

UNIVERSIDADE ESTADUAL DE CAMPINAS - UNICAMP
INSTITUTO DE BIOLOGIA
DEPARTAMENTO DE GENÉTICA E EVOLUÇÃO

BC/20773
IB/81385

**CLONAGEM E CARACTERIZAÇÃO DOS GENES QUE
CODIFICAM AS γ -PROLAMINAS DE COIX E SORGO**

Este exemplar corresponde à redação final
da tese defendida pelo (a) candidato (a)
Fernando Augusto
de Freitas
e aprovada pela Comissão Julgadora. *10/12/93*

Adilson Leite

FERNANDO AUGUSTO DE FREITAS
ORIENTADOR: Dr. ADILSON LEITE

Tese apresentada ao Instituto de Biologia da
Universidade Estadual de Campinas (UNICAMP) para
obtenção do grau de Mestre em Ciências Biológicas,
área de concentração Genética de Plantas.

1993

UNIDADE	IB		
N.º CHAMADA:	T		
	F884c		
V.	Ex.		
TOMBO BC/	20773		
PROC.	28.6/94		
C	<input type="checkbox"/>	D	<input checked="" type="checkbox"/>
PREÇO	CR\$ 800,00		
DATA	05/02/94		
N.º CPD			

CM00092597-7

Dedico e ofereço esta tese
à querida Cristina, pelo apoio constante, e aos meus
pais, Elza e Sebastião, que muito colaboraram
durante todos esses anos.

AGRADECIMENTOS

Ao Dr. Adilson Leite, pela orientação durante o desenvolvimento deste trabalho e contribuição para a minha formação científica.

Ao Prof. Dr. Paulo Arruda, pela oportunidade de trabalho, concedida ainda durante minha graduação, o que foi de grande importância para a minha formação acadêmica e escolha profissional.

À Dra. Laura M. M. Ottoboni e aos Profs. Dr. Louis Bernard Klaczko e Dra. Maricilda Palandi de Mello, pela participação no exame de qualificação e importantes sugestões apresentadas.

Às Profas. Dra. Yoko Bomura Rosato, Dra. Laudénir M. Prioli e Dra. Maricilda Palandi de Mello, pela valiosa contribuição prestada com a análise prévia do trabalho.

Ao Dr. Pedro Puigdomènech, por ceder o cDNA da γ -zeína de 28 kDa.

Ao Dr. Claude Créatin do Laboratoire Physiologie Végétale Moléculaire, Université Paris XI, por ceder o banco genômico de sorgo, bem como as sementes de *Sorghum vulgare* c.v. INRA 450.

Ao Prof. Dr. José Camillo Novello do Departamento de Bioquímica (UNICAMP), pela inestimável ajuda prestada com o sequenciamento da região N-terminal da γ -kafirina.

À Dra. Yoko Bomura Rosato, pelas inestimáveis ajudas prestadas.

À Dra. Laura M. M. Ottoboni, pela prontidão, sempre que preciso, e pela amizade.

Ao amigo Márcio, pela amizade e valiosa ajuda na confecção das fotos deste trabalho.

Aos queridos amigos Andrés e Sílvia, pela leitura minuciosa deste trabalho, pela grande amizade e agradável convivência nesses anos todos.

Aos excelentes amigos cantores Gonçalo e Chico, por tornarem o ambiente de trabalho ainda melhor.

Aos grandes amigos Germano, Edson (grandão), Marymar, Maria Elena,

Edson (baiano), Teresa, Emi, Patrícia, Luzão, Dante, Sibov, Felipe, André, Marília e Agostinho, pelo ótimo e agradável convívio.

Aos amigos Ricardão, David, Isa (Maria Luísa P. N. Targon), Anete e Bia, pela amizade e companheirismo.

À querida amiga Geórgia, pela grande e inestimável amizade.

Aos amigos Edmilson, Regina, Susi, Sandra e à grande Lyrian, pela amizade e alegre convivência durante todo o tempo que nos conhecemos.

Ao Maurício, pela grande ajuda no preparo da vidraria e pela disposição em cooperar em todos os momentos.

À amiga Berenice, pela prontidão na confecção dos géis de seqüenciamento e, principalmente, pela amizade.

Aos amigos da secretaria, Sandra, Tânia e Marciano, pela amizade e solicitude com que sempre me atenderam.

A todos aqueles que, direta ou indiretamente, colaboraram na execução deste trabalho.

À CAPES pelo apoio financeiro.

E à querida Cristina, por sua amizade, compreensão, apoio e grande incentivo em todos os momentos da realização desta tese.

Antes que seja tarde

*Com força e com vontade
A felicidade há de se espalhar
Com toda intensidade*

*Há de molhar o seco
De enxugar os olhos
De iluminar os becos
Antes que seja tarde*

*Há de assaltar os bares
E retomar as ruas
E visitar os lares
Antes que seja tarde*

*Há de rasgar as trevas
E abençoar o dia
E de guardar as pedras
Antes que seja tarde*

*Com força e com vontade
A felicidade há de se espalhar
Com toda intensidade*

*Há de deixar sementes
No mais bendito fruto
Na terra e no ventre
Antes que seja tarde*

*Há de fazer alarde
E libertar os sonhos
Da nossa mocidade
Antes que seja tarde*

*Há de mudar os homens
Antes que a chama apague
Antes que a fé se acabe
Antes que seja tarde*

*Com força e com vontade
A felicidade há de se espalhar
Com toda intensidade*

(Ivan Lins - Vitor Martins)

SUMÁRIO

As γ -prolaminas de milho (zeína), *Coix* (coixina) e sorgo (kafirina) foram isoladas a partir de sementes maduras e analisadas por SDS-PAGE e "Western blot" utilizando anticorpos produzidos contra γ -coixina e γ -zeína de 28 kDa. As γ -prolaminas de *Coix* e sorgo apresentaram peso molecular aparente igual a 22 kDa e 27 kDa, respectivamente, quando submetidas à eletroforese em SDS-PAGE. A existência de alto grau de homologia entre essas proteínas, foi verificada através do reconhecimento das mesmas pelos dois anticorpos nos experimentos de "Western blot".

Após os estudos de caracterização das proteínas, procedeu-se o isolamento de clones recombinantes de γ -coixina e γ -kafirina a partir de bibliotecas de cDNA de *Coix* e genômica de sorgo, respectivamente. O isolamento e sequenciamento do clone recombinantes de cDNA de γ -coixina, clone pBCX22.5, mostra uma região 5' não traduzida constituída por 27 nucleotídeos, seguido por uma fase de leitura constituída por 600 nucleotídeos, enquanto que a região 3' não traduzida é composta por 207 nucleotídeos. O clone genômico de γ -kafirina teve 2938 pb seqüenciada, compreendendo 1513 pb da região 5' não traduzida, 636 da região estrutural e 789 da região 3' não traduzida. Através do sequenciamento da região N-terminal da proteína madura de γ -kafirina, foi possível demonstrar que os primeiros 19 aminoácidos correspondem ao peptídeo sinal, o qual possui alto grau de homologia com os 19 primeiros aminoácidos da γ -coixina e γ -zeínas.

A região N-terminal da γ -kafirina e γ -coixina possui as mesmas seqüências repetidas que estão presentes na γ -zeína de 28 kDa. A primeira região, na γ -zeína de 28 kDa, consiste de 8 repetições do hexapeptídeo PPPVHL, com excessão da sétima repetição, onde o aminoácido Leu é trocado por Val. Na γ -kafirina este hexapeptídeo aparece quatro vezes, sendo que a quarta repetição possui a troca do aminoácido Leu para Val. A γ -coixina, por sua vez, apresenta três repetições completas e uma incompleta do hexapeptídeo. Porém, com relação a γ -zeína de 16 kDa, verifica-se que esta possui apenas três repetições modificadas do hexapeptídeo PPPVHL. A segunda seqüência repetida presente na região codificadora da γ -zeína de 28 kDa, é composta por dois octapeptídeos, QPHPCPCQ e QPHSPCQ. A γ -kafirina possui somente um octapeptídeo, PPHSPCH, enquanto a γ -coixina possui duas versões desse octapeptídeo,

QPHPCPYQ e PQHPSPFQ. Ao contrário, verifica-se que a γ -zeína de 16 kDa possui somente uma versão modificada desta seqüência.

A comparação das seqüências de nucleotídeos das regiões 5' das γ -prolaminas de milho e sorgo mostra a existência de aproximadamente 80% de homologia. Duas seqüências equivalentes ao "endosperm box" foram encontradas no promotor da γ -kafirina. A primeira cópia é encontrada na posição -340 com relação ao ATG inicial, sendo que as regiões flanqueadoras 5' e 3' desta seqüência são aproximadamente iguais às correspondentes seqüências da γ -zeína de 28 kDa. A diferença consiste em uma transição presente na região 3' (G para A). A segunda cópia encontra-se localizada na posição -620.

A seqüência GCN4, envolvida com a ligação de proteína regulatória GCN4 em levedura, também foi encontrada nos promotores de γ -prolaminas e estão distribuídas em duas regiões. A cópia proximal em relação ao ATG inicial, localizada na posição -192 do promotor de γ -kafirina, encontra-se 100% conservada nos promotores de γ -kafirina e γ -zeína de 16 kDa, sendo que o promotor de γ -zeína de 28 kDa apresenta uma única substituição. A segunda cópia é localizada na posição mais distal (-476) do promotor de γ -kafirina.

As propriedades regulatórias do promotor de γ -kafirina foram testadas *in vivo* em experimentos de expressão transitória através do bombardeamento de endosperma de milho, *Coix* e sorgo, com micropartículas co-precipitadas com DNA correspondente à construção pPKGUS (contendo o gene indicador GUS sob o controle do promotor completo da γ -kafirina). A análise histoquímica de GUS mostrou que o promotor da γ -kafirina é ativo nestes tecidos, ao contrário do que ocorreu quando uma construção contendo este promotor em orientação invertida foi utilizado, visto que neste caso, não ocorreu atividade de GUS. Estes dados, somado ao fato da polinização em *Coix* e sorgo não serem controladas, levou-nos a utilizar o milho para a continuidade dos trabalhos.

A atividade relativa correspondente ao bombardeamento da construção pPKGUS em endosperma, foi aproximadamente cinco vezes maior do que a obtida para a construção pPK285GUS (contendo o gene indicador GUS sob o controle do promotor da γ -kafirina deletado), indicando a existência de seqüências importantes para a otimização do processo de transcrição na região acima de -285. Da mesma forma, observou-se uma queda de 60 % na atividade relativa do clone pPK285GUS quando bombardeado em coleóptilo, comparando-se com a construção pPKGUS no mesmo tecido.

Para verificar se a queda da atividade relativa correspondente ao bombardeamento do clone pPK285GUS, em diferentes tecidos, quando comparada com o clone pPKGUS, era causada pela ausência de seqüências regulatórias presentes na região compreendida entre -285 e -1192 do promotor da γ -kafirina, mais duas construções foram feitas. A região compreendida entre -285 e -1192 do promotor da γ -kafirina, foi adicionada aos 90 primeiros nucleotídeos do promotor CaMV35S digirindo a expressão de GUS (pPK90GUS). Endosperma bombardeado com esta construção apresentou um aumento da atividade de GUS quando comparado com a construção pRT90GUS (contendo os 90 primeiros nucleotídeos do promotor CaMV35S fundido ao gene *gus*). Esse aumento de expressão também foi observado quando coleóptilo foi bombardeado com esses clones.

Com a finalidade de identificar regiões conservadas envolvidas no processamento da região 3' dos mRNAs de γ -prolaminas, foram alinhadas as seqüências flanqueadoras 3' dos genes de γ -prolaminas de milho, *Coix* e sorgo. A partir deste alinhamento foram reconhecidas 5 regiões contendo possíveis sinais de poliadenilação que foram denominadas de SP-1, SP-2, SP-3, SP-4, SP-5. O sinal de poliadenilação de mamíferos AATAAA está presente apenas na região SP-2 da seqüência flanqueadora do gene de γ -coixina, sendo que os outros sinais de poliadenilação encontrados constituem versões modificadas desta seqüência.

ABSTRACT

The γ -prolamins of maize (zein), *Coix* (coixin), and Sorghum (kafirin) were isolated from mature seeds and analyzed by SDS-PAGE and Western blot using polyclonal antibodies raised against γ -coixin and 28 kDa γ -zein. The γ -prolamins of *Coix* and Sorghum presented in SDS-PAGE polypeptides with apparent molecular weight of 22 kDa and 27 kDa, respectively. The Western blot analysis showed a strong cross reaction between the polyclonal antibodies and the γ -prolamins.

A γ -coixin cDNA clone and a γ -kafirin genomic clone were isolated using a 28 kDa γ -zein cDNA clone (pME119) as probe. The nucleotide sequence of the γ -coixin cDNA clone (pBCX22.5) revealed a nucleotide 5' untranslated sequence followed by a 600 nucleotide open reading frame and a 3' noncoding sequence of 207 nucleotides. The γ -kafirin clone (10S) had 2938 bp completely sequenced. Its 5' and 3' flanking regions consist of 1513 bp and 789 bp, respectively. This clone has a 636 nucleotide open reading frame, coding for a 212 amino acid protein. The first 19 amino acids of γ -kafirin represent the signal peptide, which was determined by N-terminal sequencing. This signal peptide shares high homology with the first 19 amino acids of γ -coixin and γ -zeins.

The N-terminal region of γ -kafirin and γ -coixin has the same repetitive sequences that were found in the 28 kDa γ -zein. In the 28 kDa γ -zein, the first homologous region consists of 8 repetitions of the hexapeptide PPPVHL, with a variation in the 7th repeat where a Leu was replaced by a Val. Four complete repeats of the hexapeptide PPPVHL were found in γ -kafirin where the amino acid Leu was replaced by Val at the 4th repeat. The γ -coixin presents three complete and one incomplete repetitions of the hexapeptide. In the 16 kDa γ -zein, only three modified versions of this hexapeptide were found. The second common region was composed by two octapeptide repeats, QPHPCPCQ and QPHSPCQ, with a unique amino acid substitution in the 28 kDa γ -zein. Two octapeptide repeats, QPHPCPYQ and PQHSPFQ, were also found in γ -coixin. Only one octapeptide was found in γ -kafirin (PPHSPCH) and in the 16 kDa γ -zein. The 16 kDa γ -zein octapeptide was a modified version of the one found in 28 kDa γ -zein.

The γ -prolamin promoter regions of maize and sorghum share about 80% of homology. Two endosperm boxes were found in the γ -kafirin promoter. The first box was located at -340 bp from the translation initiation site (ATG), and its

5' and 3' flanking regions were identical to the correspondent sequences of the 28 kDa γ -zein promoter, exception is made for the G to A transition in the second nucleotide at the 3' region. The second putative endosperm box was located at -620, and its 5' and 3' flanking regions were different from that of the 28 kDa γ -zein promoter.

