

BEATRIZ DE ALMEIDA BARROS

**CONSTRUÇÃO DE CASSETES DE EXPRESSÃO PARA
SILENCIAMENTO GÊNICO DE FATORES ANTINUTRICIONAIS DA
SOJA, VIA INTERFERÊNCIA POR RNA**

Dissertação apresentada à
Universidade Federal de Viçosa,
como parte das exigências do
Programa de Pós-Graduação em
Genética e Melhoramento, para
obtenção do título de *Magister
Scientiae*.

**VIÇOSA
MINAS GERAIS – BRASIL
2006**

**Ficha catalográfica preparada pela Seção de Catalogação e
Classificação da Biblioteca Central da UFV**

T

B277c
2006

Barros, Beatriz de Almeida, 1981-

Construção de cassetes de expressão para silenciamento
gênico de fatores antinutricionais da soja, via interferência
por RNA / Beatriz de Almeida Barros. – Viçosa : UFV, 2006.
ix, 58f. : il. ; 29cm.

Orientador: Everaldo Gonçalves de Barros.

Dissertação (mestrado) - Universidade Federal de Viçosa.
Inclui bibliografia.

1. Soja - Melhoramento genético. 2. Genética molecular.
3. Clonagem molecular. 4. Regulação da expressão gênica.
I. Universidade Federal de Viçosa. II. Título.

CDD 22.ed. 633.342

BEATRIZ DE ALMEIDA BARROS

CONSTRUÇÃO DE CASSETES PARA SILENCIAMENTO GÊNICO DE
FATORES ANTINUTRICIONAIS DA SOJA, VIA INTERFERÊNCIA POR
RNA

Dissertação apresentada à
Universidade Federal de Viçosa,
como parte das exigências do
Programa de Pós-Graduação em
Genética e Melhoramento, para a
obtenção do título de *Magister
Scientiae*.

APROVADA: 29 de setembro de 2006.

Prof. Luiz Orlando de Oliveira

Prof. Wagner Campos Otoni

Profa. Andréa de Oliveira Barros Ribon

Profa. Andreia Barcelos Passos Lima

Prof. Everaldo Gonçalves de Barros
(Orientador)

AGRADECIMENTOS

A Deus, pelo dom e pela alegria de viver.

À Universidade Federal de Viçosa e ao Programa de Pós-Graduação em Genética e Melhoramento por toda a minha formação acadêmica.

Ao Instituto de Biotecnologia Aplicada à Agropecuária pelo ambiente de trabalho.

Aos meus pais, Tadeu e Maria Amélia, pelo apoio e compreensão nos momentos de ausência. Obrigado por me lembrarem que eu sempre poderei voltar pra casa... Vocês são meu exemplo de vida.

À minha irmã e afilhada pelo companheirismo e amizade.

Ao Leandro pelo cuidado, amor e por me ensinar a ser uma pessoa melhor.

Ao meu orientador, professor Everaldo, pela confiança, apoio, incentivo, por ter contribuído para a minha formação profissional e pelo exemplo de profissionalismo.

Aos professores Elza e Maurilio pela amizade e confiança.

Aos professores Wagner, Andréa Ribon e Luiz Orlando por terem aceitado participar da minha banca examinadora.

À Andreia Barcelos por me auxiliar no início dos experimentos (quando tudo ainda era uma caixinha de surpresas), pelo incentivo e ajuda durante a condução dos mesmos e, finalmente, por também participar da minha banca examinadora.

Às meninas do Laboratório de Sequenciamento, por transformarem o dia-a-dia do laboratório em momentos muito agradáveis. Um agradecimento especial à Poly e Cassi, pela amizade, carinho, conselhos, incentivo, e, principalmente pela presença durante as risadas e “encrencas”. Afinal de contas, divas unidas...

Aos colegas dos Laboratórios Genética Molecular de Plantas I e II e ao Márcio Mendes pela amizade.

Às companheiras de República: Lygia, Tati, Marinne, Rejane e Cinthia pela paciência, amizade e pela convivência que nos torna quase uma família e ameniza a saudade de casa.

E a todos que de alguma forma contribuíram para a realização deste trabalho.

BIOGRAFIA

BEATRIZ DE ALMEIDA BARROS, filha de Tadeu Quirino de Almeida e Maria Amélia Barros, nasceu em 23 de outubro de 1981, em Aiuruoca, Minas Gerais.

Em 1999, concluiu o Ensino Médio na Escola Estadual Maria do Carmo Lima Pinto, em Alagoa, Minas Gerais.

Em 2001, iniciou o curso superior na Universidade Federal de Viçosa, onde em 28 de janeiro de 2005, se graduou Bacharel em Ciências Biológicas. Durante a graduação desenvolveu atividades de Iniciação Científica na linha de pesquisa de Genética Molecular de Plantas (2002/2004).

Em março de 2005, iniciou o curso de Mestrado em Genética e Melhoramento na Universidade Federal de Viçosa, submetendo-se a defesa em 29 de setembro de 2006.

SUMÁRIO

RESUMO	vi
ABSTRACT	vii
1. INTRODUÇÃO GERAL	1
2. REFERÊNCIAS BIBLIOGRÁFICAS	5

CAPÍTULO 1

Isolamento, clonagem e análise *in silico* de um segmento da região promotora do gene da subunidade α da β -conglitinina de soja.

RESUMO	7
1. INTRODUÇÃO	9
2. MATERIAL E MÉTODOS	
2.1. Material Vegetal	11
2.2. Construção de <i>primers</i> específicos	11
2.3. Extração de DNA genômico	11
2.4. Amplificação e clonagem no vetor pGEM-TEasy	12
2.5. Transformação de <i>E. coli</i> DH5 α ultracompetentes e análise de transformantes	13
2.6. Caracterização <i>in silico</i> da região promotora da subunidade α da β -conglitinina	14
2.7. Clonagem da região promotora da subunidade α da β -conglitinina no vetor pKANNIBAL	15
3. RESULTADOS E DISCUSSÃO	16
4. CONCLUSÕES	20
5. PERSPECTIVAS	20
6. REFERÊNCIAS BIBLIOGRÁFICAS	21

CAPÍTULO 2

Construção de cassetes de expressão para silenciamento gênico, via interferência por RNA, de fatores antinutricionais da semente de soja.

RESUMO	25
1. INTRODUÇÃO	27
2. MATERIAL E MÉTODOS	
2.1. Material Vegetal	32
2.2. Extração de RNA total e RT-PCR	32
2.3. Construção de <i>primers</i> específicos e isolamento dos fragmentos dos genes que codificam a P34 e os inibidores BBI	34
2.4. Amplificação e clonagem no vetor pGEM-TEasy	36
2.5. Transformação de <i>E. coli</i> DH5 α ultracompetentes e análise dos transformantes	36
2.6 Clonagem em vetores pKANNIBAL e pBKN	38
2.7 Clonagem no vetor pCAMBIA 3301	39
3. RESULTADOS E DISCUSSÃO	40
4. CONCLUSÕES	53
5. PERPECTIVAS	53
6. REFERÊNCIAS BIBLIOGRÁFICAS	54

RESUMO

BARROS, Beatriz de Almeida, M. Sc. Universidade Federal de Viçosa, setembro de 2006.
Construção de cassetes de expressão para silenciamento gênico de fatores antinutricionais da soja, via interferência por RNA. Orientador: Everaldo Gonçalves de Barros. Co-orientadores: Maurilio Alves Moreira e Elza Fernandes de Araújo.

A soja é uma das mais importantes culturas do mundo e como apresenta um alto teor protéico, ela tem sido muito utilizada na alimentação animal e humana. Apesar disso, as proteínas encontradas na semente desta leguminosa não são consideradas ideais por apresentarem baixo teor de metionina e lisina, além de fatores alergênicos e de inibidores de proteases. Dentre estes fatores antinutricionais estão a proteína P34 e os inibidores de protease do tipo Bowman-Birk. A ausência destes fatores aumentaria a qualidade da semente de soja, bem como a tornaria disponível para uma faixa mais ampla de consumidores. Recentes avanços na biotecnologia vegetal têm possibilitado o silenciamento completo ou em altos níveis de genes específicos. Uma destas técnicas é conhecida como interferência por RNA ou *RNA interference*. Este método é altamente eficiente e a clivagem do mRNA alvo é induzida pela presença de pequenos RNAs dupla fita na célula. O objetivo deste trabalho foi a construção de cassetes de expressão para silenciamento gênico, via interferência por RNA, dos inibidores de protease do tipo Bowman-Birk e da proteína P34 em sementes de soja. Para a expressão semente-específica dos transgenes, *primers* específicos foram desenhados para a amplificação da região promotora do gene da subunidade α da β -conglucininina. A estes oligonucleotídeos foram adicionados sítios de restrição – *Sac* I e *Xho* I - adequados para sua clonagem em vetores de expressão em plantas. Foram clonados 634 pb no vetor pGEM-TEasy. A clonagem foi confirmada por PCR, clivagem enzimática e sequenciamento do clone isolado. A sequência clonada foi analisada *in silico* para a identificação de cis-elementos relacionados à expressão semente específica de genes colocados sob controle deste promotor. Os elementos TATA *box*, CAAT *box*, RY LEG *box* e RY FLEB *box* encontrados. O fragmento de interesse foi, então, isolado do vetor pGEM-TEasy por clivagem com as enzimas *Sac* I e *Xho* I e inseridos em vetores de expressão em plantas (pKANNIBAL) originando o vetor pBKN. Para a

construção dos cassetes de expressão, a escolha da região do cDNA de cada gene de interesse a ser amplificada foi determinada com o auxílio do programa BLOCK-iTTM RNAi Designer (INVITROGEN), e os *primers* específicos foram desenhados de acordo seqüências depositadas no banco de dados público – *GenBank*. Inicialmente foi feita, por meio de RT-PCR, uma análise da expressão dos genes de interesse durante o desenvolvimento da semente e em demais órgãos da planta (raiz, caule e folha). Essa análise evidenciou a expressão contínua e abundante desses genes durante todo o desenvolvimento da semente bem como a presença de transcritos correspondentes em todos os órgãos vegetais analisados. Após sua amplificação, todos os fragmentos obtidos a partir de cDNA de semente foram sequenciados e sua identificação foi confirmada por meio de alinhamento utilizando o programa BLAST. Para a clonagem da seqüência *sense*, os vetores pKANNIBAL e pBKN, além dos fragmentos amplificados referentes aos genes de interesse, foram clivados com as enzimas *Xho* I e *Kpn* I. A reação de ligação foi realizada utilizando T4 DNA ligase e células de *Escherichia coli* ultracompetentes foram transformadas por meio de choque térmico. Os transformantes foram analisados por PCR e reação de clivagem enzimática. A clonagem do fragmento *antisense* foi realizada da mesma forma, utilizando as enzimas *Xba* I e *Cla* I e os clones contendo o fragmento *sense* como vetores. Em seguida, as construções foram transferidas para o vetor binário pCAMBIA 3301 e confirmadas por PCR.

ABSTRACT

BARROS, Beatriz Almeida. M. Sc., Universidade Federal de Viçosa, September 2006.
Construction of expression cassettes for RNAi-based silencing of genes encoding antinutritional factors in soybean seeds. Adviser: Everaldo Gonçalves de Barros.
Co-advisers: Maurilio Alves Moreira and Elza Fernandes de Araújo.

Soybean is one of the most important crops in the world extensively used as a food and feed source. However, the proteins present in soybean seeds are not considered ideal because they contain low amounts of the essential amino acids methionine and lysine. Adverse nutritional and other effects following consumption of raw soybean meal have been attributed to the presence of endogenous inhibitors of digestive enzymes and allergenic factors. Among these proteins are the P34 protein and the Bowman-Birk protease inhibitors. The absence of these factors would increase the nutritional quality of the soybean seeds, as well as the availability of this product to large number of consumers. Recent advances in plant biotechnology have enabled the silencing of specific genes. One of the techniques used to achieve this goal is RNA interference (RNAi). This method of posttranscriptional gene silencing is highly efficient, especially when the construct results in the formation of a intron-spliced hairpin RNA. It is proposed that RNAi works by enzymatic RNA degradation induced by double-stranded RNA. The main goal of this work was to construct expression cassettes for RNAi-based silencing of genes encoding antinutritional factors in soybean seeds. To achieve seed-specific expression, specific primers were designed for the amplification of the promoter region of the alpha-subunit of the beta-conglycinin gene. Appropriate restriction sites - *Sac* I and *Xho* I - were added to the primers which were used to amplify a 634 bp fragment. This product was cloned into a pGEM-TEasy vector and cloning was verified by PCR, enzymatic digestion and sequencing. The sequenced clone was analyzed *in silico* for identification of cis-elements related to seed specific expression. The elements TATA box, CAAT box, RY LEG box and RY FLEB box were found. The fragment of interest was isolated from vector pGEM-TEasy by cleavage with enzymes *Sac* I and *Xho* I and inserted in the expression vector pKANNIBAL. This new vector was designated pBKN. For the construction of the expression cassettes, primers were designed according to published sequences deposited in the *GenBank*. The fragments of interest were isolated by the RT-PCR method using RNA isolated from soybean seeds.

RT-PCR was also used to analyze the expression of the gene of interest in the seed and in other organs of the plant (root, stem and leaves). These analyses showed that the expression of these genes was continuous and abundant during all stages of seed development, and the corresponding transcripts were present in all organs analyzed. The fragments obtained by RT-PCR were sequenced and their identity was confirmed by BLAST analysis. To clone the sense sequence, the vectors pKANNIBAL and pBKN, and the fragments of interest were cleaved with enzymes *Xho* I and *Kpn* I. The ligation reaction was catalyzed by T4 DNA ligase and competent cells of *Escherichia coli* were transformed by heat shock. Cloning was verified by PCR and enzymatic digestion. The cloning of the antisense fragment was done using restriction enzymes *Xba* I and *Cla* I and vectors containing the sense fragments. The complete expression cassettes were cloned into the binary vector pCAMBIA 3301 and verified by PCR.

1. INTRODUÇÃO GERAL

A soja (*Glycine max* L. Merrill) é uma das mais importantes culturas do mundo e a maior do Brasil em área plantada. O Brasil é o segundo maior produtor mundial desta leguminosa e estima-se que a safra 2005/2006 atinja uma área cultivada de 22,138 milhões de hectares – 4,7% inferior à safra anterior - e uma produção de 53,8 milhões de toneladas, superando as 51,4 toneladas produzidas na safra 2004/2005. Esse incremento é decorrência de um aumento de 9,8% na produtividade, que passou de 2.208 para 2.424 kg/ha (www.conab.gov.br). Neste contexto, o cultivo de soja geneticamente modificada, a *Roundup Ready*[®], tem sido significativo. Em 2005, com a aprovação da nova lei de Biossegurança (Lei Nº 11.105) que regulamentou a liberação do plantio e comércio da soja RR no país, a área de plantio desta cultura atingiu 9,4 milhões de hectares (James, 2006) – cerca de 42,5% do total de área cultivada.

A soja constitui um bom alvo para o melhoramento genético, via transformação, por representar uma cultura de grande interesse econômico e nutricional. Modificações usando técnicas de engenharia genética podem facilitar o rápido desenvolvimento de novas variedades com características de interesse agrônomo e industrial (Passos-Lima, 2005). Esforços diretos são agora empregados para a utilização da engenharia de vias metabólicas para alterar a constituição nutricional das plantas. Isto depende de um detalhado conhecimento do metabolismo vegetal e de suas enzimas constituintes.

A principal estratégia empregada, que utiliza a transgenia, envolve o aumento da expressão de genes biossintéticos sob controle de promotores constitutivos. Para a redução dos níveis de produtos gênicos indesejáveis, dois métodos têm sido comumente utilizados: a interrupção (*knockout*) e o silenciamento de genes. Na interrupção de genes, a sequência alvo é mutada para eliminar sua expressão ou função, enquanto que no silenciamento, a destruição do transcrito ou a inibição da transcrição são induzidas. Esta última via apresenta maior facilidade de manipulação genética e análise das plantas transgênicas, além de poder ser ativada de forma temporal e espacial por meio da utilização de promotores específicos (Tang & Galili, 2004).

Das metodologias de silenciamento gênico induzido por RNA dupla fita (dsRNA), a interferência por RNA (*RNA interference*), é a mais eficiente (Smith, 2000).

O silenciamento gênico por meio de RNA é um fenômeno característico de organismos eucariotos e tem importantes papéis em diversos processos biológicos incluindo regulação do desenvolvimento (Baulcombe, 2004), defesa antiviral (Waterhouse *et al.*, 2001; Zerbini *et al.*, 2005) e remodelamento da cromatina (Lippman & Martienssen, 2004; Qi & Hannon, 2005).

Em plantas, existem três rotas de silenciamento via RNA (Baulcombe, 2004). A primeira compreende um silenciamento gênico pós-transcricional (PGTS) mediado por pequenos RNAs de interferência (siRNAs) que são processados a partir de longos RNAs dupla-fita (dsRNAs). A origem desses dsRNAs inclui a replicação de RNAs intermediários de vírus que infectam plantas, seqüências repetidas invertidas de transgenes e produtos das RNA polimerases dependentes de RNA (RdRps). A segunda via envolve uma classe de pequenos RNAs endógenos conhecidos como micro-RNAs (miRNAs) que são gerados a partir de precursores, RNAs dupla fita em estrutura de grampo, que são transcritos de genes miRNAs. Os miRNAs regulam negativamente a expressão gênica pelo pareamento de bases com o mRNA alvo levando à degradação do mesmo, à inibição da tradução ou a ambos. A terceira via envolve o silenciamento gênico transcricional (TGS) mediante metilação de histonas e DNA e, conseqüentemente, remodelamento da cromatina (Qi & Hannon, 2005).

De uma forma geral, o passo inicial da via de silenciamento estimulada por RNA está ligado ao aparecimento de dsRNAs na célula, similares ou perfeitamente iguais em seqüência ao gene a ser silenciado. Os dsRNAs podem ser endógenos ou exógenos (seqüências transgênicas ou virais).

Em plantas, quando um RNA transgênico, contendo seqüências invertidas, é gerado e assume a estrutura de grampo, ele é reconhecido pela DCL 2 (*Dicer-like 2*) (Jones-Rhoades *et al.*, 2006) e clivado em pequenos (21-22 nt) e longos (24-26 nt) siRNAs (Zamore, 2004). Proteínas geralmente conhecidas como *Dicer* são enzimas, altamente conservadas, compostas de dois domínios modulares de RNase III, um domínio de helicase, um domínio PAZ de provável transferência de substrato e um domínio de função desconhecida (DUF283) e um motivo de ligação a RNA fita-dupla (Bernstein *et al.*, 2001). A *Dicer* atua de forma dimérica, embora um dos sítios ativos de cada enzima seja defectivo (Hannon, 2002), e processam dsRNAs em pequenos RNAs

(siRNAs) de 21-25 nucleotídeos de tamanho, dependendo da espécie. Após serem originados, os siRNAs são metilados e transportados ao citosol, por meio de uma proteína transportadora HASTY de maneira Ran-dependente (Jones-Rhoades *et al.*, 2006).

