirsutum* para produção de plantas geneticamente modificadas que sejam capazes de resistir ao ataque de pragas, tal como bicudo-do-algodoeiro (*Anthonomus grandis*). Nomeadamente, a presente invenção descreve construções de DNA que contenham os promotores dos genes de interesse, para criação de plantas geneticamente modificadas, bem como um método para modificação da expressão gênica.



Relatório Descritivo da Patente de Invenção para
“COMPOSIÇÕES E MÉTODOS PARA MODIFICAR A EXPRESSÃO DE GENES
USANDO PROMOTOR DE GENES ESPECÍFICOS DE FLOR E FRUTO DE
PLANTAS”

5

CAMPO TÉCNICO DA INVENÇÃO

A presente invenção refere-se a promotores de expressão de genes específicos em plantas.

10

Mais especificamente, a presente invenção proporciona uma sequência de nucleotídeos para a expressão de genes de interesse em flores e frutos de *Gossypium hirsutum* para produção de plantas geneticamente modificadas que sejam capazes de resistir ao ataque de pragas, tal como bicudo-do-algodoeiro (*Anthonomus grandis*).

15

Nomeadamente, a presente invenção descreve construções de DNA que contenham os promotores dos genes de interesse, para criação de plantas geneticamente modificadas, bem como um método para modificação da expressão gênica.

ANTECEDENTES DA INVENÇÃO

20

O algodão é a principal fonte de fibra têxtil e a segunda cultura mais importante de semente oleaginosa do mundo (KHAN et al, 2000). Segundo o Departamento de Agricultura Norte-Americano (USDA) 35% da fibra utilizada mundialmente provêm da cultura algodoeira (<http://www.ers.usda.gov/Briefing/Cotton/>, em 01/05/2011), sendo que a espécie *Gossypium hirsutum* é responsável por mais de 95% da produção mundial da fibra (GUO et al, 2008), fatos que tornam essa cultura economicamente importante no que tange a indústria têxtil tanto quanto a agroindústria em geral, devido a extração de óleo para biodiesel.

25

No Brasil, a cotonicultura vem ganhando cada vez mais importância no agronegócio.

30

Segundo o Ministério da Agricultura, Pecuária e Desenvolvimento, as lavouras brasileiras alcançaram o índice de produtividade 60% superior ao dos Estados Unidos, sendo que Mato Grosso e Bahia são os estados responsáveis por 82% da produção nacional e se destacam pelo investimento na área (<http://www.agricultura.gov.br/vegetal/culturas/algodao>, em 01/05/2011).

35

Contudo, a produtividade desta cultura ainda enfrenta grande barreira, devido às pragas que ainda são fator limitante.

O bicudo do algodoeiro (*Anthonomus grandis*) é a praga de maior incidência no cerrado brasileiro e que proporciona os maiores danos (DEGRANDE, 1998).

As fêmeas deste coleóptero ovipositam nos botões florais e maçãs do algodão e as larvas, quando eclodem, se alimentam dos mesmos promovendo sua queda sendo que podem atacar cerca de 50% de botões florais e maçãs de uma plantação de algodão, causando sua inviabilidade na maioria das vezes (RIBEIRO 2007).

O hábito endofítico do bicudo-do-algodoeiro diminui a qualidade da fibra, dificulta o controle químico, e por conseqüência, causa danos econômicos consideráveis.

Dessa forma, o método mais promissor em termos de custo e sustentabilidade para o controle do bicudo-do-algodoeiro é o desenvolvimento de linhagens de algodão, geneticamente resistentes, que possam suprimir o desenvolvimento larval do inseto.

A manipulação de características de uma determinada cultura se tornou possível mediante o advento da biotecnologia e da engenharia genética.

Já é de conhecimento do estado da técnica que diversos genes provenientes do *Bacillus thuringiensis* (Bt) já foram utilizados para a produção de plantas transgênicas resistentes a uma série de insetos.

A bactéria *B. thuringiensis*, proveniente do solo, forma esporos durante sua fase estacionária de crescimento. Esses esporos contêm cristais formados por proteínas Cry de um ou mais tipos que possuem uma potente e específica atividade inseticida. A toxina Bt é ingerida pelo inseto na forma de protoxina que, em contato com o pH alcalino do aparelho digestivo do animal é clivada na sua forma ativa. Uma vez ativada, a toxina se liga a parede do intestino promovendo a abertura de poros. O movimento de solutos através da membrana do tecido epitelial é desbalanceado, causando a morte do inseto (SANAHUJA et al., 2011). Dentre as proteínas Cry, a maioria do tipo-Cry1I quando expressa em um sistema homólogo (por exemplo *Bacillus*) ou em sistema heterólogo (por exemplo *E. coli* e planta transgênica) exibe atividade principalmente contra o grupo de insetos Lepidóptera e raramente contra o grupo Coleóptera, o qual pertence o bicudo. No entanto, a descoberta e caracterização de um novo gene, o gene cry1Ia12, proveniente da cepa S811 de *B. thuringiensis* tornou-se um potencial agente biotecnológico para o controle desta praga (SANAHUJA et al, 2011). A proteína recombinante deste gene, Cry1Ia12, expressa em *Escherichia coli* demonstrou atividade contra larvas de *A. grandis*, demonstrando o potencial deste gene para o uso em transgenia, visando o controle deste importante inseto-praga (GROSSI-DE-SÁ et al., 2007). Muitas culturas já foram geneticamente manipuladas para expressarem proteínas Cry no combate a pragas, inclusive algodão (HOFFMANN et al. 1992). Os genes codificantes dessas proteínas foram

majoritariamente expressos sob o controle do promotor constitutivo CaMV35S, originalmente isolado do vírus do mosaico da couve-flor.

5 A linhagem de algodão portando o gene Bt (US7345229) é resistente ao ataque de *Heliothis zea*, *Pectinophora gossypiella*, *Heliothis virescens* e *Trichoplusia ni*. Apesar de abrangente, tal linhagem não confere resistência ao *A. grandis*, a principal praga de algodão, destacadamente no Brasil.

10 Dessa forma, se faz necessária a construção de uma nova linhagem de plantas, dotadas de flores e frutos, resistente ao ataque de insetos e pragas, principalmente a obtenção de uma nova linhagem de algodão, resistente ao ataque do bicudo-do-algodoeiro, um dos objetivos da presente invenção.

15 O promotor é o principal responsável pelo mecanismo de regulação da expressão de um gene, pois possui sítios de ligação para fatores transcricionais (TF) e para a própria RNA polimerase, responsável pela transcrição do gene. Diversas regiões de promotores, denominadas elementos cis, foram caracterizadas como sítio de ligação de diferentes TFs, dentre elas a mais bem caracterizada é o elemento TATA-box. A pluralidade desses elementos permite a modulação temporal e espacial da expressão de um gene. Poranto, a região promotora consiste em uma importante ferramenta biotecnológica, pois permite o direcionamento da expressão do gene de interesse em uma linhagem transgênica.

20 Um exemplo dessa aplicabilidade da região promotora seria o uso do promotor com expressão específica em fruto de tomate (WO1991001324), para direcionar a expressão de genes capazes de melhorar a composição protéica e o valor nutricional, bem como facilitar a estocagem, manuseamento, cozimento e aspecto do fruto.

25 Promotores constitutivos como o CaMV35S são capazes de conduzir a expressão de um transgene em todos os tecidos da planta. Entretanto, o seu nível de expressão não é necessariamente semelhante em todas as células vegetais (SUNIKUMAR et al., 2002), há alterações na expressão da proteína ao longo do desenvolvimento da planta. Esse efeito tem sido relacionado ao processo de metilação do promotor durante o desenvolvimento da planta já, que este permanece ativo desde o início do desenvolvimento. Referido resultado tem sido relatado especialmente quando se inicia a fase reprodutiva, quando ocorre o ataque do *A. grandis* (CHEN et al., 2000). Além disso, níveis variáveis de expressão são observados em diferentes tecidos apesar de o promotor ser considerado constitutivo.

30 Por exemplo, expressão de Cry1A em ovários de algodão está entre os menores encontrados nos diferentes tecidos do algodoeiro. Isso pode ser especialmente crítico para insetos que atacam esse tecido durante o desenvolvimento, como foi observado

35

para o *Hefiothis zea* e que também pode ocorrer para o *A. grandis* (KRANTHI et al., 2005). E, finalmente, impactos negativos em insetos não alvo podem ser minimizados pela diminuição material vegetal expressando Bt, apesar de que até o momento não existe uma prova conclusiva de que o algodão Bt pode ter impacto em insetos não alvo

5 (DALE et al., 2002).

Além disso, a expressão constitutiva de uma proteína heteróloga em uma variedade transgênica demanda um alto custo energético para a planta. Outros promotores constitutivos também foram caracterizados em algodão (EP1953231), em café (US6441273), em *A. thaliana* (US20020115850) e em milho

10 (US6670467).

Dessa forma, o método mais promissor em termos de custo e sustentabilidade para o controle do bicudo-do-algodoeiro é o desenvolvimento de linhagens de algodão geneticamente resistentes que possam suprimir o desenvolvimento larval do inseto.

Com o advento da biotecnologia e da engenharia genética tornou-se possível a manipulação de características de uma cultura economicamente importante como o algodão através da inserção de transgenes. Diversos genes provenientes do *Bacillus thuringiensis* (Bt) já foram utilizados para a produção de plantas transgênicas resistentes a uma série de insetos.

15

Linhagens de plantas geneticamente modificadas para controle de pragas vem sendo amplamente comercializadas no Brasil, contudo, nenhuma das variedades atualmente disponíveis demonstra a capacidade de controlar esse importante inseto-praga, que é considerado um problema, pronunciadamente no Brasil.

20

As linhagens geneticamente modificadas comercializadas no Brasil fazem uso de proteínas tóxicas com atividade inseticida sob controle do promotor constitutivo CaMV35S. Essa região regulatória é capaz de conduzir a expressão de um transgene em todos os tecidos da planta. Entretanto, o seu nível de expressão não é necessariamente semelhante em todas as células vegetais, variando

25

30

ao longo do desenvolvimento da planta e nos diferentes tecidos apesar de o promotor ser considerado constitutivo. Também podemos citar possíveis impactos negativos em insetos não-alvo e, por fim, a expressão constitutiva de uma proteína heteróloga em uma variedade transgênica demanda um alto custo energético para a planta.

Portanto, de acordo com o estado da técnica da presente invenção, verifica-se que sequências regulatórias de três genes específicos de flor e fruto em plantas, mais precisamente em algodão (*Gossypium hirsutum*) são capazes de iniciar e acionar a transcrição de polinucleotídeos. Estas sequências promotoras

35

podem ser utilizadas na modificação da transcrição de polinucleotídeos endógenos e/ou heterólogos na produção de polipeptídeos de interesse, como proteínas tóxicas a insetos, mas ainda não possuem uma descrição ou ensinamentos prévios. Conseqüentemente, pelos motivos já apresentados nesta descrição, persiste a
5 necessidade de linhagens geneticamente modificadas de algodão que possam ser resistentes ao ataque do bicudo do algodoeiro.

Desta forma, o presente invento pretende apresentar construções de DNA que contenham promotores de genes, que desempenhem função específica no tecido floral e no fruto de variedades de plantas algodão, para criação de
10 organismos geneticamente modificados, resistentes ao ataque de insetos e pragas, principalmente novas linhagens de algodão, que sejam resistentes ao bicudo do algodoeiro.

SUMÁRIO DA INVENÇÃO

A invenção compreende a descrição de sequências regulatórias de nucleotídeos, promotoras da expressão de genes específicos, tal como genes estruturais, em plantas, incluindo monocotiledôneas e dicotiledôneas.
15

Mais especificamente a invenção refere-se a sequências regulatórias de polinucleotídeos isoladas de plantas de algodoeiro, como sendo capazes de iniciar e acionar a transcrição de polinucleotídeos, e ao uso destas
20 sequências regulatórias na modificação de transcrição de polinucleotídeos endógenos e/ou heterólogos e produção de polipeptídeos.

Além disso, a invenção diz respeito ao uso destas construções na forma de vetores de expressão, vetores recombinantes e em plantas, células vegetais ou protoplastos transgênicos.
25

Figura como um objetivo da presente invenção, sequências de DNA, tais como SEQ ID NO 1, SEQ ID NO 2 e SEQ ID NO 3, como sendo, respectivamente, as sequências regulatórias localizadas a montante 5' de cada um dos genes específicos GhPGFS3, GhPGFS4 e GhPGFS8, em tecidos florais e de fruto de plantas. Refere-se também ao uso destas sequências na forma de vetores de
30 expressão para geração de novas linhagens de plantas transgênicas, células vegetais e protoplastos, com a capacidade de direcionar a expressão do gene de interesse somente em alguns tecidos da planta como tecidos florais e de frutos.

Vantajosamente a presente invenção supera o estado da técnica. Inicialmente com relação ao método provido para modificar a expressão de genes em um organismo, como uma planta, incluindo a incorporação estável no
35 genoma do organismo pela utilização de outro vetor recombinante,

Adicionalmente a presente invenção permite o direcionamento

de expressão gênica, como proteínas Cry tóxicas, nos locais específicos do ataque de pragas, o que é de extrema importância para a obtenção de plantas transgênicas com níveis adequados de proteínas que confirmam a resistência da planta.

5 Assim, para suprir a falta de promotores, disponíveis no mercado atualmente, eficientes em direcionar expressão em flores e frutos de plantas, mais especificamente em algodão, é que este pedido de patente está sendo direcionado.

10 Vale ressaltar, novamente, que as sequências regulatórias descritas pelo estado da técnica podem ser utilizadas na nova geração de plantas transgênicas que estão sendo desenvolvidas pela utilização de promotores tecido-específico, como por exemplo, em flores e frutos.

15 O presente avanço descreve que a construção de DNA que contém o promotor de cada um dos genes específicos de flor e fruto, *GhPGFS3*, *GhPGFS4* e *GhPGFS8*, é capaz de direcionar a expressão do gene repórter (GUS) de forma bastante elevada em flores e frutos de plantas. Dessa forma estes promotores podem ser utilizados para direcionar outros genes de interesse, como os que codificam proteínas Cry tóxicas para vários grupos de insetos praga, especificamente em tecidos reprodutivos como flores e frutos para a geração de plantas resistentes a insetos praga.

20 Ademais, as sequências ora propostas se mostraram eficientes também na inibição do ataque de pragas, tais como Broca-dos-Frutos, *Lobiopa insularis*, (Coleoptera: Nitidulidae), que ataca frutos como morango, tomate, pêssego, goiaba, maçã, laranja, melão e melancia; a vassoura-de-bruxa, doença fúngica típica dos frutos de cacauzeiros, ocasionada pelo basidiomiceto *Moniliophthora perniciosa*; lagarta do cartucho do milho, *Spodoptera frugiperda*, que ataca predominantemente as maçãs; adultos da lagarta das maçãs, *Heliothis virescens*, que depositam seus ovos nas folhas e botões florais e lagarta rosada, *Pectinophora gossypiella*, que atacam botões e flores, causando perda total de maçãs pequenas e dano parcial em maçãs mais velhas.