Besides the TATA box, CATC box and the endosperm box, another putative regulatory sequence, GCN4-like sequences, were found in the 5' flanking region of γ -prolamin genes. The proximal copy of this sequence was located at -192 in the γ -kafirin promoter and was completely conserved in γ -kafirin and 16 kDa γ -zein genes. The 28 kDa γ -zein gene presented one substitution. The second GCN4-like sequence was located 476 bp from the ATG. The 16 kDa γ -zein gene presented a substitution in this sequence.

In order to investigate γ -kafirin promoter, immature seeds from sorghum, *Coix* and maize were bombarded with micro projectiles carrying a construction containing the *gus* reporter gene fused to 1190bp of the γ -kafirin promoter (pPKGUS). The GUS histochemical analysis of the bombarded seeds showed that the γ -kafirin promoter was able to drive GUS expression in the seed endosperm. No expression was observed in seeds bombarded with a construction having the promoter in the inverted orientation. Based on this observation and on the fact that maize pollination occurs as a synchronized event further experiments using endosperm were performed only in maize seeds.

To address the issue of tissue specificity, coleoptile, embryo, leaf and maize endosperm were bombarded with pPKGUS and a deleted version of the γ -kafirin promoter (pPK285GUS). Maximum relative GUS activity was observed in the endosperm bombarded with pPKGUS. Bombardment with the deleted promoter resulted in a GUS activity reduction of approximately 80% in endosperm and 60% in coleoptile. No activity was detected in embryos bombarded with both constructions.

To verify whether the reduction of GUS activity observed for the deleted promoter was due to the absence of *cis*-acting elements, two additional constructions were made. In one of those constructions the γ -kafirin promoter region between -285 and -1192 was fused to the first 90 bp of the CaMV35S promoter to produce pPK90GUS. Endosperm bombarded with this construction presented higher GUS activity than the ones bombarded with pRT90GUS (construct containing the first 90 bp of CaMV35S fused to GUS). This

enhancement of expression was also observed in coleoptile bombarded with pPK90GUS.

Comparison among the non-translated 3' flanking regions of the γ -prolamins revealed high degree of similarity in the upstream but not in the downstream region relative to the poly-A site. 16 to 42 bp upstream to the poly-A site was found a very conserved region that included three different versions of the AATGAA motif, recently identified as the poly-A signal for the 28 kDa γ -zein. This proximal motif was found in all γ -prolamins, except for the 16 kDa γ -zein that had the modified version AATGGA. The distal copy (AATGAC) was conserved in all the prolamin sequences. Following that distal copy, an ATTGAA, a GATGAA, and the canonical sequence (AATGAA) were found for γ -kafirin, γ -coixin and γ -zeins, respectively. Upstream to this triplicated motif was found a 14 bp conserved sequence that includes the SR-1, which is essential for the mRNA 3' end processing of 28 kDa γ -zein.

ÍNDICE

SUMÁRIO	vi
ABSTRACT	ix
ÍNDICE	xii
ÍNDICE DE FIGURAS E TABELAS	xv
ABREVIACÕES	xvi
1. INTRODUÇÃO	1
1.1. Proteínas de reserva de cereais	2
1.1.1. Prolaminas de milho	4
1.1.2. Prolaminas de <i>Coix</i>	4
1.1.3. Prolaminas de sorgo	4
1.1.4. Características das prolaminas de milho, sorgo e <i>Coix</i>	5
1.1.4.1. α -prolaminas	5
1.1.4.2. β -prolaminas	6
1.1.4.3. γ -prolaminas	7
1.1.4.4. δ -prolaminas	9
1.1.5. Expressão tecido-específica das prolaminas	9
2. OBJETIVOS	11
3. MATERIAIS E MÉTODOS	12
3.1. Material Vegetal	12
3.2. Caracterização das γ -prolaminas de milho, sorgo e <i>Coix</i>	12
3.2.1. Extração das γ -prolaminas	12
3.2.2. Determinação do teor de proteína	13
3.2.2.1. Digestão da matéria orgânica	13
3.2.2.2. Determinação colorimétrica do teor de nitrogênio	13
3.2.3. Análise das γ -prolaminas por SDS-PAGE	14
3.2.3.1. Aplicação das amostras e corrida do gel	14
3.2.4. Coloração do gel de poliacrilamida com prata	14
3.2.5. Western blot	15
3.3. Linhagens de bactérias	16
3.3.1. XL1-Blue	16
3.3.2. DH5 α	16
3.3.3. DL538	16
3.4. Biblioteca de cDNA de <i>Coix</i>	17
3.4.1. Seleção de clones de γ -coixina	17
3.4.2. Excisão dos clones recombinantes de γ -coixina	17
3.5. Biblioteca genômica de sorgo	18
3.5.1. Purificação de DNA do bacteriófago λ	18

3.5.2. Mapeamento dos clones recombinantes de γ -kafirina.....	20
3.5.2.1. Cinética de digestão parcial.....	20
3.5.2.2. Marcação da extremidade 5' das sondas de oligonucleotídeos.....	21
3.5.2.3. Preparo do marcador de peso molecular.....	21
3.5.2.4. Digestão do material a ser mapeado.....	22
3.6. Características dos plasmídios utilizados.....	22
3.7. Isolamento de DNA nuclear de planta.....	22
3.8. Isolamento de DNA de plasmídio em larga escala.....	24
3.9. Purificação de fragmentos de DNA a partir de géis de agarose.....	25
3.10. Experimentos de reconstrução dos genes de γ -prolaminas de <i>Coix</i> e sorgo.....	26
3.11. Seqüenciamento de DNA.....	27
3.12. Construção dos plasmídios utilizados nos experimentos de expressão transitória.....	28
3.13. Experimentos de expressão transitória.....	29
3.14. Métodos computacionais.....	30
3.15. Soluções.....	30
3.16. Meios de cultura.....	32
4. RESULTADOS.....	33
4.1. Caracterização das γ -prolaminas.....	33
4.2. γ -coixina.....	33
4.2.1. Seleção de clones de cDNA de γ -coixina.....	33
4.2.2. Seqüenciamento dos clones de γ -coixina.....	35
4.3. γ -kafirina.....	37
4.3.1. Seleção de clones genômicos de γ -kafirina.....	37
4.3.2. Mapeamento do fago recombinante 10S.....	39
4.3.3. Seqüenciamento do clone 10S.....	39
4.4. Experimentos de reconstrução.....	45
4.5. Estudo comparativo das γ -prolaminas de <i>Coix</i> , sorgo e milho.....	45
4.5.1. Comparação das seqüências de aminoácidos das γ -prolaminas.....	45
4.5.2. Composição de aminoácidos.....	49
4.6. Comparação das seqüências flanqueadoras 5' dos genes de γ -prolaminas de milho e sorgo.....	50
4.7. Comparação das regiões flanqueadoras 3' dos genes de γ -prolaminas de milho, <i>Coix</i> e sorgo.....	51
4.8. Efeito do promotor da γ -kafirina na expressão do gene indicador <i>gus</i> em tecidos de milho.....	55
5. DISCUSSÃO.....	63
6. CONCLUSÕES.....	72
7. BIBLIOGRAFIA.....	75
APÊNDICE I.....	85
APÊNDICE II.....	87

APÊNDICE III	90
APÊNDICE IV	91
APÊNDICE V	96
APÊNDICE VI	97

ÍNDICE DE FIGURAS E TABELAS

Figura 1. Caracterização das γ -prolaminas de milho, <i>Coix</i> e sorgo por SDS-PAGE e Western blot	34
Figura 2. Mapa de restrição do cDNA de γ -coixina	36
Figura 3. Seqüência de nucleotídeos do cDNA de γ -coixina	38
Figura 4. Análise de restrição do clone genômico de γ -kafirina.....	41
Figura 5. Mapeamento do fago contendo o inserto de γ -kafirina	42
Figura 6. Mapa de restrição do clone genômico de γ -kafirina	43
Figura 7. Seqüência de nucleotídeos do clone genômico de γ -kafirina	46
Figura 8. Representação do clone genômico de γ -kafirina	47
Figura 9. Reconstrução dos genes de γ -coixina e γ -kafirina.....	48
Figura 10. Alinhamento das seqüências de aminoácidos das γ -prolaminas de milho, <i>Coix</i> e sorgo.....	52
Figura 11. Alinhamento das regiões 5' dos genes de γ -prolaminas de sorgo e milho	54
Figura 12. Alinhamento das regiões 3' dos genes de γ -prolaminas de sorgo, milho e <i>Coix</i>	56
Figura 13. Esquema das construções utilizadas nos experimentos de bombardeamento.....	57
Figura 14. Bombardeamento da construção pPKGUS em endosperma de sorgo, milho e <i>Coix</i>	60
Figura 15. Gráficos de atividade relativa resultantes dos experimentos de bombardeamento em diferentes tecidos de milho	61
Figura 16. Endosperma imaturo de milho bombardeado com as construções pPKGUS, pRT103GUS e pPK285GUS.....	62
Figura 17. Mapa do vetor λ -ZAPII	88
Figura 18. Mapa do bacteriófago EMBL3A	89
Figura 19. Mapa de restrição do vetor pUC19	92
Figura 20. Mapa de restrição do vetor pBluescript KS+/-	94
Figura 21. Mapa de restrição do vetor de expressão pRT103.....	95
Figura 22. Esquema do aparelho utilizado nos experimentos de bombardeamento com micropartículas.....	97
Tabela I. Relação dos subclones utilizados no seqüenciamento total do clone genômico de γ -kafirina	44
Tabela II. Comparação de aminoácidos das γ -prolaminas de sorgo, <i>Coix</i> e milho	53

ABREVIACOES

DAP	dias aps a polinizao
EDTA	cido etilenodiaminotetractico
IPTG	isopropil-tio- β -D-galactosdeo
2-ME	2-mercaptoetanol
PEG	polietileno glicol
pb	pares de base
rpm	rotaes por minuto
SDS	dodecil sulfato de sdio
SDS-PAGE	eletroforese em gel de poliacrilamida contendo SDS
Tris	tris-(hidroximetil)-aminometano
X-gal	5-bromo-4-cloro-3-indolil- β -D-galactopiranosdeo
X-gluc	cido 5-bromo-4-cloro-3-indolil- β -D-glucurnico
EMBL	European Molecular Biology Laboratories
DDBJ	DNA Data Bank of Japan
GenBank	Genetic Sequence Data Bank
$\times g$	fora centrfuga em relao  acelerao gravitacional
DO ₆₀₀	absorbncia em 600 nm
pfu	plaque forming unit
CaMV35S	transcrito 35S do vrus do mosaico da couve flor

1. INTRODUÇÃO

Várias espécies de cereais originaram-se por um processo de domesticação iniciado pelo homem há cerca de 10.000 anos, no início do desenvolvimento da agricultura. O processo de domesticação dos cereais, foi responsável pela modificação dos hábitos nômades dos povos primitivos, através da substituição gradativa da caça e coleta de alimentos pela agricultura.

Achados arqueológicos sugerem que gramíneas selvagens tiveram um importante papel na nutrição dos povos primitivos no início do desenvolvimento da agricultura. Isso foi possível porque os cereais são facilmente cultiváveis, seus grãos contém alta proporção de carboidratos, proteínas e gorduras, possuindo alto valor alimentar. Além disso, os grãos dos cereais são facilmente armazenados por longos períodos. Entre os cereais, alguns tiveram especial importância, tal como o milho, que foi a base da alimentação dos astecas, maias e incas, e o trigo, que foi largamente utilizado pelos egípcios e romanos. Entre os povos africanos, o sorgo foi amplamente consumido, bem como o trigo e arroz, que desempenharam papel importante entre os povos asiáticos (Small, 1984).

Atualmente, os cereais suplementam aproximadamente 70% das necessidades humanas de proteínas nos países subdesenvolvidos. Nas regiões tropicais e sub-tropicais, o milho (*Zea mays*) tem sido usado como o cereal responsável pelo fornecimento dessa proteína, sendo que em países como os EUA e outros países da Europa, o consumo deste ocorre indiretamente na forma de proteína animal (Tosello, 1987). Entre as espécies cultivadas nos EUA, o sorgo (*Sorghum bicolor*), cereal originário da savana africana, ocupa a sétima posição, sendo ainda muito utilizado como alimentação e forragem em regiões onde o clima é árido e quente para o plantio do milho. A importância do sorgo tem aumentado muito nos últimos anos com a produção de plantas híbridas (DeRose et al., 1989).

No sudeste asiático, o *Coix* (*Coix lacryma-jobi*), apesar de ser considerado um cereal de pouca importância econômica, foi usado durante muito tempo como alimento, ornamento e na produção de bebidas alcoólicas (Venkateswarlu & Chaganti, 1973; Arora, 1977; Diaz, 1985). Atualmente, o *Coix* é encontrado em todas as regiões tropicais. Dentre as variedades de *Coix*, a variedade Adlay foi introduzida no Brasil por volta de 1938 através de sementes trazidas das Filipinas, sendo submetido a processos de melhoramento genético com o objetivo de obter sementes de melhor qualidade e aumento de produção (Shaafhawsen, 1952).

O estudo dos genes que codificam proteínas de reserva, fornecem valiosas informações sobre os processos de expressão gênica, tornando possível a manipulação desses genes a nível molecular visando a obtenção de plantas com valor nutricional superior. Dessa maneira, informações filogenéticas que evidenciam as relações de parentesco existentes entre os vários cereais da tribo Andropogoneae, que inclui milho, sorgo e *Coix*, bem como informações a nível molecular que corroboram essas evidências (Leite et al., 1990, Ottoboni et al., 1993), levaram-nos a estudar os genes que codificam proteínas de reserva de sorgo e *Coix*. Estes estudos possibilitam a melhor compreensão dos mecanismos de regulação desses genes nas diferentes espécies da tribo Andropogoneae.

1.1. Proteínas de reserva de cereais

A fração protéica da semente contém uma enorme variedade de proteínas presentes em pequenas quantidades, sendo essenciais para a manutenção do metabolismo celular. Uma outra categoria é composta por um número relativamente pequeno de proteínas que são armazenadas em grandes quantidades, sendo que estas possuem a função de fornecer nitrogênio e aminoácidos à plântula em germinação, sendo hidrolizadas com rapidez (Higgins, 1984).

Landry & Moureaux (1980) propuseram uma classificação para as proteínas das sementes baseada na divisão proposta por Osborne & Mendel (1914), onde elas são agrupadas de acordo com sua solubilidade na extração seqüencial, que diferencia quatro classes:

- 1) albuminas: solúveis em água;
- 2) globulinas: solúveis em solução salina;
- 3) prolaminas: solúveis em solução alcoólica;
- 4) glutelinas: solúveis em soluções alcalinas e ácidas diluídas, e em soluções salinas ou alcoólicas que contenham um agente redutor.