No citosol, esses siRNAs de 21 nt são incorporados aos complexos efetores RISC (complexo de silenciamento induzido por RNA) que induzem a clivagem e posterior degradação do mRNA alvo (Meins *et al.*, 2005). RISCs são complexos ribonucleoproteicos que contêm membros da família das proteínas Argonautas (com domínios PAZ-PIWI), siRNAs e mRNAs complementares; além de muitos outros fatores importantes para a sua funcionalidade (Hannon, 2002). As proteínas da família das Argonautas parecem garantir a especificidade do RISC e, juntamente com outras proteínas e enzimas existentes no complexo, determinam a especificidade da resposta a um siRNA particular.

É provável que o domínio PAZ das Argonautas esteja relacionado à ligação aos miRNAs ou siRNAs gerados ou liberados pela *Dicer* (Song *et al.*, 2003) e o domínio PIWI deve interagir com e inibir os domínios de RNase III e de ligação a dsRNA da *Dicer*. Isto foi interpretado como uma interação de transferência de substrato e, ainda especulado, que as interações do domínio PIWI com a *Dicer* estimularia a liberação dos siRNAs (Tahbaz *et al.*, 2004). Após sua incorporação ao complexo RISC, o mRNA alvo é clivado em um único ponto e liberado, ficando exposto ao ataque de endonucleases endógenas. Em *Arabidopsis*, o terminal 3' do mRNA clivado é degradado pela AtXRN4 exorribonuclease 5'→3', enquanto que o fragmento 5' é clivado por um exossomo contendo uma atividade de exonuclease 3'→5' (Souret *et al.*, 2004).

Um siRNA incorporado ao complexo RISC pode ser utilizado em novos ciclos de reconhecimento e clivagem, o que protege esta molécula de uma rápida degradação (Hutvagner *et al.*, 2001); mas é improvável que os siRNAs sejam reciclados após serem liberados do complexo RISC (Nelson *et al.*, 2004).

Em plantas e *C. elegans*, uma amplificação do sinal de silenciamento e a sua transferência célula-a-célula têm sido observadas e RNA polimerases dependentes de RNA (RdRp) foram encontradas (Bushman, 2003). Essa enzima é responsável pela geração e amplificação de siRNAs em dsRNAs. Estes siRNAs são usados como *primers*

pela RdRp para a geração de novos dsRNAs que servem como novos alvos para a *Dicer* para a produção de mais siRNAs. Graças a este sistema, nesses organismos, um número específico de moléculas de dsRNAs pode inativar várias cópias de transcritos endógenos (Ahlquist, 2002). Além do sistema de amplificação do sinal, existe também um sistema de transmissão de sinal célula-a-célula conhecido como supressão sistêmica. Em plantas, o movimento do sinal móvel, ainda desconhecido, pode ser realizado por meio dos plasmodesmos e/ou via floema (Hannon, 2002).

A tecnologia de silenciamento gênico, via interferência por RNA, tem sido empregada em várias culturas constituintes na alimentação para aumentar sua qualidade nutricional. Por exemplo, o conteúdo de cafeína em café é consideravelmente reduzido em plantas transgênicas por supressão do gene que codifica a enzima cafeína sintase (Ogita *et al.*, 2003). Em outro estudo, o silenciamento gênico estimulado por RNA foi utilizado com sucesso para gerar variante dominante de milho com alto conteúdo de lisina por meio do *knockout* do gene que codifica uma proteína de reserva, zeína de 22 kD, que é pobre em lisina (Segal *et al.*, 2003). Outra grande aplicação da interferência por RNA tem sido a modificação genética do teor de ácidos graxos no óleo. Em algodão, por exemplo, a composição de ácidos graxos foi geneticamente modificada por meio do silenciamento de enzimas chave no metabolismo de ácidos graxos como a ω 6-dessaturase e a estearoil-acil Δ 9-dessaturase (Liu *et al.*, 2002). A baixa expressão dessas proteínas leva à produção de óleos com alto conteúdo de ácido oléico e ácido esteárico que são ácidos graxos essenciais e saudáveis ao coração humano. Isto evidencia o poder desta tecnologia como ferramenta molecular para gerar variedades com características agronomicamente importantes.

O programa de Melhoramento da Qualidade da Soja do BIOAGRO/UFV, que desenvolveu, por métodos convencionais, um germoplasma especial visando à criação de variedades de soja para a indústria de alimentos, investe agora no silenciamento de genes específicos que codificam fatores antinutricionais, tais como a proteína P34 e inibidores do tipo BBI, na semente da soja.

REFERÊNCIAS BIBLIOGRÁFICAS

- Ahlquist, P. **RNA-dependent RNA polymerases, viruses, and RNA silencing.** *Science*, **296**:1270:1273, 2002.
- Baulcombe, D. **RNA silencing in plants.** *Nature*, **431**: 356-363, 2004.
- Bernstein, E.; Caudy, A. A.; Hammond, S. M.; Hannon, G. J. **Role for a bidentate ribonuclease in the initiation step of RNA interference.** *Nature*, **409**: 363-366, 2001.
- Bushman, F. **RNA interference: Applications in Vertebrates.** *Mol. Ther.*, **7**: 9-11, 2003.
- Davis, W. P. **An historical perspective from the green revolution to the gene revolution.** *Nutr. Rev.*, **61**: 124-134, 2003.
- Hannon, G. J. **RNA interference.** *Nature*, **418**: 244-251, 2002.
- Hutvagner, G.; McLachlan, J.; Pasquinelli, A. E.; Balint, E.; Tuschl, T.; Zamore, P. D. **A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA.** *Science*, **293**: 834-838, 2001.
- James, C. **Global status of commercialized transgenic crops.** *ISAAA Briefs*, **34**, 2006.
- Jones-Rhoades, M. W.; Bartel, D. P.; Bartel, B. **MicroRNAs and their regulatory roles in plants.** *Annu. Rev. Plant Biol.*, **57**: 19-53, 2006.
- Lippman, Z.; Martienssen, R. **The role of RNA interference in heterochromatic silencing.** *Nature*, **431**: 364-370, 2004.
- Meins, F.; Si-Ammour, A.; Blevins, T. **RNA silencing systems and their relevance to plant development.** *Annu. Rev. Cell. Dev. Biol.*, **21**: 297-318, 2005.
- Nelson, P. T.; Hatzigeorgiu, A.G.; Mourelatos, Z. **miRNP:mRNA association in polyribosomes in human neuronal cell line.** *RNA*, **10**: 360-365, 2004.

- Nykänen, A.; Haley, B.; Zamore, P. D. **ATP requirements and small interfering RNA structure in the RNA interference pathway.** *Cell*, **107**: 309-321, 2001.
- Passos-Lima, A. B. **Construção de cassete para a co-supressão do gene da oleoil dessaturase e transformação genética de embriões somáticos de soja.** *Tese Doutorado*. 117 p. Universidade Federal de Viçosa. Viçosa, MG. 2005.
- Qi, Y.; Hannon, G. J. **Uncovering RNAi mechanisms in plants: Biochemistry enters the foray.** *FEBS letters*, **579**: 5899-5903, 2005.
- Smith, N. A. **Total silencing by a intron-spliced hairpin RNA.** *Nature*, **407**: 319-320, 2000.
- Song, J. J.; Lui, J.; Toli, N. H.; Schneiderman, J.; Smith, S. K.; Martienssen, R. A. Hannon, G. J.; Joshua-Tor, L. **The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes.** *Nat. Struct. Biol.*, **10**: 1026-1032, 2003
- Souret, F. F.; Kastenmayer, J. P.; Green, P. J. **AtXRN4 degrades mRNA in *Arabidopsis* and its substrates include selected miRNA targets.** *Mol. Cell*, **15**: 173–83, 2004.
- Tahbaz, N.; Kolb, F. A.; Zhang, H.; Jaronczyk, K.; Filipowics, W.; Hobman, T. C. **Characterization of the interactions between mammalian PAZ PIWI domain proteins and Dicer.** *EMBO Rep.*, **5**: 189-194, 2004.
- Tang, G.; Galili, G. **Using RNAi to improve plant nutritional value: from mechanism to application.** *Trends Biotechnol.*, **22**: 463-469, 2004.
- Waterhouse, P.M.; Wang, M.B.; Lough, T. **Gene silencing as an adaptative defense against viruses.** *Nature*, **411**: 834-842, 2001.
- Zamore, P. D. **Plant RNAi: how a viral silencing suppressor inactivates siRNA.** *Curr. Biol.*, **14**: 198-200, 2004.

CAPÍTULO 1
ISOLAMENTO, CLONAGEM E ANÁLISE *IN SILICO* DE UM SEGMENTO DA
REGIÃO PROMOTORA DO GENE DA SUBUNIDADE α DA β -
CONGLICININA DE SOJA

RESUMO

A expressão semente-específica de transgenes é requerida para várias aplicações que utilizam a engenharia genética, entre elas a melhoria da qualidade nutricional de grãos por engenharia metabólica. Proteínas de reserva estão presentes em altos níveis durante desenvolvimento da semente e sua expressão é fortemente controlada espacial e temporalmente. Portanto, seqüências regulatórias presentes nos genes que codificam estas proteínas representam uma valiosa ferramenta que pode ser usada para dirigir a expressão de transgenes de uma maneira semente-específica. Os promotores dos genes da β -conglitinina de soja são exemplos de promotores específicos para sementes bem caracterizados e bastante utilizados para esse fim em dicotiledôneas. A β -conglitinina, um dos principais componente das proteínas de reserva de sementes de soja, é uma proteína trimérica composta pelas subunidades, α , α' , e β . Essa proteína corresponde a cerca de 30% do conteúdo protéico total da semente. Os genes que codificam estas subunidades são transcritos de maneira coordenada durante o desenvolvimento da semente. Os objetivos deste trabalho foram isolar e caracterizar *in silico* uma seqüência da região promotora do gene que codifica a subunidade α da β -conglitinina que possa ser utilizada para expressão de transgenes em sementes de soja. Foram desenhados *primers*, de acordo com uma seqüência pública do gene de interesse (AY605542.1) e um fragmento (~ 600 bp) foi amplificado a partir do DNA de soja da variedade Monarca. O fragmento foi clonado no vetor pGEM-T Easy e confirmado por seqüenciamento. A seqüência obtida, de 634 pb, foi analisada com o auxílio do software PLACE (*Plant Cis-Acting Regulatory Elements* database). Foram encontrados o TATA box e sete CAAT box. A análise desta seqüência promotora também revelou vários motivos putativos de DNA que podem estar envolvidos na regulação transcricional semente-específica dos genes que codificam a β -conglitinina (subunidade α). Existem quatro motivos CATGCAC (*RY LEG* box), dois CATGCATG (*RY LEB* box) e um *Vicilin Box* (esta

seqüência tem sido encontrada em membros da família das globulinas 7S). Acredita-se que estes *cis*-elementos estejam envolvidas na expressão semente-específica dos genes sob o controle deste promotor. A seqüência de 634 pb, isolada e caracterizada, será utilizada para a construção de cassetes de expressão para controle dos genes que codificam a proteína P34 e os inibidores de protease do tipo BBI visando a transformação de soja e obtenção de linhagens com expressão semente-específica de transgenes.

1. INTRODUÇÃO

A transformação genética de plantas é uma ferramenta poderosa que tem permitido o desenvolvimento de cultivares com qualidade nutricional melhorada ou com capacidade de síntese de compostos de interesse farmacêutico ou industrial. Apesar dos potenciais benefícios advindos desta tecnologia, ainda há muitos questionamentos relativos à segurança dos consumidores e do meio ambiente. Para minimizar possíveis efeitos adversos das plantas geneticamente modificadas, a expressão do transgene deve ser fortemente regulada, tanto temporalmente quanto espacialmente. A submissão do cassete de expressão sob o controle de promotores adequados pode constituir parte da resposta para a questão sobre a segurança em relação a plantas geneticamente modificadas (Potenza *et al.*, 2004).

Promotores são regiões do DNA, que precedem a região codificante de um gene e que contêm seqüências que são reconhecidas por proteínas envolvidas no início da transcrição. Ao longo dos anos, vários promotores foram isolados, caracterizados e aplicados em experimentos de transformação genética de plantas. Como a atividade dos promotores afeta qualitativa e quantitativamente a transcrição, o sucesso da tecnologia de transferência gênica, desde a pesquisa básica até o desenvolvimento de novos organismos, depende da eficiência de seleção e uso de tais seqüências reguladoras (Potenza *et al.*, 2004).

A escolha do promotor a ser utilizado em uma construção depende fundamentalmente dos objetivos aos quais se quer atingir. Muitas vezes é desejável a utilização de promotores que conferem uma expressão temporal ou tecido-específica, mas promotores que dirigem uma expressão constitutiva e em altos níveis têm sido uma poderosa ferramenta na engenharia genética de plantas. O promotor mais comumente usado para este fim, tanto em dicotiledôneas como em monocotiledôneas, é o promotor CaMV35S do vírus do mosaico da couve-flor (Odell *et al.*, 1985). Uma desvantagem do uso de promotores de origem viral em uma construção para transformação de plantas é que essa seqüência poderia ser reconhecida como “invasora” e inativada por meio de um mecanismo conhecido por silenciamento gênico transcricional (TGS) desencadeado pela metilação de DNA e formação de heterocromatina. É impossível prever se haverá ou não

o silenciamento do transgene, mas quando ocorre, ele pode diminuir a atividade gênica em toda a planta ou afetar a especificidade de promotores (Kloti *et al.*, 2002). O risco de silenciamento pode ser evitado por meio do uso de promotores endógenos. Para plantas, aqueles que controlam os genes que codificam actina e ubiquitina são os mais utilizados. Um aspecto importante que deve ser considerado quando se utiliza um promotor forte e de expressão constitutiva é que, de uma forma geral, a expressão constitutiva de um transgene pode alterar o metabolismo da planta levando ao aparecimento de fenótipos indesejáveis. O isolamento e a utilização de promotores que dirijam uma expressão temporal ou tecido específica são importantes para o desenvolvimento de cultivares transgênicos com uma expressão controlada e uma maior aceitação do público alvo.

A expressão semente-específica de transgenes tem sido usada para aumentar a qualidade nutricional e funcional (Keeler *et al.*, 1997; Shintani & DellaPenna, 1998; Goto *et al.*, 1999; Ye *et al.*, 2000; Barro *et al.*, 1997) de grãos e produção de compostos farmacêuticos ou industriais (Cahoon & Shanklin, 2000). Em plantas produtoras de grãos, proteínas de reserva são altamente expressas durante o desenvolvimento da semente e os promotores destes genes têm sido isolados e caracterizados para serem utilizados em experimentos de engenharia genética. Em dicotiledôneas, os promotores mais utilizados são o promotor do gene da β -faseolina de feijão (Bustos *et al.*, 1989) e os promotores dos genes da β -conglícinina de soja (Chen *et al.*, 1989). A β -conglícinina, um dos componentes de reserva mais abundantes, é uma proteína trimérica composta pelas subunidades α (72 kDa), α' (76 kDa) e β (53 kDa), que constituem cerca de 30% do teor protéico total desta leguminosa (Yoshino *et al.*, 2001). Os genes que codificam esta proteína são organizados em uma família multigênica e transcritos de uma maneira coordenada e embrião-específica durante o desenvolvimento da semente (Harada *et al.*, 1989).

Diante disso, o objetivo deste trabalho foi isolar, clonar e caracterizar *in silico* um fragmento da região promotora do gene que codifica a subunidade α da β -conglícinina, visando o desenvolvimento de vetores de expressão semente-específica para futuros ensaios de transformação genética em soja.

2. MATERIAIS E MÉTODOS

2.1 Material Vegetal

Para o isolamento e clonagem do segmento referente à região promotora do gene que codifica a subunidade α da β -conglucina, amostras de DNA foram extraídas de folhas jovens de soja da variedade Monarca. As plantas foram cultivadas em casa de vegetação do Instituto de Biotecnologia Aplicada à Agropecuária (BIOAGRO) da Universidade Federal de Viçosa, Minas Gerais.

2.2 Construção de *primers* específicos

Os oligonucleotídeos específicos foram construídos a partir de uma sequência de DNA genômico depositadas no banco de dados público – *GenBank* (*National Center for Biotechnology Information* – NCBI) (Nº Acesso: AY605542.1). Os *primers* foram desenhados com auxílio do “Primer3 Input Program” (<http://www.genome.wi.mit.edu/cgi-bin/primer/primer3.cgi>) e sítios de restrição reconhecidos pelas enzimas *Sac* I e *Xho* I foram adicionados à extremidade 5’ dos *primers forward* e *reverse*, respectivamente. Parâmetros que atestam a viabilidade de uso destes oligonucleotídeos foram determinados com o auxílio do programa *QuickPrimer* do pacote DNASTAR (DNASTAR INC.). As sequências estão descritas na Tabela 1.

2.3 Extração de DNA genômico

A extração de DNA genômico de folhas jovens de soja foi feita de acordo com Doyle e Doyle (1990). O conteúdo de DNA foi estimado em espectrofotômetro ($A_{260\text{ nm}}$) e a sua qualidade foi verificada por eletroforese em minigéis de agarose 0,8%.

Tabela 1: Sequência dos *primers* desenhados para amplificação do fragmento referente à região promotora do gene da subunidade α da β -conglucininina.

Nº Acesso	Primers	Sequência	Sítio de Restrição
AY60554 2.1	BetaF	5` GGC <u>GAG CTC</u> GCA AAA GCC ATG CAC AAC AAC A 3`	<i>Sac</i> I
	BetaR	5`GGC <u>CTC GAG</u> GCG CTC TCA TCA TAG TAT 3`	<i>Xho</i> I

2.4 Amplificação e clonagem no vetor pGEM-TEasy

As reações de amplificação foram conduzidas no termociclador da GeneAmp® PCR System 9700 (APPLIED BIOSYSTEMS), com período inicial de desnaturação de 94°C por 2 min, seguido por 35 ciclos (94°C por 30 s; 52°C por 30 s; e 72°C por 45s) e um período adicional de extensão a 72°C, por 4 min. Cada reação continha, em 25 μ l, 30 ng de DNA genômico, Tris-SO₄ 60 mM (pH 8,9), (NH₄)₂SO₄ 18 mM, MgSO₄ 2 mM, dNTPs 0,2 mM, 0,4 μ M de *primers* e 1U de Platinum® *Taq* DNA Polymerase High Fidelity (INVITROGEN).