30 Por fim, o objeto da presente invenção divulga o uso de regiões promotoras específicas de órgãos florais e frutos que ao serem fusionadas a quaisquer genes codificantes, principalmente os codificantes de proteínas tóxicas, podem ser utilizados para geração de plantas transgênicas resistentes a pragas que atacam culturas durante o florescimento e frutificação.

35 BREVE DESCRIÇÃO DAS FIGURAS

Figura 1. Gráfico adaptado de *Genevestigator*. Cada ponto retangular exibe o valor do potencial de expressão (EP) de experimentos de

microarranjo representado por uma variação de azul. EP é definido como o valor médio absoluto de 1% dos maiores valores de todos os microarranjos do banco de dados do *Genevestigator*. À direita o total de experimentos de microarranjo que possuem o gene indicado na parte superior da figura. Os genes submetidos a esta

5 análise foram selecionados devido a sua alta expressão no banco de dados do PAVE. Os três genes mostrados foram selecionados para análise de seu padrão de expressão devido à observada expressão em tecidos florais.

Figura 2. Análise de expressão relativa dos genes *GhPGFS3*, *GhPGFS4* e *GhPGFS8* em tecidos vegetativos de algodão. No eixo y está representada a expressão relativa dos prováveis genes específicos de flor em relação à expressão dos genes de referência *GhPP2A1* e *GhUBQ14* em diferentes tecidos de algodão (eixo x). Estes valores de expressão relativa foram gerados pelo programa *qBase* que emprega o modelo de quantificação relativa delta-CT através da fórmula $Q = E^{\Delta CT}$, onde E é a eficiência do gene e ΔCT é a amostra com menor valor de

10 expressão menos o valor de expressão em questão.

Figura 3. Análise de expressão relativa dos genes *GhPGFS3*, *GhPGFS4* e *GhPGFS8* em botões florais de algodão. No eixo y está representada a expressão relativa dos genes selecionados em relação à expressão dos genes de referência *GhACT4* e *GhUBQ14* em estágios de desenvolvimento dos botões florais de 2- 12 mm (eixo x).

15

20

Figura 4. Análise de expressão relativa dos genes *GhPGFS3*, *GhPGFS4* e *GhPGFS8* em órgãos florais de algodão. No eixo y está representada a expressão relativa dos genes *GhPGFS3*, *GhPGFS4* e *GhPGFS8* em relação à expressão dos genes de referência *GhACT4* e *GhFBX6* em diferentes órgãos florais de algodão (eixo x).

25

Figura 5. Análise de expressão relativa dos genes *GhPGFS3*, *GhPGFS4* e *GhPGFS8* ao longo do desenvolvimento de maçã de algodão. No eixo y está representada a expressão relativa dos genes *GhPGFS3*, *GhPGFS4* e *GhPGFS8* em relação à expressão dos genes de referência *GhMZA* e *GhPTB* em diferentes

30 estágios de maçã de algodão (eixo x).

Figura 6. Eletroforese em gel de agarose 1,2 % contendo os fragmentos resultantes da reação de PCR utilizando os oligonucleotídeos *GhPGSF3_GSP1/GSP2_GW1*, *GhPGSF4_GSP1/GSP2_GW1* e *GhPGSF8_GSP1/GSP2_GW1* em A, B e C respectivamente através da técnica de RACER para definição do Sítio de Início de Transcrição desses genes. Para o gene *GhPGFS3* obteve-se um fragmento de 200 pb (A) para os genes *GhPGFS4*(B) e *GhPGFS8* (C) um fragmento 400 pb.

35

Figura 7. Eletroforese em gel de agarose 1,2 % contendo os fragmentos resultantes do isolamento da região promotora dos genes *GhPGFS3*, *GhPGFS4* e *GhPGFS8*. A) Observa-se a banda de 380 pb esperada para o gene *GhPGFS4* (*pGhPGFS4*). B) É possível visualizar o fragmento de 514 pb esperado para o promotor do gene *GhPGFS8* (*pGhPGFS8*). C) Estão em destaques os fragmentos de promotor para o gene *GhPGFS3* (*pGhPGFS3*) de 505pb, 827pb e 1250pb, respectivamente.

Figura 8. Coloração histoquímica de GUS da construção *pGhPGFS4::GUS* em botões florais e plântulas de *Arabidopsis*. Forte coloração GUS pode ser observada em estames (D, E, F, G) desde estágios iniciais até estágios tardios do desenvolvimento de botões florais. No estigma de flores abertas (F, G) e na porção apical do fruto (J, L). Em plântulas é possível observar uma fraca coloração na porção apical da folha (H, I).

Figura 9. Análise histoquímica do promotor do gene *GhPGFS8* em plantas da *Arabidopsis* transformadas com a construção *pGhPGFS8::GUS*. A expressão órgão específica do promotor *pGhPGFS8* foi examinada durante o desenvolvimento dos botões florais e pode ser visualizada nos estames (B) e essa coloração permanece forte e específica nos estames até botões no estágio tardio de desenvolvimento como pode ser observado na figura (C, D, E, F, G, H). Forte coloração GUS é observada também na porção apical do estigma (G), na pétala (H) e na região apical do fruto (J). Em plântulas é possível observar atividade de GUS pela ativação do promotor *pGhPGFS8* nos primórdios foliares (I).

Figura 10. Expressão transiente dos três fragmentos de promotor de *GhPGFS3* por biobalística em botões florais de *G. hirsutum*. Padrão de expressão similar para os três fragmentos de promotor, observado através do gene repórter GUS em estames. A. *p505_GhPGFS3*; B. *p827_GhPGFS3*; C. *p1250_GhPGFS3*.

Figura 11. Expressão transiente dos fragmentos de promotor dos genes *GhPGFS4* e *GhPGFS8* por biobalística em botões florais de *G. hirsutum*. A. Padrão de expressão em estames e carpelos observada para o promotor do gene *GhPGFS4*. B. Padrão de expressão em estames e grãos de pólen observado para o promotor do gene *GhPGFS8*.

Figura 12. Resultado do BLASTp indicando a família gênica a qual o gene *GhPGFS3* pertence.

Figura 13. Resultado do BLASTn indicando a família gênica a qual o gene *GhPGFS4* pertence.

Figura 14. Resultado do BLASTn indicando a família gênica a

qual o gene *GhPGFS8* pertence.

DESCRIÇÃO DETALHADA DA INVENÇÃO

O desenvolvimento de plantas transgênicas, dotadas de níveis de proteínas heterólogas esta ligado a sequências nucleotídicas regulatórias, tal como promotores, que serão responsável pelo direcionamento de altos níveis de expressão nos tecidos alvo específico. Estes promotores são determinados de acordo com a identificação de genes que serão expressos em um determinado tecido ou condição fisiológica.

Desta maneira, figura como um primeiro objetivo da presente invenção a identificação de genes de algodão, expressos sobretudo em flores e frutos, mediante o uso de informação derivada do PAVE com a do Cotton Genome Database.

Os genes *GhPGFS3*, *GhPGFS4* e *GhPGFS8* foram selecionados na biblioteca de flores inteiras de *Gossypium raimondii* devido a sua alta expressão (seis ou mais etiquetas de expressão no PAVE) e foram comparados com a sequência dos genes de *G. hirsutum* do Cotton Genome Database usando o programa tBLASTN (*GhPGFS*, *Gossypium hirsutum* putative gene flower specific).

Tabela 1 Resultado da comparação entre os 3 Contigs de *G. raimondii* identificados no banco de dados PAVE como específicos de flor e fruto com o banco de dados do Cotton Genome Database para genes expressos em *G. hirsutum*. Similaridade entre os genes está descrita em número de nucleotídeos.

Gene	Contig PAVE	Nº leituras PAVE	Hit no Blastn <i>G. hirsutum</i>	Similaridade
<i>GhPGFS3</i>	00001_156	36	>gi 78325063 gb DT545337.1	834/1174 (71%)
<i>GhPGFS4</i>	00002_23	21	>gi 45758291 gb AY438035.1	1438/1464 (98%)
<i>GhPGFS8</i>	00111_01	38	>gi 193217225 gb EY197012.1 EY197012	793/825 (96%)

Com a sequência destes três genes foi realizado um BLASTn no Banco de Dados de *Arabidopsis* com objetivo de encontrar os possíveis ortólogos, os quais foram então analisados na plataforma Genevestigator (Figura 1).

Esta análise possibilitou visualizar uma elevada expressão potencial em tecidos florais em *Arabidopsis* e estes genes tiveram seu padrão de expressão avaliado de forma detalhada através de reações de PCR em tempo real (qPCR) em plantas de algodão da espécie *Gossypium hirsutum* da variedade "BRS Cedro" de três meses de idade. As plantas foram cultivadas sob temperatura controlada (21 ± 4 °C) e sob fotoperíodo natural. Os possíveis genes específicos de flor e fruto tiveram seu perfil de expressão gênica avaliado em quatro condições

diferentes: entre diferentes órgãos vegetativos e reprodutivos, entre distintos estágios do desenvolvimento de botões florais, de maçãs (maçã é a denominação usual para os frutos de algodão) e entre os diferentes órgãos florais de botões de *G. hirsutum*.

5 Duas replicatas biológicas para cada amostra foram usadas para análise por qPCR e três replicatas técnicas foram analisadas para cada replicata biológica, sendo que todas replicatas biológicas apresentaram uma reprodutibilidade adequada.

10 A normalização dos dados obtidos com a análise por qPCR requer genes que apresentem transcritos com expressão uniforme na maioria das células do organismo ou entre as espécies que estão sendo analisadas, assim como durante várias fases do desenvolvimento e sob diferentes condições ambientais. Estes genes são usualmente denominados genes constitutivos ou "Housekeeping Genes". Por isso foi analisado os perfis de expressão gênica de nove genes candidatos a genes de referência em algodão por qPCR em um conjunto de 22 diferentes amostras

15 de tecidos de algodão agrupados em 5 diferentes conjuntos experimentais com o intuito de identificar os melhores genes para a normalização dos dados de expressão gênica durante o desenvolvimento floral de algodão (Artico et al., 2010).

20 A primeira análise do padrão de expressão dos possíveis genes específicos de flor e fruto foi feita entre diferentes órgãos de algodão, incluindo órgãos vegetativos como folha, caule, raiz e ramo e órgãos reprodutivos como coletânea de flor (botões florais de 2, 4, 6, 8 e 10 mm de diâmetro) e coletânea de maçã (maçãs de 10-15, 16-20, 21-30 e >30 mm de diâmetro).

25 Surpreendentemente foi observado que os genes *GhPGFS4* e *GhPGFS8* tiveram uma expressão relativa em botões florais mais alta que nos demais tecidos, enquanto que *GhPGFS3* além da elevada expressão em botões florais também é altamente expresso em maçã, evidenciando um acúmulo de transcritos desses genes em tecidos de flor e fruto, enquanto sua expressão é menor ou em níveis quase indetectáveis nos outros tecidos analisados como raiz, folha, caule, ramo. Isso sugere que esses genes apresentam uma expressão específica em tecidos

30 florais.

O gene *GhPGFS4* também apresentou expressão em maçãs, mas menor que a do gene *GhPGFS3* (Figura 2).

35 Em seguida, analisamos a expressão dos genes *GhPGFS3*, *GhPGFS4* e *GhPGFS8* durante o estágio de desenvolvimento dos botões florais de 2, 6, 7, 8, 10 e 12 mm de diâmetro. Foi observado que os genes *GhPGFS4* e *GhPGFS8* são altamente expressos no botão de 6 mm e esse nível de expressão se mantém elevado também nos botões de 7, 8 e 10 mm, enquanto que *GhPGFS3* apresenta um

padrão de expressão constante durante os estágios de 2, 4, 6 e 8mm e decai em 10mm (Figura 3) Esses resultados foram comprovados com replicatas biológicas para cada botão floral analisado.

5 Na análise da expressão dos genes *GhPGFS3*, *GhPGFS4* e *GhPGFS8* nos diferentes órgãos florais do botão de algodão. *GhPGFS3* apresentou alta expressão em carpelos e estames, enquanto os dois últimos genes tiveram uma expressão relativa bastante elevada em estames e apresentaram uma menor expressão em pétala, sépala e carpelos (Figura 4).

10 Ao longo da última análise de expressão foi verificado a quantidade de transcritos dos genes *GhPGFS3*, *GhPGFS4* e *GhPGFS8* nos diferentes estágios de desenvolvimento de maçã. O gene *GhPGFS3* apresenta um padrão constante e elevado de expressão a partir do estágio de 16-20mm. Já o gene *GhPGFS4* demonstra um nível relativamente alto de expressão nos estágios 10-15mm e 16-20mm, decaindo nos estágios posteriores de desenvolvimento. Por fim,
15 *GhPGFS8* apresenta níveis basais de expressão com um pico de transcritos no estágio 21-30 mm de desenvolvimento (Figura 5). Esses resultados foram similares nas replicatas biológicas para cada tecido analisado.

A sequência de EST desses genes foi analisada no programa BLAST para análise de sua possível função biológica. Este programa identifica regiões
20 similares entre sequências e pode inferir sobre as relações funcionais e evolutivas entre elas, podendo inclusive, identificar membros de famílias gênicas.

A sequência de aminoácidos do gene *GhPGFS3* apresentou alta similaridade através de análises no programa BLASTp à sequência da superfamília do citocromo P450 (Figura 12). Enzimas dessa família já foram descritas
25 catalisando diversos processos biológicos, dentre eles pode-se citar a biossíntese do homoterpeno TMTT, importante na composição de compostos voláteis que compõem odores florais e atuam contra herbivoria (LEE et al. 2010). Entretanto, análises mais detalhadas são necessárias para caracterizar de maneira mais refinada a função desse gene de *G. hirsutum*.

30 A sequência do gene *GhPGFS4* apresentou alta similaridade através de análises no programa tBLASTN à sequência de uma β -galactosidase de algodão (*GhGal1*) expressa em flores e fibras (ZHANG e LIU, 2005) (Figura 13). Os transcritos do gene *GhGal1*, foram abundantes em flores e em fibras de 5 a 10 dpa (dias após antese) declinando em fibras de 20 dpa, isto é, está envolvido nos estágios
35 iniciais do desenvolvimento da fibra em algodão e por isso sua expressão foi detectada em frutos de até 30 mm de diâmetro em nossos experimentos.

A sequência do gene *GhPGFS8* apresentou a partir do

tBLASTN alta similaridade à sequência de pectase liase (PLL) de *A. thaliana* (Figura 14). Os transcritos destes genes em *Arabidopsis* são altamente expressos em órgãos reprodutivos principalmente em anteras e pólen (KULIKAUSKAS e MCCORMIC, 1997).

5 Confirmada a especificidade da expressão gênica em tecidos florais de algodão via qPCR, iniciou-se o isolamento da região posicionada a montante (5') da sequência de EST destes genes (região promotora) empregando o kit Universal Genome Walker (GW) (Clontech) que permite "andar" pelo genoma identificando regiões desconhecidas.