As funções de cada classe parecem estar divididas como segue:

a) Função metabólica: o papel metabólico é a função principal de muitas albuminas (ex: amilases) e globulinas. Em alguns casos, uma enzima pode ocorrer em quantidades elevadas, constituindo uma fração significativa de nitrogênio reduzido presente na semente (Su & Preiss, 1978).

b) Função estrutural: muitas proteínas presentes no endosperma desempenham funções estruturais. Estas proteínas não estão bem caracterizadas e não se conhece nenhum componente que se encontre em quantidades significativas

na semente. Muitos peptídeos estruturais são extraídos na fração de glutelinas e contribuem para a multiplicidade de bandas separadas em SDS-PAGE (Wilson et al., 1981).

c) Função de armazenamento: certas proteínas de semente parecem ter como única função, a de constituir reservas de nitrogênio, carbono e enxofre durante o desenvolvimento da semente, fornecendo estes nutrientes ao embrião durante a germinação. Essas proteínas são encontradas em grandes quantidades, podendo ocupar até 10% do peso seco da semente em cereais, e 25% a 35% em leguminosas. As proteínas de reserva têm sido estudadas nos cereais e leguminosas importantes do ponto de vista agrícola. As proteínas de reserva dos cereais são ricas em aminoácidos como asparagina, glutamina, arginina e prolina, e deficientes em lisina, treonina e triptofano (Pernollet & Mossé, 1983). Esta característica faz com que os cereais de uma maneira geral, apresentem um baixo valor nutritivo.

Aproximadamente 50% das proteínas de semente dos cereais são prolaminas, que são assim denominadas pelo seu alto teor em prolina e glutamina (Osborne & Mendel, 1914), 40% são glutelinas, sendo que o arroz e a aveia são exceções e possuem 85% de glutelinas e 80% de globulinas, respectivamente (Larkins, 1981). Ao contrário, nas leguminosas e outras dicotiledôneas, as proteínas de reserva majoritárias são do tipo globulinas, ainda que contenham albuminas (Larkins, 1981). As prolaminas recebem nomes específicos que indicam a sua origem, assim sendo, em milho, *Coix* e sorgo são denominadas de zeínas, coixinas e kafirinas, respectivamente. A denominação de kafirina deve-se à variedade *Sorghum kafir* cultivada no sul da África (Harlan, 1992). O *Coix* (*Coix lacryma-jobi* var. *Adlay*) contém aproximadamente o dobro da proteína encontrada nas sementes de milho (Ottoboni et al., 1990a).

A extração diferencial das prolaminas pode ser feita com soluções alcoólicas de várias composições, separando-as em termos de suas propriedades de solubilidade (Esen, 1986). Desta forma, as prolaminas são coletadas em três frações com composição de polipeptídeo própria, sendo que as α -prolaminas são solúveis em isopropanol 40% a 95%, as β -prolaminas são solúveis em isopropanol 30% a 80% / 2-ME 1% e as γ -prolaminas são solúveis em isopropanol 0% a 90% / 2-ME 1%.

1.1.1. Prolaminas do milho

As prolaminas do milho (*Zea mays*), zeínas, são subdivididas em cinco classes de peso molecular aparente de 10, 14, 16, 19, 22 e 28 kDa, quando analisadas em SDS-PAGE. Essas proteínas são denominadas de α -, β -, γ - e δ -zeína, correspondendo aos polipeptídeos de 19 e 22 kDa, 14 kDa, 16 e 28 kDa e finalmente, pelo polipeptídeo de 10 kDa, respectivamente.

A composição de aminoácidos das diversas classes de zeínas é similar, no sentido de que todas são ricas em glutamina, leucina, alanina e prolina. As zeínas de 28 , 16 e 10 kDa são caracterizadas pelo alto teor de metionina e cisteína (Gianazza et al.,).

1.1.2. Prolaminas de Coix

A análise das coixinas em SDS-PAGE mostra cinco bandas de pesos moleculares diferentes, sendo denominadas C1 (27 kDa), C2 (25 kDa), C3 (22 kDa), C4 (17 kDa) e C5 (15 kDa) (Ottoboni et al., 1990b), tal que, por extração diferencial, elas podem ser classificadas em dois grupos principais: 1) As α -coixinas, solúveis em 40-95% de isopropanol, são constituídas por 4 classes de proteínas de diferentes pesos moleculares aparentes, C1, C2, C4 e C5 e 2) a γ -coixina, solúvel em 0-80% de isopropanol na presença de agentes redutores (2-ME, DTT), é constituída por uma única classe de proteína, C3, apresentando peso molecular aparente de 22 kDa. A análise seqüencial através de isoeletrofocalização, SDS-PAGE e "Western blot", revelou que C1 é composta, no mínimo, por dois polipeptídeos, C2 por cinco polipeptídeos, C3 e C4 por apenas um polipeptídeo, enquanto que C5 é constituída por 2 polipeptídeos (Leite et al., 1990).

1.1.3. Prolaminas de sorgo

Quando separadas por SDS-PAGE, uma mistura heterogênea de polipeptídeos é resolvida, sendo que as α -kafirinas são representadas pelos polipeptídeos de peso molecular aparente igual a 22, 21, 20 e 19 kDa, enquanto que as β -kafirinas são representadas pelos polipeptídeos de 16 e 14 kDa (Shull & Kirleis, 1989). Da mesma forma, um polipeptídeo com peso molecular aparente de 27 kDa foi denominado de γ -kafirina (Barros et al., 1991).

1.1.4. Características das prolaminas de milho, sorgo e *Coix*

1.1.4.1. α -prolaminas

As α -zeínas constituem a maior fração das proteínas de reserva do milho, compreendendo 80% da zeína total (Shewry & Tatham, 1990), enquanto que em *Coix*, as α -coixinas representam 85% da coixina total (Targon et al., 1992). Esta classe de proteínas é representada por uma família multigênica, contendo aproximadamente 100, 20 e 40 cópias dos genes em milho (Hagen & Rubenstein, 1991), sorgo (DeRose et al., 1989), e *Coix* (Ottoboni et al., 1993), respectivamente. Inicialmente, a homologia existente entre as proteínas dessa classe foi demonstrada através de experimentos de SDS-PAGE e "Western blot", ocorrendo reações cruzadas com anticorpos produzidos contra coixinas e zeínas e as proteínas de milho, *Coix* e sorgo (Shull & Kirleis, 1989; Leite et al., 1990). Essa homologia foi posteriormente confirmada através de comparações de seqüências de aminoácidos deduzidas de clones genômicos de α -kafirina de 22 kDa (DeRose et al., 1989), α -coixina de 25 kDa (Ottoboni et al., 1993), α -zeínas de 19 kDa (Pedersen et al., 1986) e 22 kDa (Wandelt & Feix, 1989), revelando seqüências muito conservadas. Esta classe de proteínas é caracterizada pela presença de 10 repetições de um domínio flanqueado por um número variável de resíduos de glutaminas, sendo que cada domínio corresponde a uma α -hélice. Baseado no alinhamento das repetições presentes nas α -prolaminas de milho, sorgo e *Coix*, e também no padrão de distribuição dos aminoácidos hidrofóbicos e hidrofílicos, Garrat et al. (1993) descreveram um modelo estrutural para esta classe de proteínas. Esse modelo sugere que cada par de α -hélices esteja arranjado de maneira antiparalela e que cada hélice, da molécula como um todo, esteja arranjada em uma rede hexagonal. A maioria das α -hélices mostram seis faces de resíduos polares e hidrofóbicos alternados, o que permite um denso empacotamento dessas proteínas, característica importante para o armazenamento destas no interior do corpúsculo protéico.

Experimentos de imunocitocalização têm permitido que as diferentes classes de prolaminas sejam localizadas no interior do corpúsculo protéico. No milho, foi demonstrado que a distribuição das α -zeínas ocorre, principalmente, na parte mais interna do corpúsculo protéico (Lending et al., 1988). Ao contrário, as α -coixinas apresentam-se distribuídas por todo o corpúsculo protéico (Targon et al., 1992), não sendo confinada preferencialmente em uma região do mesmo, como ocorre com α -zeína em milho .

Baseado em suas características de solubilidade, a coixina de 17 kDa foi previamente descrita como pertencendo à classe das α -coixinas (Leite et al., 1990). Entretanto, através da clonagem e seqüenciamento de cDNA, verificou-se grande homologia com a β -zeína de 14 kDa (Leite et al., 1992).

A α -zeína de 22 kDa tem recebido muita atenção, pois além de constituir o principal componente da fração total de zeínas, esta tem sua síntese reduzida pela mutação *opaco-2* (o2) (Kodrzycki et al., 1989). Esse mutante produz um endosperma com alto teor de lisina, mas apresenta-se com aspecto farináceo e macio, o que prejudica o seu armazenamento e comercialização. O alto teor de lisina é devido à redução da síntese da α -zeína de 22 kDa, o que provoca uma queda de 50-70% no conteúdo de zeína. A diminuição no acúmulo das zeínas é acompanhada por um aumento do conteúdo das frações albumina, globulina e glutelina (Burr & Burr, 1982). Recentemente, o gene o2 foi clonado e caracterizado (Motto et al., 1988; Schmidt et al., 1987), verificando-se que esta proteína pertence à classe de proteínas ativadoras de transcrição denominadas "leucine zipper (bZIP)". Aparentemente, a proteína *opaco-2* (O2) é hábil para reconhecer mais de uma seqüência no DNA. Foram feitos dois experimentos independentes, onde foi demonstrado que a proteína O2 reconhece a seqüência GATGApypuTG no promotor de uma albumina de 32 kDa, denominada B-32 (Lohmer et al., 1991), e da mesma forma, demonstrou-se que ocorre o reconhecimento da seqüência TCCACGTAGA presente no promotor da α -zeína de 22 kDa (Schmidt et al., 1992). Foi demonstrado também, que a proteína O2 de milho reconhece a seqüência GACATGTC, presente no promotor da α -coixina de 25 kDa (Yunes et al., 1993). Experimentos realizados em nosso laboratório, incluindo o isolamento de clones de cDNA e genômico de *Coix* apresentando grande similaridade com o gene o2 de milho (resultados ainda não publicados), sugerem um mecanismo de regulação conservado entre milho e *Coix* após a divergência das espécies. Da mesma forma, a estrutura conservada das α -prolaminas confirma esta hipótese.

1.1.4.2. β -prolaminas

A β -zeína compreende 10% a 15% da fração total de zeínas. Ao contrário do que ocorre com as α -zeínas, a β -zeína é representada somente por uma ou duas cópias desse gene no genoma haplóide do milho (Kiriwara et al., 1988). A homologia entre as β -prolaminas foi demonstrada através de experimentos de

SDS-PAGE e "Western blot", onde anticorpos contra β -zeína mostraram reação cruzada com as proteínas pertencentes à classe das β -kafirinas (Shull & Kirleys, 1989). A β -zeína encontra-se distribuída na periferia do corpúsculo protéico, aparentemente desempenhando funções estruturais. Esta idéia é corroborada pelo alto conteúdo de cisteína presente nesta proteína, visto que este aminoácido está envolvido com a formação de pontes dissulfeto (Lending et al., 1988).

A seqüência de aminoácidos deduzida de cDNAs de β -zeína (Pedersen et al., 1986; Marks et al., 1985) não mostra similaridade com seqüências de α -zeínas. A proteína β -zeína não contém seqüências de aminoácidos repetidas que possam caracterizar um domínio estrutural bem definido. Experimentos de dicroísmo circular e dispersão ótica rotatória, mostraram uma composição deficiente em α -hélices, tendo principalmente, estruturas β -pregueadas (33% a 55%, dependendo da metodologia empregada) (Shewry & Tatham, 1990).

Muito pouco é conhecido sobre a regulação da expressão das β -prolaminas. Porém, através de experimentos de retardamento em gel, foi identificado um fator protéico, presente somente em endosperma, que se liga ao promotor da β -zeína de 14 kDa. Uma análise posterior através de "footprinting" revelou que o sítio de reconhecimento dessa proteína (chamada Beta-1) está localizado próximo ao TATA box. O pico na produção da proteína Beta-1 ocorre em torno de 16 DAP e parece estar ligada à expressão tecido-específica do gene de β -zeína (So & Larkins, 1991).

Conforme descrito anteriormente, em *Coix* foi caracterizado uma proteína apresentando solubilidade característica de α -prolamina. Porém, a análise posterior de um clone de cDNA que codifica esta proteína, revelou a existência de grande homologia com a β -zeína (Leite et al., 1992).

1.1.4.3. γ -prolaminas

As γ -prolaminas possuem também um elevado teor de cisteína, motivo pelo qual esta classe de proteínas é solúvel somente na presença de agentes redutores. Em milho, as γ -prolaminas constituem 5% a 10% da fração total de zeínas, enquanto que em *Coix* estas representam aproximadamente 15% da fração total de coixinas. As γ -zeínas são representadas apenas por um ou dois genes (Gallardo et al., 1988, Kirihara et al., 1988), sendo que para o gene de γ -zeína de 28 kDa, foram descritos dois alelos resultantes de um processo de duplicação gênica (Das et al., 1991a; Das et al., 1991b).

Estudos de imunocitocalização têm demonstrado que as γ -zeínas acumulam-se, principalmente, na periferia dos corpúsculos protéicos do endosperma (Lending et al., 1988), aparentemente desempenhando uma função estrutural, já que apresentam alto teor de cisteína (Ludevid et al., 1984). O início do desenvolvimento do corpúsculo protéico é marcado pela deposição da β - e γ -zeínas, enquanto que somente no decorrer do desenvolvimento do mesmo, é que começa a deposição das α -zeínas no seu interior (Lending et al., 1988; Lending & Larkins, 1989).

A homologia existente entre a γ -coixina e γ -zeína foi também demonstrada através de experimentos "Western blot" e "Southern blot" (Leite et al., 1990). Posteriormente, os estudos estenderam-se às γ -prolaminas de sorgo (γ -kafirinas), demonstrando também a existência de homologia com esta proteína.

Após a clonagem e caracterização de clones genômico e de cDNA de γ -zeína de 28 kDa (Prat et al., 1985, 1987; Reina et al., 1990a; Wang & Esen, 1986), clone genômico de γ -zeína de 16 kDa (Reina et al., 1990b), c-DNA de γ -kafirina (Barros et al., 1991) e o clone de cDNA de γ -coixina (cujo seqüenciamento é mostrado neste trabalho e em Leite et al., 1991), verificou-se que essas proteínas possuem um domínio repetitivo formado pelo hexapeptídeo conservado PPPVHL, o qual pode ser responsável pela solubilidade dessas proteínas em soluções aquosas, na presença de agentes redutores de pontes dissulfeto, tais como 2-ME ou DTT. Através de dispersão óptica rotatória e dicroísmo circular, foi demonstrado que a γ -zeína de 28 kDa apresenta conteúdo de α -hélice variando entre 19% e 32% e de β -pregueada variando entre 11% e 34%, dependendo do procedimento utilizado (Shewry & Tatham, 1990).