Os produtos de amplificação foram analisados em gel de agarose 0,8% em TBE (Tris-Borato 90 mM, EDTA 1 mM, pH 8,0) contendo 0,2 μ g/mL de brometo de etídeo. O padrão de bandas foi visualizado sob luz ultravioleta, fotodocumentado com o auxílio de equipamento *Eagle Eye™ II* (STRATAGENE). Para a eliminação de bandas inespecíficas, o fragmento de interesse foi excisado do gel por meio do kit *QIAquick Gel Extraction* (QIAGEN INC.) e ligado ao vetor pGEM T-Easy (Promega), de acordo com as recomendações do fabricante.

2.5 Transformação de *Escherichia coli* DH5 α ultracompetentes e análise dos transformantes

Células de *Escherichia coli* DH5 α ultracompetentes foram transformadas por choque térmico segundo Sambrook *et al.* (1989). Em 100 μ l de células ultracompetentes foram adicionados 5 μ l da reação de ligação e esta mistura foi incubada em gelo por 30 min. O choque térmico foi realizado a 42°C por 45 s seguido de uma incubação em gelo por 2 min. Para a recuperação das células, foram adicionados 800 μ l de meio SOC (0,5% de extrato de levedura, 2% de triptona, NaCl 10mM, KCl 2,5mM, MgCl₂ 10 mM, MgSO₄ 10mM) e a cultura foi incubada a 37°C por uma hora a 150 rpm. Após esse período, as células foram concentradas a 1500 x g por 5 min em centrífuga 5415D (*Eppendorf*), ressuspensas em 200 μ l de meio SOC, dos quais 100 μ l foram plaqueados em meio LB sólido contendo ampicilina 50 μ g/ml, 1 mg de X-Gal e 10 mM de IPTG. As placas foram incubadas a 37°C por cerca de 16 horas.

Para a análise dos transformantes, 14 colônias isoladas foram repicadas em meio sólido e analisadas por PCR. Cada reação continha Tris-HCl 10mM pH 8,3, KCl 50 mM, MgCl₂ 2,0 mM, 0,2 mM de dNTPs, 0,4 mM de *primers* específicos e 1,5 μ l de *Taq* DNA polimerase. As colônias que mostraram a amplificação do produto esperado foram inoculadas em meio LB líquido contendo ampicilina 50 μ g/ml e incubadas a 37°C durante 16 horas sob agitação. Após o crescimento, as bactérias foram aliqüotadas em glicerol 25%, congeladas em nitrogênio líquido e armazenadas a -80°C. O DNA plasmidial foi extraído com o auxílio do kit Wizard[®] Plus SV Minipreps (Promega), de acordo com as recomendações do fabricante. A clonagem foi confirmada utilizando-se amostras de DNA plasmidial em reações de PCR realizadas como descrito anteriormente, clivagem por enzimas de restrição e seqüenciamento.

A clivagem inicial foi realizada com 500 ng de DNA plasmidial, 5U da enzima *Sac* I e tampão de clivagem (Tris-HCl 10 mM pH 7,5, MgCl₂ 7 mM, KCl 50 mM e DTT 1 mM) por 1 hora a 37°C. Em seguida, foram adicionados tampão de clivagem (Tris-HCl 50 mM pH 8,0, MgCl₂ 10 mM e NaCl 50 mM), 5U de *Xho* I e água em quantidade suficiente para dobrar o volume da reação inicial. A reação foi novamente incubada a 37°C por uma hora. O resultado da clivagem foi visualizado em gel de agarose 0,8%.

O seqüenciamento dos clones recombinantes foi realizado utilizando-se o kit ABI PRISM BigDye[®] III v 3.1 Terminator Cycle Sequencing Ready Reaction (APPLIED BIOSYSTEMS) e o seqüenciador automático ABI PRISM 377 Genetic Analyzer (APPLIED BIOSYSTEMS). As reações de seqüenciamento foram conduzidas de acordo com a técnica de seqüenciamento por terminação de cadeia por dideoxynucleotídeos (ddNTPs), descrita por Sanger *et al.* (1977). As reações de amplificação linear foram realizadas usando o termociclador (APPLIED BIOSYSTEMS), GeneAmp[®] PCR System 9.700, programado para um período inicial de desnaturação a 96°C, por 2 min, seguido por 25 ciclos (96°C, por 30 s; 50°C, por 20 s; e 60°C, por 4 min). Foram utilizados os *primers* universais: M13F e M13R. Os produtos de extensão, produzidos na reação de seqüenciamento, foram purificados pela adição de 0,1 volumes de Acetato de Sódio 3M (pH 4,8), 2 volumes de etanol 95% (v/v) e incubação à temperatura ambiente por 15 min. Os fragmentos foram coletados por centrifugação (10.000 x g por 20 min) e o precipitado foi lavado com etanol 70% (v/v), seco à temperatura ambiente e, em seguida, ressuspensão em tampão de corrida (17% azul de bromofenol:83% de formamida). Finalmente, os fragmentos foram desnaturados a 94°C por 5 min e separados por eletroforese em gel desnaturante de poliacrilamida 4,5%.

A seqüência do clone isolado foi analisada pelo Programa *Seqman* do pacote DNASTAR (DNASTAR Inc.) e o alinhamento com outras seqüências depositadas no *GenBank* foi feito usando o BLAST (Altschul *et al.*, 1997).

2.6 Caracterização *in silico* da região promotora da subunidade α da β -conglícinina

A seqüência clonada foi analisada pelo software PLACE (Plant Action *Cis* Element), disponível no endereço <http://www.dna.affrc.go.jp/PLACE>. (Prestridge, 1991; Higo *et al.*, 1999), para a identificação, por meio de alinhamento, de *cis*-elementos presentes nesta seqüência que garantam a expressão semente-específica de genes colocados sob controle deste promotor.

2.7 Clonagem da região promotora da subunidade α da β -conglícinina no vetor pKANNIBAL

Após ser seqüenciado e caracterizado, o fragmento correspondente à região promotora do gene da subunidade α de β conglícinina de soja foi transferido para o vetor pKANNIBAL (Wesley *et al.*, 2001), 6063 pb, gentilmente doado pelo Dr. Francisco Aragão (EMBRAPA/CENARGEN), em substituição ao promotor CaMV35S do vírus do mosaico da couve-flor. Este vetor contém, além do promotor CaMV35S, a região de clonagem para a seqüência *sense*, o íntron PDK, a região de clonagem da seqüência *antisense* e o terminador OCS. A marca de seleção deste plasmídeo é o gene *npt*, cujo produto confere resistência à canamicina para a bactéria hospedeira.

Células de *Escherichia coli* contendo o vetor pKANNIBAL foram inoculadas, crescidas em meio seletivo e o DNA plasmidial foi extraído conforme descrito anteriormente. A clivagem com *Sac* I e *Xho* I foi realizada e visualizada em gel de agarose 0,8%. A banda referente ao vetor foi excisada do gel e purificada como já descrito. O fragmento referente à região promotora da subunidade α da β conglícinina foi liberado do vetor pGEM-T Easy mediante clivagem com as mesmas enzimas de restrição e purificado da mesma forma.

A reação de ligação foi feita para um volume final de 10 μ l, utilizando 10 ng do vetor clivado, 50 ng do inserto clivado, tampão de reação (Tris-HCl 50 mM pH 7,6, MgCl₂ 10 mM, ATP 1 mM, DTT 1 mM e 25% de polietileno glicol-8000) e 1U de T4 DNA ligase (INVITROGEN). A reação foi incubada a 4°C por 16 horas.

A transformação de células ultracompetentes e a análise dos transformantes foi realizada conforme o descrito no item 2.5.

3. RESULTADOS E DISCUSSÃO

A amplificação do DNA genômico da variedade de soja Monarca com *primers* específicos para o início da região promotora do gene que codifica a subunidade α da β -conglucininina, gerou três fragmentos. Para a eliminação das bandas inespecíficas, o fragmento de interesse foi excisado do gel e ligado ao vetor pGEM T-Easy. Foram analisadas, por PCR, 14 colônias de *Escherichia coli* DH5 α transformantes. Somente uma apresentou o fragmento esperado e o clone foi denominado pPBC (Figura 1A). A clonagem foi confirmada por PCR, ensaios de clivagem por meio de enzimas de restrição (Figura 1B) e seqüenciamento do clone pPBC. A reação de PCR e o seqüenciamento evidenciaram também que o aparecimento da banda de 400 pb ocorre devido ao anelamento interno inespecífico do *primer* F desenhado. A seqüência obtida, de 652 pb, foi alinhada com outras seqüências genômicas depositadas no *GenBank*, com o auxílio do BLAST, o que confirmou a identidade da seqüência clonada com o gene que codifica a subunidade α da β -conglucininina (Figura 02).

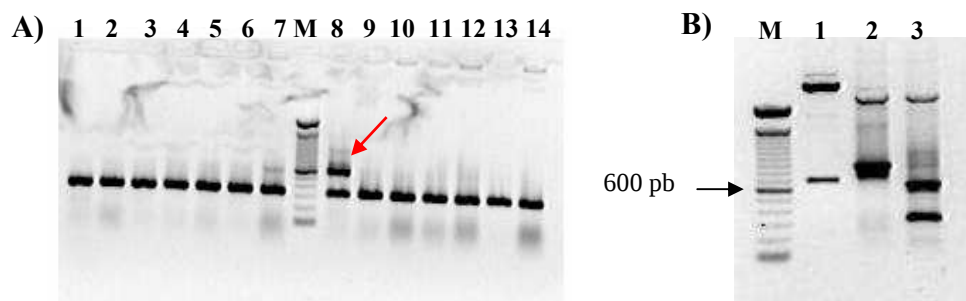


Figura 1: Análise de transformantes para confirmação da clonagem da região promotora do gene que codifica a subunidade α da β -conglucininina de soja. (A) representa a amplificação obtida a partir das 14 colônias transformantes selecionadas, representadas pelos números de 1 a 14. A seta indica a banda esperada obtida a partir de uma única colônia. (B) apresenta a reação de clivagem com as enzimas *Sac* I – *Xho* I (1) e o resultado da reação de PCR do DNA plasmidial do clone pPBC com *primers* M13 (2) e *primers* específicos (3). M refere-se ao marcador de tamanho de 100 pb (INVITROGEN).

Clone pPBC 1	AAAAGCCATGCACAACAACACGTACTCACCAGGTGCAATCGTGCTGCCCAAAAACATTC	60
AY605542.1 2298	AAAAGCCATGCACAACAACACGTACTCACCAGGTGCAATCGTGCTGCCCAAAAACATTC	2357
Clone pPBC 61	ACCAACTCAATCCATGATGAGCCACACATTTGTTGTTTGTAAACCAATCTCAAACGCGG	120
AY605542.1 2358	ACCAACTCAATCCATGATGAGCCACACATTTGTTGTTTGTAAACCAATCTCAAACGCGG	2417
Clone pPBC 121	TGTTCTCTTTGGAAAGCAACCATATCAGCATATCACACTATCTAGTCTCTTGATCATGC	180
AY605542.1 2418	TGTTCTCTTTGGAAAGCAACCATATCAGCATATCACACTATCTAGTCTCTTGATCATGC	2477
Clone pPBC 181	ATGCGCAACCAAAAGACAACACATAAAGTATCCTTTGAAAGCAATGTCCAAGTCCATCA	240
AY605542.1 2478	ATGCGCAACCAAAAGACAACACATAAAGTATCCTTTGAAAGCAATGTCCAAGTCCATCA	2537
Clone pPBC 241	AATAAAATTGAGACAAAATGCAACCTCACCCCACTTCACTATCCATGGCTGATCAAGATC	300
AY605542.1 2538	AATAAAATTGAGACAAAATGCAACCTCACCCCACTTCACTATCCATGGCTGATCAAGATC	2597
Clone pPBC 301	GCCGCGTCCATGTAGGTCTAAATGCCATGCACATCAACACGTACTCAACATGCAGCCAA	360
AY605542.1 2598	GCCGCGTCCATGTAGGTCTAAATGCCATGCACATCAACACGTACTCAACATGCAGCCAA	2657
Clone pPBC 361	ATTGCTCACCATCGCTCAACACATTTCTTGTTAATTTCTAAGTACACTGCCTATGCGACT	420
AY605542.1 2658	ATTGCTCACCATCGCTCAACACATTTCTTGTTAATTTCTAAGTACACTGCCTATGCGACT	2717
Clone pPBC 421	CTAACTCGATCACAACCATCTTCCGTCACATCAATTTGTTCAATTCAACACCCGTCAAA	479
AY605542.1 2718	CTAACTCGATCACAACCATCTTCCGTCACATCAATTTGTTCAATTCAACACCCGTCAAA	2777
Clone pPBC 480	CTTGATGCCACCCCATGCATGCAAGTTAACAAGAGCTATATCTCTTCTATGACTATAAA	539
AY605542.1 2778	CTTGATGCCACCCCATGCATGCAAGTTAACAAGAGCTATATCTCTTCTATGACTATAAA	2837
Clone pPBC 540	TACCCGCAATCTCGGTCCAGGTTTTTCATCATCGAGAACTAGTTCAATATCCTAGTATACC	599
AY605542.1 2838	TACCCGCAATCTCGGTCCAGGTTTTTCATCATCGAGAACTAGTTCAATATCCTAGTATACC	2897
Clone pPBC 600	TTAATAAATAATTTAAGATACTATGATGAGAGC	632
AY605542.1 2898	TTAATAAATAATTTAAGATACTATGATGAGAGC	2930

Figura 2: Alinhamento, realizado com o auxílio do BLAST, entre a sequência clonada no vetor pGEM T-Easy e a sequência do gene da subunidade α da β -conglucininina de soja.

A região amplificada, de acordo com o alinhamento dos genes codificantes das subunidades da β -conglucininina e o gene que codifica a β -faseolina em feijão (Yoshino *et al.*, 2001), abrange do ponto +1 da tradução até cerca de -565 em relação ao ponto +1 da transcrição. Estes mesmos autores identificaram, nesta região, uma sequência muito similar ao *vicilin box* (Box I), descrito por outros autores em outros genes relacionados a

membros da família das globulinas 7S (Brown *et al.*, 1988; Higgins *et al.*, 1988; Newbigin *et al.*, 1990). Esta sequência está relacionada com a expressão semente específica destas proteínas.

A busca por outros *cis*-elementos foi realizada com o auxílio do programa PLACE e essa análise identificou os potenciais elementos: TATA *box*, CAAT *box*, quatro motivos CATGCAC (*RY LEG box*) e dois CATGCATG (*RY FLEB box*). Os elementos CAAT estão envolvidos na expressão tecido-específica de genes que codificam leguminas. Já os elementos *RY LEG box* e *RY FLEB box* estão relacionados, respectivamente, com expressão específica (Fujiwara & Beachy, 1994) e quantitativa (Baumlein *et al.*, 1992) em sementes (Figura 3).

```

AAAGC CATGCAC AACAACACGTA CTACCAAGGTG CAAT CGTGCTGCCCAAAACATT CACCAACT CAAT CCATGATGA
      RYLEGBOX
GCCCAACACATTTGTTGTTTGTAAACCAATCTCAAACGCGGTGTTCTCTTTGGAAAGCAACCATATCAGCATATCACACTA
TCTAGTCTCTTGGAT CATGCAT GCGCAACCAAAAGACAACACATAAAGTATCCTTTGAAAG CAAT GTCCAAGTCCATCA
      RYLEGBOX RYFLEB4
AATAAAATTGAGACAAAATGCAACCTCACCCACTTCACTATCCATGGCTGATCAAGATCGCCGCGTCCATGTAGGTCTA
AATGC CATGCAC ATCAACACGTA CTCAACATGCAGCCCAAATTGCTCACCATCGCTCAACACATTTCTTGTAATTTCTA
      RYLEGBOX
AGTACACTGCCTATGCGACTCTAACTCGATCACAACCATCTTCCGTCA CAT CAATTTTGT CAAT TCAACACCGTCAAC
      Box I
TTGCATGCCACCC CATGCAT GCAAGTTAACAAGAGCTATATCTTCTATGAC TATAAATACCCG CAAT CTCGGTCCAGG
      RYLEGBOX RYFLEB4 TATA box
TTTTC ATCATCGAGAACTAGTT CAAT ATCCTAGTATACCTTAATAAATAATTTAAGATACT ATG ATGAGAGCGC
      +1 da transcrição +1 da tradução

```

Figura 3: Resumo das análises de caracterização realizadas com o auxílio do programa PLACE mostrando o posicionamento dos *cis*-elementos encontrados na sequência clonada.

A presença de todos os *cis*-elementos identificados indica, com certa segurança, a funcionalidade desta região para controle da transcrição de transgenes em sementes. Experimentos realizados por Chen *et al.* (1986), com o gene que codifica a subunidade α' , demonstraram que a região compreendida entre os pontos +14 a -257, em relação ao ponto mais +1 da transcrição é suficiente para garantir um alto nível de expressão semente específica. A adição de 8 kb a essa região não aumentou significativamente a síntese de proteínas repórteres na semente. Outro ponto a ser testado seria a atividade

deste promotor durante o desenvolvimento da semente. De acordo com Harada *et al.* (1989), a quantidade de mRNA referente aos genes que codificam a β -conglucina oscila durante o desenvolvimento do embrião, atingindo o máximo por volta de 80 dias após o florescimento. Apesar dessa oscilação, a utilização deste promotor em cassetes de expressão para silenciamento não se torna inviável porque mesmo um baixo nível de expressão pode ser suficiente para direcionar a clivagem de mRNAs que não se acumulam em níveis tão altos quanto àqueles relacionados a proteínas de reserva.

Após estas análises, o fragmento correspondente à região promotora do gene que codifica a subunidade α da β -conglucina foi clonado no vetor pKANNIBAL, em substituição ao promotor 35S, produzindo o vetor denominado pBKN (5334 pb).

A funcionalidade *in vivo* deste promotor está sendo verificada por meio de sua ligação com uma fusão de dois genes repórteres (Gus/GFP) no vetor binário pCAMBIA 1304 para posterior expressão transiente em cotilédones de soja. A atividade deste promotor durante a formação da semente será avaliada em cotilédones de plantas transformadas coletados durante os dias após o florescimento descritos por Harada *et al.* (1989) adaptados para a variedade CAC-1.

4. CONCLUSÕES

- Foi possível isolar e clonar um fragmento de 652 pb referente à região promotora do gene da subunidade α da β -conglícinina de soja.

- A análise *in silico* do fragmento clonado demonstra a presença de *cis*-elementos relacionados à expressão semente-específica além de toda a sequência líder, o que permite a utilização deste promotor tanto para o silenciamento gênico como para a super-expressão de genes na semente de soja.

5. PERSPECTIVAS

- A funcionalidade deste promotor está sendo verificada pela sua ligação com uma fusão de dois genes repórteres (Gus/GFP) no vetor binário pCAMBIA 1304 e posterior expressão transiente em cotilédones de soja.