10 Uma região promotora de 1451 pb foi obtida para o gene *GhPGFS3*, 380 pb para o gene *GhPGFS4* e de 514 pb para o gene *GhPGFS8*. Os fragmentos em questão foram purificados, clonados em vetor pGEM®-T Easy (Promega) e inseridos em *E. coli* cepa DH5 α para posterior sequenciamento. Após a identificação das regiões promotoras através do sequenciamento foram realizadas
15 buscas por elementos regulatórios nas sequências amplificadas por GW utilizando-se o banco de dados PLACE (<http://www.dna.affrc.go.jp/htdocs/PLACE/>) (HIGO et al., 1999) para identificação de possíveis sítios de ligação a fatores de transcrição já caracterizados em plantas.

20 O ponto de início de tradução do gene *GhPGFS3* foi determinado pelo alinhamento com outros genes de espécies vegetais relacionadas. O sítio de início de transcrição foi determinado para o mesmo gene utilizando o GeneRacer™ Kit e identificado na posição -86 pb (Figura 6). Para a busca por elementos cis-regulatórios presentes na região promotora desse gene, utilizando-se de informações disponíveis no banco de dados do PLACE, que revelou a presença de
25 inúmeros motivos comumente encontrados em regiões promotoras de plantas. Como por exemplo, foi observada a presença de motivos típicos de regiões promotoras de eucariotos, como TATAbox2 (TATAAAT, posição -116 pb) necessário para o início da transcrição do gene da β -faseolina, a maior proteína de estocagem de *Phaseolus vulgaris* (GRACE et al., 2004).

30 Nas posições -4, -169, -274, -601 pb encontra-se o elemento CAAT Box1, característicos de sequências promotoras de eucariotos (STEPHEN e JAMES, 2003). Elementos característicos por direcionar a expressão em pólen (POLLEN1LELAT52, AGAAA, posições -289, -1065 pb), em fruto (ERELEE4, AWTTCAAA, na posição -1123 pb) e em semente (RYREPEATBNNAPA, CATGCA
35 nas posições -965, -1004 pb) também foram encontrados.

Tabela 2: Elementos cis relevantes encontrados na região promotora do gene GhPGFS3 envolvidos em várias respostas, principalmente na expressão específica em pólen e fruto.

Tabela de elementos regulatórios de GhPGFS3			
Elemento	Sequência	Função	Posição
CAATBOX1	CAAT	Sequência consenso	-4, -9, -92, -122, -169, -194, -203, -224, -230, -274, -492, -496, -515, -601, -632, -820, -830, -961, -988, -1033, -1109
CARGCVW8GAT	CWW WWWWWWWG	Sítio de ligação AgL-15	-553 e -692
ERELEE4	AWTTCAA	Elemento de fruto responsivo a etileno	-1123

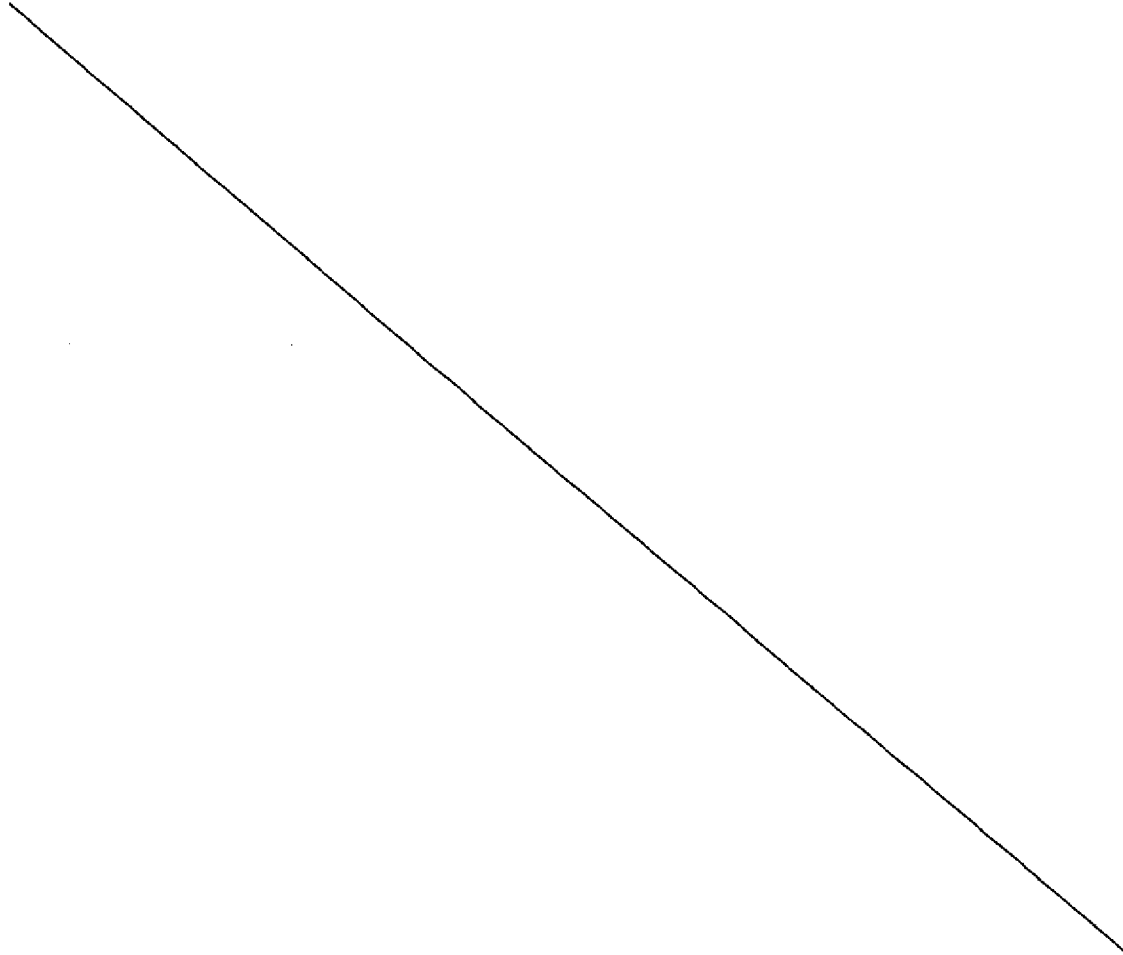
GATABOX	GATA	Requerido por alto nível de expressão tecido específica	-139, -269, -424 e -568
GTGANTG10	GTGA	Elemento específico de pólen	-14, -76, -82, -298, -348, -499, -673, -791, -824, -866, -1077, -1102, -1137 e -1248
MYBPZM	CCWACC	Biosíntese de flavonóides em órgãos florais	-472 e -896
POLLEN1LELAT52	AGAAA	Elemento responsável pela ativação específica de pólen em tomate	-289, -711, -730, -751, -811, -1020, -1065 e -1704
SITEIIATCYTC	TGGGCY	Encontrado na região promotora de citocromos de <i>A. thaliana</i>	-1243
TATABOX2	TATAAAT	Elemento crítico para iniciação de transcrição	-116
TATABOX5	TTATT	Elemento crítico para iniciação da transcrição em <i>P. sativum</i>	-113, -399, -904, -912, -1186
MYB.ph3	WAACNRWYW	Sítio de ligação de fator de transcrição específico de epiderme de <i>Petunia hybrida</i>	-270 e -672

O fragmento promotor do gene *GhPGFS4* tem um tamanho de 380 nucleotídeos. O ponto de início de tradução também foi determinado pelo alinhamento com outros genes de espécies vegetais relacionadas. O sítio de início de transcrição foi determinado para o mesmo gene utilizando o GeneRacer™ Kit e identificado na posição -124 pb (Figura 6). A análise no PLACE revelou a presença de motivos típicos de regiões promotoras de eucariotos, como TATAbox5 (AAATAA, posição 14, -153 e -313 pb) (GRACE et al., 2004). Nas posições -157, -254, -307 pb encontra-se o elemento CAAT Box1 (STEPHEN e JAMES, 2003). Diversos outros elementos especificamente presentes em regiões promotoras de genes expressos em pólen (GTGANTG10, GTGA, posições -87, -269 pb e POLLEN1LELAT52, AGAAA, posições -24, -251, -311, -352 pb), em fruto (TGTCACACMCUCUMISIN, TGTCACA, nas posições -270 pb) e em raiz (RHERPATEXPA7, GCACGA na posição -325 pb) foram encontrados (YAMAGATA et al., 2002). Foi possível identificar também um

5

10

motivo TGAC (WRKY71OS) comumente presente em promotores de genes ativados por giberelina e o motivo TTGAC (WBOXATNPR1) de genes ativados por ácido salicílico (ZHANG et al., 2004).



5 **Tabela 3:** Possíveis elementos cis relevantes encontrados na região promotora do gene *GhPGFS4* envolvidos em várias respostas, principalmente na expressão específica em pólen e fruto.

Tabela de elementos regulatórios de <i>GhPGFS4</i>			
Elemento	Sequência	Função	Posição
TATAbbox5	AAATAA	Início da transcrição do gene da glutamina sintase (<i>Pisum sativum</i>)	-14, -153 e -313 pb
CAAT Box1	CAAT	Elemento característicos de sequências promotoras de eucariotos	-157, -254 e -307 pb
GTGANTG10	GTGA	Elementos presentes em regiões promotoras de genes expressos em pólen	-87 e -269 pb

POLLEN1LELAT52	AGAAA	Elementos presentes em regiões promotoras de genes expressos em pólen	-24, -251, -311 e -352 pb
TGTCACACMCUCU MISIN	TGTCACA	Elemento presentes em regiões promotoras de genes expressos em fruto	-270 pb
RHERPATEXPA7	GCACGA	Elementos presentes em regiões promotoras de genes em raiz	-325 pb
WRKY71OS	TGAC	Motivo comumente presente em promotores de genes ativados por giberelina	-268 pb
WBOXATNPR1	TTGAC	Motivo comumente presente em promotores de genes ativados por ácido salicílico	-156 pb

Na sequência de 514 pb da região promotora do gene *GhPGFS8* foram encontrados vários elementos TATAbox nas posições -14pb do ponto de início de tradução. Outros elementos consenso de região promotora também foram encontrados, como o CAAT, repetido 4 vezes nesse fragmento e servem como elementos promotores basais para o início da transcrição (WU et al., 2005).

A região promotora do gene *GhPGFS8* também contém potenciais motivos com homologia a elementos regulados de maneira tecido específica, como por exemplo GTGANTG10 (GTGA, posição -134 pb) e POLLEN1LELAT52 (AGAAA, posição -214 e -458 pb). Encontra-se também a -85 pb do ponto de início da tradução o elemento TGTCTC, ARFAT, que é um fator de resposta a auxina e nas posições -78, -232 e - 287 encontra-se o elemento CCAAT, denominado CCAATBOX1, responsável por processos como indução do florescimento pelo fotoperíodo, sinalização em resposta a luz e regulação do ritmo circadiano (WENKEL et al., 2006).

Tabela 4: Possíveis elementos cis relevantes encontrados na região promotora do gene *GhPGFS8* envolvidos em várias respostas, principalmente na expressão específica em pólen e fruto.

Tabela de elementos regulatórios de <i>GhPGFS8</i>			
Elemento	Sequência	Função	Posição
TATAbox5	AAATAA	Início da transcrição do gene da glutamina síntase (<i>Pisum</i>)	-14 pb

		<i>sativum</i>)	
CAAT Box1	CAAT	Elemento característicos de sequências promotoras de eucariotos	-77, -107, -230, -476 pb
GTGANTG10	GTGA	Elementos presentes em regiões promotoras de genes expressos em pólen	-134 pb
POLLEN1LELAT52	AGAAA	Elementos presentes em regiões promotoras de genes expressos em pólen	-214 e -458 pb
ARFAT	TGTCTC	Fator de resposta a axina	-85 pb
CCAATBOX1	CCAAT	Responsável por processos como indução do florescimento pelo fotoperíodo, sinalização em resposta a luz e regulação do ritmo circadiano	-78, -232 e -287 pb

O sítio de início de transcrição foi determinado para o mesmo gene utilizando o GeneRacer™ Kit e identificado na posição -154 pb do ATG inicial (Figura 6).

5 Visando avaliar a funcionalidade e especificidade das regiões promotoras isoladas no presente trabalho, análise do padrão de expressão do promotor destes genes foi realizada através de experimentos histoquímicos de coloração GUS. Para isso fragmentos de promotor de 505, 827 e 1250pb para o gene *GhPGFS3*, 350 pb do promotor do gene *GhPGFS4* e 440 pb do promotor do *GhPGFS8* foram amplificados e clonados no vetor Gateway®, pKGWFS7, contendo o gene repórter uidA (GUS) ou GFP (KARIMI et al., 2007) (Figura 7). Plantas transgênicas de *Arabidopsis thaliana* foram transformadas através do sistema de infiltração da inflorescência mediado por *Agrobacterium tumefaciens* (floral-dip) (DESFEUX et al., 2000). Três linhagens independentes de plantas transgênicas contendo os promotores dos genes *GhPGFS3*, *GhPGFS4* e do gene *GhPGFS8*, segregando 3:1 foram selecionadas em meio de cultura contendo o antibiótico Canamicina (50µg/mL).

A seleção das plantas transgênicas até a terceira geração foi

realizado somente através dos ensaios de segregação e meio com antibiótico. Sementes e plantas pertencentes às linhagens T3, produzidas pelas plantas T2 homocigotas, foram então utilizadas em ensaios histoquímicos de coloração GUS para análise do padrão de expressão tecido específico dos promotores isolados. Nesses testes foram utilizadas plântulas com 15 dias de idade e inflorescência de plantas com 60 dias.

Para o gene *GhPGFS3* 15 a 20 linhagens independentes foram obtidas para cada construção. O referido estudo continua em andamento através da seleção de linhagens homocigotas (contendo dois alelos) com apenas uma cópia do transgene inserido no genoma até a terceira geração, momento em que será realizado o ensaio fluorimétrico para a avaliação do padrão de expressão do gene repórter GUS.

A análise histoquímica em plântulas de *Arabidopsis* contendo o promotor do gene *pGhPGFS4::GUS* demonstrou uma leve coloração de GUS na margem da folha e na região apical da siliqua. Já a expressão de GUS detectada em flores foi forte e específica em estames a partir do estágio 12 de desenvolvimento em *Arabidopsis* até flores tardias. É possível observar também coloração na porção superior do carpelo, mais especificamente na região do estigma da flor (Figura 8). A análise de coloração GUS da construção *pGhPGFS8::GUS* em plântulas demonstrou uma fraca expressão no primórdio foliar. Já em flores a marcação demonstrou que esse fragmento de promotor direciona uma expressão crescente e específica em estames e no carpelo desde o estágio 12 de desenvolvimento em *Arabidopsis* até flores no estágio tardio e ainda em pétalas entre os diferentes órgãos florais. A porção apical das síliqua também apresentou forte coloração de GUS (Figura 9).