Muito pouco é conhecido sobre a regulação das γ -prolaminas. Recentemente, foi descrito uma variedade de milho *o2* denominado de QPM (Quality Protein Maize), que possui uma modificação no teor de γ -zeína de 28 kDa (Geetha et al., 1991). Essa variedade apresenta uma mutação que caracteriza-se pela recuperação do aspecto vítreo do milho normal. Além de apresentar alto teor de lisina, com baixos níveis de α -zeína de 22 kDa, característica do mutante *o2*, esta variedade apresenta 2 a 3 vezes mais γ -zeína de 28 kDa do que o milho normal. Uma análise no conteúdo de mRNA de γ -zeína de 28 kDa, indica que o aumento no acúmulo desta proteína é proporcional ao aumento na taxa de transcrição, produzindo níveis elevados do correspondente mRNA. Embora o aumento no acúmulo dessa proteína possa estar relacionado com a dureza e com o aspecto vítreo da semente, isto não explica o alto teor em lisina, pois a γ -zeína é

pobre em lisina, indicando que um efeito secundário está ocorrendo, aumentando a síntese de alguma outra proteína rica em lisina (Geetha. et al., 1991)

1.1.4.4. δ -prolaminas

A δ -zeína constitui o menor componente da fração total das zeínas, representando menos que 5% desse conteúdo. Com base nas características de solubilidade da δ -zeína, similares às características das α -zeínas, esta foi inicialmente descrita como α -zeína (Esen, 1986). Entretanto, através do seqüenciamento de um clone de cDNA, foi possível a dedução da estrutura primária, verificando-se que esta proteína deveria constituir uma nova classe, pois não apresentava a homologia esperada com a α -zeína (Kirihara et al., 1988). Essa classe de proteína é representada por apenas uma ou duas cópias desse gene no genoma haplóide do milho. Anticorpos produzidos contra δ -zeína apresentaram acentuada reação cruzada com α - e β -zeínas (Kirihara et al. 1988).

Recentemente, foi descrito que os polipeptídeos de zeínas são degradados seqüencialmente da periferia para o interior do corpúsculo protéico durante a germinação da semente, o que pode refletir na localização de cada classe dentro dessas organelas. As γ -zeínas são degradadas antes de todas as outras classes, concordando com sua localização na periferia dos corpúsculos protéicos. A β -, α - e δ -zeínas são degradadas em seguida e de forma mais gradual. Este fato sugere que essas proteínas estejam mais protegidas ao ataque de enzimas proteolíticas (Esen & Stetler, 1992)

1.1.5. Expressão tecido-específica das prolaminas

As dificuldades na obtenção de plantas transgênicas de milho, têm levado muitos autores a estudar a regulação gênica e acúmulo de zeínas, utilizando experimentos de expressão transitória e/ou transformação estável de plantas dicotiledôneas (Matzke et al., 1984; Goldsbrough et al., 1986), alga unicelular *Acetabularia mediterranea* (Brown et al., 1986) e levedura (Langridge et al., 1984). Quando Petúnia foi transformada via *Agrobacterium tumefaciens* com um clone contendo o cDNA de α -zeína de 19 kDa, flanqueado pelas regiões 5' e 3' do gene da β -faseolina (proteína de reserva do feijão), verificou-se a síntese de mRNA e acúmulo da proteína de 19 kDa nas sementes, indicando que o mRNA sintetizado é corretamente processado (Williamson et al., 1988). Ainda nesse

sentido, genes codificando as zeínas de 14 kDa, 19 kDa e 22 kDa foram introduzidos em células provenientes de cultura de tecido de girassol, utilizando-se *Agrobacterium tumefaciens* como vetor. Neste caso, houve somente a detecção dos mRNAs correspondentes, indicando que neste sistema, o processo de tradução não ocorre com eficiência (Goldsbrough et al., 1986). Ao contrário, quando genes de zeínas de 19 kDa e 22 kDa foram introduzidos em células de *Acetabularia mediterranea* através de microinjeção, verificou-se a expressão dessas proteínas (Langridge et al., 1984). Protoplastos isolados de endosperma de milho com 10 DAP, têm sido utilizados como um sistema homólogo em experimentos de expressão transitória, demonstrando que o promotor da α -zeína de 19 kDa é ativo neste tecido (Schwall & Feix, 1988). Da mesma forma, esses mesmos autores demonstraram que os promotores das zeínas de 10 kDa e 28 kDa dirigem a expressão da enzima cloranfenicol acetil transferase (CAT) em protoplasto de endosperma de milho. Quando o promotor da α -zeína de 22 kDa foi colocado dirigindo a expressão do gene GUS, para obtenção de plantas transgênicas de tabaco via *Agrobacterium tumefaciens*, verificou-se atividade de GUS restrita ao endosperma das plantas transformadas, conservando sua tecido-especificidade (Scherthner et al., 1988).

2. OBJETIVOS

Ao contrário do que ocorre com a classe das α -prolaminas, estudos realizados com γ -prolaminas de milho revelaram que, apesar da elevada proporção desta proteínas no endosperma, as γ -zeínas são expressas por um número reduzido de genes. Esta constatação indica que os promotores das γ -zeínas talvez sejam mais ativos do que aqueles descritos para α -prolaminas, fato que viabiliza a construção de vetores, a serem utilizados na obtenção de plantas transgênicas, utilizando os promotores de γ -prolaminas dirigindo a expressão de genes de interesse especificamente em endosperma.

Dada a homologia existente entre as proteínas γ -coixina, γ -zeínas e γ -kafirina, demonstrada por experimentos prévios realizados em nosso laboratório (Leite et al., 1990), tornou-se conveniente o isolamento desses genes para a melhor caracterização dos mesmos. O principal objetivo da presente tese corresponde à caracterização dos genes de γ -prolaminas de *Coix* e sorgo, o que foi alcançado através das seguintes etapas:

- 1) isolamento de clones recombinantes de γ -coixina a partir de uma biblioteca de cDNA de *Coix*.
- 2) isolamento de clones recombinantes de γ -kafirina a partir de uma biblioteca genômica de sorgo.
- 3) seqüenciamento dos clones recombinantes das γ -prolaminas de *Coix* e sorgo.
- 4) estudo comparativo das regiões promotoras dos genes de γ -prolaminas de milho e sorgo.
- 5) estudo comparativo das regiões codificadoras das γ -prolaminas de milho, *Coix* e sorgo.
- 6) estudo comparativo das regiões flanqueadoras 3'das γ -prolaminas de milho, *Coix* e sorgo.
- 7) estudo da atividade do promotor do clone genômico de γ -kafirina, através de experimentos de bombardeamento com micropartículas com o objetivo de delimitar regiões importantes na atividade do mesmo.

3. MATERIAIS E MÉTODOS

3.1. Material Vegetal

Sementes de *C. lacryma-jobi* var. Adlay, *Sorghum vulgare* c.v. INRA 450 e *Zea mays* var. Maya foram obtidas do banco de germoplasma do Departamento de Genética e Evolução, Universidade Estadual de Campinas.

3.2. Caracterização das γ -prolaminas de milho, sorgo e *Coix*

3.2.1. Extração das γ -prolaminas

As γ -prolaminas de *Coix*, sorgo e milho foram extraídas a partir de 2 g de semente. O endosperma maduro foi triturado até a obtenção de um pó fino. Em seguida, foi feita uma extração de gordura, sendo adicionado 1 volume de acetona, agitando vigorosamente. Após a agitação, foi feita uma centrifugação a 3000 rpm em centrífuga de mesa (Beckman-GPR centrifuge, rotor GH-3.7) durante 5 minutos. Essa etapa foi repetida por mais 3 vezes. Após a extração, o excesso de acetona foi retirado por liofilização. Foi feita uma extração preliminar das proteínas solúveis em soluções salinas adicionando-se 5 volumes (v/p) de NaCl 0,5 M, mantendo sob agitação durante 2 horas. Após a extração, centrifugou-se por 15 minutos em uma centrífuga de mesa refrigerada a 3000 rpm (Beckman-GPR centrifuge, rotor GH-3.7), descartando-se o sobrenadante. Em seguida, o resíduo foi lavado duas vezes com água 5:1 (v/p). Após a lavagem, foi feita uma centrifugação de 15 minutos, descartando-se o sobrenadante. Para a extração da fração total das prolaminas do endosperma, adicionou-se uma solução de isopropanol 60% contendo 2-ME 2% na proporção 5:1 (v/p). Essa suspensão foi mantida sob agitação durante a noite. Nesta etapa, as prolaminas passaram para a fase líquida, sendo recuperadas no sobrenadante após centrifugação a 3000 rpm em centrífuga de mesa (Beckman-GPR centrifuge, rotor GH-3.7) por 15 minutos. O sobrenadante foi colocado em um novo tubo, ao qual foi adicionado 1 volume de água destilada e 0,02 volume de acetato de sódio 3 M (concentração final 30 mM). Após esta etapa, a solução foi incubada durante a noite a 4°C. As α -, β - e δ -prolaminas precipitaram, enquanto que as γ -prolaminas permaneceram na fase líquida. Para a separação das γ -prolaminas, a solução foi centrifugada por 30 minutos a 3000 rpm em centrífuga de mesa (Beckman-GPR centrifuge, rotor GH-3.7) a 4°C, e o sobrenadante contendo γ -prolaminas foi armazenado a 4°C.

3.2.2. Determinação do teor de proteína.

A determinação do teor de proteína nas frações contendo as prolaminas, foi feita de acordo com a metodologia descrita por Nkonge & Balance (1982), com algumas modificações.

3.2.2.1. Digestão da matéria orgânica

Aliquotas de 100 μ l das frações contendo as γ -prolaminas, foram colocadas em tubos de ensaio para digestão. Em seguida, foi adicionado 1,5 ml de H_2SO_4 concentrado e os tubos foram colocados em placas digestoras até obtenção de um extrato claro. Após o resfriamento, colocou-se 100 μ l de H_2O_2 30%. Em seguida, os tubos foram novamente colocados na placa digestora durante aproximadamente 30 minutos para total evaporação da H_2O_2 . Os extratos foram diluídos com água destilada a um volume de 20 ml.

3.2.2.2. Determinação colorimétrica do teor de nitrogênio

Preparou-se uma curva padrão a partir de uma solução estoque contendo 100 μ g/ml de nitrogênio, da seguinte maneira: 47,16 mg de sulfato de amônio, previamente seco por 3 horas a 100°C, foram dissolvidos com 1,5 ml de ácido sulfúrico e diluídos com água destilada para 100 ml. As soluções de trabalho, contendo 2, 4, 6, 8 e 10 μ g N/ml, foram preparadas diluindo-se a solução estoque com o branco (solução 7,5% de ácido sulfúrico)

Para a determinação do teor de nitrogênio nas amostras, 0,5 ml dos extratos obtidos como descrito em 3.3.2.1 foi colocado em tubos de ensaio contendo 1,5 ml de um tampão composto de uma mistura 1:1 da solução A (Na_2HPO_4 0,2 M; NaOH 0,2 M; $KNaC_4H_4O_2 \cdot 4H_2O$ 0,36 M) e da solução B (NaOH 2,5 M). Em seguida, adicionou-se 0,4 ml da solução C (salicilato de sódio 1,2 M; nitoprussiato de sódio 1 mM). As amostras foram homogeneizadas e incubadas a 25°C por 5 minutos para equilibrar a temperatura. Em seguida, foram adicionados 0,2 ml de hipoclorito de sódio 5,0% e as amostras foram incubadas a 25°C, durante 30 minutos. Após este período, foram adicionados 10 ml de água destilada, e foi feita a leitura da absorbância a 660 nm em espectrofotômetro. O mesmo procedimento foi feito com as soluções padrão de nitrogênio. O teor de proteína foi calculado multiplicando-se o teor de nitrogênio pelo fator 6,25 (Nkonge & Balance, 1982).

3.2.3. Análise das γ -prolaminas por eletroforese em gel de poliacrilamida na presença de SDS.

A metodologia utilizada na preparação dos géis de poliacrilamida foi a descrita por Laemmli (1970).

3.2.3.1. Aplicação das amostras e corrida do gel

Nos géis a serem corados com Comassie Brilliant Blue R 250, foram utilizados aproximadamente 5 μ g de proteína, enquanto que para os experimentos de "Western blot", utilizou-se aproximadamente 0,4 μ g de cada proteína.

O tampão de corrida consistiu de: Tris-HCl 25 mM; glicina pH 8,3 0,192 M; SDS 0,1%. Após a aplicação das amostras nos géis, a corrida eletroforética foi feita à temperatura ambiente a 100 volts, até que a frente de azul de bromofenol chegasse ao final do gel. Após o processo de eletroforese, os géis foram corados durante 2 horas em uma solução contendo 0,25% Comassie Brilliant Blue R 250 em metanol: água: ácido acético (5:5:1, v/v/v), e descorados em solução aquosa contendo 7,5 % (v/v) de metanol e 6,0% (v/v) de ácido acético, até o aparecimento das bandas. Os géis foram mantidos em uma solução ácido acético 5%.

3.2.4. Coloração do gel de poliacrilamida com prata

A metodologia de coloração de géis de poliacrilamida com prata, aumenta a sensibilidade na detecção das bandas. Esta metodologia foi utilizada para a verificação da pureza das γ -prolaminas. A coloração com prata foi efetuada segundo a metodologia descrita por Moreno et al. (1985), com algumas modificações. Após a coloração do gel com Comassie Brilliant Blue R 250 e a descoloração, este foi colocado sob agitação por três vezes, durante 10 minutos, em 400 ml ácido tricloro-acético 5%. O gel foi incubado duas vezes por 10 minutos em 400 ml de metanol 40% / ácido acético 10% , e duas vezes por 10 minutos em etanol 10% / ácido acético 5%. Em seguida, o gel foi mantido durante 10 minutos, sob agitação em 200 ml de dicromato de potássio 3,4 M / ácido nítrico 0,0032 M. A lavagem do gel foi feita por três vezes durante 10 minutos em 400 ml de água deionizada, para retirar o excesso de dicromato de potássio. A partir desta etapa, o recipiente contendo o gel foi protegido da luz, cobrindo-o com papel alumínio. O gel foi mantido sob agitação em 200 ml de nitrato de prata 0,012 M, durante 30 minutos. Para a visualização das bandas, a solução anterior foi

descartada e trocada por 200 ml de carbonato de sódio 0,28 M (contendo 0,5 ml de formaldeído por litro de solução), sendo que após 1 minuto, este volume foi trocado por mais 200 ml da mesma solução. O gel foi mantido nesta solução até que a intensidade das bandas fossem adequadas. Para interromper a reação, o gel foi mantido em 400 ml de ácido acético 5%, durante 5 minutos. Ao término do processo, o gel foi lavado em água deionizada.