- A atividade deste promotor durante a formação da semente será avaliada em cotilédones de plantas transformadas coletados durante os dias após o florescimento descritos por Harada *et al* (1989).

6. REFERÊNCIAS BIBLIOGRÁFICAS

- Altschul, S. F.; Madden, T. L.; Schaffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. Gapped **BLAST** and **PSI-BLAST**: a new generation of protein database search programs. *Nucl. Acid. Res.*, **25**: 3389-3402, 1997.
- Barro, F.; Rooke, L.; Bekes, F.; Gras, P.; Tatham, A. S.; Fido, R.; Lasserri, P.; Shewry, P. R.; Barcelo, P. **Transformation of wheat with high molecular weight subunit genes results in improved functional properties**. *Nature Biotechnol.*, **15**: 1295-1299, 1997.
- Baumlein, H.; Nagy, I.; Villarroel, R.; Inze, D.; Wobus, U. **Cis-analysis of a seed protein gene promoter: the conservative RY repeat CATGCATG within the legumin box is essential for tissue-specific expression of a legumin gene**. *Plant J.* **2**:233-239, 1992.
- Brown, D.; Ellis, T. H.; Gatehouse, J. A. **The sequence of a gene encoding convicilin from pea (*Pisum sativum* L.) shows that convicilin differs from vicilin by an insertion near the N-terminus**. *Biochem. J.*, **251**: 717-726, 1988.
- Bustos, M. M.; Gultinan, M. J.; Jordano, J.; Begum, D.; Kalkan, F. A.; Hall, T. C. **Regulation of beta-glucuronidase expression in transgenic tobacco plants by an A/T-rich, cis-acting sequence found upstream of a French bean beta-phaseolin gene**. *Plant Cell*, **1**: 839-853, 1989.
- Cahoon, E. B.; Shanklin, J. **Substrate-dependent mutant complementation to select fatty acid desaturase variants for metabolic engineering of plant seed oils**. *Proc. Natl. Acad. Sci.*, **97**: 12350-12355, 2000.
- Chen, Z. L.; Naito, S.; Nakamura, I.; Beachy, R. N. **Regulated expression of genes encoding soybean beta-conglycinin in transgenic plants**. *Dev. Genet.* **10**: 112-122, 1989.
- Chen, Z. L.; Schuler, M. A.; Beachy, R. N. **Functional analysis of regulatory elements in a plant embryo-specific gene**. *Proc. Natl. Acad. Sci. USA*, **83**: 8560-8564, 1986.

- Doyle, J. J.; Doyle, J. L. **Isolation of plant DNA from fresh tissue.** *Focus*, **12**: 13-15, 1990.
- Fujiwara, T.; Beachy, R. N. **Tissue-specific and temporal regulation of a beta-conglycinin gene: roles of the RY repeat and other *cis*-acting elements.** *Plant Mol. Biol.*, **24**:261-272, 1994.
- Grace, M. L.; Chandrasekharan, M. B.; Hall, T.C.; Crowe, A. J. **Sequence and spacing of TATA box elements are critical for accurate initiation from the beta-phaseolin promoter.** *J. Biol. Chem.*, **279**:8102-8110, 2004.
- Goto, F.; Yoshihara, T.; Shigemoto, N.; Toki, S.; Takaiwa, F. **Iron fortification of rice seed by the soybean ferritin gene.** *Nature Biotechnol.*, **17**: 282-286, 1999.
- Harada, J. J.; Baker, S. J.; Goldberg, R. B. **Soybean β -conglycinin genes are clustered in several DNA regions and are regulated by transcriptional and posttranscriptional processes.** *Plant Cell*, **1**: 415-425, 1989.
- Higgins, T. J. V.; Newbigin, E. J.; Spencer, D., Llewellyn, D. J.; Craig, S. **The sequence of a pea vicilin gene and its expression in transgenic tobacco plants.** *Plant Mol. Biol.*, **11**: 683-695, 1988.
- Higo, K. Y.; Ugawa, M.; Iwamoto T.; Korenaga, T. **Plant *cis*-acting regulatory DNA elements (PLACE) database.** *Nucl. Acid. Res.*, **27**: 297-300, 1999.
- Keller, B.; Baumgartner, C. **Vascular-specific expression of the beanGrp-1.8 gene is negatively regulated.** *Plant Cell*, **3**: 1639-1646, 1991.
- Kloti, A.; He, X.; Potrykus, I.; Hohn, T.; Futterer, J. **Tissue-specific silencing of a transgene in rice.** *Proc. Natl. Acad. Sci USA.*, **99**:10881-10886, 2002.
- Sambrook, J.; Fritsch, E. F.; Maniatis, T. **Molecular cloning, a Laboratory Manual.** 2nd ed. USA: Cold Spring Harbor Laboratory, 1989

- Newbiggin, E. J.; de Lumen, B. O.; Chandler, P. M.; Gould, A.; Blagrove, R. J.; March, J. F.; Kortt, A. A.; Higgins, J. V. **Pea convicilin: structure and primary sequence of the protein and expression of a gene in the seeds of transgenic tobacco.** *Planta*, **180**: 461-470, 1990.
- Odell, J. T.; Nagy, F.; Chua, N. N. **Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter.** *Nature*, **313**: 810-812, 1985.
- Prestridge, D.S. **SIGNAL SCAN: A computer program that scans DNA sequences for eukaryotic transcriptional elements.** *CABIOS* **7**: 203-206, 1991.
- Potenza, C.; Aleman, L.; Sengupta-Copalan, C. **Targeting transgene expression in research, agricultural and environmental applications: promoters used in plant transformation.** *In Vitro Cell. Dev. Biol.-Plant*, **40**:1-22, 2004.
- Sanger, F.; Nicklen, S.; Coulson, A. R. **DNA Sequence with chain-termination inhibitors.** *Proc. Natl. Acad. Sci. USA*, **74**:5463-5467, 1977.
- Shintani, D.; DellaPenna, D. **Elevating the vitamin E content of plants through metabolic engineering.** *Science*, **282**: 2098-2100, 1998.
- Shirsat, A.; Wilford, N.; Croy, R.; Boulter, D. **Sequences responsible for the tissue specific promoter activity of a pea legumin gene in tobacco.** *Mol. Gen. Genet.*, **215**:326-331, 1989.
- Wesley, S. V.; Helliwell, C. A.; Smith, N. A.; Wang, M.; Touse, D. T.; Liu, Q.; Gooding, P. S.; Singh, S. P.; Abbott, D.; Stoutjesdijk, P. A.; Robinson, S. P.; Gleave, A. P.; Green, A. G.; Waterhouse, P. M. **Construct design for efficient, effective and high-throughout gene silencing in plants.** *Plant J.* **27**:581-590, 2001.
- Ye, X.; Al-Babili, S.; Kloti, A.; Zhang, J.; Lucca, P.; Beyer, P.; Potrykus, I. **Engineering the provitamin A (beta-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm.** *Science*, **287**:303-305, 2000.

Yoshino, M.; Kanazawa, A.; Tsutsumi, K.; Nakamura, I.; Shimamoto, Y. **Structure and characterization of the gene encoding α subunit of soybean β -conglycinin.** *Genes Genet. Syst.*, **76**: 99-105, 2001.

CAPÍTULO 2
CONSTRUÇÃO DE CASSETES DE EXPRESSÃO PARA SILENCIAMENTO
GÊNICO, VIA INTERFERÊNCIA POR RNA, DE FATORES
ANTINUTRICIONAIS DA SEMENTE DE SOJA

RESUMO

Por apresentar um alto teor protéico, a soja tem sido muito utilizada na alimentação animal e humana. Apesar disso, fração proteica encontrada na semente desta leguminosa não é considerada ideal por apresentarem baixo teor de metionina e lisina, além da presença de fatores alergênicos de inibidores de proteases. Dentre estes fatores antinutricionais estão a proteína P34 e de inibidores de protease do tipo Bowman-Birk. A eliminação destes fatores aumentaria a qualidade da proteína de soja, e também tornaria esse tipo de alimento disponível a uma faixa mais ampla de consumidores. Neste sentido, o objetivo deste trabalho foi a construção de cassetes de expressão para silenciamento gênico, via interferência por RNA, dos inibidores de protease do tipo Bowman-Birk e da proteína P34 em sementes de soja. A escolha da região a ser amplificada foi determinada com o auxílio do programa BLOCK-iTTM RNAi Designer (INVITROGEN), e os *primers* específicos foram desenhados de acordo com seqüências depositadas no banco de dados público – *GenBank*, utilizando o programa “Primer3 Input Program”. A viabilidade destes oligonucleotídeos foi testada com o programa QuickPri (DNASTAR). Antes da construção dos cassetes de silenciamento, foi feita, por meio de RT-PCR, uma análise da expressão dos genes de interesse durante o desenvolvimento da semente e em raiz, caule e folha. Essa análise evidenciou a expressão contínua e abundante desses genes durante todo o desenvolvimento da semente e bem como sua presença em todos os órgãos vegetais analisados. Após sua amplificação a partir de cDNA de soja, todos os fragmentos foram seqüenciados e sua identificação foi confirmada por meio de alinhamento utilizando o BLAST. Estes fragmentos apresentavam cerca de 350 pb quando relacionados à P34 e 200 pb quando relacionados a BBI. Para a clonagem da seqüência *sense*, os vetores pKANNIBAL e pBKN, além dos fragmentos amplificados referentes aos genes de interesse, foram clivados com as enzimas *Xho* I e *Kpn* I. A reação

de ligação foi realizada utilizando T4 DNA ligase e células de *Escherichia coli* ultracompetentes foram transformadas por meio de choque térmico. Os transformantes foram analisados por PCR e por clivagem enzimática. A clonagem do fragmento *antisense* foi realizada da mesma forma, utilizando as enzimas *Xba* I e *Cla* I e os clones contendo o fragmento *sense* como vetores. Em seguida, as construções foram transferidas para o vetor binário pCAMBIA 3301. Em uma próxima etapa, nós cotiledonares de soja serão transformados, via *Agrobacterium tumefaciens*, e as linhagens que apresentarem baixo teor ou ausência das proteínas de interesse serão incorporadas ao Programa de Melhoramento da Qualidade da Soja do BIOAGRO/UFV.

1. INTRODUÇÃO

Além de ser um dos mais importantes componentes de rações animais, a soja tem se tornado bastante importante na dieta humana por ser uma fonte protéica de boa qualidade. As sementes contêm altos teores de proteínas e óleos (Friedman & Brandon, 2001) e essa reserva protéica constitui uma mistura de proteínas - α , β e γ -conglícininas, glicínina e outras globulinas como a P34. As sementes também contêm proteínas bioativas incluindo β -amilase, citocromo c, lectinas, lipoxigenases, urease, KTI (inibidor de tripsina Kunitz) (Wolf, 1970) e os inibidores tipo BBI (Bowman-Birk) de tripsina e quimiotripsina. Alguns metabólitos secundários também estão presentes como isoflavonas e ácido fítico.

Proteínas de soja não são consideradas ideais porque são deficientes em metionina, que é um aminoácido essencial; e, apesar de possuírem um maior teor de lisina que o trigo, o teor deste aminoácido ainda é menor do que aquele presente na caseína do leite. Além disso, adversidades nutricionais e outros efeitos que se seguem à ingestão de soja crua são atribuídos à presença de inibidores endógenos de enzimas digestivas e lectinas que dificultam a digestão, além de fatores alergênicos (Friedman & Brandon, 2001).

Alergia a alimentos pode ser um sério problema nutricional em crianças e adultos, e alguns alimentos contêm proteínas com o potencial para elicitar reações alérgicas em uma considerável porcentagem da população. Evitar estes alimentos é a única alternativa disponível, limitando seriamente a dieta e a qualidade de vida dos indivíduos alérgicos (Bush & Hefle, 1996).

Três proteínas presentes na semente, Gly mBd 60K, Gly mBd 30k e Gly mBd 28k, constituem os principais alérgenos da soja (Ogawa *et al.*, 2000). Estudos de ligação de IgE mostraram que mais de 65% das reações alérgicas são provocadas pela proteína Gly mBd 30K (Helm *et al.*, 1998).

A identidade molecular dos genes que codificam estas proteínas já foi descrita. Gly mBd 60K é a subunidade α da β -conglícinina, uma das mais abundantes proteínas de reserva de soja (Ogawa *et al.*, 1995). Gly mBd 28 K é similar à classe de globulinas

MP27-MP33 (Tsuji *et al.*, 1995). A última, Gly mBd 30K, mostrou-se idêntica a uma proteína vacuolar de 34 kDa, P34 (Kalinski *et al.*, 1992, Ogawa *et al.*, 1995). Esta proteína é um membro da superfamília das papaínas, que são *cisteíno*-proteases (Kalinski *et al.*, 1992). Como outros membros desta família, a Gly mBd 30K é sintetizada como um precursor e pós-traducionalmente processada (Kalinski *et al.*, 1992). Ao contrário das demais *cisteíno*-proteases que sofrem um autoprocessamento em condições ácidas, o processamento da Gly mBd 30K se dá em um único passo, num resíduo de aspartato (Kalinski *et al.*, 1992), por meio da mesma enzima que processa as proteínas de reserva 11S (Hara-Nishimura *et al.*, 1995). A diferença desta proteína em relação a outras proteases do tipo papaína é a substituição de uma *cisteína* conservada no sítio ativo por uma glicina (Kalinski *et al.*, 1992).

A função desta proteína ainda não é clara e nenhuma atividade enzimática associada a ela foi descrita (Kalinski *et al.*, 1992; Herman *et al.*, 2003). P34 pode ter uma função na defesa contra a infecção por *Pseudomonas* por se ligar ao elicitor secretado por essa bactéria (Cheng *et al.*, 1998). P34 é muito abundante em sementes (Kalinski *et al.*, 1992), mas ela é também encontrada em células vegetativas que estão sujeitas à infecção bacteriana (Cheng *et al.*, 1998). Cultivares de soja resistentes a *Pseudomonas* possuem níveis mais altos de P34 em tecidos foliares quando comparados com cultivares suscetíveis (Herman, 2005).

Considerando que o gene que codifica a P34 apresenta baixo número de cópias (Kalinski *et al.*, 1990), vários métodos podem ser utilizados para desenvolver um cultivar de soja hipoalergênico. Cultivares livres de alérgenos poderiam ser prospectados em coleções de germoplasma e, então, cruzados com cultivares elites. Mas, de acordo com Helm *et al.* (1998), a P34 apresenta 14 epítomos e a presença desta gama de epítomos indica que a probabilidade de se encontrar um variante natural com alterações suficientes para romper a alergenicidade é extremamente pequena. O engenhieramento do gene que codifica esta proteína poderia ser feito para alterar a sua seqüência de aminoácidos, interrompendo a seqüência alergênica, mas numerosos variantes deveriam ser ensaiados para a seleção de um que não fosse reconhecido pela população de IgEs. A modificação dos epítomos tem sido estudada para amendoim, mas a remoção dos epítomos pode alterar o dobramento da proteína, e conseqüentemente, sua estabilidade e

acúmulo intracelular. Uma outra alternativa, que tem se mostrado a mais promissora, seria a supressão do gene que codifica a proteína alergênica por meio de silenciamento gênico (Herman *et al.*, 2003).

Uma segunda classe de fatores antinutricionais da semente de soja corresponde aos inibidores endógenos de enzimas digestivas. Plantas contêm uma variedade de inibidores de protease que são divididos em 16 classes (Ryan, 1990); sendo que o inibidor de tripsina (Kunitz) e os inibidores Bowman-Birk de soja, além dos inibidores I e II de batata, são os mais bem estudados. Inibidores de protease do tipo Bowman-Birk foram primeiramente isolados e caracterizados em sementes de soja (Birk *et al.*, 1963; Bowman, 1946) e, mais tarde, em outras leguminosas. Um homólogo cíclico de BBI com 14 resíduos de aminoácidos foi caracterizado em girassol e considerado como o mais potente BBI já descoberto (Korsinczky *et al.*, 2001).

Inibidores de protease do tipo BBI possuem baixo peso molecular (8-20 kDa) e um alto conteúdo de *cisteína* (Odani & Ikenaka, 1976). Esses resíduos, unidos por ligações dissulfeto, são importantes na manutenção da conformação do inibidor, permitindo a abertura do sítio ativo e o acesso da enzima alvo (Prakash *et al.*, 1996). A estrutura tridimensional dos inibidores BBI de amendoim, soja e cevada revelou que as pontes dissulfeto ocorrem entre dois resíduos de *cisteína* conservados nas posições (C₁-C₁₄, C₂-C₆, C₃-C₁₃, C₄-C₅, C₇-C₉, C₈-C₁₂, C₁₀-C₁₁). Os inibidores BBI inibem tripsina e quimiotripsina em monocotiledôneas e tripsina, quimiotripsina e elastase em dicotiledôneas (Mello *et al.*, 2003).

Diversas funções foram propostas para BBIs, incluindo a regulação de proteinases endógenas durante a germinação, reserva de aminoácidos sulfurados durante o período de dormência, proteção da planta contra insetos e microrganismos (Mayer *et al.*, 1974; Ryan, 1973). Há, também, interesse no uso desses inibidores como agentes antitumorais (Kennedy, 1998).

Cinco tipos de BBI - A, B, CII, DII, EI - foram identificados em soja (Odani & Ikenaka, 1977) e os cDNAs para BBI-A, CII e DII foram clonados (Baek *et al.*, 1994). BBI-B é muito similar a BBI-A em sua especificidade inibitória e sequência aminoacídica (Odani & Ikenaka, 1977) e, provavelmente, é codificado por um segundo gene BBI-A2 (Deshimaru *et al.*, 2004). BBI-EI está relacionado ao BBI-DII e é obtido

por uma proteólise de DII (Odani & Ikenaka, 1978). Desta forma os inibidores BBI foram agrupados em três classes – A, C e D – com distintas especificidades; BBI-A, -C e -D inibem tripsina/quimiotripsina, elastase/tripsina, e tripsina/tripsina, respectivamente, em seus sítios ativos duplos.

Deshimaru *et al.* (2004) identificaram, em *Glycine soja*, um terceiro gene do tipo A e um do tipo C que seriam pseudogenes e um tipo C2, cujo produto ainda não foi isolado.

Considerando que os inibidores BBI são codificados por uma família multigênica, sua eliminação de sementes se torna extremamente difícil e demorada por metodologias tradicionais. Tecnologias de silenciamento podem ser aplicadas, mas devem ser primeiro aprimoradas para que se obtenha um silenciamento eficaz e estável de toda a família.