Com o objetivo de avaliar o padrão tecidual de expressão em *G. hirsutum* através da fusão do gene repórter *uidA*, regulado por tais regiões promotoras foi realizado o ensaio de expressão transiente por biobalística de acordo com o descrito por RECH *et al.* (2008). Botões florais de *G. hirsutum* de 5 a 8mm foram submetidos a bombardeamento e todos os fragmentos de promotor de *GhPGFS3* (505, 827 e 1250pb), *GhPGFS4* (380pb) e *GhPGFS8* (440pb). O ambos os promotores do gene *GhPGFS3*, assim como o promotor do gene *GhPGFS4* e do gene *GhPGFS8* apresentaram um padrão de expressão semelhante para todos os três fragmentos em estames e pólen (Figura 10 e 11).

De maneira geral *GhPGFS3* apresenta expressa padrão específico em flores e frutos de algodão. As análises de expressão por qPCR revelaram um alto nível de transcritos durante grande parte do desenvolvimento dos botões florais e de maçãs, especialmente em estames e carpelos.

Tal resultado foi corroborado com o padrão obtido pela

expressão transiente em botões florais de *G. hirsutum* e mais dados provenientes da análise em *A. thaliana* ainda estão sendo aguardados e considerados promissores.

5 A sequência promotora isolada de *GhPGFS3* apresenta diversos elementos regulatórios de plantas que seriam responsáveis pelo controle e atividade específicos do padrão de expressão observado.

O gene *GhPGFS4* apresenta um padrão de expressão específico em flores e frutos de algodão, principalmente em botões florais de 6 mm e em tecidos florais como estames, carpelo, pétalas e sépala.

10 O gene *GhPGFS8* apresenta um padrão de expressão flor-específica, especialmente em estame, pétala, sépala e carpelo entre os órgãos florais assim como também em frutos pela análise de GUS.

A sequência promotora isolada de ambos os genes apresenta inúmeros elementos regulatórios de plantas sendo que os mais interessantes são os responsáveis pela coordenação e ativação específica no desenvolvimento das anteras e do grão de pólen, AGAAA e GTGA e o elemento ativador necessário para a expressão específica em fruto, TGTCACA.

20 O resultado da análise desses elementos em um fragmento do promotor na construção com o gene repórter GUS em *Arabidopsis* corroborou com a expressão específica original desses genes em algodão vistas por qPCR, o que sugere que fragmentos pequenos destes promotores já contem motivos cis importantes para a manutenção da expressão temporal e espacial destes genes.

Assim, destaca-se também como objetivo da presente invenção, a produção de plantas transgênicas pertencentes a diferentes espécies de interesse comercial.

25 Cabe ressaltar que a utilização de promotores da própria espécie na produção de plantas transgênicas constitui uma vantagem do ponto de vista tecnológico, já que permitirá um controle da expressão do transgene de forma mais forte e específica.

30 EXEMPLO 1 - SELEÇÃO IN SILICO DE ESTS COM PADRÃO DE EXPRESSÃO ÓRGÃO/TECIDO ESPECÍFICO EM FLOR DE ALGODÃO

Para identificação dos genes de algodão dois banco de dados distintos foram usados, o PAVE (<http://www.agcol.arizona.edu/cgi-bin/pave/Cotton/index.cgi>) e o Cotton Genome Database (U.S. Department of Agriculture, Agricultural Research Service CottonDB - <http://cottondb.org/cdbhome.html>). Assim selecionamos genes que apresentaram alta expressão (seis ou mais etiquetas de expressão no PAVE) exclusiva ao desenvolvimento floral de *G. raimondii* e comparamos a sequência destes genes com

aqueles genes de *G. hirsutum* do Cotton Genome Database usando o programa tBLASTN (ALTSCHUL et al., 1997) e o parâmetro matriz BLOSUM62. Esta matriz é construída a partir do alinhamento local de seqüências com similaridade igual ou maior que 62%. Com a seqüência desses genes também foi realizado BLASTn no Banco de
5 Dados de Arabidopsis com objetivo de encontrar seus possíveis ortólogos, os quais foram então analisados na plataforma Genevestigator que permite avaliar o padrão de expressão a partir de centenas de experimentos de microarranjo (Figura 1). Genevestigator (<https://www.genevestigator.ethz.ch>) é uma ferramenta Web, porém com aplicativo JavaScript instalado no computador (ZIMMERMANN et al., 2004). A
10 plataforma de *A. thaliana* com microarranjos Affymetrix ATH1 22K foi selecionada nesta análise. Nossos resultados foram obtidos pela visualização dos gráficos de Anatomia (Anatomy) com visualização no mapa de intensidade (heat map) em escala logarítmica. Apenas seqüências que apresentaram elevada expressão em tecidos florais foram selecionadas para o desenho dos iniciadores.

15 EXEMPLO 2- ANÁLISE DO PADRÃO DE EXPRESSÃO DOS GENES GhPGFS4 E GhPGFS8 POR qPCR

a) Material Vegetal

Em todos os experimentos descritos no presente trabalho foram utilizadas plantas da espécie *Gossypium hirsutum* de três meses de idade da
20 variedade "BRS Cedro", cultivadas sob temperatura controlada (21 ± 4 °C) e sob fotoperíodo natural na Embrapa CENARGEM em Brasília (DF, Brasil). Os órgãos utilizados foram botões florais, frutos, folhas, caules, ramos e raízes. Também foram incluídos sete estágios de desenvolvimento floral (botões florais com os seguintes diâmetros: 2, 4, 6, 7, 8, 10 e 12 mm) e quatro estágios de desenvolvimento de frutos
25 (frutos com os seguintes diâmetros: 10-15, 16-20, 21-30 e maior que 30 mm) (GREENBERG *et al.*, 2004). Além disso, órgãos florais (sépala, pétala, estame, carpelo e pedicelo) foram dissecados a partir do botão de 6 mm e armazenados. O material foi isolado a partir de, pelo menos, cinco diferentes plantas de algodão para a obtenção de um pool para cada tecido. O procedimento foi repetido com outras cinco
30 plantas a fim de se obter um segundo pool que representasse a replicata biológica. Todas as amostras foram imediatamente congeladas em nitrogênio líquido e armazenadas à -80°C até serem utilizadas para a extração do RNA.

b) Desenho dos iniciadores

Os iniciadores para as análises da PCR quantitativa em tempo
35 real (qPCR) foram desenhados utilizando-se a ferramenta Primer3 (ROZEN e SKALETSKY, 2000) e avaliados usando o programa NetPrimer (Premier Biosoft International). Foram selecionados apenas aqueles resultantes em amplicons entre 80

pb e 200 pb e com temperaturas médias de hibridização em torno de 60°C. Todos os iniciadores foram inicialmente testados via qPCR por visualização em gel de agarose e a especificidade da amplificação foi verificada pela análise das curvas de dissociação geradas pelo programa Opticon Monitor 3 (DNA Engine Opticon Real-Time PCR Detection System, Bio-Rad). A normalização dos dados relativos de expressão foi realizada utilizando o conjunto dos melhores genes constitutivos definidos por Artico et al., 2010.

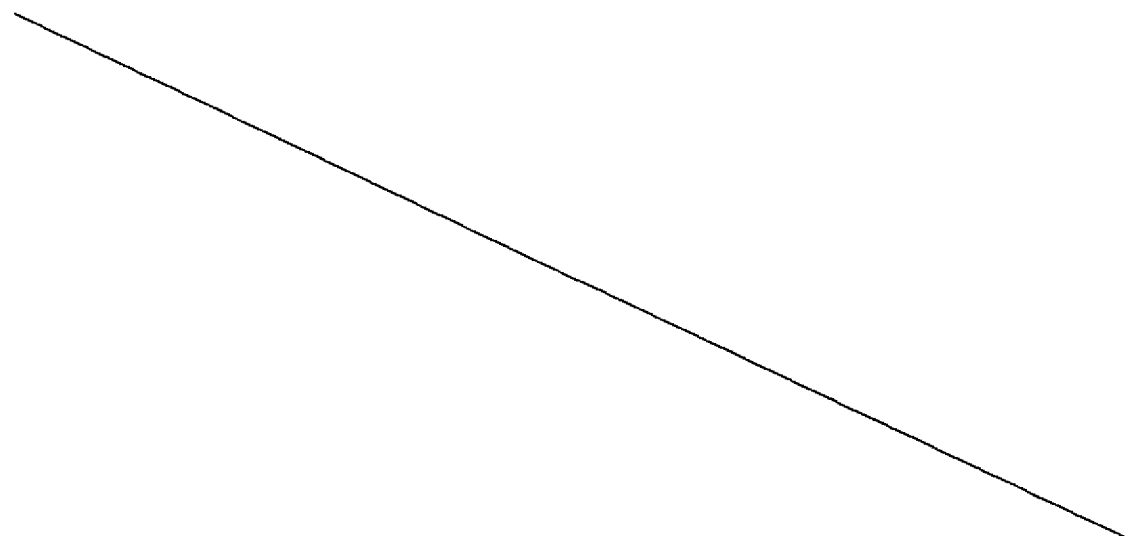


Tabela 5: Relação de Sequências de iniciadores, gerada pelo programa *Primer 3*, para supostos genes específicos de flores e para os genes constitutivos de *G. hirsutum*.

Nome	Lócus homólogo descrito em <i>A. thaliana</i>	Processo biológico descrito em <i>A. thaliana</i>	Iniciadores direto/reverso
GhPGFS3	At4g02530	Proteína do lúmen do tilacóide	TGAGTGTGTTTCAGGGTCCA/ AGACGATTGGAGAGGTGTGG
GhPGFS4	At4g35010	β -Galactosidase	GGGTCCAGATCAACAGGCTA/ AGGGAGGACGCTAATGGAAC
GhPGFS8	At3g01270	Pectase liase	GACGGCAAGTTTTACGTGGT/ ATCCAAAGTGGCCTGAGTTG
GhPP2A1	At1g59830	Subunidade catalítica da proteína fosfatase 2ª	GATCCTTGTGGAGGAGTGG/ GCGAAACAGTTCGACGAGAT
GhUBQ14	Atg02890	Poli-ubiquitina	CAACGCTCCATCTTGTCTT/ TGATCGTCTTCCCGTAAGC
GhACT4	At5g09810	Membro da família gênica da Actina	TTGCAGACCGTATGAGCAAG/ ATCCTCCGATCCAGACACTG
GhFBX6	At5g15710	Membro da família protéica F-box	TGCCTGCAGTAAATCTGTGC/ GGTGAAAGGGTTTCCAAAT

c) Extração de RNA e síntese de cDNA

As amostras de RNA total foram preparadas a partir dos tecidos de algodão (*G. hirsutum*, variedade Cedro). Os tecidos utilizados de órgãos

vegetativos da planta foram: folha, caule, ramo, raiz, coletânea de flor (botões de 2, 4, 6, 8 e 10 mm) e coletânea de maçã (10-15, 16-20, 21-30 e >30 mm de diâmetro); de órgãos reprodutivos foram: botão floral de 2, 4, 6, 7, 8, 10 e 12 mm e sépala, pétala, estame, carpelo, pedicelo isolados do botão floral de algodão de 6 mm. Todos os tecidos foram coletados de diferentes indivíduos. Cada tecido foi dividido em dois pools a fim de obtermos replicatas biológicas. Os tecidos foram mantidos congelados à -80°C e macerados em nitrogênio líquido para a extração de RNA.

As extrações de RNA total foram realizadas a partir de 100 mg de cada tecido vegetal macerado em nitrogênio líquido, utilizando o Invisorb Spin Plant RNA Mini kit (Invitex), de acordo com o protocolo do fabricante. As amostras foram dosadas no NanoDrop ND-1000 (Thermo Fisher Scientific), com as leituras das absorvâncias na faixa 230-280 nm e utilizadas na síntese de cDNAs. A integridade do RNA foi avaliada também em eletroforese em gel de agarose 1% com coloração em brometo de etídio. A presença de produtos espúrios de amplificação causados pelo DNA genômico foi também continuamente verificado pelo perfil de dissociação da qPCR. Ambos os testes mostraram que o Invisorb Spin Plant RNA Mini kit remove eficientemente DNA contaminante das amostras de RNA em algodão.

A primeira fita de cDNA de cada amostra foi sintetizada a partir de 1 µg de RNA total, ao qual adicionou-se 50 µM de oligonucleotídeos (dT24V) e 10 mM de cada deoxirribonucleosídeo 5'-trifosfato (dNTPs). Essas reações foram incubadas a 65°C por 5 minutos e resfriadas no gelo. Em seguida foram adicionados 20 mM de ditiotretol (DTT), 200 unidades de Superscript III (Invitrogen) e o respectivo tampão First-Strand totalizando o volume de 20µL. Esta reação foi incubada a 50°C durante 1h, segundo instruções do fabricante. A inativação da transcriptase reversa foi feita por incubação da reação a 70°C por 15 minutos.

d) PCR quantitativa em tempo real (qPCR)

Para os experimentos de qPCR os cDNAs utilizados foram diluídos 50 vezes em água MilliQ® estéril ultrapura e as mesmas alíquotas foram utilizadas em todos os experimentos. Cada reação foi realizada em triplicata para cada par de iniciador. As qPCRs foram efetuadas no sistema de detecção Chromo 4 (Bio-Rad), e os resultados visualizados pelo software Opticon Monitor 3 (DNA Engine Opticon Real-Time PCR Detection System, Bio-Rad). As reações de amplificação dos cDNAs foram realizadas em um volume final de 20 µL contendo 2,00 µL do detector SYBR Green I (1X) (Molecular Probes), 10 µL do cDNA diluído (1:50), 0,80 µL dos iniciadores direto e reverso (10 µM cada), 0,05 µL de dNTP (10 mM), 2,00 µL de tampão PCR sem magnésio (10x), 1,20 µL de cloreto de magnésio (50 mM), 0,05 µL de Platinum® Taq DNA Polymerase (5 U/µL) e 3,90 µL de água. Na reação controle

negativa, o volume do cDNA foi substituído por água MilliQ® estéril. Para a análise dos genes nos tecidos vegetativos juntamente com a coletânea de flor e a coletânea de maçã, o volume final de cada reação foi aumentado para 40 µL a fim de melhorar os resultados.

5 Os parâmetros da reação da PCR foram 94°C por 5 minutos para a ativação da enzima polimerase, seguida por 40 ciclos de desnaturação a 94°C por 15 segundos, anelamento dos iniciadores específicos para cada gene a 60°C por 10 segundos e polimerização a 72°C por 15 segundos.