3.2.5. Western blot

Após a eletroforese, o gel foi incubado 2 vezes durante 40 minutos à temperatura ambiente em um tampão de transferência composto de: Tris-HCl 48 mM; glicina 39 mM; SDS 0,04% (p/v); metanol 20% (v/v). Em seguida, as proteínas foram transferidas para membranas de nitrocelulose (Hybond-C extra, Amersham). A transferência foi feita no aparelho Multiphor II Novablot (LKB-Produkter, Bromma, Sweden), o qual é formado por duas placas de grafite conectadas a uma fonte de voltagem através de eletrodos. Para a transferência, montou-se o sistema na placa de grafite conectada ao eletrodo positivo da seguinte maneira (tomando-se o cuidado para retirar todas as bolhas existentes entre as folhas e o gel): a) três folhas de papel 3MM cortadas do tamanho exato do gel e umidecidas no tampão de transferência, b) a membrana de nitrocelulose umidecida no mesmo tampão, c) o gel sobre a membrana de nitrocelulose, d) três folhas de papel 3MM umidecidas sobre o gel e e) por cima desta montagem, colocou-se a segunda placa de grafite, conectando-a ao eletrodo negativo. A transferência das proteínas para as membranas foi feita a 0,8 mA/cm² durante 2 horas. Após a transferência, procedeu-se da seguinte maneira: a) incubou-se a membrana durante toda a noite a 4°C no tampão TBS (Tris-HCl 20 mM ; NaCl 137 mM, pH acertado para 7,6) contendo 5% de leite desengordurado Molico; 0,1% Tween 20 e 0,25% de azida sódica, b) em seguida, incubou-se as membranas durante, no mínimo, 2 horas à temperatura ambiente com agitação, em tampão TBS contendo 5% de leite Molico e o anticorpo (contra γ -zeína de 28 kDa ou γ -coixina) na diluição de 1:5000, c) após este período, a membrana foi lavada 2 vezes por 1 minuto e 3 vezes por 5 minutos à temperatura ambiente, em tampão TBS contendo 0,1% Tween 20 (TBS-T), d) depois de lavada, a membrana foi incubada durante 1 hora à temperatura ambiente, sob agitação, com o complexo proteína A-peroxidase diluído na proporção 1:5000 em tampão TBS, e) em seguida, a membrana foi lavada 2 vezes por 1 minuto e 3 vezes por 5 minutos à temperatura ambiente, com

o tampão TBS contendo 0,1% Tween 20 (TBS-T), f) a membrana foi incubada em uma solução de revelação contendo 100 ml de tampão acetato 50 mM pH 4,5; 50 mg de 3-amino-9-etilcarbazol; 3 ml de N,N-dimetilformamida; 100 µl de H₂O₂ 30%. Para preparar esta solução, os 50 mg de 3-amino-9-etilcarbazol foram dissolvidos em 3 ml de N,N-dimetilformamida. O tampão acetato e 100 µl de H₂O₂ 30% foram adicionados somente após a dissolução. Quando a coloração desejada foi obtida, a membrana foi colocada em água destilada para interromper a reação e depois deixada à temperatura ambiente para secar.

3.3. Linhagens de Bactérias

Em seguida são descritas as linhagens de bactérias (*Escherichia coli*) utilizadas durante a execução dos trabalhos. A descrição detalhada de cada linhagem é dada no Apêndice I.

3.3.1. XL1-Blue

Genótipo relevante: *supE44 endA1 thi-1 lac⁻Δ(lac-proAB) gyrA96 hsdR17 recA1 F' proAB⁺ lacI^s lacZΔM15 Tn10 (tet^r)* (Bullock et al., 1987)

Uso: Hospedeira do bacteriófago λ-ZAP durante a seleção de clones recombinantes do banco de cDNA de *Coix*.

3.3.2. DH5α

Genótipo relevante: *F⁻ supE44 Δ (lacZYA - argF)U169 (φ80 lacZΔM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1* (Hanahan, 1983).

Uso: Preparação de células competentes para os experimentos de transformação realizados durante as subclonagens.

3.3.3. DL538

Genótipo relevante: *hsdR mcrA mcrB recD sbcC* (Whittaker et al., 1988).

Uso: Hospedeira do bacteriófago λ EMBL4 durante a seleção de clones recombinantes da biblioteca genômica de sorgo.

3.4. Biblioteca de cDNA de *Coix*

A biblioteca de cDNA foi construída em λ -Zap (Stratagene), utilizando-se mRNA isolado de corpúsculos protéicos de sementes imaturas de *Coix* (Leite et al., 1992). A descrição detalhada deste vetor é dada no Apêndice II.

3.4.1. Seleção de clones de cDNA de γ -coixina.

Após a etapa de titulação (Sambrook et al., 1989), colocou-se em um tubo de ensaio, um volume da suspensão de fago contendo 1000 ou 10000 unidades formadoras de placa (pfu) para a placa de Petri de 8,5 cm ou 13,5 cm, respectivamente. Adicionou-se ao tubo, 0,1 ml da suspensão da bactéria XL1-Blue, mantendo sob leve agitação a 37°C durante 20 minutos. Em seguida, colocou-se 3,5 ml ou 6 ml de "top" agarose aquecida (60°C) na placa contendo meio LB, pequena ou grande, respectivamente, vertendo rapidamente sob cada placa. As placas foram incubadas a 37°C durante a noite. As placas foram então mantidas em geladeira por 30 min, e em seguida, colocou-se o filtro Hybond-N (Amersham) sobre a placa, fazendo-se 3 furos em posições assimétricas com uma agulha contendo tinta nanquim. O filtro foi deixado sobre a placa durante 1 minuto, sendo retirado, em seguida, com uma pinça e deixado por 1 minuto em solução de denaturação (NaOH 0,5 M / NaCl 1,5 M). Após a denaturação, o filtro foi deixado durante 5 minutos em solução de neutralização (NaOH 0,25 M / NaCl 1,5 M) sendo em seguida, lavado em SSC 2x por 15 minutos, secado à temperatura ambiente e posteriormente incubado por 2 horas na estufa a 80°C. Essas membranas foram hibridizadas com o inserto contendo o cDNA de γ -zeína de 28 kDa, clone pME119 (Prat et al., 1985), marcado radioativamente.

3.4.2. Excisão dos clones recombinantes de γ -coixina

A técnica de excisão *in vivo* tem como objetivo recuperar o inserto (cDNA) clonado no vetor λ -ZAP. A placa de lise de interesse foi transferida para Eppendorf contendo 0,5 ml de tampão SM e 20 μ l de clorofórmio. Esse Eppendorf foi agitado em vortex e incubado 1 ou 2 horas à temperatura ambiente. Em um tubo Falcon de 50 ml, foram colocados 200 μ l do sobrenadante do tubo Eppendorf para onde a placa de lise foi previamente transferida (volume contendo $> 1.10^5$ pfu), 200 μ l de bactéria XL1-Blue $DO_{600} = 1,0$ e mais 1 μ l de fago "helper" R408 (contendo $> 1.10^6$ pfu/ml). Essa mistura foi incubada por 15 minutos a 37°C. Após esse período, adicionou-se 5 ml de 2x YT, incubando-se sob agitação (300 rpm) a

37°C durante 3 horas. Em seguida, foi feita uma incubação a 70°C por 20 minutos, sendo posteriormente centrifugado por 5 minutos a 4000xg. O sobrenadante foi coletado em um tubo esterilizado. Este estoque continha o "phagemid" empacotado como fago filamentosos, podendo ser guardado a 4°C por 1 ou 2 meses. Para o plaqueamento do "phagemid", foram preparados dois tubos Eppendorf como segue: a) 10 µl do estoque de "phagemid" mais 200 µl de bactéria XL1-Blue $DO_{600} = 1,0$ e b) 20 µl de uma diluição 10^{-2} do estoque de "phagemid" mais 200 µl de bactéria XL1-Blue $DO_{600} = 1,0$. Cada tubo foi incubado a 37°C durante 15 minutos. Foram plaqueados 100 µl de cada cultura em placas de LB/ampicilina, as quais foram incubadas durante a noite a 42°C.

3.5. Biblioteca genômica de sorgo

A biblioteca genômica de sorgo foi construída em fago λ EMBL4 (Frischauf et al. 1983), a qual foi cedida pelo Dr. Claude Crétin do Laboratoire Physiologie Végétal Moléculaire, Université Paris XI, para a seleção de um clone genômico de γ -kafirina. O banco genômico foi obtido utilizando-se fragmentos de DNA genômico parcialmente digeridos com MboI, que foram posteriormente ligados em braços do bacteriófago λ EMBL4 preparados através de digestão com BamHI. Para o isolamento dos clones recombinantes, os fagos foram plaqueados em 30 placas de 135 mm x 17 mm. Após a transferência das placas de lise para membranas de nylon, as mesmas foram hibridizadas com o cDNA de γ -zeína de 28 kDa marcado radioativamente, clone pME119 (Prat et al., 1985). Os clones recombinantes de γ -kafirina foram isolados através de três seleções sucessivas, conforme descrito anteriormente. Maiores detalhes do vetor utilizado nesta etapa pode ser encontrado no Apêndice II.

3.5.1. Purificação de DNA do bacteriófago λ

A metodologia empregada foi a descrita por Sambrook et al. (1989), com algumas modificações.

Inoculou-se 15 ml de meio LB/MgSO₄ 10 mM com a bactéria *E. coli*, linhagem DL-538, incubando-se durante a noite a 37°C, sob agitação vigorosa. A DO_{600} do cultivo (diluição 1/10 em meio LB) foi medida para determinar o número de células ($DO_{600} = 1$ corresponde a 8×10^8 células/ml). Em um tubo de cultivo, misturou-se 1×10^{10} células com 5×10^7 fagos, incubando-se durante 25 minutos a 37°C, sob vigorosa agitação. O conteúdo deste tubo foi inoculado em um

Erlenmeyer contendo 500 ml de meio LB/MgSO₄ 10 mM, previamente aquecido a 37°C, e mantido sob vigorosa agitação nesta temperatura. A DO₆₀₀ do cultivo foi acompanhada a partir de 3 horas após a inoculação, com intervalos de 30 minutos, sendo que a ocorrência da lise foi indicada pela observação de uma queda brusca da absorbância. Após essa etapa, adicionou-se 10 ml de clorofórmio ao cultivo, incubando o mesmo a 37°C durante 10 minutos, sob vigorosa agitação. O cultivo foi deixado em gelo durante 15 a 20 minutos, sendo em seguida, transferido para garrafas de polipropileno de 250 ml e centrifugado a 10.000 rpm em rotor Beckman JA-14 por 10 minutos a 4°C. O sobrenadante foi transferido para um Erlenmeyer de 1 litro, ao qual adicionou-se 50 µl de RNase A (10 mg/ml) e 500 µl de DNase A (1 mg/ml). Após essa etapa, o lisado foi incubado entre 30 minutos e 1 hora em banho-maria a 37°C, sendo que após esse tempo, adicionou-se 27,2 g de NaCl, o qual foi dissolvido suavemente por agitação com barra magnética estéril. Após a dissolução do NaCl, o lisado foi mantido em gelo por, no mínimo, 1 hora, transferido para garrafas de polipropileno de 250 ml e centrifugado a 10.000 rpm (rotor Beckman JA-14) durante 10 minutos a 4°C. O sobrenadante foi transferido para um Erlenmeyer de 1 litro, ao qual foi adicionado 50 g de PEG-6000, agitando suavemente para dissolver. Essa solução foi deixada em gelo durante toda a noite. O lisado foi transferido para garrafas de polipropileno de 250 ml e centrifugado a 10.000 rpm (rotor Beckman JA-14), a 4°C durante 30 minutos. O sobrenadante foi descartado e os frascos foram deixados secando invertidos por 10 a 15 minutos à temperatura ambiente. O precipitado foi ressuspensionado em 4 ml de tampão SM. A suspensão foi transferida para tubos Falcon de 15 ml. Adicionou-se 4 ml de clorofórmio, seguida de uma rápida agitação. O lisado foi centrifugado em centrífuga de mesa por 12 minutos a 1500 rpm. A fase superior foi transferida para um tubo estéril e conservada a 4°C até o momento da ultracentrifugação. Para proceder a ultracentrifugação, adicionou-se 0,5 g de CsCl por ml de suspensão de bacteriófago, agitando-se suavemente a mistura. O material foi ultracentrifugado utilizando-se o rotor Beckman SW41 durante 4 horas a 4°C, a 22.000 rpm em um gradiente de CsCl (em SM) montado nas seguintes faixas de densidade: 1,45 g/ml (índice de refração 1,3768), 1,50 g/ml (i.r. 1,3815) e 1,7 g/ml (i.r. 1,3990); sendo que as duas primeiras fases continham um volume de 1,5 ml e a última, continha somente 1 ml. Após a centrifugação, a banda branca contendo o vírus foi retirada com o auxílio de um pipetador P1000. O bacteriófago foi dialisado durante 1 hora em 1,5 litro da seguinte solução: NaCl 10 mM, Tris-HCl 50 mM (pH 8,0) e MgCl₂ 10 mM.

A diálise foi repetida por mais 1 hora com uma nova solução. Após a diálise, o conteúdo da membrana foi transferido para tubos de microcentrífuga, sendo adicionado EDTA, a partir de uma solução estoque 0,5 M (pH 8,0), para uma concentração final de 20 mM. Foi adicionado Pronase-E para uma concentração final de 0,5 mg/ml e SDS, a partir de uma solução estoque de 10%, para uma concentração final de 0,5%. A solução foi misturada por inversão várias vezes e em seguida, incubada por 1 hora a 37°C. Após este período, a solução foi deixada resfriar à temperatura ambiente e em seguida, adicionou-se igual volume de fenol equilibrado com Tris-HCl 50 mM (pH 8,0). A solução foi misturada diversas vezes por inversão, até a formação de uma completa emulsão. A separação de fases foi feita por centrifugação, sendo que em seguida, a fase aquosa foi transferida para um outro tubo com o auxílio de uma pipeta. Foi realizada uma nova extração com igual volume de fenol equilibrado com clorofórmio/álcool isoamílico, centrifugando-se em seguida para a separação das fases. A fase aquosa foi transferida para outro tubo e uma nova extração foi feita. Após essa última extração, foi adicionado acetato de sódio a partir de uma solução estoque 3 M (pH 7,0), para uma concentração final de 0,3 M, misturando-se por inversão. Nesta etapa, 2 volumes de etanol gelado (-20°C) foram acrescentados, misturando-se por inversão. A suspensão foi deixada à temperatura ambiente por 30 minutos e, em seguida, centrifugada por 30 minutos a 4°C. O sobrenadante foi descartado. Adicionou-se 1 ml de etanol 70% (-20°C) e centrifugou-se novamente por 30 minutos a 4°C. O sobrenadante foi novamente descartado. O precipitado foi liofilizado por 4 minutos e ressuspendido em TE contendo RNase (40 µg/ml).