A tecnologia de silenciamento gênico, via interferência por RNA, tem sido empregada em várias culturas constituintes na alimentação para aumentar sua qualidade nutricional. Por exemplo, o conteúdo de cafeína em café é consideravelmente reduzido em plantas transgênicas por supressão do gene que codifica a enzima cafeína sintase (Ogita *et al.*, 2003). Em outro estudo, o silenciamento gênico estimulado por RNA foi utilizado com sucesso para gerar variante dominante de milho com alto conteúdo de lisina por meio do *knockout* do gene que codifica uma proteína de reserva, zeína de 22 kD, que é pobre em lisina (Segal *et al.*, 2003). Outra grande aplicação da interferência por RNA tem sido a modificação genética do teor de ácidos graxos no óleo. Em algodão, por exemplo, a composição de ácidos graxos foi geneticamente modificada por meio do silenciamento de enzimas chave no metabolismo de ácidos graxos como a ω 6-dessaturase e a estearoil-acil Δ 9-dessaturase (Liu *et al.*, 2002). A baixa expressão dessas proteínas leva à produção de óleos com alto conteúdo de ácido oléico e ácido esteárico que são ácidos graxos essenciais e saudáveis ao coração humano. Além desses, muitos outros exemplos de culturas com uma melhor qualidade nutricional, produzidas por meio de interferência por RNA, são reportados na literatura. Isto evidencia o poder desta tecnologia como ferramenta molecular para gerar variedades com características agronomicamente importantes.

Assim, o objetivo deste trabalho foi construir cassetes de expressão para silenciamento gênico, via interferência por RNA, da proteína P34 e dos inibidores do tipo Bowman-Birk em sementes de soja. As linhagens transgênicas a serem obtidas em uma próxima etapa, com baixo teor ou ausência de P34 e BBI, serão introduzidas no Programa de Melhoramento da Qualidade da Soja do BIOAGRO/UFV.

2. MATERIAIS E MÉTODOS

2.1 Material Vegetal

Foram utilizadas sementes de soja da variedade comercial CAC-1 cultivada sob aquecimento controlado e fotoperíodo de 14 horas de luz em casa de vegetação da Universidade Federal de Viçosa, Viçosa – Minas Gerais.

As sementes foram coletadas em um período de aproximadamente três meses (plantio escalonado), separadas em oito estágios de desenvolvimento, que compreenderam praticamente todo o período de enchimento do grão, congeladas em nitrogênio líquido e armazenadas a -80°C. Esses estágios foram determinados com base no peso de matéria fresca da semente: 1^o – 0 a 75 mg; 2^o – 76 a 150 mg; 3^o – 151 a 225 mg; 4^o – 226 a 300 mg; 5^o – 301 a 375 mg; 6^o – 376 a 450 mg; 7^o – 451 a 525 mg e 8^o – representado por sementes maduras.

2.2 Extração de RNA total e RT-PCR

A extração de RNA total das sementes foi conduzida de acordo com Sambrook *et al.* (1989), com algumas adaptações ao protocolo original. Todas as etapas de extração do RNA total foram realizadas a 4°C e em condições livres de RNases.

Foram macerados cerca de 4 g de sementes na presença de nitrogênio líquido e a cada amostra foram adicionados 18 mL de tampão NTES (NaCl 0,1 M; Tris-HCl 0,01 M, pH 7,5; EDTA 1mM e SDS 1%), 6 mL de fenol e 6mL de clorofórmio:álcool isoamílico (24:1). Os tubos foram agitados vigorosamente em agitador vórtex por 15 min e centrifugados a 4°C por 10 min a 8.000 x g. A fase aquosa foi transferida para um novo tubo e nele foi adicionado 0,1 do volume de acetato de sódio 2M e 2,5 volumes de etanol 96% e, em seguida, a mistura foi incubada por 1 hora a -20°C. Foi realizada uma nova centrifugação a 8.000 x g por 15 min e o precipitado resultante foi lavado com etanol 70% e ressuspenso em 2,5 mL de água DEPC (água deionizada tratada com dietil pirocarbonato). Após uma centrifugação de 5 min a 5.000 x g, o sobrenadante foi

transferido para novo tubo onde foram adicionados 2,5 mL de cloreto de lítio 4M. O tubo foi incubado por cerca de 12 horas a 4°C para promover a precipitação do RNA. A amostra foi então centrifugada a 8.000 x g por 30 min e o precipitado foi ressuspenso em 1,8 mL de água DEPC e acrescido de 0,2 mL de acetato de sódio 2M e 3,6 mL de etanol 96%. Após precipitação de 6 horas a -20°C, a solução foi centrifugada a 8.000 x g por 10 min e em seguida, o precipitado foi lavado em etanol 70% e depois de seco, ressuspenso em água DEPC.

As amostras de RNA total foram quantificadas em espectrofotômetro (A_{260nm}) e a sua qualidade foi avaliada por eletroforese em gel de agarose 1,5% em tampão TBE (Tris-borato 90mM e EDTA 1mM, pH 8,0) contendo 0,2 µg/mL de brometo de etídeo. O padrão de bandas do RNA foi visualizado sob luz ultravioleta e fotografado com o sistema de fotodocumentação *Eagle Eye II* (STRATAGENE).

Para as reações de RT-PCR todas as amostras foram tratadas com *RQ1 DNase RNase-free* (PROMEGA), conforme as recomendações do fabricante. As amostras foram incubadas em tampão de reação (Tris-HCl 40 mM, pH 8,0; MgSO₄ 10 mM e CaCl₂ 1 mM) por 45 min a 37°C e extraídas com igual volume de fenol e de clorofórmio: álcool isoamílico após centrifugação a 11.000 x g por 2 min a 4°C. A fase aquosa foi transferida para um novo tubo e realizada uma outra extração com clorofórmio:álcool isoamílico (24:1). A fase aquosa foi submetida à precipitação com acetato de sódio 3M e etanol 96% por 1 hora a -20°C e depois centrifugada a 12.000 x g por 15 min. O precipitado foi lavado com etanol 70%, seco e ressuspenso em água DEPC. As amostras foram novamente quantificadas em espectrofotômetro.

A primeira fita de cDNA foi sintetizada usando o kit *SuperScript™ First-Strand Synthesis System* (INVITROGEN) de acordo com as recomendações do fabricante. As amostras de RNA total (2 µg) foram incubadas com 1 µL de oligo (dT)₁₂₋₁₈ a 70°C por 10 min e em seguida, incubadas no gelo. Foram adicionados tampão de PCR (Tris-HCl 20 mM, pH 8,4 e KCl 50 mM), MgCl₂ 5 mM, 0,5 mM dos quatro desoxirribonucleotídeos (dATP, dCTP, dGTP e dTTP), e DTT 5 mM, e estas foram incubadas a 42°C por 5 min. Em seguida, foram adicionadas 200 U da enzima transcriptase reversa e as amostras foram incubadas a 42°C por 50 min e a 70°C por 15 min. Após a síntese da primeira fita de cDNA, o RNA foi degradado pela adição de 1U de RNase H sob incubação a 37°C por 20

min. Para cada amostra foi feito um controle negativo que continha todos os reagentes menos a transcriptase reversa. Nesta etapa foi feita uma reação controle usando os *primers* de actina 3 de soja, com a finalidade de verificar a contaminação das amostras de RNA com DNA genômico e normalizar as concentrações de cDNA molde. A reação de amplificação a partir do DNA genômico de soja com os *primers* da actina 3, produz um fragmento de 520 pb. A reação a partir do cDNA gera um fragmento de 440 pb devido à remoção de um íntron de 80 pb.

2.3 Construção de *primers* específicos e isolamento dos fragmentos dos genes que codificam a P34 e os inibidores BBI

A escolha da região do cDNA dos genes que codificam a proteína P34 e os inibidores do tipo BBI que seria amplificada foi determinada com base em análises *in silico* com o auxílio dos programas TACG (<http://workbench.sdsc.edu>) e BLOCK-iTTM RNAi Designer (INVITROGEN) (<https://rnaidesigner.invitrogen.com/sima/>). O programa TAGC foi utilizado para a construção do mapa de restrição de todas as seqüências. Esse mapa permitiu a escolha das enzimas de restrição adequadas para os experimentos de clonagem. Já o programa BLOCK-iTTM RNAi Designer forneceu os potenciais siRNAs que poderiam se formar a partir de cada seqüência de interesse numa ordem de classificação. Assim, a região escolhida para amplificação compreende os melhores siRNAs obtidos.

Os oligonucleotídeos específicos foram construídos, a partir de seqüências de cDNA depositadas em banco de dados público – *GenBank* (*National Center for Biotechnology Information* – NCBI), com auxílio do programa “Primer3 Input Program” (<http://www.genome.wi.mit.edu/cgi-bin/primer/primer3.cgi>). Sítios de restrição reconhecidos pelas enzimas *Xho* I - *Kpn* I e *Xba* I – *Cla* I foram adicionados em seus terminais opostos, para a amplificação dos fragmentos *sense* e *antisense*, respectivamente. Parâmetros que atestam a viabilidade de uso destes oligonucleotídeos foram determinados com o auxílio do programa *QuickPrimer* do pacote DNASTAR (DNASTAR INC.). As seqüências estão descritas na Tabela 1.

Tabela 1: Sequência dos *primers* desenhados para amplificação dos fragmentos *sense* e *antisense* dos genes

Nº Acesso	Gene alvo	Primers	Sequência	Sítio de Restrição
J05560.1	P34	P34 sense-F	5' GGC <u>CTC GAG</u> GCA CAT GCA ATA GCA ACA GG 3'	<i>Xho</i> I
		P34 sense-R	5' GGC <u>GGT ACC</u> TCC CCC GGT GTA TAA ATG AA 3'	<i>Kpn</i> I
		P34 antisense-F	5' GGC <u>TCT AGA</u> GCA CAT GCA ATA GCA ACA GG3'	<i>Xba</i> I
		P34 antisense-R	5' GGC <u>ATC GAT</u> TCCCCCGGTGTATAAATGAA 3'	<i>Cla</i> I
X68704.1	BBI-A	BBIA sense-F	5' GGC <u>CTC GAG</u> GAC GAA CAT GGT GGT GCT AA 3'	<i>Xho</i> I
		BBIA sense-R	5' GGC <u>GGT ACC</u> TAT CTG AAC AGC GGC ATT GA 3'	<i>Kpn</i> I
		BBIA-antisense-F	5' GGC <u>TCT AGA</u> GAC GAA CAT GGT GGT GCT AA 3'	<i>Xba</i> I
		BBIA-antisense-R	5' GGC <u>ATC GAT</u> TAT CTG AAC AGC GGC ATT GA 3'	<i>Cla</i> I
X68705.1	BBI-CII	BBIC sense-F	5' GGC <u>CTC GAG</u> CGC ATG GAA CTG AAC CTC TT 3'	<i>Xho</i> I
		BBIC sense-R	5' GGC <u>GG TACC</u> TGT AGC AGA AGT CGG TGG TG 3'	<i>Kpn</i> I
		BBIC-antisense-F	5' GGC <u>TC TAGA</u> CGC ATG GAA CTG AAC CTC TT 3'	<i>Xba</i> I
		BBIC-antisense-R	5' GGC <u>ATC GAT</u> TGT AGC AGA AGT CGG TGG TG 3'	<i>Cla</i> I
X68706.1	BBI-DII	BBID sense-F	5' GGC <u>CTC GAG</u> TTT CCT TGT GGG GGT TAC AG 3'	<i>Xho</i> I
		BBID sense-R	5' GGC <u>GGT ACC</u> GCA GAA GTC GTT GGT GTC AA 3'	<i>Kpn</i> I
		BBID-antisense-F	5' GGC <u>TCT AGA</u> TTT CCT TGT GGG GGT TAC AG 3'	<i>Xba</i> I
		BBID-antisense-D	5' GGC <u>ATC GAT</u> GCA GAA GTC GTT GGT GTC AA 3'	<i>Cla</i> I

2.4 Amplificação e clonagem no vetor pGEM-TEasy

As reações de amplificação a partir de cDNA de sementes de soja maduras dos fragmentos *sense* e *antisense* foram conduzidas no termociclador GeneAmp® PCR System 9700 (APPLIED BIOSYSTEMS), com período inicial de desnaturação a 94°C por 2 min, seguido por 35 ciclos (94°C por 30 s; 60°C por 30 s; e 72°C por 20s) e um período adicional de extensão a 72°C, por 3 min. Cada reação continha, em 25 µl, 0,5 µl da reação de síntese de cDNA, Tris-SO₄ 60 mM (pH 8,9), (NH₄)₂SO₄ 18 mM, MgSO₄ 2 mM, dNTPs 0.2 mM, 0,4 µM de *primers* e 1U de Platinum® Taq DNA Polymerase High Fidelity (INVITROGEN).

Após a amplificação as amostras foram purificadas por meio do kit *QIAquick PCR Purification* (QIAGEN INC) e ligado ao vetor pGEM T-Easy (PROMEGA), de acordo com as recomendações do fabricante.

2.5 Transformação de *Escherichia coli* DH5α ultracompetentes e análise dos transformantes

Células de *Escherichia coli* DH5α ultracompetentes foram transformadas por choque térmico segundo Maniatis *et al.* (1989). Em 100 µl de células ultracompetentes foram adicionados 5 µl da reação de ligação e esta mistura foi incubada em gelo por 30 min. O choque térmico foi realizado a 42°C por 45 s seguido de uma incubação em gelo por 2 min. Para recuperação das células, foram adicionados 800 µl de meio SOC (0,5% de extrato de levedura, 2% de triptona, NaCl 10mM, KCl 2,5mM, MgCl₂ 10 mM, MgSO₄ 10 mM) e a cultura foi incubada a 37°C por uma hora a 150 rpm. Após esse período, as células foram concentradas a 1500 x g por 5 min em centrífuga 5415D (*Eppendorf*), ressuspensas em 200 µl de meio SOC, dos quais 100 µl foram plaqueados em meio LB sólido contendo ampicilina 50 µg/ml, 1 mg de X-Gal e 100 µM de IPTG. As placas foram incubadas a 37°C por cerca de 16 horas.

Para análise dos transformantes, 10 colônias isoladas foram repicadas em meio sólido e analisadas por PCR. Cada reação continha Tris-HCl 10mM pH 8,3, KCl 50 mM, MgCl₂ 2,0 mM, 0,2 mM de dNTPs, 0,4 mM de *primers* específicos e 1,5 µl de *Taq* DNA Polimerase. As colônias que mostraram a amplificação do produto esperado foram inoculadas em meio LB líquido contendo ampicilina 50 µg/ml e incubadas a 37°C durante 16 horas sob agitação. Após o crescimento, as bactérias foram aliquotadas em glicerol 25%, congeladas em nitrogênio líquido e armazenadas a -80°C. O DNA plasmidial foi extraído por meio do kit Wizard[®] Plus SV Minipreps (Promega) de acordo com as recomendações do fabricante. A clonagem foi confirmada utilizando-se amostras de DNA plasmidial em reações de PCR realizadas como descrito anteriormente, clivagem por enzimas de restrição e seqüenciamento.

A clivagem inicial do fragmento *sense* foi realizada com 500 ng de DNA, 5U da enzima *Kpn* I e tampão de clivagem (Tris-HCl 10 mM pH 7,5, MgCl₂ 7mM, KCl 50 mM e DTT 1 mM) por 1 hora a 37°C. Em seguida, foram adicionados tampão de clivagem (Tris-HCl 50 mM pH 8,0, MgCl₂ 10mM e NaCl 50 mM), 5U de *Xho* I e água em quantidade suficiente para dobrar o volume da reação inicial. A reação foi novamente incubada a 37°C por uma hora. Já para os fragmentos *antisense*, a clivagem foi realizada com 500 ng de DNA, 5U da enzima *Cla* I e tampão de clivagem (Tris-HCl 50 mM pH 8,0, MgCl₂ 10mM e NaCl 50 mM) por 1 hora a 37°C. Em seguida, foram adicionados 1,7 µl de NaCl 0,5 M, 5U de *Xba* I e a reação foi incubada a 37°C por mais 1 hora. O resultado da clivagem foi visualizado em gel de agarose 0,8%.

O seqüenciamento dos clones recombinantes foi realizado utilizando-se o kit ABI PRISM BigDye[®] III v 3.1 Terminator Cycle Sequencing Ready Reaction (APPLIED BIOSYSTEMS) e o seqüenciador automático ABI PRISM 377 Genetic Analyzer (APPLIED BIOSYSTEMS). As reações de seqüenciamento foram conduzidas de acordo com a técnica de seqüenciamento por terminação de cadeia por dideoxynucleotídeos (ddNTPs), descrita por Sanger *et al.* (1977). As reações de amplificação linear foram realizadas usando termociclador (APPLIED BIOSYSTEMS), GeneAmp[®] PCR System 9.700, programado para um período inicial de desnaturação a 96°C, por 2 min, seguido por 25 ciclos (96°C, por 30 s; 50°C, por 20 s; e 60°C, por 4 min). Foram utilizados os *primers* universais: M13F e M13R. Os produtos de extensão, produzidos na reação de

seqüenciamento, foram purificados pela adição de 0,1 volumes de Acetato de Sódio 3M (pH 4,8), 2 volumes de etanol 95% (v/v) e incubação à temperatura ambiente por 15 min. Os fragmentos foram coletados por centrifugação (10.000 x g por 20 min) e o precipitado foi lavado com etanol 70% (v/v), seco à temperatura ambiente e, em seguida, ressuspenso em tampão de corrida (17% azul de bromofenol:83% de formamida). Finalmente, os fragmentos foram desnaturados a 94°C por 5 min e separados por eletroforese em gel desnaturante de poliacrilamida 4,5%.

A seqüência dos clones isolados foram analisadas pelo Programa *Seqman* do pacote DNASTAR (DNASTAR Inc.) e o alinhamento com outras seqüências depositadas no *GenBank* foi realizado com o auxílio do BLAST (Altschul *et al.*, 1997).

2.6 Clonagem em vetores pKANNIBAL e pBKN

Para a clonagem em vetores pKANNIBAL e pBKN, os fragmentos *sense* e *antisense* foram liberados do vetor pGEM T-Easy com as enzimas de restrição adequadas. A clivagem foi realizada conforme já descrito no item 2.5 e os produtos excisados do gel e purificados com o auxílio do kit *QIAquick Gel Extraction* (QIAGEN INC.)

A seqüência *sense* foi ligada em vetores pKANNIBAL e pBKN, também digeridos conforme descrito anteriormente, e para cada 10 µl da reação de ligação utilizou-se 10 ng do vetor clivado, 50 ng do inserto clivado, tampão de reação (Tris-HCl 50 mM pH 7,6, MgCl₂ 10 mM, ATP 1 mM, DTT 1 mM e 25% de polietileno glicol-8000) e 1U de T4 DNA ligase (INVITROGEN). A mistura foi incubada a 4°C por 16 horas.

A análise dos transformantes e a confirmação da clonagem foi feita utilizando-se PCR de colônias e amostras de DNA plasmidial em reações de PCR e clivagem por enzimas de restrição realizadas como descrito anteriormente.