10 As medidas dos produtos de amplificação foram realizadas pela incorporação do marcador fluorescente SYBR Green na dupla fita de DNA. Para avaliação da eficiência do iniciador utilizamos o programa online Real-time PCR Miner (ZHAO e FERNALD, 2005). Este programa utiliza os pontos iniciais e finais da fase exponencial da PCR juntamente com os dados brutos de fluorescência de cada iniciador em cada ciclo e estima para cada gene os valores de eficiência das
15 amplificações e o ciclo de corte CT (Cycles threshold) sem a necessidade de uma curva padrão a partir de um algoritmo de regressão não linear. Estes valores de CT gerados pelo Miner foram utilizados no programa qBase, versão v 1.3.5 (HELLEMANS et al., 2007) que emprega o modelo de quantificação relativa delta-CT para transformar o valor de expressão em valores relativos não-normalizados, que serão
20 corrigidos pela eficiência da PCR através da fórmula $Q = E^{\Delta CT}$, onde E é a eficiência do gene e ΔCT é a amostra com menor valor de expressão menos o valor de expressão em questão.

Os resultados foram submetidos a dois diferentes programas que classificam os melhores genes constitutivos, o geNorm (VANDESOMPELE, et al. 2002) e
25 NormFinder (ANDERSEN, et al. 2004). Estes programas se baseiam em algoritmos diferentes que utilizam como informação a análise do ciclo limite (Cycles threshold) dos genes nos diferentes tecidos em duas replicatas biológicas e a eficiência do iniciador em cada amostra biológica. Com esta análise identificamos os genes que apresentam a menor variação quanto ao número de transcritos em cada tecido.

30 Para a avaliação da expressão dos possíveis genes específicos de flor, os valores de eficiência das amplificações de cada iniciador e o ponto de corte CT (Cycles threshold) dos genes nos diferentes tecido foram determinados pelo programa online Real-time PCR Miner, utilizados no programa qBase versão 1.3.5 e normalizados com os genes de referência indicados pelo
35 programa NormFinder.

EXEMPLO 3 - ISOLAMENTO E IDENTIFICAÇÃO DAS REGIÕES PROMOTORAS DOS GENES GhPGFS3, GhPGFS4 E GhPGFS8

A região do DNA localizada a montante (5') dos ESTs validados por qPCR, correspondente portanto a região regulatória dos genes em análise, foi amplificada empregando a técnica de Genome Walkertm (GW) (Clontech) a partir do DNA genômico de *Gossypium hirsutum*. Bibliotecas foram construídas a partir de quatro grupos de fragmentos oriundos da digestão de DNA genômico com diferentes enzimas de restrição de corte abrupto (DraI, EcoRV, PvuII e StuI). Esses fragmentos foram então ligados a adaptadores fornecidos pelo kit. A técnica —genome walking— permite amplificar, através de PCR, regiões localizadas à montante de uma seqüência conhecida (DEVIC et al., 1997) empregando-se oligonucleotídeos complementares aos adaptadores e à seqüência conhecida (no nosso caso do EST validado) (Tabela 2). A primeira etapa do procedimento de Genome Walker consiste de um PCR primário empregando oligonucleotídeos mais externos relacionados ao adaptador (AP1) e à seqüência do EST validado (GSP1). Uma alíquota desse PCR primário é então usada como molde para a realização do PCR secundário que utiliza oligonucleotídeos mais internos (AP2 e GSP2) à região amplificada. Para tal foram realizados os seguintes procedimentos como descrito a seguir.

a) Isolamento de DNA total de folha de algodão

A extração de DNA genômico foi realizada a partir de folhas de *G. hirsutum* segundo o kit Genomic DNA from Plant (Macherey-Nagel). As amostras de DNA foram quantificadas em NanoDrop ND-1000 (Thermo Fisher Scientific) com as leituras das absorbâncias na faixa 230-280 nm, analisadas em gel de agarose 1% corado com brometo de etídio e acondicionadas em freezer à -20°C.

b) Desenho dos oligonucleotídeos gene-específicos (GSPs)

Oligonucleotídeos complementares à porção 5' terminal da seqüência do EST representativo de cada gene validado foram desenhados com o intuito de amplificar produtos a montante da referida região. Os critérios adotados foram os seguintes: tamanho do oligonucleotídeo (mínimo: 25 bases; ótimo: 26 bases; máximo: 28 bases); temperatura de anelamento (mínimo: 66°C; ótimo: 67°C; máximo 68°C) e conteúdo em GC (mínimo: 40%; máximo: 60%).

Tabela 6: Pares de iniciadores utilizados para o isolamento da região promotora dos genes GhPGFS3, GhPGFS4 e GhPGFS8 através da técnica de Genome Walker.

Gene	Pares de iniciadores(GSP1/GSP2)
GhPGFS3_GSP1/GSP2_GW1	CACTATGGCCGGCTTATTCCCAAGT/ CTTCATGGAGACGATTGGAGAGGTGT
GhPGFS3_GSP1/GSP2_GW2	TCCTAGTTCAAATTCGGCAATTACAC/

	CGGCAATTACACCAATACACAAGATG
GhPGFS4_GSP1/GSP2_GW1	ATATCTGGGGTGCTACGTGGGTAAT/ CTGGTTCATGTCCATTCCAGAAAACA
GhPGFS8_GSP1/GSP2_GW1	GCAAACAGGACCTTGATAGGATTGGA/ TAGTGTTGGGATTAGGGCGAAAATG

c) Construção das bibliotecas de GW e realização da PCR

nested

Para a construção das bibliotecas foram utilizadas alíquotas de DNA genômico de algodão (2,5 µg) digeridas durante 16 horas a 37 °C com as enzimas de restrição DraI, EcoRV, PvuII e StuI separadamente. Em seguida o DNA digerido foi purificado via extração clorofórmio/fenol e ligado aos adaptadores. Tanto a purificação quanto a ligação dos adaptadores foram processadas seguindo as instruções do fabricante do kit. Após a ligação dos adaptadores, as amplificações das seqüências promotoras foram realizadas segundo os parâmetros da reação de PCR: 7 ciclos de 94°C por 25 segundos para a ativação da enzima polimerase e 72°C para anelamento dos iniciadores específicos e polimerização e 32 ciclos de 94°C por 25 segundos e 67°C por 3 minutos, condições utilizadas tanto na PCR primária quanto na secundária.

Para *GhPGFS3* o procedimento recomendado pelo fabricante foi realizado até a etapa do PCR primário, mas devido à dificuldade de obtenção de fragmentos únicos, 2µl do produto desta reação foi utilizado como molde para a PCR secundária, realizada de acordo com as condições: 94°C por 5 minutos para desnaturação inicial, seguido de 30 ciclos a 94°C por 5 minutos, 94°C por 30 segundos, 67°C por 1 minuto para anelamento dos oligonucleotídeos específicos, 72°C por 3 minutos para a polimerização, seguido de 7 minutos a 72°C para extensão final. Tais reações isolaram um fragmento de aproximadamente 500pb, que correspondem ao que chamamos de GW1. Novos primers foram desenhados para continuar o isolamento da região promotora (GW2). A PCR primária seguiu as condições: desnaturação inicial a 94°C por 2 minutos, seguida de 30 ciclos a 94°C por 30 segundos, 65°C para anelamento dos iniciadores por 30 segundos e extensão a 72°C por 2,5 minutos. Extensão final foi de 72°C por 5 minutos. Desta reação foram usados 2,5µl para a PCR secundária, realizada sob as mesmas condições que a PCR primária. Nesta reação foram isolados cerca de 1300pb.

d) Isolamento, purificação e clonagem das regiões promotoras

Os fragmentos únicos amplificados na etapa de GW foram purificados e concentrados utilizando-se o Kit DNA Clean & Concentrator (Zymo

Research) de acordo com especificações fornecidas pelo fabricante.

Os produtos de amplificação purificados foram em seguida inseridos no vetor pGEM® -T Easy (Promega). Alíquotas de aproximadamente 50 ng de DNA foram dispostas em microtubo de 1 ml juntamente com 50% de Tampão de Ligação 2X, 20 ng de vetor e 1,2U de T4 DNA ligase. A reação de ligação permaneceu por 1 hora à temperatura ambiente e o/n a 4°C. Após esse período, uma alíquota do produto de ligação foi utilizada para transformar *Escherichia coli* cepa DH5α por choque térmico.

e) Minipreparação de DNA plasmidial e sequenciamento dos insertos

Foram selecionados 5 clones de cada evento de transformação para validação quanto à presença do plasmídeo recombinante. Foi realizado uma PCR utilizando os primers M13 Foward (5'-CAGGAAACAGCTATGAC-3') e M13 Reverse (5'-GTAAAACGACGGCCAG-3') para identificação das colônias recombinantes. As que foram identificadas como positivas por conter um fragmento do tamanho esperado para cada gene foram repicadas em 4 ml de meio LB líquido acrescido de ampicilina 100 mg/l e mantidas sob agitação constante de 300 rpm a 37°C durante 16 horas. Após o período de incubação as amostras foram centrifugadas e sucessivas minipreparações de DNA plasmidial foram realizadas utilizando-se o GeneJet™ Plasmid Miniprep Kit (Fermentas).

Os plasmídeos foram digeridos com a enzima EcoRI para a confirmação da presença do inserto, e o DNA plasmidial (200 ng) foi submetido a uma reação de sequenciamento com os oligonucleotídeos M13 senso e antisense empregando o MegaBace 1000 DNA Analysis Systems (GE Healthcare). Foram seqüenciados 2 clones de cada evento. As seqüências foram analisadas utilizando-se o programa DNA Baser e um consenso a partir dos clones seqüenciados foi obtido.

f) Análise das seqüências e busca de domínios regulatórios

As seqüências obtidas foram submetidas a análises comparativas com seqüências depositadas no banco de dados do NCBI usando Blastn. Ao mesmo tempo foram realizadas buscas por motivos regulatórios normalmente presentes em regiões promotoras de genes de plantas utilizando o banco do PLACE database (<http://www.dna.affrc.go.jp/htdocs/PLACE/>) (HIGO et al., 1999).

EXEMPLO 4 – IDENTIFICAÇÃO DO SÍTIO DE INÍCIO DE TRANSCRIÇÃO DO GENE *GhPGFS3*, *GhPGFS4* E *GhPGFS8*

A identificação do sítio de início de transcrição (SIT) para o gene *GhPGFS3* foi feita através do kit GeneRacer™ (Invitrogen). Os procedimentos foram feitos seguindo as instruções do fabricante utilizando RNA total extraído de

botões florais de algodão de 6 mm de diâmetro. Primeiramente, foi utilizada a enzima CIP (calf intestinal phosphatase), que remove os grupamentos fosfato da extremidade 5' de mRNA truncados e outras moléculas de RNA não-mensageiros. Após purificação, restam somente mRNAs completos e a enzima TAP (tobacco acid pyrophosphatase) retira a estrutura CAP5', deixando livre uma extremidade fosfato na qual se ligará o adaptador fornecido pelo kit. Após a ligação do adaptador foi feita a transcrição reversa utilizando oligonucleotídeos do tipo Random Primer (fornecido pelo kit). Para obter a extremidade 5' do cDNA, isto é, o primeiro nucleotídeo a ser transcrito foram utilizados dois pares de oligonucleotídeos específicos a cada um dos genes em duas rodadas de amplificação (oligonucleotídeos utilizados GhPGSF3_GSP1/GSP2_GW1, GhPGSF4_GSP1/GSP2_GW1, GhPGSF8_GSP1/GSP2_GW1, Tabela 2).

No PCR primário se utilizou oligonucleotídeos mais externos relacionados ao adaptador (GeneRacer 5' primer) e ao cDNA (GSP1). O resultado da amplificação deste PCR primário é utilizado como molde para um PCR secundário que utiliza primers mais internos em relação ao adaptador (GeneRacer 5' Nested primer) e ao cDNA (GSP2). O fragmento resultante deste PCR secundário foi clonado, sequenciado e analisado de acordo com as orientações do fabricante (Figura 6).

EXEMPLO 5 – ANÁLISE FUNCIONAL DAS REGIÕES PROMOTORAS DOS GENES GhPGFS3, GhPGFS4 E GhPGFS8

a) Clonagem e subclonagem do promotor dos genes GhPGFS3, GhPGFS4 e GhPGFS8 em vetores bacterianos e em vetores binários de expressão em plantas

A seqüência dos promotores dos genes GhPGFS4 e GhPGFS8 foi clonada através da amplificação por PCR a partir do DNA genômico extraído a partir de folhas de *G. hirsutum*. Os pares de iniciadores pGhPGFS4-FW/pGhPGFS4-RV e pGhPGFS8-FW/pGhPGFS8-RV, descritos na Tabela 3, foram utilizados para a amplificação da região promotora dos genes GhPGFS4 e GhPGFS8, respectivamente. As reações de PCR foram realizadas em um total de 50µL, contendo 1µL de DNA (200 ng/µL), MgCl 1mM, 0,4mM de cada dNTP, 0,1µM de cada iniciador, 10µL do tampão de PCR Pfu 10X concentrado, e 1,25 U da enzima Pfu DNA Polimerase (Fermentas). Cada reação foi incubada por 3 minutos a 94°C, seguidos por 30 ciclos de amplificação de 30 segundos, a 94°C, 60 segundos a 47°C para a região promotora de GhPGFS4 e 48°C para a região promotora de GhPGFS8, e 2 minutos a 72°C. Finalmente, uma etapa final de extensão a 72°C, por 10 minutos, foi realizada. As bandas obtidas apresentaram os tamanhos esperados de aproximadamente 350 e 440 pb para as regiões promotoras dos genes GhPGFS4 (pGhPGFS4) e GhPGFS8 (pGhPGFS8), respectivamente (Figura 7, A e B).

Devido a dificuldade de amplificação da região promotora de GhPGFS3 (pGhPGFS3) em DNA genômico, a amplificação dos fragmentos de promotor desse gene foi realizada através da técnica de PCR de fusão (adaptado de SZEWCZYK et al, 2007). Foram utilizados como molde para a PCR os dois clones das duas etapas de isolamento da região promotora através de GenomeWalker (GW1 e GW2). Em ambos os casos os insertos foram previamente isolados, sejam por amplificação de PCR ou por digestão por ER para serem utilizados como molde na PCR de fusão. Os pares de iniciadores para o isolamento dos três fragmentos de promotor estão descritos na tabela 3. A PCR de fusão foi realizada num volume total de 20µl, contendo 50ng de cada molde GW1 e GW2, 4µl do tampão de PCR 5x HF Phusion, 0,2mM de cada dNTP, 0,5µM de cada iniciador e 0,4U de Phusion High-Fidelity DNA polymerase (Finnzymes). Cada reação foi incubada por 30 segundos a 98°C, seguidos de 30 ciclos a 98°C por 10 segundos, 60°C para o anelamento dos iniciadores específicos dos promotores por 30 segundos, seguidos de 1 minuto e 15 segundos de extensão da enzima a 72°C. A extensão final foi feita a mesma temperatura por 5 minutos. As bandas obtidas de aproximadamente 500, 800 e 1200pb correspondem aos tamanhos esperados para os três fragmentos de pGhPGFS3 (Figura 7, C). Os clones obtidos através de PCR de fusão foram confirmados por seqüenciamento.

Tabela 7: Pares de iniciadores utilizados para a reação de PCR para isolamento da região promotora dos genes GhPGFS3, GhPGFS4 e GhPGFS8 para clonagem no vetor de entrada pENTR-D-TOPO.