3.5.2. Mapeamento dos clones recombinantes de γ -kafirina

O mapeamento do clone genômico de γ -kafirina foi realizado utilizando-se o sistema λ Mapping system, RPN1721 (Amersham). Os conceitos relacionados com a metodologia são descritos no Apêndice III.

3.5.2.1. Cinética de digestão parcial

Foram feitas três diluições de enzima de restrição nas seguintes concentrações: 1, 0,1 e 0,01 unidades/µl, utilizando-se tampão de restrição 1x. Para cada diluição foi preparado, em gelo, a seguinte reação: 0,25 µg do DNA clonado em bacteriófago λ ; 1 µl da diluição da enzima de restrição; 1,5 µl de tampão de restrição 10x e água até 15 µl.

As reações foram incubadas em banho-maria a 20°C, transferindo, a cada 15 minutos, 5 µl da reação para um Eppendorf contendo 3 µl de EDTA 100 mM. Cada alíquota retirada foi aplicada em gel de agarose 0,5% após a adição de tampão de amostra. Uma digestão total foi utilizada como controle. Após a corrida, foi possível escolher o melhor tempo e concentração, observando as condições que produziram o maior número de bandas em quantidades equimolares.

3.5.2.2. Marcação da extremidade 5' das sondas de oligonucleotídeos

1 µl de cada oligonucleotídeo foi diluído com a adição de 49 µl de tampão TE para uma concentração final de 1 pmol/ml. Cada oligonucleotídeo foi marcado separadamente através da seguinte reação, em dois tubos diferentes: 4,5 µl de Oligon. diluído (ON-L ou ON-R) ; 3,0 µl de [γ -³²P]ATP, 10mCi/ml, > 5000Ci/mmol; 1,5 µl de Tampão Kinase (Tris-HCl 0,1 M; MgCl₂.6H₂O 10 mM; DTT 7 mM pH 8,0); 1,0 µl de Polinucleotídeo Kinase (3 unidades/µl) e 5 µl de água. A reação foi agitada suavemente, incubada em banho-maria a 37°C durante uma hora e em seguida, aquecida em banho-maria a 65°C por 3 minutos. As diluições foram feitas separadamente, retirando-se 2 µl de cada preparação de oligonucleotídeo e adicionando-os a uma mistura contendo 125 µl de tampão TE, 75 µl de tampão de amostra e 50 µl de cloreto de sódio 1 M. A sonda pode ser armazenada por, no máximo, 2 semanas a 20°C.

3.5.2.3. Preparo do marcador de peso molecular

Os marcadores de peso molecular correspondem a fragmentos de DNA de bacteriófago λ contendo as extremidades coesivas *cos* (R e L), de modo a formarem um gradiente de peso molecular quando submetidos à eletroforese.

Para a marcação da extremidade esquerda (*cos* L), a seguinte reação foi preparada: 1 µl do "size marker" do "kit" de mapeamento; 9 µl de H₂O e 5 µl do oligonucleotídeo L marcado. Da mesma forma, a marcação da extremidade direita (*cos* R) foi feita como segue: 2 µl do "size marker"; 8 µl de H₂O e 5 µl do oligonucleotídeo R marcado. Antes da aplicação no gel, as reações foram aquecidas a 70°C por 3 minutos e 42°C por 30 minutos, como para os produtos da digestão do material a ser mapeado (item 3.5.2.4).

3.5.2.4. Digestão do material a ser mapeado

0,25 µg de DNA (para cada tempo escolhido na digestão parcial) foi digerido nas condições determinadas através do experimento de cinética, interrompendo a reação com 9 µl de EDTA 100 mM. Foi retirado um volume de cada tempo de reação escolhido, para uma determinada enzima, e transferido para um novo Eppendorf, de modo a obter 27 µl no final. Dessa mistura final, 16 µl foram transferidos para um novo tubo, no qual ocorreria a reação com o oligonucleotídeo R e 8 µl foram transferidos para outro tubo, no qual ocorreria a reação do oligonucleotídeo L. Os volumes foram completados para 16 µl com água destilada. Para as reações R e L, foram adicionados 7,5 µl de oligonucleotídeos R e L marcados, respectivamente. Para as digestões completas, que serviram como controle, colocou-se a mesma quantidade de oligonucleotídeos R e L. As reações foram aquecidas em banho-maria a 70°C durante 3 minutos e em seguida, a 42°C durante 30 minutos. As reações foram submetidas à eletroforese em gel de agarose 0,5%, a 2 volts/cm durante 20 horas. Ao final da corrida, o gel foi secado a vácuo, sendo exposto a um filme de raio-X.

3.6. Características dos plasmídios utilizados

Nas etapas de subclonagem do cDNA de γ -coixina e clone genômico de γ -kafirina, foram utilizados os plasmídios pBluescript KS +/- (Short et al., 1988) ou pUC19 (Messing, 1983; Norrander et al., 1983; Yanisch-Perron et al., 1985), conforme as conveniências com relação ao posicionamento dos sítios enzimáticos na região de policlonagem dos mesmos. Para os experimentos de bombardeamento com micropartículas, o vetor de expressão pRT103 (Töpfer et al., 1987) foi utilizado como base na construção dos plasmídios bombardeados. A descrição detalhada dos plasmídios pUC19, pBluescript KS +/- e pRT103 é dada no Apêndice IV.

3.7. Isolamento de DNA nuclear de planta

Esta técnica pode ser empregada no isolamento de DNA a partir de folhas, raízes e plântulas (Rivin et al., 1982).

As folhas foram lavadas com água destilada e incubadas por 15 minutos em solução de hipoclorito de sódio 5%, lavadas novamente com água estéril e secas ao ar. No caso de plântulas obtidas em condições estéreis, não houve a necessidade da esterilização em hipoclorito de sódio. Para o início da extração, um

almofariz foi resfriado com N₂ líquido. Somente duas amostras foram extraídas por vez. Cerca de 1 a 2 g de amostra foram triturados com ajuda de pistilo até a obtenção de um pó fino, sem permitir o descongelamento do material. O pó foi transferido do almofariz para um frasco em gelo e após o descongelamento do material, adicionou-se 10 a 20 ml de tampão de homogeneização. A homogeneização foi feita duas vezes em homogeneizador do tipo "polytron" por 30s na maior velocidade. O homogenato foi filtrado em 3 camadas de gaze estéril em funil de Buchner com ajuda de bomba de vácuo. Neste ponto, foi feita a sedimentação dos núcleos através de centrifugação do filtrado a 350xg por 10 minutos a 4°C (2000 rpm em rotor Beckman JA-20). O sobrenadante foi descartado e o sedimento foi ressuscitado em tampão de lise gelado com a ajuda de um bastão de vidro. Adicionou-se Sarkosyl (a partir de uma solução estoque 20% (m/v)) suficiente para obter-se uma concentração final de 2%. A partir de então, o material foi incubando por 1 hora a 50°C. Após a incubação, foi adicionado 0,95 g de CsCl por ml do lisado e misturado gentilmente por inversão até completa dissolução do CsCl. As proteínas insolúveis e polissacarídeos foram removidos por centrifugação a 17.000xg (12.000 rpm em rotor Beckman JA-20), por 15 minutos a 4°C, sendo que o sobrenadante foi coletado. Foi adicionado brometo de etídio ao sobrenadante, até a concentração de 300 µg/ml. Para o ajuste da concentração de CsCl, o índice de refração da solução foi acertado até 1,3895. O material foi transferido para tubos de ultracentrífuga e centrifugado até o equilíbrio a 20°C (rotor VTi-80 Beckman 50.000 rpm por 20 h). A banda de DNA cromossômico foi removida cuidadosamente, sob exposição à luz UV e com o auxílio de uma seringa munida de agulha 30x12. O brometo de etídio foi removido por sucessivas extrações com n-butanol equilibrado com água. As extrações foram realizadas até o desaparecimento de cor na fase alcoólica. A solução contendo o DNA foi dialisada exaustivamente contra 1.000 volumes de TE, trocando-se o tampão de diálise 2 a 3 vezes. Após a diálise, foi adicionado 0,1 volume de acetato de sódio 3 M pH 5,2 e 2 volumes de etanol gelado. A precipitação foi feita por 2 horas a -20°C, sendo que, em seguida, o DNA foi coletado por centrifugação, seguida de lavagem do sedimento com etanol 70%. O DNA foi ressuscitado em TE e armazenado a 4°C. A concentração de DNA foi determinada espectrofotometricamente.

3.8. Isolamento de DNA de plasmídio em larga escala

A técnica utilizada é descrita por Sambrook et al. (1989), com algumas modificações.

Devido ao fato dos plasmídios pUC 19, pBluescript KS+/-, e pRT103, apresentarem origem de replicação derivada de ColE1, não foi necessário a etapa de amplificação, visto que esta origem de replicação apresenta um controle mais relaxado de replicação plasmidial, resultando em um alto número de cópias do plasmídio por célula.

A bactéria hospedeira (DH5 α) foi inoculada em 25 ml de meio líquido LB/ampicilina e crescida durante a noite. Esse cultivo foi utilizado para inocular 500 ml de LB/ampicilina líquido, o qual foi deixado sob vigorosa agitação durante a noite a 37°C. Em seguida, esse cultivo foi centrifugado a 7.000 rpm em rotor Beckman JA-14, por 10 minutos a 4°C e o sedimento foi ressuspensionado em 4 ml de solução I (Tris-HCl 25 mM pH 8,0; glicose 50 mM; EDTA 10 mM). A suspensão foi transferida para um tubo de centrifuga próprio para rotor Beckman JA-20. Foi adicionado 1 ml de solução de lisozima (25 mg/ml em solução I). A suspensão foi misturada cuidadosamente e mantida à temperatura ambiente por 10 minutos. Em seguida, foram adicionados 10 ml de solução II (NaOH 0,2 M; SDS 1%), recentemente preparada, misturando-se por repetidas inversões. A suspensão foi mantida em gelo por 10 minutos. A cada tubo, adicionou-se 7,5 ml de solução gelada de acetato de potássio 3 M pH 4,8, misturando-se por inversão até desaparecimento das diferentes fases líquidas. A suspensão foi mantida em gelo por 10 minutos. O precipitado floculoso formado é constituído de DNA cromossômico, RNA de alto peso molecular e complexos de potássio/SDS/proteína/complexos de membrana. A separação de fases foi feita por centrifugação a 12.000 rpm por 20 minutos a 4°C em rotor Beckman JA-20, sendo o sobrenadante coletado por filtração em gaze. Falha na formação do precipitado compacto, geralmente ocorre devido a uma inadequada mistura do lisado com a solução de acetato de potássio. O sobrenadante foi transferido para novo tubo de rotor Beckman JA-20. Foi adicionado 0,6 vol de isopropanol ao sobrenadante, o qual foi misturado várias vezes por inversão e mantido à temperatura ambiente por 10 minutos. Após esse tempo, foi feita uma centrifugação a 11.000 rpm em rotor Beckman JA-20 por 10 minutos à temperatura ambiente. O sedimento foi lavado com etanol 70% à temperatura ambiente, tendo sido levemente seco em liofilizador. Após a secagem, o sedimento foi ressuspensionado em 4 ml de TE e em

seguida, 4,4 g de CsCl foram adicionados, misturando-se até a completa dissolução. Foi adicionado também, 0,4 ml de solução de brometo de etídio (10 mg/ml). Foi feita uma nova centrifugação a 8.000 rpm, durante 5 minutos, à temperatura ambiente em rotor Beckman JA-20 para retirar complexos formados entre brometo de etídio e proteínas bacterianas. A fase líquida foi transferida para tubo Falcon, até que o acerto do índice de refração fosse feito ($\eta = 1,3860$). Uma vez acertado o índice de refração, a solução foi transferida para tubos de ultracentrifuga, próprios para rotor Beckman VTi-80. Para a formação do gradiente, a centrifugação foi feita a 60.000 rpm por 20 horas a 20°C. A retirada da banda correspondente ao DNA plasmidial e a remoção do brometo de etídio foi feita conforme já descrito no item 3.7.

3.9. Purificação de fragmentos de DNA a partir de géis de agarose.

Após a digestão, o DNA foi submetido à eletroforese em gel de agarose contendo 0,5 $\mu\text{g/ml}$ de brometo de etídio. A concentração do gel de agarose variou de acordo com os tamanhos dos fragmentos produzidos na digestão do DNA. Os insertos foram retirados dos géis utilizando-se papel de DEAE (DE-81, Whatman), de acordo com modificações da metodologia descrita por Dretzen et al. (1981) e Danner (1982). Para a retirada dos insertos, foi feito um corte no gel com o auxílio de um bisturi abaixo das bandas de interesse e, com a ajuda de pinças, o papel DE-81 de tamanho adequado foi inserido até o fundo da placa sob a qual o gel foi montado. O gel foi novamente submetido à eletroforese, que foi monitorada com luz UV e interrompida quando todo o DNA entrou no papel DE-81. A recuperação do DNA foi feita incubando-se o papel DE-81 em 800 ml de tampão de eluição (Tris-HCl 20 mM pH 8,0; EDTA 2 mM; NaCl 1,5 M) em Eppendorf durante 1 hora em banho-maria a 45°C. Após este período o papel DE-81 foi retirado com o auxílio de uma pinça, e o Eppendorf foi centrifugado por 30 segundos para a retirada dos fragmentos de papel remanescentes. Nos casos em que a banda a ser retirada migrou juntamente com o indicador azul de bromofenol, foi necessário a extração do mesmo com n-butanol. Para isso, foram feitas três extrações como segue: 1) n-butanol saturado com água destilada, 2) n-butanol e 3) n-butanol saturado com água destilada. Em seguida, adicionou-se 2,5 volumes de etanol gelado (-20°C), incubando-se a solução a -20°C durante a noite. Para fragmentos menores que 100 pb, além do etanol gelado, foi necessário adicionar MgCl_2 até a obtenção de uma solução de concentração final igual a 0,01 M.