Para a clonagem da seqüência *antisense*, os clones contendo a seqüência *sense* foram clivados, conforme descrito anteriormente, com *Xba* I e *Cla* I e utilizados como

vetores. A reação de ligação, a transformação e a análise de transformantes foram realizadas como o descrito para a clonagem da sequência *sense*.

2.7 Clonagem no vetor pCAMBIA 3301

As construções, contidas nos clones pKNXi e pBKNX, foram transferidas para o vetor binário pCAMBIA 3301. Esse vetor, de 11307 pb, contém uma região de T-DNA (5077 pb) e apresenta o promotor 35SCaMV dirigindo a expressão do gene *gus* (que codifica a β -glucuronidase), o terminador do gene que codifica a nopalina sintase (NOS), além do gene *bar* para a seleção em plantas. Esse gene codifica a enzima fosfonitricina acetiltransferase, que confere resistência a herbicidas que contenham glufosinato de amônio como composto ativo. Os cassetes foram liberados dos vetores intermediários por meio de clivagem com as enzimas *Sac* I e *Pst* I. As condições de clivagem, purificação dos fragmentos, ligação, transformação e análise de transformantes foram realizadas de acordo com o descrito anteriormente.

3. RESULTADOS E DISCUSSÃO

Para a construção dos cassetes de expressão, o cDNA foi sintetizado a partir de RNA total de sementes de soja e os fragmentos obtidos por PCR. O tamanho dos fragmentos visualizados correspondeu ao tamanho esperado para o produto amplificado a partir dos *primers* sintetizados para este trabalho. Estes tamanhos eram, aproximadamente, 400 pb de bases para os fragmentos referentes à P34 e 200 pb para aqueles referentes ao BBI. O uso de seqüências relativamente pequenas para silenciamento tem sido bastante descrito em trabalhos semelhantes àquele aqui reportado (Ogita *et al.*, 2004; Gilissen *et al.*, 2005; Miles *et al.*, 2005). Além de ser suficiente para garantir um silenciamento eficaz, pequenos fragmentos são de fácil manipulação e favorecem a eficiência de clonagem e transformação já que a construção de cassetes de expressão geralmente envolve vetores de alto peso molecular.

Antes da construção dos cassetes de expressão, foram realizados ensaios de RT-PCR para a análise da expressão dos genes de interesse em sementes de soja coletadas ao longo de todo o período de enchimento do grão (Figura 1A) e em outros órgãos vegetais - raiz, caule e folha (Figura 1B). Embora a utilização da técnica de RT-PCR como método para análise quantitativa de expressão gênica seja discutível (Bustin, 2000), este método foi escolhido porque o objetivo maior dos experimentos seria a análise qualitativa da expressão gênica. Outra vantagem se refere à facilidade de execução da metodologia. Como controle interno foi utilizado o gene de actina de soja que é expresso constitutivamente.

A análise por RT-PCR mostrou que os genes analisados são expressos em todos os órgãos vegetais e durante todos os estágios de desenvolvimento da semente. O padrão de expressão revelou-se uniforme em todos os órgãos analisados. Kalinski *et al.* (1992) analisaram a presença de mRNA e da proteína P34 em partes da semente (cotilédones, aleurona, etc) e em raízes, caules e folhas de soja. Embora o mRNA tenha sido detectado em toda a planta, a proteína só foi encontrada na semente. Segundo esses autores, como a P34 pertence a uma superfamília de tiol-proteases, o mRNA detectado não codificaria a proteína P34 mas outras proteases relacionadas. No entanto, Cheng *et al.* (1998) detectaram a proteína P34 em folhas, raiz, caule e suspensão celular de dois

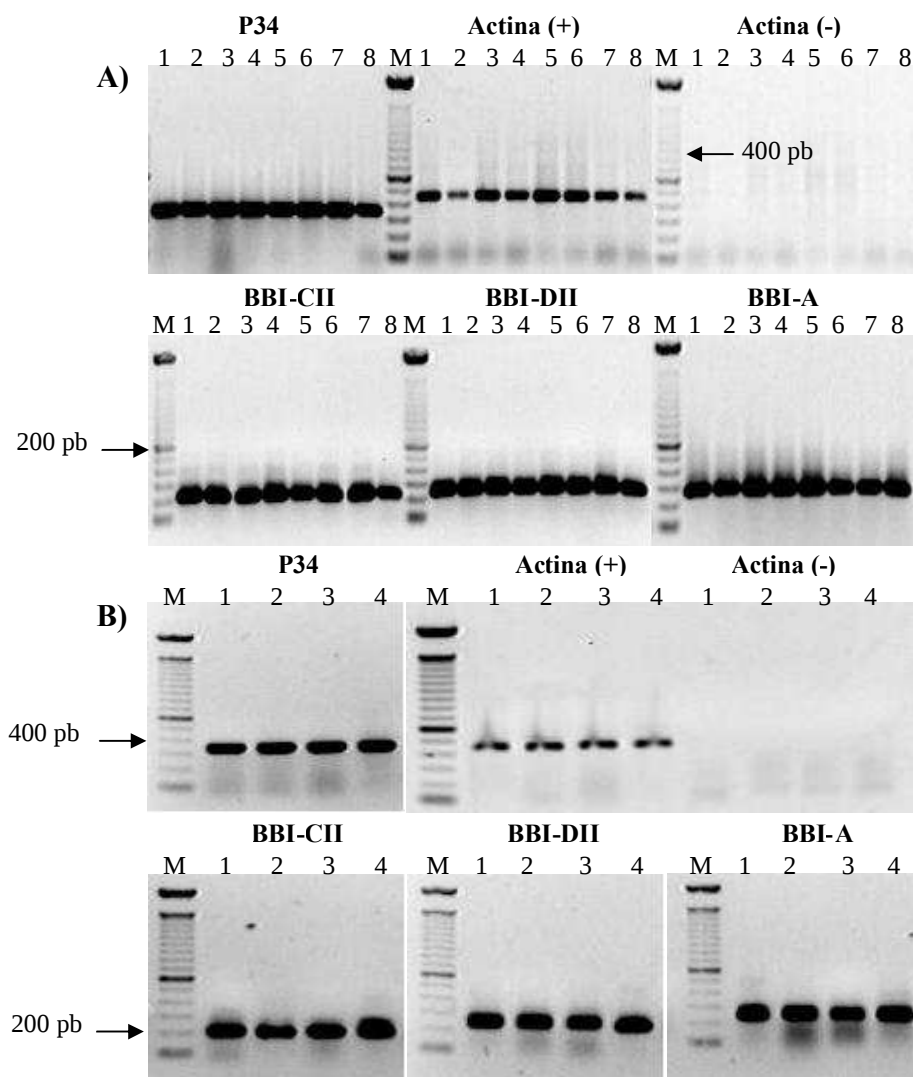


Figura 1: Ensaios de RT-PCR para análise da expressão gênica dos genes que codificam a proteína P34 e os inibidores de protease do tipo Bowman-Birk. (A) apresenta os resultados dos ensaios realizados a partir de cDNA obtido de RNA extraído de sementes coletadas durante o desenvolvimento da semente da soja. Os números de 1 a 8 representam os oito estágios de desenvolvimento da semente. Em (B) está representado a análise da expressão dos genes de interesse em outros órgãos vegetais. Os números de 1 a 4 se referem à raiz, caule, folha e semente, repectivamente. Os controles, positivo e negativo, de cada ensaio foram feitos com actina e M refere-se ao marcador de tamanho de 100 pb (INVITROGEN).

cultivares de soja. De acordo com Herman (2005), a presença de P34 em células vegetativas está relacionada com o gene de resistência *RPG4*. Ensaio com *imunoblot* mostraram que a proteína é mais abundante em linhagens de soja resistentes a *Pseudomonas*. Qu *et al.* (2003) demonstraram, em arroz, que inibidores do tipo BBI também são expressos em folhas, raiz e órgãos florais. Além disso, a super-expressão de BBI em plantas transgênicas de arroz resultou na resistência ao fungo patogênico *Pyricularia oryzae*, indicando que inibidores de protease podem ter um papel na defesa sistêmica de plantas de arroz. Diante disso, o silenciamento destes genes não deve ser governado por promotor constitutivo, já que sua presença em soja também pode estar relacionada com resistência a doenças.

A clonagem dos fragmentos correspondentes aos genes de interesse foi feita a partir de cDNA de sementes maduras. Os fragmentos *sense* e *antisense* de cada gene foram sintetizados por PCR, purificados e individualmente clonados no vetor pGEM T-Easy. Após a transformação de células ultracompetentes de *Escherichia coli*, 10 colônias recombinantes, para cada tipo de inserto, foram analisadas por meio de PCR usando *primers* M13 e todas continham o fragmento esperado. Uma colônia de cada experimento foi escolhida e a clonagem foi confirmada por PCR com *primers* específicos para cada fragmento (Figura 2A), por clivagem com enzimas de restrição (Figura 2B) e por meio de seqüenciamento dos DNAs plasmidiais dos clones isolados.

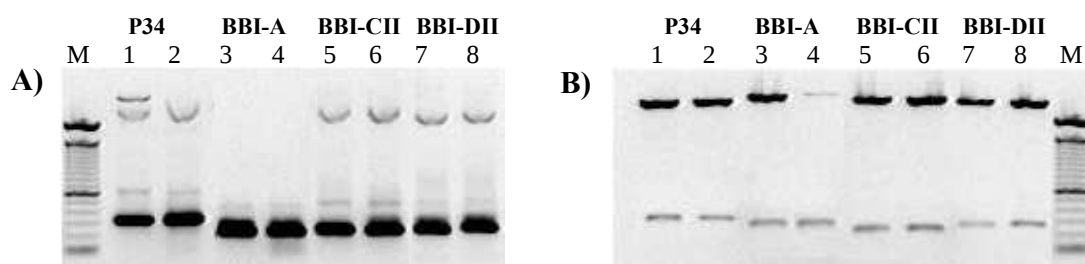


Figura 2: Ensaio para confirmação da clonagem dos fragmentos *sense* e *antisense* de cada gene de interesse no vetor pGEM-TEasy. (A) representa a reação de PCR e (B) a reação de clivagem com DNA plasmidial dos clones isolados. As canaletas ímpares representam os fragmentos *sense* e as pares, os fragmentos *antisense*. M refere-se ao marcador de tamanho de 100 pb (INVITROGEN).

As seqüências apresentaram 349 pb (P34), 224 pb (BBI-A), 242 pb (BBI-CII) e 263 pb (BBI-DII). Cada uma delas foi alinhada com demais seqüências depositadas no *GenBank*, com o auxílio do BLAST, o que confirmou a similaridade de cada fragmento com as seqüências dos genes aos quais eles se referem (Figura 3).

Posteriormente, os fragmentos *sense* foram liberados do vetor pGEM-TEasy por meio de clivagem com as enzimas *Xho* I e *Kpn* I, purificados e ligados aos vetores pKANNIBAL e pBKN. A clonagem foi confirmada por PCR e clivagem enzimática do DNA plasmidial de clones isolados. Os clones positivos foram denominados pKNXs e pBKNXs quando foram utilizados os vetores pKANNIBAL e pBKN, respectivamente. X se refere ao fragmento de interesse: P34, BBI-A, BBI-CII ou BBI-DII.

A)

Clone P34	1	GTTCTTAAGCGCCATCCTTGAGCAACCAATTAGTGTCTCCATTGGATGCAAAAGATTTTC	60
J05560.1	785	GTTCTTAAGCGCCATCCTTGAGCAACCAATTAGTGTCTCAATTG - ATGCAAAAGATTTTC	843
Clone P34	61	ATTTATACACCGGGGGAATTTATGATGGAGAAACTGTACAAGTCCGTATGGGATTAAT	120
J05560.1	844	ATTTATACACCGGGG - GAATTTATGATGGAGAAACTGTACAAGTCCGTATGGGATTAAT	902
Clone P34	121	CACTTTGTTTTACTTGTGGGTTATGGTTCAGCGGATGGTGTAGATTACTGGATAGCGAAA	180
J05560.1	903	CACTTTGTTTTACTTGTGGGTTATGGTTCAGCGGATGGTGTAGATTACTGGATAGCGAAA	962
Clone P34	181	AATTCATGGGGAGAAGATTGGGGAGAAGATGGTTACATTTGGATCCAAAGAAACACGGGT	240
J05560.1	963	AATTCATGGGGAGAAGATTGGGGAGAAGATGGTTACATTTGGATCCAAAGAAACACGGGT	1022
Clone P34	241	AATTTATTAGGAGTGTGTGGGATGAATTATTTTCGCTTCATACCCAACCAAGAGGARTCA	300
J05560.1	1023	AATTTATTAGGAGTGTGTGGGATGAATTATTTTCGCTTCATACCCAACCAAGAGGAATCA	1082
Clone P34	301	GAAACACTGGTGTCTGCTCGCGTTAAAGGTC	331
J05560.1	1083	GAAACACTGGTGTCTGCTCGCGTTAAAGGTC	1113

B)

Clone BBIA	1	GACGAACATGGTGGTGCTAAAGGTGTGTTTGGTGCTACTTTTCCTTGTGGGGGGTACTAC	60
X68704.1	18	GACGAACATGGTGGTGCTAAAGGTGTGTTTGGTGCTACTTTTCCTTGTGGGGGGTACTAC	77
Clone BBIA	61	TAGTGCCAACCTTGAGGCTGAGTAAGCTTGGCCTGCTCATGAAAAGTGATCATCATCAACA	120
X68704.1	78	TAGTGCCAACCTTGAGGCTGAGTAAGCTTGGCCTCCTCATGAAAAGTG - - - ATCATCAACA	134
Clone BBIA	121	CTCAAATGATGATGAGTCTTCAAACCATGCTGTGATCAATGCGCGTGC GCGCGTCAAA	180
X68704.1	135	CTCAAATGATGATGAGTCTTCAAACCATGCTGTGATCAATGCGCATGCACAAAGTCAAA	194
Clone BBIA	181	CCCTCCTCAATGCCGCTGTTTCAGATA	206
X68704.1	195	CCCTCCTCAATGCCGCTGTTTCAGATA	220

C)

Clone BBIC	1	CGCATGGAAGTGAAGTCTTTTCAAAAGTGGATCACTCATCAAGTGGATGATGAGTCTTCA	60
X68705.1	32	CGCATGGAAGTGAAC-CTCTTCAAAAGTG-ATCACTCATCAAGTG-ATGATGAGTCTTCA	88
Clone BBIC	61	AAAYCATGCTGTGCATCTCTGCATGTGCACAGCCTCAATGCCACCTCAATGCCATTGTGC	120
X68705.1	89	AAACCATGCTGTG-ATCTCTGCATGTGCACAGCCTCAATGCCACCTCAATGCCATTGTGC	147
Clone BBIC	121	AGATATTAGGTTGAATTCATGTCACTCAGCTTGTGATCGCTGTGCGTGCACACGCTCGAT	180
X68705.1	148	AGATATTCGGTTGAATTCATGTCACTCAGCTTGTGATCGCTGTGCGTGCACACGCTCGAT	207
Clone BBIC	181	GCCAGGCCAGTGTGCTTGCCTTGACACCACCGACTTCTGCTACA	224
X68705.1	208	GCCAGGCCAGTGTGCTTGCCTTGACACCACCGACTTCTGCTACA	251

D)

Clone BBID	1	TTTCCTTGTGGGGGTTACAGCTGCACGCATGGAAGTGAAGCTTCTTCAAAAGTGA---GTC	57
X68706.1	59	TTTCCTTGTGGGGGTTACAGCTGCACCGATGGAAGTGAAGCTTCTTCAAAAGTGATCAGTC	118
Clone BBID	58	AGCCAGTTATGCATGATGATGAGTATTCAAACCATGCTGTGATCTCTGCATGTGCACAC	117
X68706.1	119	ATCAAGTTATG-ATGATGATGAGTATTCAAACCATGCTGTGATCTCTGCATGTGCACAC	177
Clone BBID	118	GCTCAATGCCTCCTCAATGCAGCTGGGAAGATATTAGGCTGAATTCATGCCACTCAGATT	177
X68706.1	178	GCTCAATGCCTCCTCAATGCAGCTGTGAAGATATTAGGCTGAATTCATGCCACTCAGATT	237
Clone BBID	178	GTAAGAGCTGTATGTGCACACGCTCACAGCCAGGACAGTGGCGTTGTCTTGACACCAACG	237
X68706.1	238	GTAAGAGCTGTATGTGCACACGCTCACAGCCAGGACAGTGTGCTTGTCTTGACACCAACG	297
Clone BBID	238	ACTTCTGC	245
X68706.1	298	ACTTCTGC	305

Figura 3: Alinhamento entre as seqüências clonadas e as seqüências de cDNA, depositadas no *GenBank*, que dos genes que codificam as proteínas P34 (A), BBI-A (B), BBI-CII (C), BBI-DII (D). As barras em azul indicam as bases idênticas no alinhamento

Para a clonagem dos fragmentos *antisense*, estes também foram liberados do vetor pGEM-TEasy e os clones pKNXs e pBKNXs foram utilizados como vetores. A clonagem foi confirmada da mesma forma utilizada para os fragmentos *sense*. Os clones foram denominados pKNXi e pBKNXi e as construções apresentaram o tamanho esperado (Figuras 4-6).

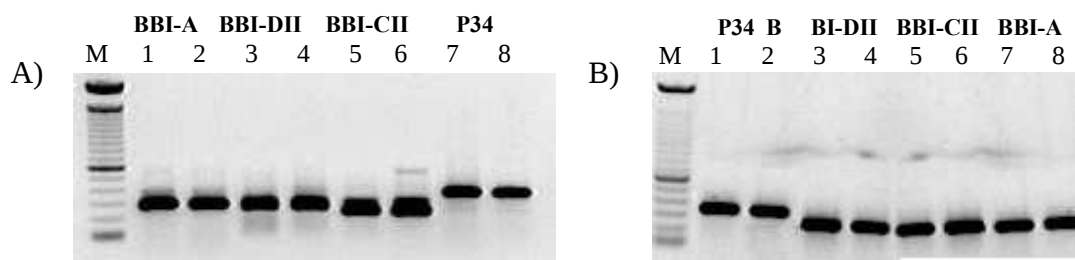


Figura 4: Reação de PCR para confirmação da clonagem dos fragmentos *sense* e *antisense* de cada gene de interesse nos vetores pKANNIBAL (A) e pBKN (B). As canaletas ímpares representam os fragmentos *sense* e as pares, os fragmentos *antisense*. M refere-se ao marcador de tamanho de 100 pb (INVITROGEN).