Gene	Iniciadores direto/reverso
p505_GhPGFS3	CACCTGCCGTGTGATTGATTGAA/ ATTGCCAATGTCACTTCCTCCCTTC
p827_GhPGFS3	CACCGCCGTGAATTGAAGGAAGAA/ ATTGCCAATGTCACTTCCTCCCTTC
p1250_GhPGFS3	CACCTTGTGACAGCCCAAATGAGA/ ATTGCCAATGTCACTTCCTCCCTTC
pGhPGFS4_FW/pGhPGFS4_RV	CACCCTAAAGAAACAGTAAATAGC/ TGTTATATATACCAGTAGAAGAAATT
pGhPGFS8_FW/pGhPGFS8_RV	CACCGAATTGAACTTTTAGTTT/ AGTGGTTTTTATTTAACTAGGAAAA

As reações de PCR foram purificadas utilizando-se do kit DNA Clean & Concentrator™ (Zymo Research), de acordo com as instruções do fabricante. Os fragmentos dos promotores foram então primeiramente clonados no vetor de entrada pENTR-D-TOPO (Invitrogen), para posterior recombinação em

vetores de expressão do sistema Gateway® (Invitrogen). As reações de ligação foram realizadas contendo 4µL da reação de PCR purificada, 1µL de solução salina (1,2M NaCl, 0,06M MgCl₂) diluída 1:4, e 1µL do vetor TOPO®, em um volume final de 6µL. Após 30 minutos à temperatura ambiente (22-23°C), 2µL da reação de ligação foi utilizado para transformar células de *Escherichia coli* TOP10 quimiocompetentes. As bactérias transformantes foram selecionadas em meio de cultura LB sólido (peptona 10g/L, extrato de levedura 5g/L, NaCl 5g/L, ágar 15g/L, pH7,0) contendo o antibiótico canamicina na concentração final de 50 µg/mL. Em seguida, a confirmação dos clones positivos para os três fragmentos de GhPGFS3 e pGhPGFS4 e pGhPGFS8 foi realizada através de PCR a partir do plasmídeo isolado das colônias positivas, utilizando-se dos mesmos pares de iniciadores utilizados na reação de isolamento do fragmento do promotor.

A clonagem dos promotores dos genes GhPGFS3, GhPGFS4 e GhPGFS8 no vetor binário Gateway® pKGWFS7 (Karimi et al., 2002) foi realizada através de recombinação. Nesse vetor, o promotor dos genes foram clonados a frente da sequência codificante do gene *gus*, que codifica a enzima β-glucuronidase (GUS), através da recombinação entre as regiões attL1/attR1 e attL2/attR2 dos vetores de entrada e de destinação, pENTR-D-TOPO/pKGWFS7, respectivamente.

As reações de recombinação foram realizadas nas seguintes condições: 1µL da enzima LR Clonase™ II (Invitrogen), 150ng do vetor de entrada pENTR-D-TOPO (contendo o promotor dos genes GhPGFS3, GhPGFS4 e GhPGFS8), 150ng do vetor de destinação pKGWFS7 e tampão TE (10mM Tris-HCl, 1mM EDTA) em um volume final de 5µL. Essa mistura foi incubada por 1 hora a 25°C. Em seguida, 1 µL de uma solução de proteinase K (2µg/µL) foi adicionado à cada reação anterior e uma nova incubação por 10 minutos a 37°C foi realizada. 2µL de reação foram então utilizados para a transformação de células de *E.coli* TOP10 quimiocompetentes, e os clones positivos foram desta vez selecionados em meio LB sólido contendo o antibiótico espectinomicina na concentração final de 50µg/µL. Os clones positivos foram em seguida confirmados através de reações de PCR de colônia, utilizando os mesmos pares de iniciadores descritos anteriormente neste item. O DNA plasmidial destes clones foi extraído e purificado com o kit GeneJet™ Plasmid Miniprep Kit (Fermentas) segundo instruções do fabricante, e 1µL de DNA (1 µg/µL) foi utilizado para a transformação de células de *Agrobacterium tumefaciens* GV3101 eletrocompetentes. A seleção dos clones positivos foi realizada em meio LB sólido, contendo 100µg/mL dos antibióticos rifampicina e spectinomicina. Após a confirmação através de reações de PCR de colônia, um clone contendo cada um dos promotores dos genes de algodão foi selecionado para a transformação de plantas da espécie *A. thaliana*.

b) Obtenção de plantas da espécie *Arabidopsis thaliana* expressando os promotores dos genes *GhPGFS3*, *GhPGFS4* e *GhPGFS8*

A obtenção de plantas expressando os promotores dos genes *GhPGFS3*, *GhPGFS4* e *GhPGFS8* foi realizada através do sistema de infiltração da inflorescência mediado por *Agrobacterium tumefaciens* (floral-dip) (DESFEUX et al., 2000). Para isso, uma colônia isolada de *A. tumefaciens*, contendo cada clone obtido nos item A desta seção, foi crescida por 48 horas a 28°C sob agitação de aproximadamente 200rpm, em 2mL de meio LB líquido (peptona 10g/L, extrato de levedura 5g/L, NaCl 5g/L, pH7,0), contendo 100µg/mL do antibiótico rifampicina e 100µg/mL do antibiótico espectinomicina. Essa cultura foi utilizada para inocular 200mL de meio LB líquido, contendo os mesmo antibióticos presentes na cultura anterior. Após 16 horas de crescimento a 28 °C sob agitação, a cultura foi centrifugada por 15 minutos a 4000rpm. O sobrenadante foi descartado, e as células foram ressuspensas em 200mL de uma solução contendo 5% de sacarose e 0,01% do surfactante Silwet L-77. As flores de plantas adultas da espécie *A. thaliana* ecotipo Erecta (LER) foram imersas nessa solução agitando-se levemente. Após 1 minuto, as plantas foram colocadas na posição horizontal em uma bandeja, e cobertas por um filme plástico com o objetivo de manter a umidade. No dia seguinte, o filme plástico foi retirado, e as plantas foram colocadas na posição vertical novamente. Cerca de 10 plantas foram transformadas para cada construção.

As sementes produzidas pelas plantas transformadas foram esterilizadas em uma solução de EtOH 70% e TWEEN 20 0,05% por 10 minutos, e semeadas em meio MS sólido (sais MS 4,6g/L, sacarose 20g/L, glicina 2mg/L; ácido nicotínico 5mg/L, piridoxina- HCl 0,5 mg/L, tiamina-HCl 0,1 mg/L, ágar 8g/L, pH5,8) (Murashige & Skoog, 1962) contendo o o antibiótico canamicina na concentração final de 50µg/mL. As linhagens transgênicas resistentes foram transferidas para potes contendo o substrato comercial Multiplant® em uma proporção de 2:1:0,5 (substrato:vermiculita:perlita) e dessa forma cultivadas sob condições de luz (fotoperíodo de 18 horas de luz/6 horas de escuro) e temperatura (22°C, ± 2°C) controlada. Em seguida as sementes da geração T0 de plantas contendo o promotor dos genes *GhPGFS3*, *GhPGFS4* e *GhPGFS8* no vetor pKGWFS7 foram submetidas novamente à seleção com canamicina, com o objetivo de identificar aquelas linhagens apresentando uma segregação de 3:1 entre plantas resistentes e sensíveis, indicando a presença de apenas uma inserção de T-DNA. 10-15 plantas resistentes da geração T1 de cada linhagem independente previamente selecionada foram transferidas para potes contendo substrato, e um novo teste de segregação foi realizado com as sementes produzidas por essas plantas, com o objetivo de identificar as linhagens T2 heterozigotas e homozigotas.

As plantas T2 cultivadas em placas apresentando aproximadamente 100% de resistência ao herbicida ou antibiótico utilizado foram transferidas para potes contendo substrato, segundo as condições mencionadas anteriormente, e as sementes T3 produzidas foram então utilizadas para os testes de análise da atividade promotora dos genes isolados através de ensaio histoquímico da coloração GUS.

c) Transformação transiente por biobalística

A transformação transiente por biobalística (modificado de RECH et al, 2008) foi realizada no Laboratório de Interação Molecular Planta-Praga na Embrapa Cenargen – DF. A preparação do DNA para bombardeamento foi feita utilizando 50µl de tungstênio, 50µl de cloreto de cálcio 2,5M e 5µl das construções contendo o promotor inserido do vetor pKGWFS7 (1µg/µl). Após agitação foram adicionados 20 µl de espermidina 2,5M, retornando a agitação em velocidade baixa por 10 minutos a fim de precipitar as partículas. As amostras foram centrifugadas a velocidade máxima por 15 segundos, o sobrenadante foi descartado e uma série de três lavagens com 150µl de etanol absoluto foi iniciada, seguida de agitação e nova centrifugação nas mesmas condições anteriores. Após as três lavagens foram adicionados 24µl de etanol absoluto. As amostras foram então submetidas a sonicação por 3 segundos, agitadas e 3,2µl foram depositados na membrana carreadora e dispostas em câmara de sílica por 15 minutos para secar.

Botões florais de *G. hirsutum* de 5-8 mm foram seccionados longitudinalmente e dispostos de maneira circular num raio de cerca de 5cm a partir do centro da placa por duas horas em meio osmótico de bombardeamento: meio MS (Aldrich-Sigma Chemical Co. Ltd.), 9g de sacarose e 2,4g de Fitigel em pH5,7.

O bombardeamento foi feito no equipamento PDS-1000/He system (Bio-Rad) sob pressão de gás hélio 1200psi e vácuo de 25polHg. Cada construção foi bombardeada em duplicata. Após o bombardeamento o material biológico foi deixado por duas horas no meio osmótico para posterior coloração de GUS.

d) Ensaio histoquímico de Coloração GUS

O ensaio histoquímico para detecção de atividade GUS foi realizado em inflorescências e plântulas segundo Jefferson (Jefferson et al., 1987). As inflorescências e plântulas foram inicialmente imersos em acetona 90% e deixadas a -20°C por 30 minutos para tirar clorofila e descolorir o material. Em seguida, o material foi lavado 2x com tampão fosfato 0,1M pH7.0. Após a lavagem, o material foi incubado em uma solução contendo 400µL de tampão fosfato 0,1M pH7.2; 4µL do substrato ácido 5-bromo-4-cloro-3-indolil-β-D-glicuronídeo (X-Gluc) 200mM; 1µL de K₄Fe(CN)₆ 0,5 mM e 4µL de K₃Fe(CN)₆ 0,05 mM para uma solução. O material foi deixado por 16h-24h em uma estufa a 37°C para o desenvolvimento da coloração. A atividade da

enzima foi interrompida retirando a solução de reação e adicionando etanol 70%. O material foi deixado no etanol por no mínimo 72hs para clareamento. Após esse tempo, as inflorescências foram analisadas e fotografadas em lupa e microscópio ótico.

5

Os exemplos e realizações aqui apresentados possuem caráter meramente ilustrativo, não sendo, portanto, limitativos à invenção, restando evidente para os especialistas na matéria que outras concentrações poderão ser empregadas, sem fugir ao escopo da invenção.

REIVINDICAÇÕES

1. Sequência de DNA caracterizada por apresentar SEQ ID NO 1, SEQ ID NO 2 e SEQ ID NO 3 para expressão dos genes GhPGFS3, GhPGFS4 e GhPGFS8, respectivamente, em tecidos florais e frutos de plantas para inibir o ataque de insetos e pragas.
2. Sequência de DNA, de acordo com a reivindicação 1, caracterizada por promover a expressão dos genes GhPGFS3, GhPGFS4 e GhPGFS8 em tecidos florais e frutos de *Gossypium hirsutum* para inibir o ataque do bicudo-do-algodoeiro.
3. Sequência de DNA, de acordo com as reivindicações 1 e 2, caracterizada por serem ligadas a genes heterólogos e/ou homólogos.
4. Sequência de DNA, de acordo com as reivindicações 1 a 3, caracterizada pelas sequências SEQ ID NO 1, SEQ ID NO 2 e SEQ ID NO 3 serem as regiões regulatórias localizadas a montante 5' dos genes GhPGFS3, GhPGFS4, e GhPGFS8 respectivamente.
5. Sequência de DNA, de acordo com as reivindicações 1 a 4, caracterizada pela SEQ ID NO 1 compreender a região regulatória do gene GhPGFS3 com similaridade com a família do *citocromo P450*
6. Sequência de DNA, de acordo com as reivindicações 1 a 5, caracterizada pela SEQ ID NO 2 compreender a região -regulatória do gene GhPGFS4 com similaridade com a sequência de β -galactosidase, *GhGal1*
7. Sequência de DNA, de acordo com as reivindicações 1 a 6, caracterizada pela SEQ ID NO 3 compreender a região -regulatória do gene GhPGFS8 com similaridade à sequência pectase liase (PLL).
8. Sequência de DNA, de acordo com as reivindicações 1 a 7, caracterizada pelo fragmento promotor do gene GhPGFS3 ter um tamanho de 1451 nucleotídeos, o fragmento promotor do gene GhPGFS4 ter um tamanho de 380 nucleotídeos e o fragmento promotor do gene GhPGFS8 ter um tamanho de 514 nucleotídeos.
9. Sequência de DNA, de acordo com as reivindicações 1 a 8, caracterizada
- a) pelo fragmento promotor da expressão do gene GhPGFS3 ter seu sítio de início de transcrição na posição -86 pb, contendo os elementos específicos de flor e fruto, combinados segundo a descrição abaixo: CARGCW8GAT (-553 e - 692 pb); ERELEE4 (-1123 pb); GTGANTG10 (-14, -76, -82, -298, -348, -499, -673, -791, -824, -866, -1077, -1102, -1137 e -1248 pb); MYBPZM (-

472 e -896 pb); POLLEN1LELAT52 (-289, -711, -730, -751, -811, -1020, -1065 e -1074 pb) e MYB.ph3 (-270 e -672 pb).

b) pelo fragmento promotor da expressão do gene GhPGFS4 ter seu sítio de início de transcrição na posição -124pb, contendo os
5 elementos específicos de flor e fruto combinados segundo a descrição abaixo: GTGANTG10 (-87 e- 269 pb); POLLEN1LELAT52 (-24, -251, -311 e- 352 pb) e TGTCACACMCUCUMISIN (-270 pb). .

c) pelo fragmento promotor da expressão do gene GhPGFS8 ter seu sítio de início de transcrição nas posições -14pb, -153pb e -313pb,
10 contendo os elementos específicos de flor e fruto combinados segundo a descrição abaixo: GTGANTG10 (-134 pb) e POLLEN1LELAT52 (-214 e -458 pb).