3.10. Experimentos de reconstrução dos genes de γ -prolaminas de *Coix* e sorgo

Os experimentos de reconstrução foram conduzidos com a finalidade de estimar o número de cópias do gene de γ -coixina e γ -kafirina nos genomas correspondentes. Para isso, o DNA genômico de *Coix* foi digerido com HindIII, montando-se o seguinte sistema de reação: 10 μ g de DNA genômico; 5 μ l de espermidina 0,1 M; 20 μ l de tampão de restrição 10x; 50 unidades de HindIII e H₂O q.s.p. 200 μ l. A digestão foi feita durante a noite a 37°C, sendo interrompida por adição de 0,1 volume de EDTA 0,5 M pH 8,0 e 0,1 volume de acetato de sódio 3 M pH 5,2. Em seguida, adicionou-se 800 μ l de etanol gelado (-20°C), sendo que a precipitação foi feita a -70°C durante 1 hora. Após essa etapa, o DNA foi centrifugado por 30 minutos a 4°C, lavado com etanol 70% e centrifugado por mais 15 minutos. A secagem foi feita por liofilização. Ao final, o DNA foi ressuscitado em 25 μ l de água destilada, sendo que foi necessário manter em banho-maria a 37°C durante 2 horas.

2 μ g do DNA a ser utilizado como controle do número de cópias, clone pBCX22.5, foram linearizados com EcoRI, seguindo-se as condições de reação da enzima de restrição utilizada. O procedimento empregado para interromper a reação foi o mesmo seguido para o DNA genômico. O material digerido foi ressuscitado em 20 μ l de água destilada. Os cálculos envolvidos na determinação da quantidade de DNA a ser utilizada como controle do número de cópias do gene de γ -coixina por genoma haplóide do *Coix*, são descritos no Apêndice V.

O plasmídeo pBCX22.5 linearizado com EcoRI, em quantidades equivalentes a 3, 2 e 1 cópia do gene de γ -coixina no genoma haplóide de *Coix* e o DNA genômico de *Coix* digerido com HindIII, foram submetidos à eletroforese em gel de agarose 0,8% (25 cm de comprimento) a 25 volts. Em seguida, foi feita a transferência dos fragmentos de DNA para membrana de nylon, a qual foi hibridizada com a sonda obtida através de marcação radioativa do inserto EcoRI/XhoI originário do clone pBCX22.5.

Um experimento semelhante foi realizado com o objetivo de determinar o número de cópias do gene de γ -kafirina no genoma haplóide de sorgo, sendo que neste caso, foi utilizado como controle do número de cópias o plasmídeo pUCTP (Tabela I), o qual contém uma parte da região codificadora do gene de γ -kafirina. O DNA genômico de sorgo foi digerido com HindIII.

As intensidades das bandas nos filmes autoradiográficos foram medidas

através da leitura em um densitômetro a laser (Ultrosan XL-LKB).

3.11. Seqüenciamento de DNA

Os experimentos de seqüenciamento foram realizados com DNA dupla fita de plasmídios recombinantes de pUC 19 ou pBluescript KS +/- . Para isso, 8 µg de DNA plasmidial preparado através de minipreparação e dissolvidos em 40 µl de TE, foram denaturados adicionando-se 40 µl da solução NaOH 0,4 M / EDTA 0,4 mM, e incubados por 5 minutos à temperatura ambiente. A denaturação foi interrompida pela adição simultânea de 8µl de acetato de amônio 2 M pH 4,5 e 176 µl de etanol gelado. As amostras foram mantidas à -20°C por 1 hora e centrifugadas em microcentrífuga por 20 minutos. O sedimento foi lavado com etanol 70% e centrifugado novamente por 15 minutos. Após a secagem do sedimento, o mesmo foi redissolvido em 20 µl de água, sendo que foram utilizados 10 µl para cada reação de seqüenciamento iniciada com "primer" direto (universal) ou reverso (Boehringer Mannheim). As reações de seqüenciamento foram feitas seguindo-se as instruções fornecidas pelo fabricante do "kit" T7 Sequencing Kit (Pharmacia), utilizando-se (³⁵S)dATPαS (Amersham). As eletroforeses dos produtos de reação foram realizadas carregando-se 1,5-2,0 µl/canaleta em dois géis denaturantes de poliacrilamida nas concentrações de 5% e 6% (Sambrook et al., 1989). Para a confecção dos géis, foram utilizadas placas de 50 cm de altura (LKB), sendo uma delas tratada com Repel-silane (Pharmacia) e a outra com Bind-silane (Pharmacia), de acordo com as recomendações do fabricante. As amostras foram incubadas por 2 minutos a 75-80°C antes de serem aplicadas nos géis. Os dois géis foram submetidos a aproximadamente 2000 volts, sendo que a potência máxima foi estabelecida para 50 W para o conjunto dos géis, os quais foram montados no sistema de eletroforese Macrophor (LKB). O gel para leitura da seqüência localizada próxima ao "primer", gel de 6%, foi submetido à eletroforese até a saída do corante azul de bromofenol, enquanto que o gel para leitura da seqüência distal, gel de 5%, até 2 horas e 30 minutos após a saída do corante xileno cianol (2ª banda). Após a eletroforese, os géis aderidos às placas foram mantidos por 30 minutos sob leve agitação em uma solução contendo 10% de ácido acético e 10% de metanol, e finalmente lavados em água corrente. Os géis foram secados a 45°C durante a noite, e submetidos à autoradiografia através de exposição de filmes sensíveis à radiação (Hyperfilm - Amersham) por 1 a 4 dias.

3.12. Construção dos plasmídios utilizados nos experimentos de expressão transitória

Todas as posições indicadas na construção dos clones abaixo são com relação ao ATG inicial. A descrição detalhada dos plasmídios pBluescript e pRT103 utilizados nas construções é dada nos Apêndices VIII e IX, respectivamente. A Figura 13 apresenta o mapa dos plasmídios construídos.

pPKGUS - A região 5' flanqueadora do gene de γ -kafirina foi isolada do clone genômico 10S. O fragmento TaqI compreendendo a região de -1198 a -8, foi clonado no sítio ClaI do plasmídio pBluescript KS+ (Stratagene), resultando o clone pKTT1 (Tabela I). A região promotora da γ -kafirina, foi retirada do clone pKTT1 através de digestão do mesmo com SmaI/SalI e posteriormente inserida no plasmídio pRT103 digerido com HincII/XhoI. Este clone foi designado pPK103. Em seguida, a região codificadora do gene da β -glucuronidase (GUS), excisado do plasmídio pJII140 (Töpfer et al., 1987) por digestão com NcoI/BamHI, foi fundida ao pPK103 no correspondente sítio, entre o promotor da γ -kafirina e a região 3' do CaMV35S.

pPK285GUS - O fragmento -285/-8 do promotor da γ -kafirina foi excisado do clone pKTT1 por digestão com as enzimas EcoRV/SalI. Este inserto foi ligado ao plasmídio pRT103 digerido com as enzimas HincII/XhoI. Este clone foi designado pPK285. Em seguida, a região codificadora do gene da β -glucuronidase foi obtida e clonada no sítio de NcoI/BamHI do clone pPK285, produzindo o clone pPK285GUS, da mesma forma como foi descrito para o clone pPKGUS.

pRT103GUS - Este clone foi obtido pela inserção da região codificadora do gene da β -glucuronidase nos sítios das enzimas NcoI/BamHI do plasmídio pRT103.

pRT90GUS - O plasmídio pRT103 foi digerido com as enzimas EcoRV/HincII e recircularizado, resultando o clone pRT90. O pRT90 foi então digerido com NcoI/BamHI e ligado ao inserto correspondente à região codificadora do gene *gus* da mesma forma como foi descrito para o clone pPKGUS.

pPK90GUS - O fragmento -1190/-285 do promotor da γ -kafirina foi excisado do clone pKTT1 através da digestão com EcoRV. Este fragmento foi ligado ao plasmídio pRT103 previamente digerido com HincII/EcoRV, resultando o clone pPK90. Em seguida a região codificadora do gene *gus* foi inserida

conforme descrito para as outras construções.

3.13. Experimentos de expressão transitória por bombardeamento com micropartículas

Os experimentos de expressão transitória foram realizados através do bombardeamento dos tecidos com um sistema de pressão, produzida por injeção de gás hélio, construído no Laboratório de Biologia Molecular de Plantas do CBMEG por Márcio José da Silva. O esquema, bem como o princípio do sistema utilizado, encontra-se descrito no Apêndice VI.

Os diferentes tecidos foram colocados juntos em placas de Petri de 100 mm contendo 20 ml de meio MS (Murashige & Skoog, 1962), sendo previamente tratados como segue:

-endosperma imaturo de milho: As plantas de milho foram crescidas no campo experimental da Universidade Estadual de Campinas e auto polinizadas. As espigas foram colhidas com 16 dias após a polinização (DAP), esterilizadas por 15 minutos com água sanitária comercial 5% e lavadas 4 vezes com água destilada esterilizada. As sementes sofreram um corte longitudinal frontal, com a ajuda de um bisturi e foram distribuídas pela placa de Petri (135 mm x 17 mm) com as faces voltadas para cima, utilizando-se 8 endospermas para cada placa. Todas as placas receberam sementes de uma única espiga.

-embrião: Os embriões foram coletados de sementes imaturas de milho da mesma espiga utilizada para a coleta dos endospermas, sendo que 8 embriões foram utilizados para cada placa.

-coleótilo: Os coleótilos foram dissecados de sementes colocadas para germinação durante 3 dias, esterelizados por 5 minutos com água sanitária comercial 5% e lavados 4 vezes em água destilada esterilizada. Foram utilizados 8 coleótilos por placa de Petri.

-folha: As folhas foram dissecadas de sementes que foram germinadas durante 7 dias e receberam o mesmo tratamento com água sanitária comercial que o coleótilo. Foram utilizados 8 discos foliares em cada placa.

Para a preparação das micropartículas, utilizou-se 5 µg de DNA purificado com CsCl (1 µg/µl), 50 µl de CaCl₂ (2,5 M), 20 µl de espermidina (0,1 M) e 3 mg de partículas de ouro de acordo com a metodologia descrita por Ye et al. (1990).

Para cada construção, foram bombardeadas 4 placas, sendo que cada

placa foi bombardeada duas vezes com 0,5 µg de DNA. Após o bombardeamento, as placas foram incubadas durante 48 horas na ausência de luz a 27°C. Em seguida, os tecidos foram revelados para atividade de GUS pela adição do substrato da enzima, X-gluc, de acordo com a metodologia descrita por Jefferson (1987).

Os resultados obtidos com a contagem do número de pontos azuis foram tratados estatisticamente pelo método SNK (Steel & Torrie, 1980), e expressos sob a forma de atividade relativa, sendo que o cálculo foi efetuado através da razão dos resultados obtidos para o bombardeamento das construções descritas na Fig.13, nos diferentes tecidos, e aqueles obtidos para o bombardeamento do clone pRT103GUS nos respectivos tecidos.

A metodologia básica de rotina, como a) Minipreparação de plasmídios, b) Digestões enzimáticas, c) Clonagem em plasmídios, d) Southern blot, e) Transformação de *E. coli*, f) Hibridização de sondas radioativas com ácidos nucléicos em membranas de nylon e g) Marcação de sondas de DNA, entre outras, foram utilizadas de acordo com manuais de técnicas de Biologia Molecular como Sambrook et al., 1989.

3.14. Métodos computacionais

Todos "softwares" utilizados foram executados em microcomputador do tipo PC 386.

Os pesos moleculares de proteínas analisadas por SDS-PAGE, bem como os tamanhos dos fragmentos de DNA, foram estimados utilizando-se o programa Eletroforese (Adilson Leite) através de regressão não linear (Método Marquardt) conforme descrito por Plikaytis et al. (1986).

A edição e análise das seqüências de DNA foram realizadas com ajuda dos programas DNASIS (Pharmacia) e ESEE v. 1.04 (Cabot, 1988). No alinhamento das seqüências de DNA e proteínas, foi empregado o programa Clustal V (Higgins & Sharp, 1989).

3.15. Soluções

a) SM

NaCl	1,16 g
MgSO ₄ .6H ₂ O	0,457 g

Tris-HCl 1 M pH 7,5 10,0 ml
 Gelatina 20 mg
 H₂O q.s.p. 200 ml
 Esterilizar por autoclavagem

b) SSC 20x

NaCl 175,3 g
 Citrato de sódio di-hidratado 88,2 g
 H₂O q.s.p. 800 ml
 Ajustar para pH 7,0 com NaOH 10 M
 H₂O q.s.p. 1000 ml
 Esterilizar por autoclavagem.

c) TBE 5x

Ácido bórico 30,9 g
 Tris-Base 121,1 g
 EDTA dissódico 3,73 g
 H₂O q.s.p. 600 ml
 Após a completa dissolução, completar o volume para
 1000 ml.

d) TE

Tris-HCl pH 8,5 10 mM
 EDTA pH 8,0 1 mM

e) Tris-HCl 1 M

Tris-Base 121,1 g
 H₂O q.s.p. 800 ml
 Ajustar o pH para o valor desejado adicionando-se HCl
 concentrado:

pH	HCl
7,4	~70 ml
8,0	~42 ml

3.16. Meios de cultura

a) LB (Meio Luria-Bertani)

Bacto-tryptone10 g
Bacto-yeast extract5,0 g
NaCl10 g
H₂O q.s.p.800 ml
Ajustar o pH para 7,0 com NaOH 2 M
H₂O q.s.p.1000 ml
Esterilizar por autoclavagem durante 20 minutos.

b) Top agarose

agarose1,4 g
MgSO₄.6H₂O0,457 g
H₂O q.s.p.200 ml
Esterilizar por autoclavagem

c) YT 2x

NaCl10 g
Bacto-yeast extract10 g
Bacto-tryptone16 g
Titular para pH 7.5 com NaOH 2 M
H₂O q.s.p.1000 ml

4. RESULTADOS

4.1. Caracterização das γ -prolaminas

A metodologia de extração diferencial, utilizada para a extração das γ -prolaminas, não possibilita que as γ -zeínas de 16 kDa e 28 kDa sejam extraídas na mesma fração. Ao contrário, a γ -zeína de 16 kDa é extraída juntamente com a β -zeína (Esen, 1986). O gel apresentado na Fig. 1A, mostra a análise da γ -zeína de 28 kDa, γ -kafirina e γ -coixina (22 kDa) por SDS/PAGE 12,5%. Após a eletroforese, o gel foi corado com prata. Como pode ser visto na Fig. 1A, as γ -prolaminas de milho, sorgo e *Coix*, aparecem como uma única banda, o que demonstra que as mesmas apresentam elevado grau de pureza. O peso molecular da γ -kafirina foi estimado a partir da análise em SDS/PAGE como sendo aproximadamente 27 kDa. Posteriormente, as γ -prolaminas foram caracterizadas imunologicamente por "Western-blot" (Fig. 1B e 1C), utilizando-se anticorpos policlonais contra γ -coixina e γ -zeína (28 kDa) obtidos em coelhos (Leite et al., 1990). Na Fig. 1B, encontram-se os resultados obtidos utilizando-se anticorpo específico contra γ -coixina, sendo que as três γ -prolaminas foram reconhecidas com a mesma intensidade, indicando portanto a existência afinidades similares pelo anticorpo. Da mesma forma, quando utiliza-se anticorpo contra γ -zeína de 28 kDa, verifica-se o mesmo padrão de reconhecimento obtido no caso anterior. A ocorrência de reações cruzadas observadas nos experimentos de "Western bot", demonstram o alto grau de homologia existente entre as γ -prolaminas.