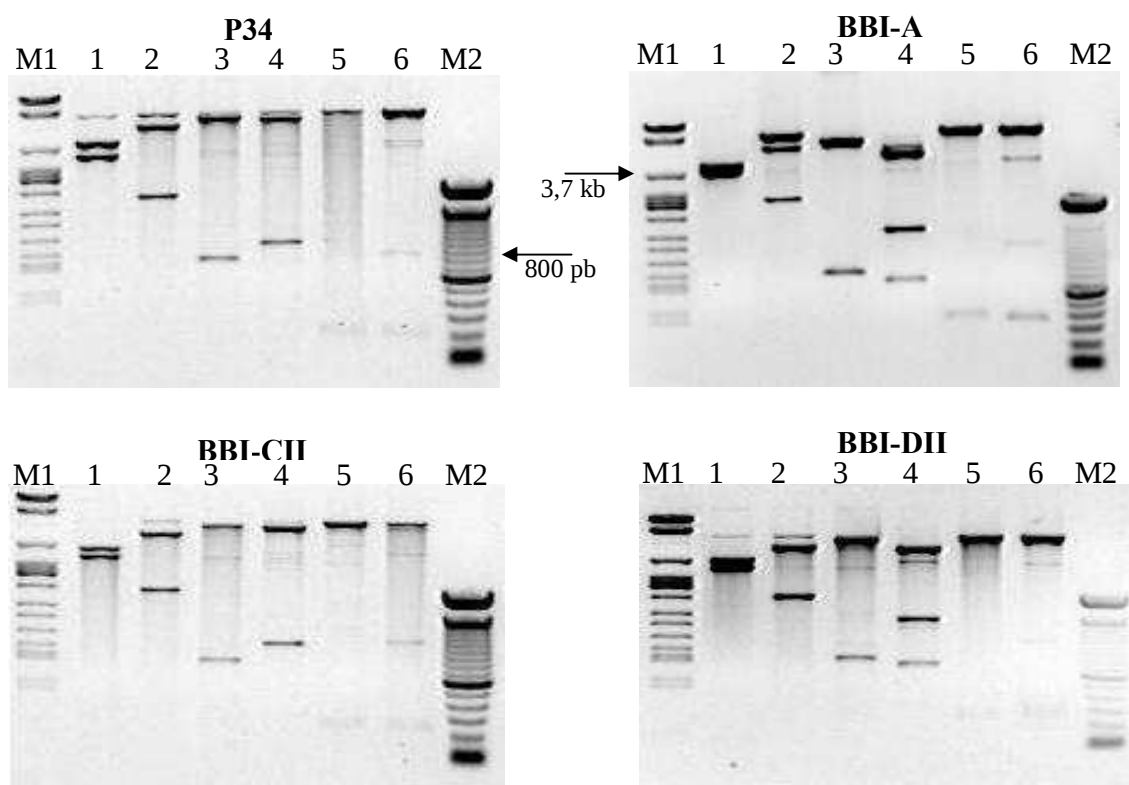


Figura 5: Reação de clivagem para confirmação da clonagem dos fragmentos *sense* e *antisense* de cada gene de interesse no vetor pKANNIBAL. As canaletas correspondem à liberação da construção inteira (1), do promotor (2), do íntron (3), do terminador (4), do fragmento *sense* (5) e do fragmento *antisense* (6). M1 refere-se ao DNA de fago λ clivado com as enzimas *Eco* RI, *Bam* HI e *Hind* III, utilizado como marcador de peso molecular. M2 refere-se ao marcador de tamanho de 100 pb (INVITROGEN).

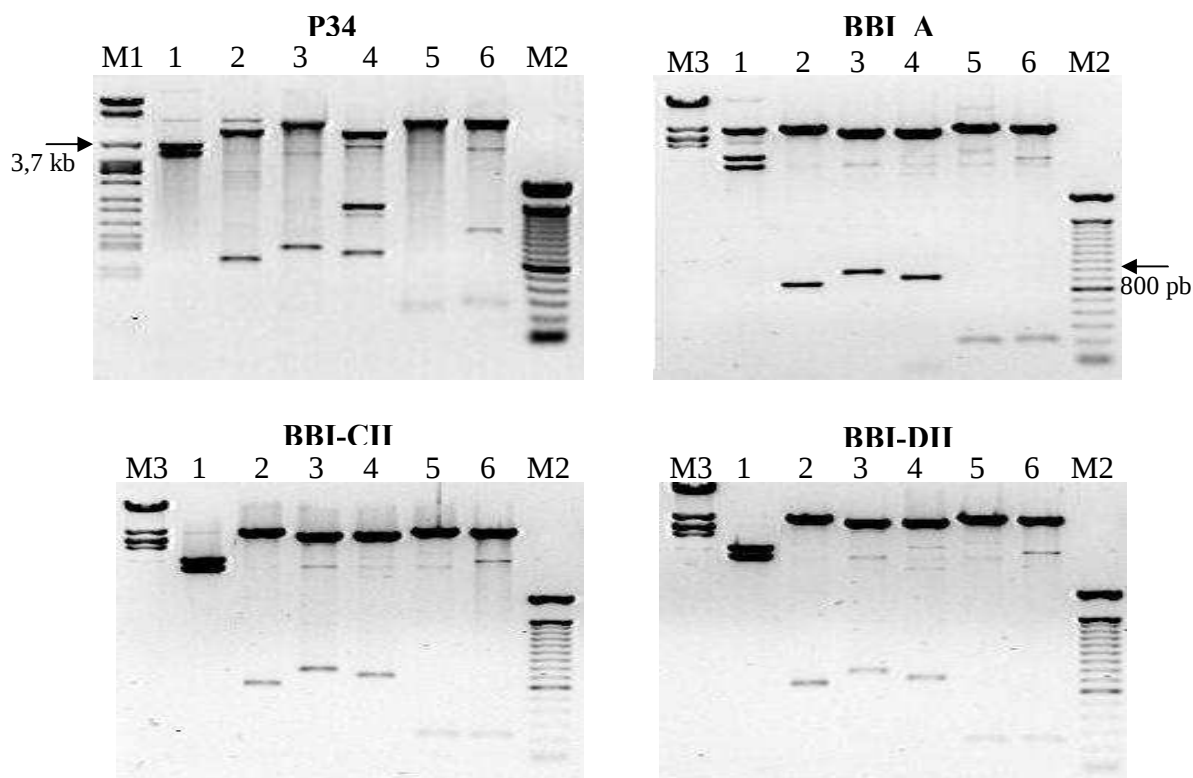


Figura 6: Reação de clivagem para confirmação da clonagem dos fragmentos *sense* e *antisense* de cada gene de interesse no vetor pBKN. As canaletas correspondem à liberação da construção inteira (1), do promotor (2), do íntron (3), do terminador (4), do fragmento *sense* (5) e do fragmento *antisense* (6). M1 e M3 referem-se ao DNA de fago λ clivado com as enzimas *Eco* RI, *Bam* HI, *Hind* III e *Eco* RI, respectivamente, utilizados como marcador de peso molecular. M2 refere-se ao marcador de tamanho de 100 pb (INVITROGEN).

Os clones pKNXi e pBKNXi contêm o cassete com a estrutura necessária para o silenciamento dos genes de interesse, mas não apresentam uma marca para seleção em plantas. Além disso, como a transformação de plantas será realizada por meio de *Agrobacterium tumefaciens*, uma estrutura de T-DNA também se faz necessária. Assim as construções foram liberadas dos clones pKNXi e pBKNXi e inseridas no vetor binário pCAMBIA 3301. Novamente, a clonagem foi confirmada por PCR e clivagem enzimática de DNA plasmidial de clones isolados, que foram denominados pXi e pBXi.

As representações esquemáticas dos clones obtidos estão apresentadas nas figuras de 7 a 10.

Para assegurar que o silenciamento gênico ocorra de forma mais eficiente o tipo de construção e o gene alvo utilizados devem ser destacados.

Cartea *et al.* (1998) mostraram o efeito do silenciamento do gene *Fad2-1* que codifica a enzima ω -6 dessaturase microssomal em *Arabidopsis*, usando construções *antisense* e co-supressão. A construção *sense* para co-supressão promoveu uma frequência de silenciamento de 10% entre as 41 plantas transgênicas obtidas. A construção *antisense* produziu aproximadamente 15% de plantas transgênicas exibindo silenciamento. Esses resultados ficam bem abaixo daqueles obtidos por Singh *et al.* (2000), utilizando construções tipo hpRNA e ihpRNA. Quando estas construções foram expressas em *Arabidopsis*, cerca de 69% das plantas transgênicas mostraram silenciamento e quando o espaçador entre as repetições invertidas foi substituído por um intron, 100% dos regenerantes mostraram silenciamento do gene da ω 6-dessaturase. Em soja, vários trabalhos são citados na literatura. Nunes e colaboradores (2006) silenciaram, via interferência por RNA, o gene *GmMIP51*, que codifica enzima a mio-inositol-1-fosfato sintase, visando a obtenção de sementes com conteúdo reduzido de fitato. Embora o silenciamento tenha sido altamente eficiente, as sementes obtidas não foram viáveis. O fitato é uma molécula reserva de grupamentos fosfatos que são importantes durante o processo de germinação das sementes. Esta construção será empregada na transformação de fruteiras para a obtenção de frutos sem sementes.

De acordo com Kerschen *et al.* (2004) a eficiência do processo de silenciamento também é dependente do gene considerado como alvo. Esses pesquisadores, por meio de um protocolo de RT-PCR quantitativo, analisaram linhagens transgênicas de *Arabidopsis* que visavam o silenciamento de 25 genes endógenos diferentes. Essa análise revelou que linhagens independentes, homozigotas, que apresentavam inserção de uma única cópia do transgene e que tinham o mesmo gene como alvo apresentaram níveis de redução do transcrito na mesma extensão. Linhagens que apresentaram inserções múltiplas diferiram no grau de redução do mRNA alvo, mas nunca excederam os níveis apresentados por linhagens com inserção única. A redução máxima dos níveis dos transcritos variou de acordo com o gene alvo. Estas observações sugerem que cada

seqüência de mRNA alvo possui um grau inerente de susceptibilidade à degradação mediada por dsRNA.

Com relação ao presente trabalho, em uma próxima etapa, nós cotiledonares de soja serão transformados e as linhagens que apresentarem baixo teor ou ausência das proteínas em questão serão incorporadas ao Programa de Melhoramento da Qualidade da Soja do BIOAGRO/UFV.

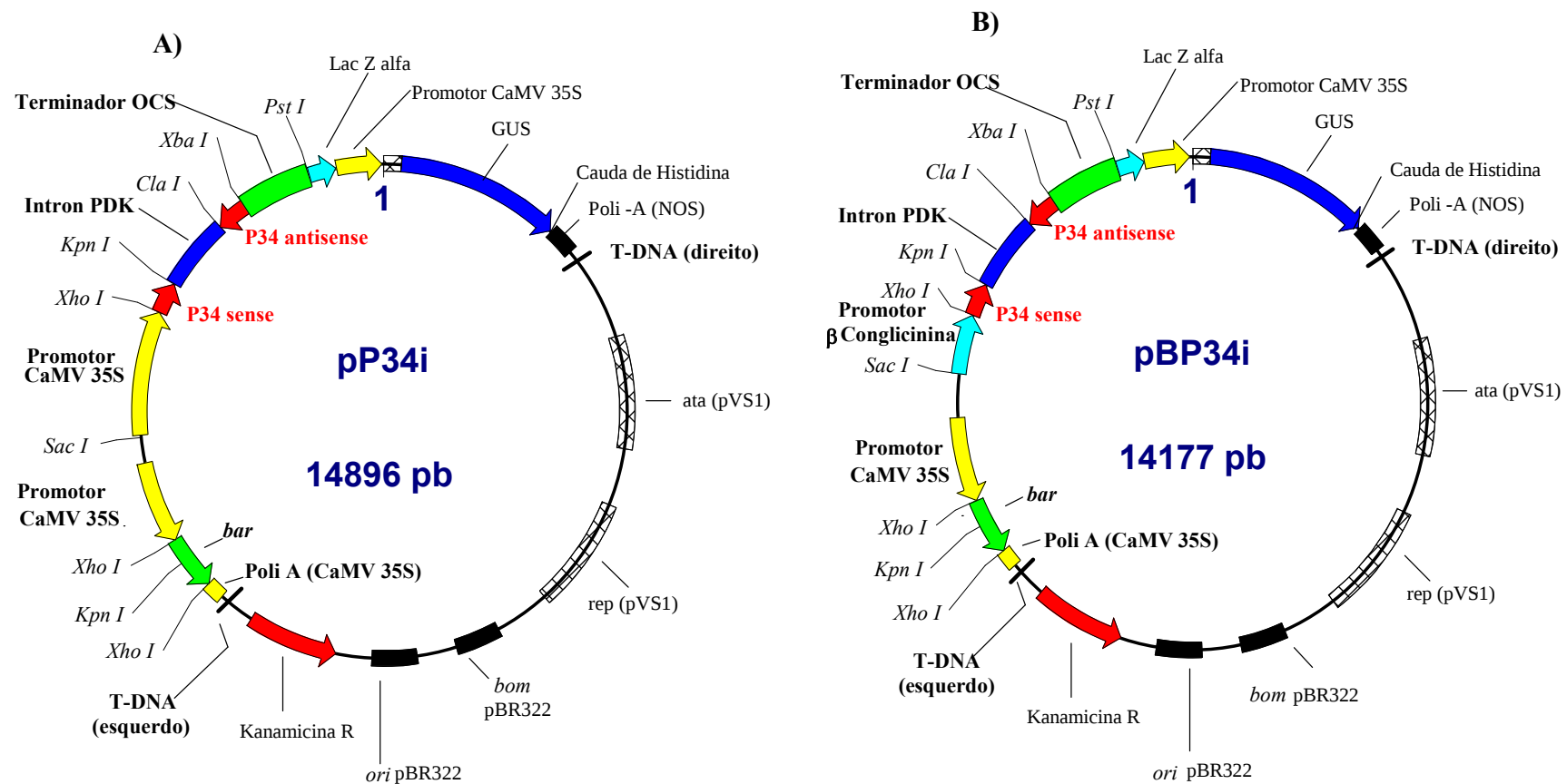


Figura 7: Representação esquemática dos clones contendo as construções para silenciamento gênico, via interferência por RNA, da proteína P34 em soja. Os cassetes estão sob controle do promotor CaMV 35S (A) e do promotor do gene da subunidade α da β -conglucininina (B).

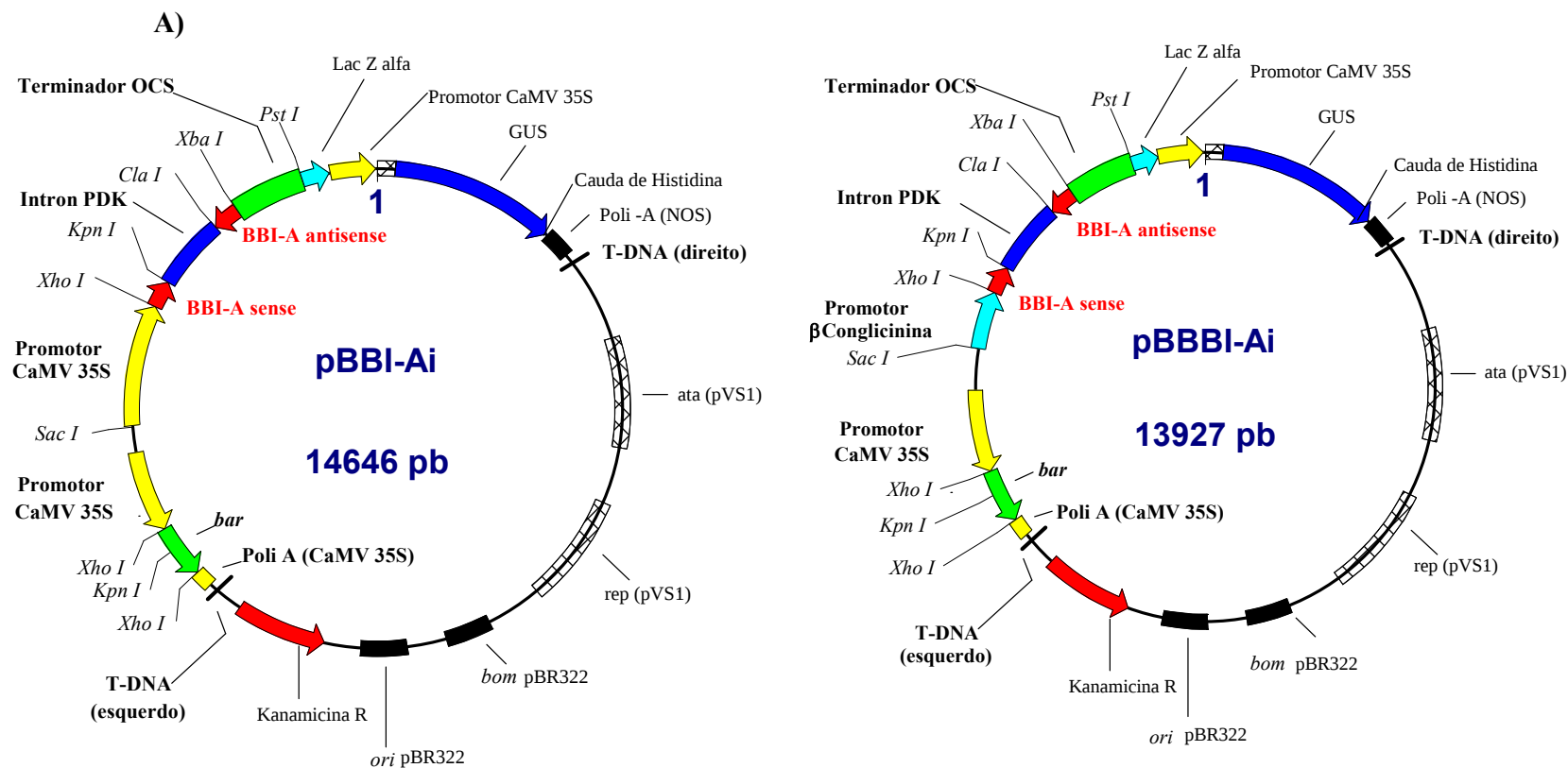


Figura 8: Representação esquemática dos clones contendo as construções para silenciamento gênico, via interferência por RNA, do inibidor BBI-A em soja. Os cassetes estão sob controle do promotor CaMV 35S (A) e do promotor do gene da subunidade α da β -conglucininina (B).