10. Polinucleotídeo com atividade promotora da expressão de genes caracterizado por compreender:

a) Sequências de DNA: SEQ ID NO 1, SEQ ID NO 2 e SEQ
15 ID NO 3 contendo promotores de expressão dos genes GhPGFS3, GhPGFS4 e GhPGFS8, respectivamente, que direcionam a expressão em tecidos florais e frutos de plantas para inibir o ataque de insetos e pragas;

b) Complementos das sequências SEQ ID NO 1, SEQ ID
NO 2 e SEQ ID NO 3;

c) Complementos reversos das sequências SEQ ID NO1,
20 SEQ ID NO 2 e SEQ ID NO 3

d) Sequências reversas das sequências SEQ ID NO 1, SEQ
ID NO 2 e SEQ ID NO 3.

11. Gene quimérico caracterizado por compreender
25 polinucleotídeos com sequências contendo promotores de expressão dos genes GhPGFS3, GhPGFS4 e GhPGFS8 caracterizados como sequências SEQ ID NO 1, SEQ ID NO 2 e SEQ ID NO 3 , com capacidade de direcionar uma sequência de polinucleotídeos de interesse na orientação sense ou antisense em tecidos florais e de frutos de plantas para inibir o ataque de insetos e pragas.

30 12. Gene quimérico de acordo com a reivindicação 11, caracterizado pelas sequências de polinucleotídeos de interesse poder ser uma região de codificação ou uma região de não codificação.

13. Gene quimérico de acordo com a reivindicação 11 a 12, caracterizado pela região de codificação ser isolada de um gene endógeno ou
35 heterólogo.

14. Gene quimérico de acordo com a reivindicação 11 a 13,

caracterizado

a) pela sequência contendo promotores de expressão dos genes GhPGFS3 ter um tamanho 1451 nucleotídeos, sítio de início de transcrição na posição -86 pb, contendo os elementos específicos de flor e fruto combinados segundo a descrição abaixo: CARGCW8GAT (-553 e - 692 pb); ERELEE4 (-1123 pb); GTGANTG10 (-14, -76, -82, -298, -348, -499, -673, -791, -824, -866, -1077, -1102, -1137 e -1248 pb); MYBPZM (-472 e -896 pb); POLLEN1LELAT52 (-289, -711, -730, -751, -811, -1020, -1065 e -1074 pb) e MYB.ph3 (-270 e -672 pb).

b) pela sequência contendo promotores de expressão dos gene GhPGFS4 ter um tamanho de 380 nucleotídeos, sítio de início de transcrição na posição -124pb, contendo os elementos específicos de flor e fruto combinados segundo a descrição abaixo: GTGANTG10 (-87 e- 269 pb); POLLEN1LELAT52 (-24, -251, -311 e- 352 pb) e TGTCACACMCUCUMISIN (-270 pb).

c) pela sequência contendo promotores de expressão dos gene GhPGFS8 ter um tamanho de 514 nucleotídeos, sítio de início de transcrição nas posições -14pb, -153pb e -313pb, contendo os elementos específicos de flor e fruto combinados segundo a descrição abaixo: GTGANTG10 (-134 pb) e POLLEN1LELAT52 (-214 e -458 pb)

15. Vetor recombinante caracterizado por conter um gene quimérico de acordo com a reivindicação 14.

16. Vetor recombinante caracterizado por compreender:

a) Polinucleotídeos com sequências SEQ ID NO 1, SEQ ID NO 2 e SEQ ID NO 3 contendo promotores de expressão dos genes GhPGFS3, GhPGFS4 e GhPGFS8 com capacidade de direcionar uma sequência de polinucleotídeos de interesse na orientação sense ou antisense em tecidos florais e frutos de plantas para inibir o ataque de insetos e pragas; e

b) Uma sequência de terminação.

17. Vetor recombinante de acordo com a reivindicação 15 caracterizado pelas sequências de polinucleotídeos de interesse poder ser uma região de codificação ou uma região de não codificação.

18. Vetor recombinante de acordo com as reivindicações 15 e 16, caracterizado pela região de codificação ser isolada de um gene endógeno ou heterólogo.

19. Vetor recombinante de acordo com a reivindicação 15 a 17, caracterizado

a) pela sequência contendo promotores de expressão dos

genes GhPGFS3 ter um tamanho 1451 nucleotídeos, sítio de início de transcrição na posição -86 pb, contendo os elementos específicos de flor e fruto combinados segundo a descrição abaixo: CARGCW8GAT (-553 e - 692 pb); ERELEE4 (-1123 pb); GTGANTG10 (-14, -76, -82, -298, -348, -499, -673, -791, -824, -866, -1077, -1102, -1137 e -1248 pb); MYBPZM (-472 e -896 pb); POLLEN1LELAT52 (-289, -711, -730, -751, -811, -1020, -1065 e -1074 pb) e MYB.ph3 (-270 e -672 pb). .

b) pela sequência contendo promotores de expressão dos gene GhPGFS4 ter um tamanho de 380 nucleotídeos, sítio de início de transcrição na posição -124pb, contendo os elementos específicos de flor e fruto combinados segundo a descrição abaixo: GTGANTG10 (-87 e- 269 pb); POLLEN1LELAT52 (-24, -251, -311 e- 352 pb) e TGTCACACMCUCUMISIN (-270 pb). .

c) pela sequência contendo promotores de expressão dos gene GhPGFS8 ter um tamanho de 514 nucleotídeos, sítio de início de transcrição nas posições -14pb, -153pb e -313pb, contendo os elementos específicos de flor e fruto combinados segundo a descrição abaixo: GTGANTG10 (-134 pb) e POLLEN1LELAT52 (-214 e -458 pb)

20. Célula transformada caracterizada por conter um vetor recombinante de acordo com qualquer um das reivindicações 15 a 19.

21. Planta, ou uma parte, ou um propágulo ou progênie da mesma caracterizada por compreender um vetor recombinante de acordo com qualquer um das reivindicações 15 a 19.

22. Método para modificar a expressão de genes em um organismo caracterizado por incorporar estavelmente no genoma do organismo um vetor recombinante de acordo com qualquer uma das reivindicações 15 a 19 ou um gene quimérico de acordo com uma das reivindicações 11 a 14.

23. Método de acordo com a reivindicação 22 caracterizado pelo organismo ser uma planta, mais especificadamente um variedade de algodão *Gossypium hirsutum*.

24. Método para produzir uma planta tendo a expressão de genes específicos de tecido floral e frutos modificados, dito método caracterizado por compreender as etapas:

a) Transformar uma célula de planta, tecido, órgão ou embrião um vetor recombinante de acordo com qualquer uma das reivindicações 15 a 18 ou um gene quimérico de acordo com qualquer uma das reivindicações 11 a 14;

b) Selecionar células transformadas, calos de células, embriões ou sementes;

c) Regenerar plantas maduras de células transformadas, calos de células, embriões ou sementes selecionados da etapa (b);

d) Selecionar plantas maduras da etapa (c) tendo a expressão do gene modificada quando comparada com uma planta transformada.

5 25. Método para inibição do ataque de insetos e pragas em plantas caracterizado por utilizar sequências SEQ ID NO 1, SEQ ID NO 2 e SEQ ID NO 3 contendo promotores de expressão dos genes GhPGFS3, GhPGFS4 e GhPGFS8, respectivamente, em tecidos florais e frutos, e cada um dos promotores dos genes, GhPGFS3, GhPGFS4 e GhPGFS8 estejam ligados a um gene heterólogo
10 e/ou endógeno.

26. Processo para obtenção de planta transgênica resistente ao ataque de insetos e pragas caracterizado por:

- isolar DNA total da folha de algodão;

15 - obtenção dos genes específicos GhPGFS3, GhPGFS4 e GhPGFS8;

- construção de sequência de DNA SEQ ID NO 1, SEQ ID NO 2 e SEQ ID NO 3;

- isolar, purificar e clonar regiões promotoras

- construção de DNA plasmidial.

20 27. Composição de fragmentos de sequência de DNA caracterizada por conter as sequências SEQ ID NO 1, SEQ ID NO 2 e SEQ ID NO 3, promotores de expressão dos genes GhPGFS3, GhPGFS4 e GhPGFS8, respectivamente em tecidos florais e frutos, de plantas para inibir o ataque de insetos e pragas.

25 28. Composição, de acordo com a reivindicação 27, caracterizada pelas sequências SEQ ID NO 1, SEQ ID NO 2 e SEQ ID NO 3 serem as regiões regulatórias localizadas a montante 5' das sequências dos genes GhPGFS3, GhPGFS4 e GhPGFS8 respectivamente.

30 29. Composição, de acordo com as reivindicações 27 e 28, caracterizada pelo fragmento promotor do gene GhPGFS3 ter um tamanho de 1451 nucleotídeos, o fragmento promotor do gene GhPGFS4 ter um tamanho de 380 nucleotídeos e o fragmento promotor do gene GhPGFS8 ter um tamanho de 514 nucleotídeos.

35 30. Composição, de acordo com as reivindicações 27 a 29, caracterizada:

a) pela sequência contendo promotores de expressão dos

genes GhPGFS3 ter seu sítio de início de transcrição na posição -86 pb, contendo os elementos específicos de flor e fruto combinados segundo a descrição abaixo: CARGCW8GAT (-553 e - 692 pb); ERELEE4 (-1123 pb); GTGANTG10 (-14, -76, -82, -298, -348, -499, -673, -791, -824, -866, -1077, -1102, -1137 e -1248 pb); MYBPZM (-472 e -896 pb); POLLEN1LELAT52 (-289, -711, -730, -751, -811, -1020, -1065 e -1074 pb) e MYB.ph3 (-270 e -672 pb). .

b) pela sequência contendo promotores de expressão dos gene GhPGFS4 ter seu sítio de início de transcrição na posição -124pb, contendo os elementos específicos de flor e fruto combinados segundo a descrição abaixo: GTGANTG10 (-87 e- 269 pb); POLLEN1LELAT52 (-24, -251, -311 e- 352 pb) e TGTCACACMCUCUMISIN (-270 pb).

c) pela sequência contendo promotores de expressão dos gene GhPGFS8 ter seu sítio de início de transcrição nas posições -14pb, -153pb e -313pb, contendo os elementos específicos de flor e fruto combinados segundo a descrição abaixo: GTGANTG10 (-134 pb) e POLLEN1LELAT52 (-214 e -458 pb).b.