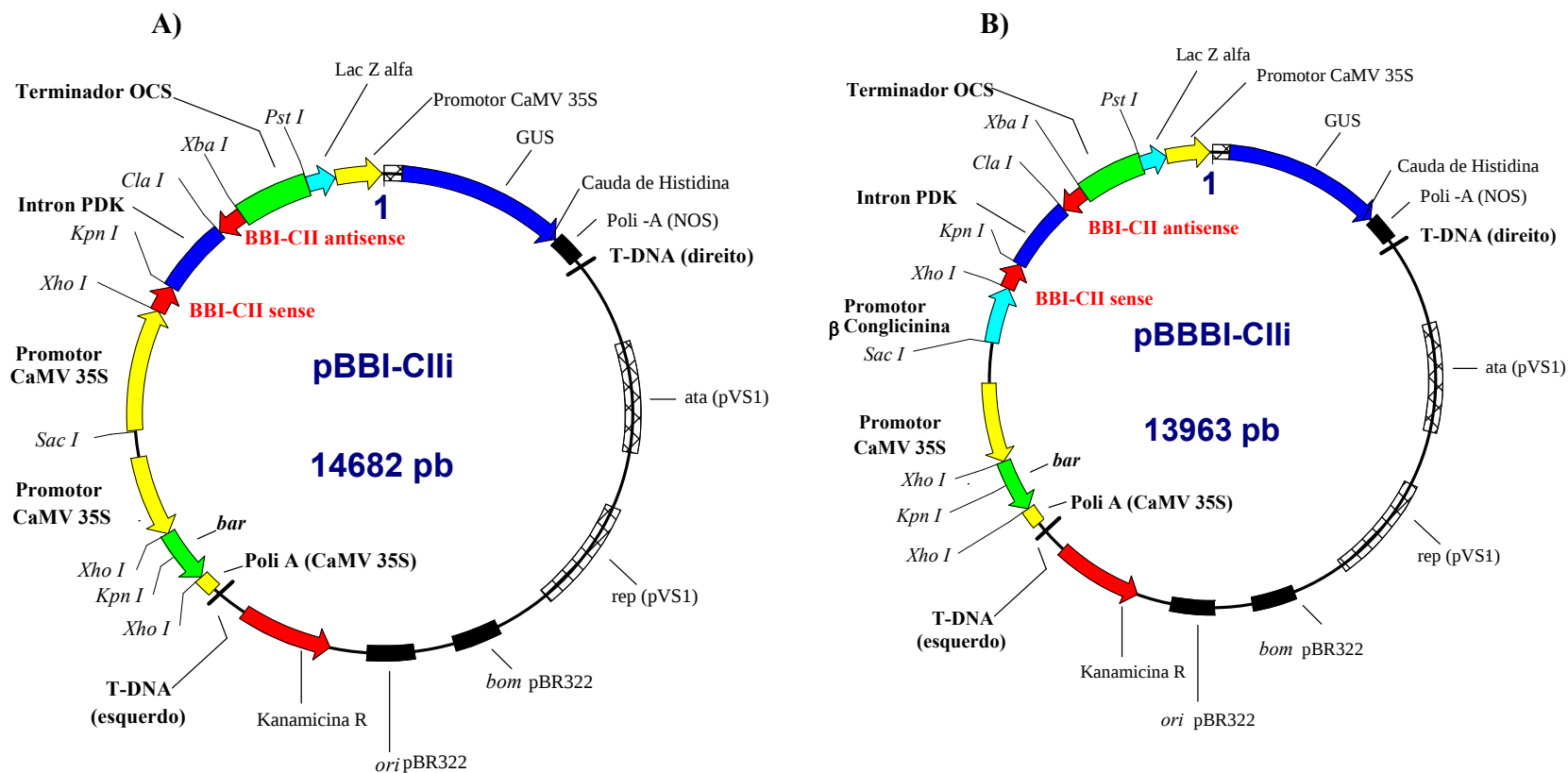


Figura 9: Representação esquemática dos clones contendo as construções para silenciamento gênico, via interferência por RNA, do inibidor BBI-CII em soja. Os cassetes estão sob controle do promotor CaMV 35S (A) e do promotor do gene da subunidade α da β -conglucina (B).

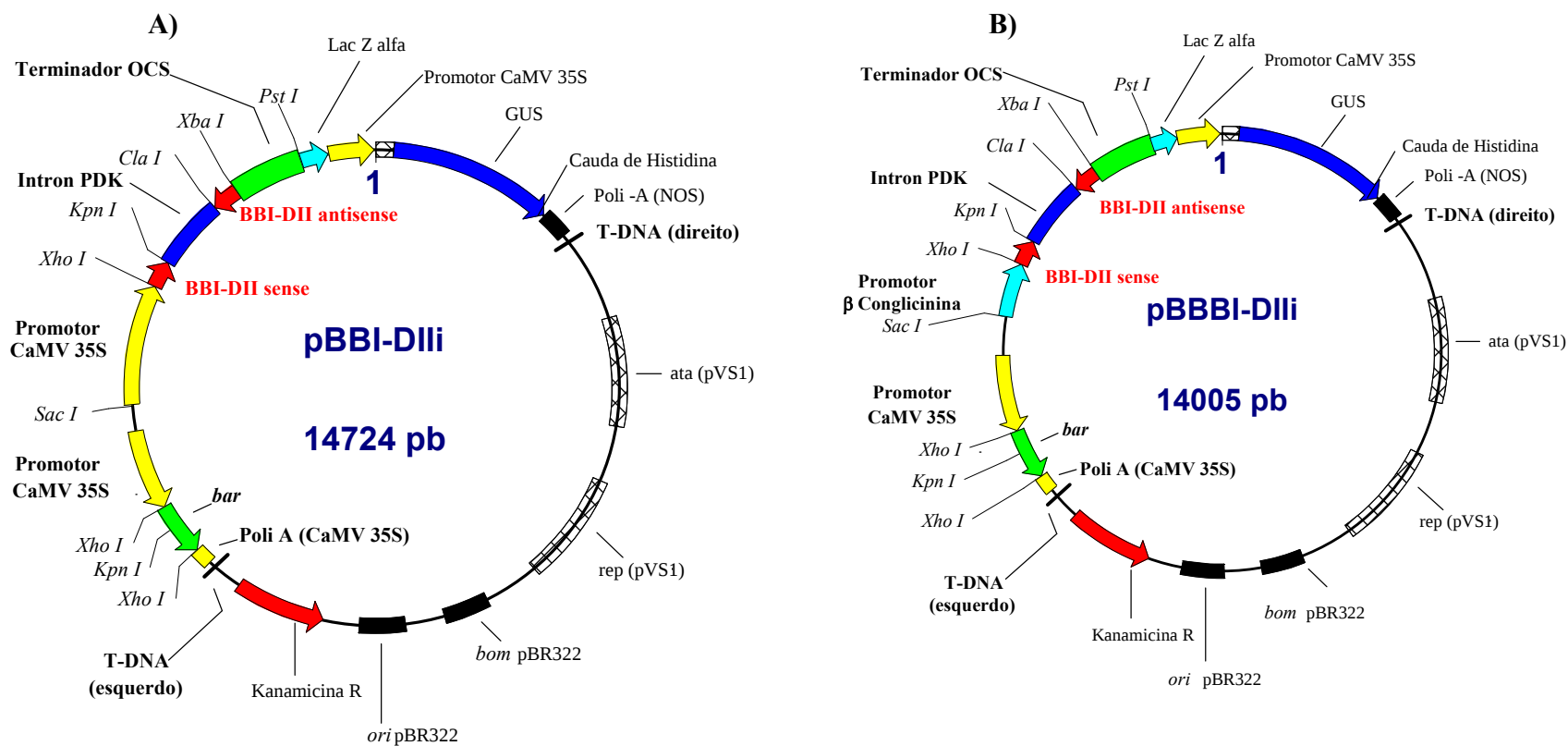


Figura 10: Representação esquemática dos clones contendo as construções para silenciamento gênico, via interferência por RNA, do inibidor BBI-DII em soja. Os cassetes estão sob controle do promotor CaMV 35S (A) e do promotor do gene da subunidade α da β -conglucininina (B).

4. CONCLUSÕES

- Foi possível isolar e clonar fragmentos referentes à região codificadora dos genes que codificam a proteína Gly mBd 30K (P34) e os inibidores de protease do tipo Bowman-Birk.

- Foram obtidos dois cassetes de expressão para silenciamento gênico, via interferência por RNA, da proteína Gly mBd 30K (P34) e seis cassetes para o silenciamento gênico de inibidores de protease do tipo Bowman-Birk.

5. PERSPECTIVAS

- *Agrobacterium tumefaciens* será transformada com esses cassetes e utilizadas para transformação genética de nós cotiledonares e/ou embriões somáticos de soja.

- As possíveis variedades com baixo teor ou ausência das proteínas de interesse serão incorporadas no Programa de Melhoramento de Qualidade da Soja do BIOAGRO/UFV.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- Altschul, S. F.; Madden, T. L.; Schaffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. Gapped **BLAST** and **PSI-BLAST**: a new generation of protein database search programs. *Nucl. Acid. Res.*, **25**: 3389-3402, 1997.
- Nunes, A. C. S.; Vianna, G. R.; Cuneo, F.; Amaya-Farfán, J.; Capdeville, G.; Rech, E. L.; Aragão, F. J. L. **RNAi-mediated silencing of the myo-inositol-1-phosphate synthase gene (GmMIPS1) in transgenic soybean inhibited seed development and reduced phytate content.** *Planta*, **224**: 125-132, 2006.
- Baek, J. M.; Song, J. C.; Choi, Y. D.; Kim, S. I. **Nucleotide sequence homology of cDNAs encoding soybean Bowman-Birk type proteinase inhibitor and its isoinhibitors.** *Biosci. Biotechnol. Biochem.*, **58**: 843- 846, 1994.
- Birk, Y.; Gertler, A.; Khalef, S. **A pure trypsin inhibitor form soya beans.** *Biochem. J.*, **87**: 281-284, 1963.
- Birk, Y. **The Bowman-Birk inhibitor.** *Int. J. Pept. Protein Res.*, **32**: 113-131, 1985.
- Bowman, D. E. **Differentiation of soybean anti-tryptic factors.** *Proc. Soc. Exp. Biol. Med.*, **63**: 547-550, 1946.
- Bush, R. K.; Hefle, S. L. **Food allergens.** *Crit. Rev. Food Sci. Nutr.*, **36**: 119-163, 1996.
- Bustin, S. A. **Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays.** *J. Mol. Endocrinol.*, **25**: 169-193, 2000.
- Cartea, M. E.; Migdal, M.; Galle, A. M.; Pelletier, G.; Guerche, P. **Comparison of sense and antisense methodologies for modifying the fatty acid composition of *Arabidopsis thaliana* oilseed.** *Plant Science*, **136**:181-194, 1998.

- Cheng, J. I.; Boyd, C.; Slaymaker, D.; Okinada, Y.; Herman, E. M.; Keen, N. T. **Purification and characterization of a 34 kDa syringolide binding protein from soybean.** *Proc. Natl. Acad. Sci.*, **95**: 3306-3311, 1998.
- Deshimaru, M.; Yoshimi, S.; Shioi, S.; Terada, S. **Multigene family for Bowman-Birk type proteinase inhibitors of wild soja and soybeans: the presence of two BBI-A genes and pseudogenes.** *Biosci. Biotechnol. Biochem.*, **68**: 1279-1286, 2004.
- Friedman, M.; Brandon, D. L. **Nutricional and health benefits of soy proteins.** *J. Agric. Food Chem.*, **49**: 1069-1084, 2001.
- Gilissen, L. W. J.; Bolhaar, S. T. H. P.; Matos, C. I.; Rouwendal, G. J. A.; Boone, M. J.; Krens, F. A.; Zuidmeer, L.; Leeuwen, A.; Akkerdaas, J.; Hoffmann-Sommergruber, K.; Knulst, A.; Bosch, D.; Weg, E.; Ree, R. **Silencing the major apple allergen Mal d 1 by using the RNA interference approach.** *J. Allergy Clin. Immunol.*, **115**: 364-369, 2005.
- Hara-Nishimura, L. Y.; Shimada, N.; Hiraiwa, N.; Nishimura, M. **Vacuolar processing enzyme responsible for maturation of seed proteins.** *J. Plant Physiol.*, **145**: 632-640, 1995.
- Helm, R. M.; Cockrell, G.; Herman, E.; Burks, A. W. Sampson, H. A.; Bannon, G. A. **Cellular and molecular characterization of a major soybean allergen.** *Int. Arch. Allergy Immunol.*, **117**: 29-37, 1998.
- Herman, E. M.; Hel, R. M.; Jung, R.; Kinney, A. J. **Genetic modification removes an immunodominant allergen from soybean.** *Plant Physiol.*, **132**: 36-43, 2003.
- Herman, E. **Soybean allergenicity and suppression of the immunodominant allergen.** *Crop Sci.*, **45**: 462-467, 2005.
- Kalinski, A. J.; Melroy, D. L.; Dwivedi, R. S.; Herman, E. M. **A soybean vacuolar protein (P34) related to thiolproteases wich is synthesized as a glycoprotein precursor during seed maturation.** *H. Biol. Chem.*, **265**: 12068-12076, 1992.

- Kalinski, A. J.; Weisemann, J.; Matthews, B. F.; Herman, E. M. **Molecular cloning of a protein associated with soybean iol bodies wich is homologous to thiol proteases of the papain family.** *J. Biol. Chem.*, **265**: 13843-113848, 1990.
- Kennedy, A. R. **The Bowman-Birk inhibitor form soybeans as an anticarcinogen agent.** *Am. J. Clin. Nutr.*, **68**:1406-1412, 1998.
- Korsinczky, M. L. J.; Schirra, H. J.; Rosengren, K. J.; West, J.; Condie, B. A.; Otvos, L.; Anderson, M. A.; Craik, D. J. **Solution structures By ¹H NMR of the novel cyclic trypsin inhibitor SFTI-1 from sunflower seeds and an acyclic permutant.** *J. Mol. Biol.*, **311**: 579-591, 2001.
- Liu, Q.; Singh, S.; Green, A. **High-oleic and high-stearic cottonseed oils: Nutricionally improved cooking oils developed using gene silencing.** *J. Am. Coll. Nutr.*, **21**:S205-S211, 2002.
- Sambrook, J.; Fritsch, E. F.; Maniatis, T. **Molecular cloning, a Laboratory Manual.** 2nd ed. USA: Cold Spring Harbor Laboratory, 1989
- Mayer, A. M.; Shain, Y. **Control of seed germination.** *Annu. Rev. Plant Physiol.*, **25**: 167-193, 1974.
- Miles, G. P.; Samuel, M. A.; Zhang, Y.; Ellis, B. E. **RNA interference-based (RNAi) suppression of AtMPK6, an *Arabidopsis* mitogen-activated protein kinase, results in hypersensitivity to ozone and misregulation of AtMPK3.** *Environm. Pollut.*, **138**: 230-237, 2005.
- Mello, M. O.; Tanaka, A. S.; Silva-Filho, M. **Molecular evolution of Bowman-Birk type proteinase inhibitors in flowering plants.** *Mol. Phylog. Evol.*, **27**: 103-112, 2003.
- Odani, S.; Ikenaka, T. **The amino acid sequence of two soybean double headed proteinase inhibitors and evolutionary consideration on the legume proteinase inhibitor.** *J. Biochem.*, **80**: 1-9, 1976.

- Odani, S.; Ikenaka, T. **Studies on soybean trypsin inhibitors. X. Isolation and partial characterization of four soybean double-headed proteinase inhibitors.** *J. Biochem.*, **82**: 1513-1522, 1977.
- Odani, S.; Ikenaka, T. **Studies on soybean trypsin inhibitors. XII. Linear sequence of two soybean double-headed proteinase inhibitors.** *J. Biochem.*, **83**: 737-745, 1978.
- Ogawa, T.; Bando, N.; Tsuji, H.; Nishikawa, K.; Kitamura, K. **Alpha-subunit of beta-conglycinin, an allergenic protein recognized by IgE antibodies of soybean-sensitive patients with atopic dermatitis.** *Biosci. Biotechnol. Biochem.*, **95**: 831-833, 1995.
- Ogawa, T.; Samoto, M.; Takahashi, K. **Soybean allergens and hypoallergenic soybean products.** *J. Nutr. Sci. Vitaminol.*, **46**: 271-279, 2000.
- Ogita, S.; Uefuji, H.; Yamaguchi, Y.; Koizumi, N.; Sano, H. **Producing decaffeinated coffee plants.** *Nature*, **423**:823, 2003.
- Prakash, B.; Selvaraj, S.; Murthy, M. R. N.; Sreerama, Y.N.; Rao, D. R.; Gowda, L. R. **Analysis of the amino acid sequence of plant Bowman-Birk inhibitors.** *J. Mol. Evol.*, **42**: 560-569, 1996.
- Qu, L.; Chen, J.; Lui, M.; Pan, N.; Okamoto, H.; Lin, Z.; Li, C.; Li, D.; Wang, L.; Zhu, G.; Zhao, X.; Gu, H.; Chen, Z. **Molecular cloning and functional analysis of a novel type of Bowman-Birk inhibitor gene family in rice.** *Plant Physiol.*, **133**: 560-570, 2003.
- Ryan, C. A. **Protease inhibitors in plants: genes for improving defenses against insects and pathogens.** *Annu. Rev. Phytopathol.*, **28**: 425-449, 1990.
- Ryan, C. A. **Proteolytic enzymes and their inhibitors in plants.** *Annu. Rev. Plant Physiol.*, **24**: 173-196, 1973.
- Sanger, F.; Nicklen, S.; Coulson, A. R. **DNA Sequence with chain-termination inhibitors.** *Proc. Natl. Acad. Sci. USA*, **74**:5463-5467, 1977.

- Segal, G.; Song, R.; Messing, J. **A new opaque variant of maize by a single dominant RNA-interference-inducing transgene.** *Genetics*, **165**:387-397, 2003.
- Singh, S.; Green, A.; Stoutjesdijk, P.; Liu, Q. **Inverted-repeat DNA: a new gene-silencing tool for seed lipid modification.** *Biochem. Soc. Transactions*, **28**: 925-927, 2000.
- Tsuji, H.; Okada, N.; Yamanishi, R.; Bando, N.; Kimoto, M.; Ogawa, T. **Measurement of Gly m Bd 30 K, a major soybean allergen, in soybean products by a sandwich enzyme-linked immunosorbent assay.** *Biosci. Biotechnol. Biochem.*, **59**: 150-151, 1995.
- Wadsworth, G. J.; Redinhbaugh, M. G.; Scandalios, J. G. **A procedure for the small-scale isolation of plant RNA suitable for RNA blot analysis.** *Anal. Biochem.*, **172**: 277-283, 1988.
- Wesley, S. V.; Helliwell, C. A.; Smith, N. A.; Wang, M.; Touse, D. T.; Liu, Q.; Gooding, P. S.; Singh, S. P.; Abbott, D.; Stoutjesdijk, P. A.; Robinson, S. P.; Gleave, A. P.; Green, A. G.; Waterhouse, P. M. **Construct design for efficient, effective and high-throughput gene silencing in plants.** *Plant J.* **27**:581-590, 2001.
- Wolf, W. J. **Soybean proteins: their functional, chemical, and physical properties.** *J. Agri. Food Chem.*, **18**: 969-976, 1970.