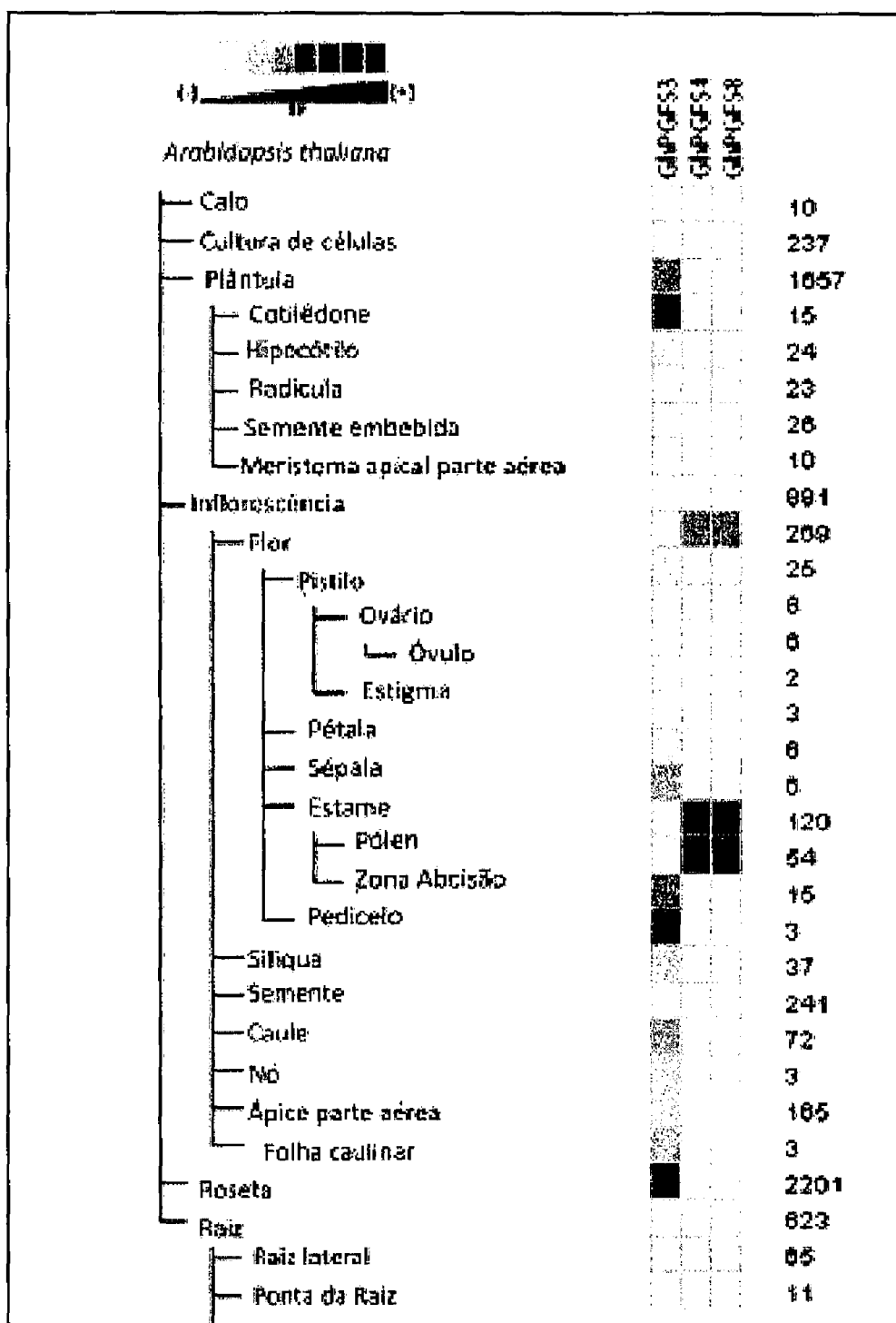


FIGURA 1

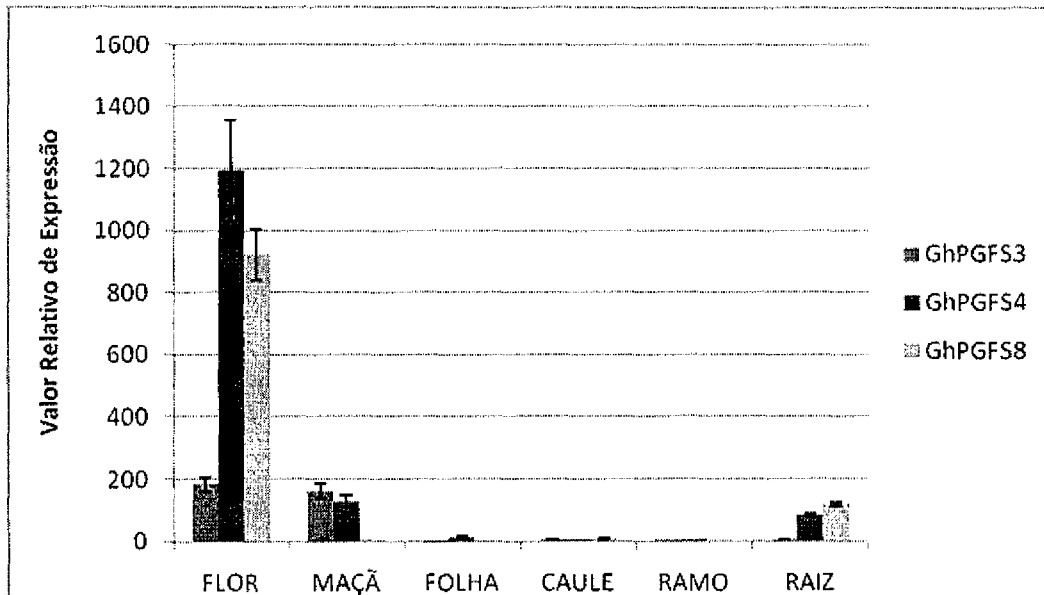


FIGURA 2

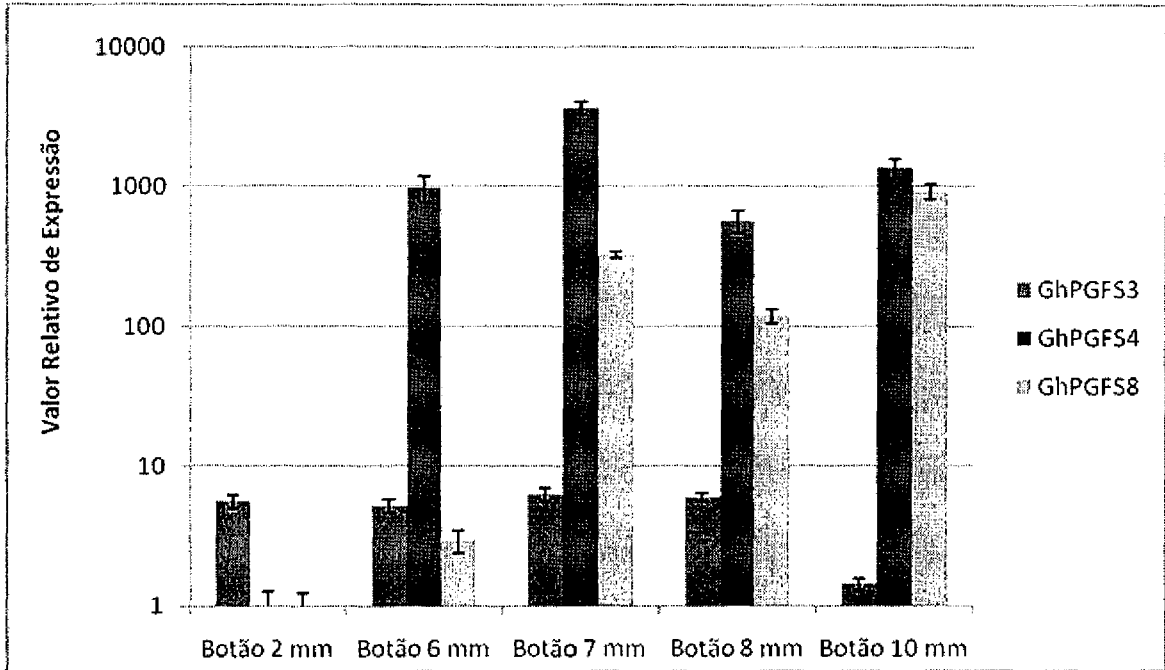


FIGURA 3

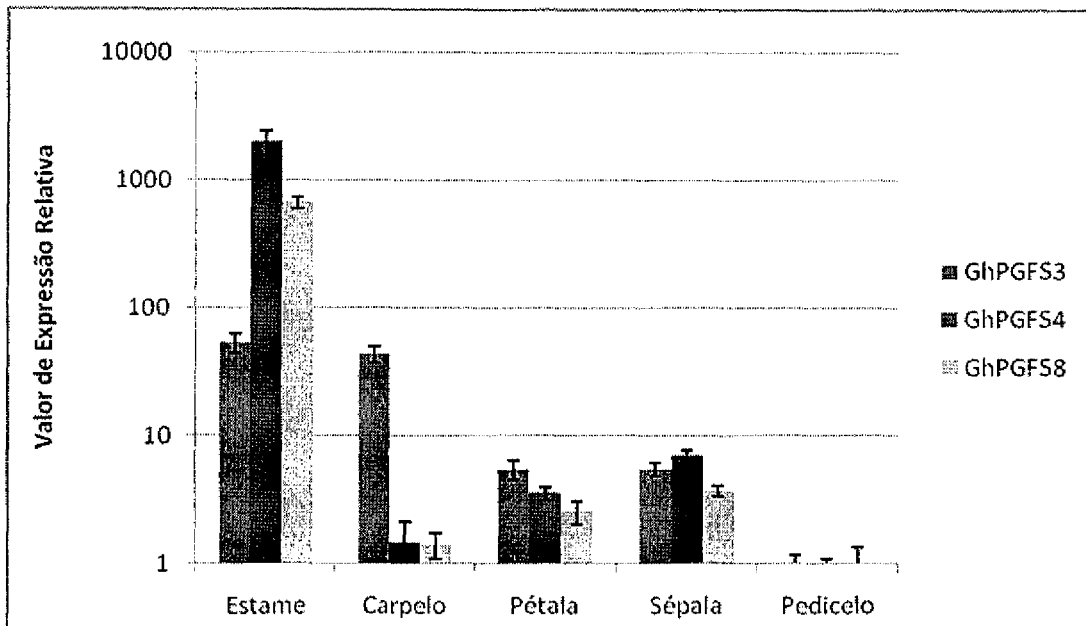


FIGURA 4

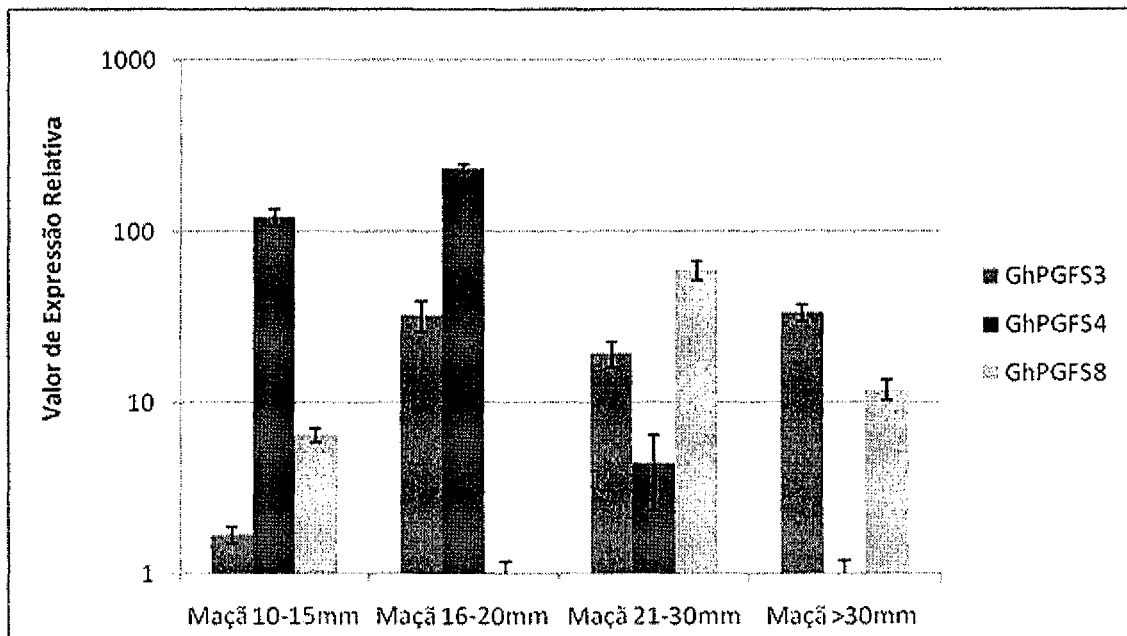


FIGURA 5

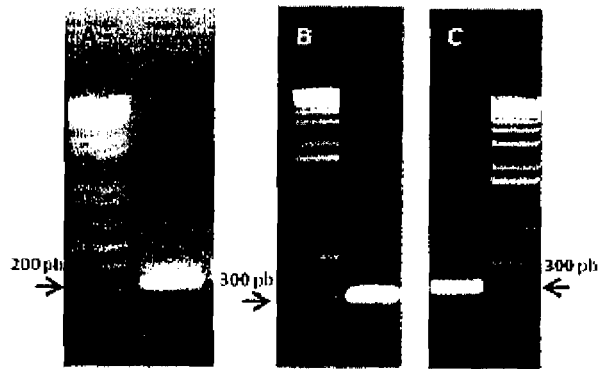


FIGURA 6

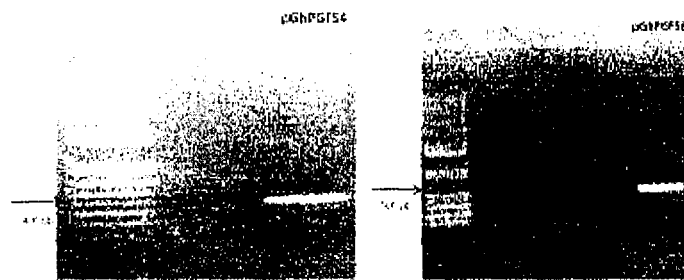


FIGURA 7

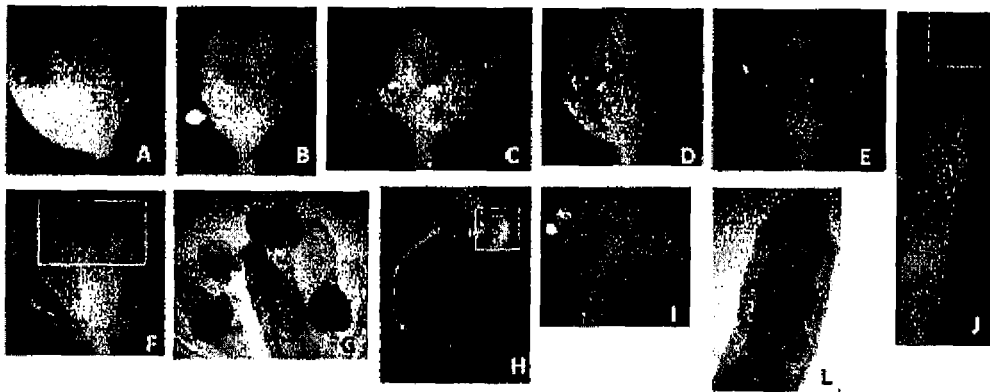


FIGURA 8

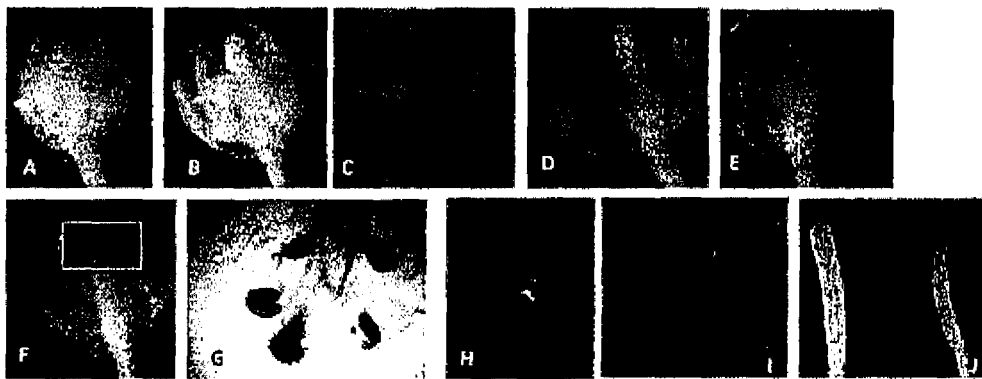


FIGURA 9



FIGURA 10

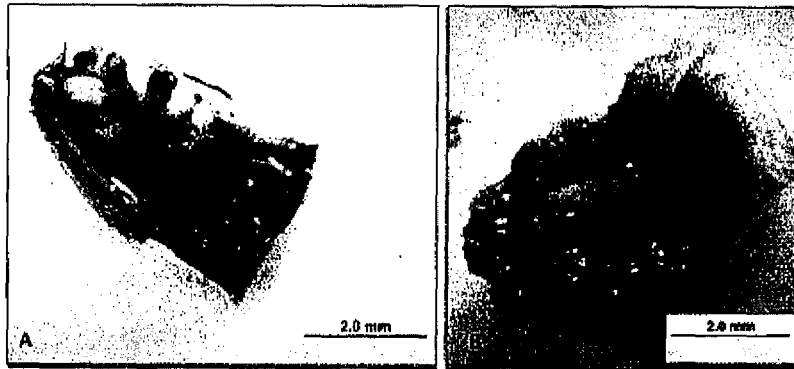


FIGURA 11

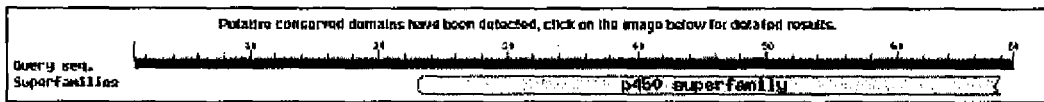


FIGURA 12

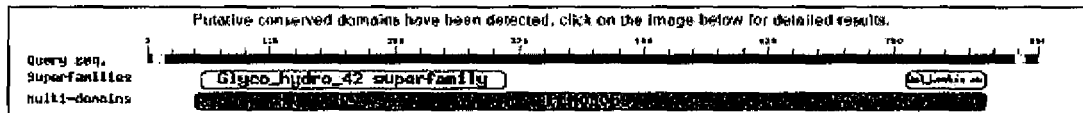


FIGURA 13

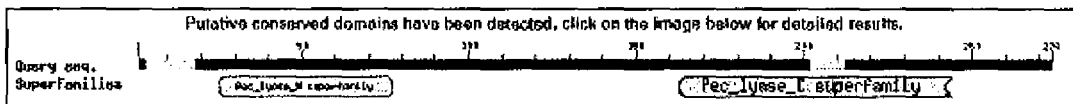


FIGURA 14

RESUMO

Patente de Invenção para "COMPOSIÇÕES E MÉTODOS PARA MODIFICAR A EXPRESSÃO DE GENES USANDO PROMOTOR DE GENES ESPECÍFICOS DE FLOR E FRUTO DE PLANTAS"

5 A presente invenção refere-se a promotores de expressão de genes específicos em plantas.

Mais especificamente, a presente invenção proporciona uma sequência de nucleotídeos para a expressão de genes de interesse em flores e frutos de *Gossypium hirsutum* para produção de plantas geneticamente modificadas que sejam capazes de resistir ao ataque de pragas, tal como bicudo-do-algodoeiro (*Anthonomus grandis*).

10 Nomeadamente, a presente invenção descreve construções de DNA que contenham os promotores dos genes de interesse, para criação de plantas geneticamente modificadas, bem como um método para modificação da expressão gênica.

15