4.2. γ -COIXINA

4.2.1. Seleção de clones de cDNA de γ -coixina

Para a identificação de clones de cDNA de γ -coixina, foram analisadas 35000 placas de lise de uma biblioteca de cDNA de *Coix*. A sonda utilizada na hibridização foi o fragmento de PstI do clone pME119, que corresponde a um cDNA de γ -zeína de 28 kDa (Prat et al., 1985). Esse fragmento possui 349 pb da região estrutural. Após dupla purificação do fragmento em gel de agarose, o inserto foi marcado por incorporação de nucleotídeos radioativos, através de extensão catalisada por fragmento Klenow da DNA Polimerase I e "primers" aleatórios,

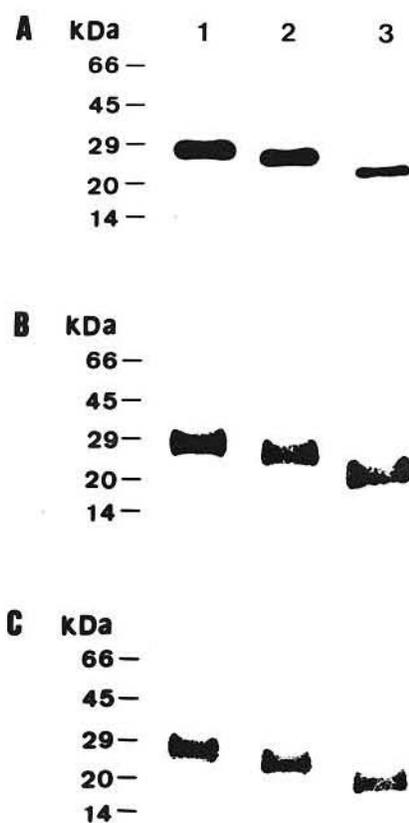


Fig. 1. Caracterização das γ -prolaminas extraídas de sementes maduras de milho, sorgo e *Coix*. (A) γ -prolaminas analisadas através de SDS-PAGE 12,5%, seguida de coloração com 0,25% Comassie Brilliant Blue R 250 em metanol/água/ácido acético (5:5:1 v/v/v) e posteriormente com prata. (B) e (C) Análise das γ -prolaminas através de Western blot, utilizando anticorpos produzidos contra γ -coixina e γ -zeína (28kDa), respectivamente. As linhas 1, 2 e 3 representam γ -zeína, γ -kafirina e γ -coixina, respectivamente.

utilizando-se o "kit" Megaprime DNA labelling system (Amersham; Feinberg & Vogelstein, 1983, 1984). Os clones positivos foram isolados através de sucessivas hibridizações, permitindo o isolamento de 11 fagos recombinantes, que foram posteriormente utilizados na excisão dos respectivos clones em pBluescript. Após a excisão dos plasmídios, foram preparadas amostras de DNAs para serem analisadas através de digestão com enzimas de restrição (EcoRI/XhoI, enzimas que liberam insertos clonados no sistema do λ -Zap). Em seguida, essas amostras foram submetidas à eletroforese em gel de agarose, e posteriormente analisadas através de "Southern-blot", utilizando-se a sonda de cDNA de γ -zeína de 28 kDa. Insertos de dois clones apresentaram intensa hibridização: clones pBCX22.4 e pBCX22.5. Dentre os vários clones analisados foram escolhidos três que apresentavam hibridização mais intensa, pBCX22.4, pBCX22.5 e pBCX22.9, para posterior sequenciamento.

4.2.2. Seqüenciamento dos clones de γ -coixina

Para a construção de um mapa de restrição do clone pBCX22.5, este foi digerido com EcoRI, XhoI e PstI, e após eletroforese em gel de agarose, os fragmentos foram transferidos para membrana de nylon (Hybond-N, Amersham), sendo que, após a transferência, a membrana foi hibridizada com a sonda de cDNA de γ -zeína de 28 kDa. Com os resultados obtidos, foi possível deduzir um mapa de restrição. A existência de um sítio de PstI na posição 483 permitiu a construção de dois subclones: pBCX22.5A e pBCX22.5B (Fig. 2). A seqüência completa do clone pBCX22.5, mostrada na Fig. 3, é composta por 840 nucleotídeos. Esta seqüência foi obtida a partir do seqüenciamento nas duas direções do clone original, bem como dos subclones obtidos a partir da digestão com PstI, pBCX22.5A e pBCX22.5B. A seqüência de nucleotídeos mostra uma região 5' não traduzida constituída por 27 nucleotídeos, sendo os primeiros 15 nucleotídeos correspondentes ao adaptador de EcoRI utilizado durante a construção do banco de cDNA. 12 nucleotídeos abaixo da seqüência do adaptador encontra-se uma fase de leitura constituída por 600 nucleotídeos, e uma região 3' não traduzida composta por 207 nucleotídeos incluindo a seqüência do "linker" de XhoI utilizado também na clonagem unidirecional em λ -ZAP. A seqüência de nucleotídeos do pBCX22.5 foi depositada nos bancos de dados EMBL, GenBank e DDBJ com o número de acesso X59850.

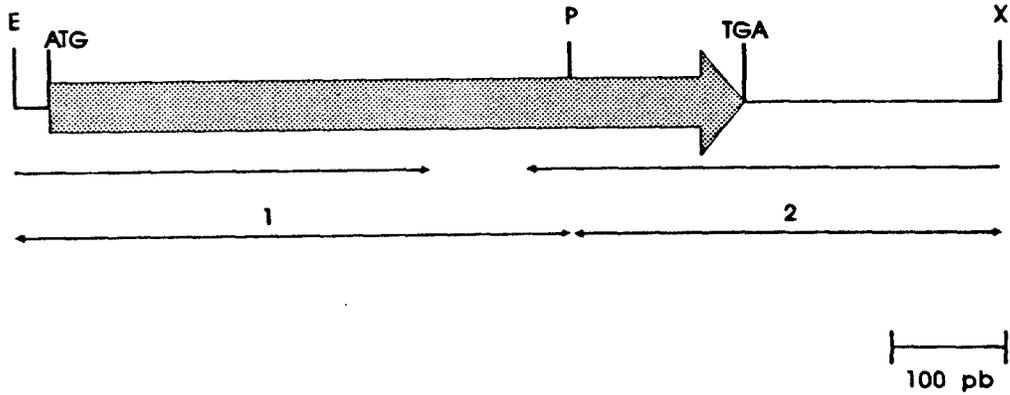


Fig. 2. Mapa de restrição do cDNA de γ -coixina, clone pBCX 22.5. São indicados o codon de iniciação, ATG, e o de terminação, TGA. As setas indicam a estratégia de seqüenciamento, sendo que 1 e 2 indicam os subclones pBCX22.5A e pBCX22.5B, respectivamente. As enzimas de restrição indicadas são E, EcoRI; P, PstI e X, XhoI, onde somente os sítios de restrição relevantes ao sequenciamento estão indicados. A região sombreada compreende a fase de leitura do gene de γ -coixina.

Três seqüências correspondentes a possíveis sinais de poliadenilação são encontradas na região 3': AATAATAAT, AATAAA e AATGAA nas posições 652, 719 e 788, respectivamente.

Os clones pBCX22.4 e pBCX22.9 foram seqüenciados utilizando-se a mesma estratégia empregada para o seqüenciamento do clone pBCX22.5. As seqüências obtidas foram similares à do clone pBCX22.5, porém resultando em cDNAs incompletos (Fig. 3). Estes clones apresentaram ainda sítios de poliadenilação localizados em posições diferentes ao descrito para o clone pBCX22.5 (Fig. 3).

4.3. γ -Kafirina

4.3.1. Seleção de clones genômicos de γ -kafirina

Nos testes de seleção de clones de γ -kafirina a partir de uma biblioteca genômica de sorgo foram utilizados $1,5 \times 10^6$ fagos, os quais foram plaqueados em 30 placas de 135 mm x 17 mm, contendo 50000 placas de lise por placa. Após a transferência para membranas de nylon, as mesmas foram hibridizadas com a sonda marcada radioativamente de pME119 (cDNA de γ -zeína de 28 kDa), conforme descrito no item 4.2.1. Os clones positivos foram isolados através de três seleções sucessivas.

Durante as etapas de seleção foram reconhecidos dois tipos de reações, sendo possível a separação dos clones positivos em dois grupos distintos: grupo 1, constituído por 5 clones apresentando forte hibridização, e grupo 2, constituído por 10 clones apresentando fraca hibridização. A análise de restrição dos clones pertencentes aos dois grupos resultou em padrões característicos para cada um dos grupos. Entretanto, todos os clones pertencentes ao grupo 1 apresentaram um mesmo padrão de restrição, o mesmo ocorrendo com os clones do grupo 2. Este fato sugeriu que os fagos recombinantes pertencentes ao grupo 1 continham insertos iguais e portanto, os mesmos sítios de restrição, o mesmo ocorrendo com os fagos do grupo 2. Isto possibilitou a escolha de somente um representante de cada grupo para a continuidade dos trabalhos. Entretanto, experimentos posteriores de "Southern blot" confirmaram a maior intensidade de hibridização da sonda pME119 com o clone pertencente ao grupo 1, indicando que este clone possuía maior homologia com a γ -zeína de 28 kDa, tendo maior probabilidade de ser o clone positivo de γ -kafirina. Sendo assim, um clone deste grupo, denominado de 10S, foi escolhido para o do seqüenciamento.

4.3.2. Mapeamento do fago recombinante 10S

O mapa de restrição do clone 10S foi feito utilizando-se as informações fornecidas pela análise de restrição mostrada na Fig. 4 e pelo experimento de mapeamento do bacteriófago λ (Materiais e Métodos).

A Fig. 4A mostra o padrão eletroforético obtido na análise de restrição do clone 10S com as enzimas BamHI, EcoRI, e PstI, enquanto que a Fig. 4B mostra o resultado da hibridização da membrana, obtida a partir da transferência do gel de agarose, com a sonda de γ -zeína de 28 kDa. A digestão do fago recombinante com EcoRI (canaleta 1, Fig 4A e 4B) além de produzir os fragmentos correspondentes aos braços direito e esquerdo do fago λ EMBL4, produziu um fragmento de 13 kb, indicando portanto a ausência de sítios internos de EcoRI.

O padrão complexo de bandas obtidas nas digestões do fago recombinante, com as enzimas BamHI e PstI (Fig 4A e 4B), induziu-nos a utilizar a técnica de mapeamento de fago lambda descrita no item 3.5.2 de Materiais e Métodos. O resultado da hibridização das digestões totais e parciais do DNA do fago recombinante é mostrado na Fig. 5. As distâncias percorridas pelos fragmentos foram determinadas através da leitura em um densitômetro a laser (Ultrosan XL-LKB) (a leitura foi feita em filmes com diferentes tempos de exposição para que fosse possível a resolução das bandas mais intensas - resultados não mostrados). Os tamanhos dos fragmentos foram estimados utilizando-se o programa Eletroforese. Os dados obtidos com essa metodologia, juntamente com a análise de restrição previamente descrita, permitiu-nos que o mapa mostrado na Fig. 6B fosse construído. Este mapa também apresenta alguns sítios de restrição obtidos por experimentos complementares.

4.3.3. Seqüenciamento do clone 10S

Para a construção dos subclones a serem seqüenciados, o clone 10S foi digerido com BamHI e PstI/EcoRI, liberando os fragmentos de 5,6 kb e 3,1 kb (Fig. 4B, linhas 2 e 5). A escolha desses fragmentos baseou-se no fato deles abrangerem toda a região estrutural do gene de γ -kafirina, como pode ser visto na Fig. 4B. Em seguida, estes fragmentos foram clonados em pBluescript KS- aberto com as respectivas enzimas de restrição. Dessa forma, os fragmentos de 5,6 kb e 3,1 kb originaram os subclones pKBB e pPKEP. Os dois subclones foram seqüenciados nas duas direções, sendo que a partir deste seqüenciamento foram deduzidos novos sítios de restrição, os quais foram posteriormente utilizados na

construção de novos subclones. As seqüências parciais destes subclones apresentaram similaridade com as seqüências descritas para os cDNAs de γ -prolaminas de milho (Prat et al., 1987) e de γ -coixina (descrita no item 4.2.2 e Leite et al., 1991), indicando que o clone 10S corresponderia ao gene de γ -kafirina. Dessa forma, com o término do seqüenciamento, foi possível construir o mapa de restrição mostrado na Fig. 6C. A Tabela I mostra a lista de subclones utilizados no seqüenciamento total do gene.

A seqüência completa do gene de γ -kafirina, incluindo 1513 pb da região 5' não traduzida, e 789 pb da região 3' não traduzida, bem como a seqüência de aminoácidos deduzida (636 pb), está representada na Fig. 7. A região flanqueadora 5' do gene de γ -kafirina apresenta seqüências regulatórias comumente encontradas em muitos genes de plantas. Nas posições -108 e -151, com relação ao ATG inicial, encontram-se as seqüências correspondentes aos possíveis "TATA box" e "CATC box", respectivamente. A seqüência TGTAAG, também chamada "endosperm box" ou elemento -300 (Kreis et al., 1986), está representada por duas cópias no promotor de γ -kafirina. Uma das cópias encontra-se localizada 340 pb acima do ATG, em posição equivalente à descrita para os genes de outras classes de prolaminas do cereais da tribo Andropogoneae, tais como α -kafirina (DeRose et al., 1989), α -coixina (Ottoboni et al., 1993), α -, β -, γ -zeínas (Boronat et al., 1986; Brown et al., 1986; Liu & Rubenstein, 1992; Pedersen et al., 1986). A segunda cópia do "endosperm box" encontra-se localizada em posição mais distal, a 618 pb do ATG.

A análise da seqüência do promotor de γ -kafirina revelou ainda a presença de várias seqüências repetidas, sendo que a mais extensa apresenta 39 pb e encontra-se nas posições -1052 e -1162. Essas seqüências possuem duas regiões altamente conservadas compostas por 13 e 12 nucleotídeos, GAGATTTTACATT e ATTTTTTATTTT, e um região intermediária menos conservada composta por 13 e 14 nucleotídeos. As diferenças existentes entre essas seqüências se restringem às regiões intermediárias, GGGAACCTAGAAA e AGGAACCCTAGAAT, as quais possuem aproximadamente 79% de homologia, sendo que os nucleotídeos grifados representam as alterações de bases existentes entre as duas regiões.

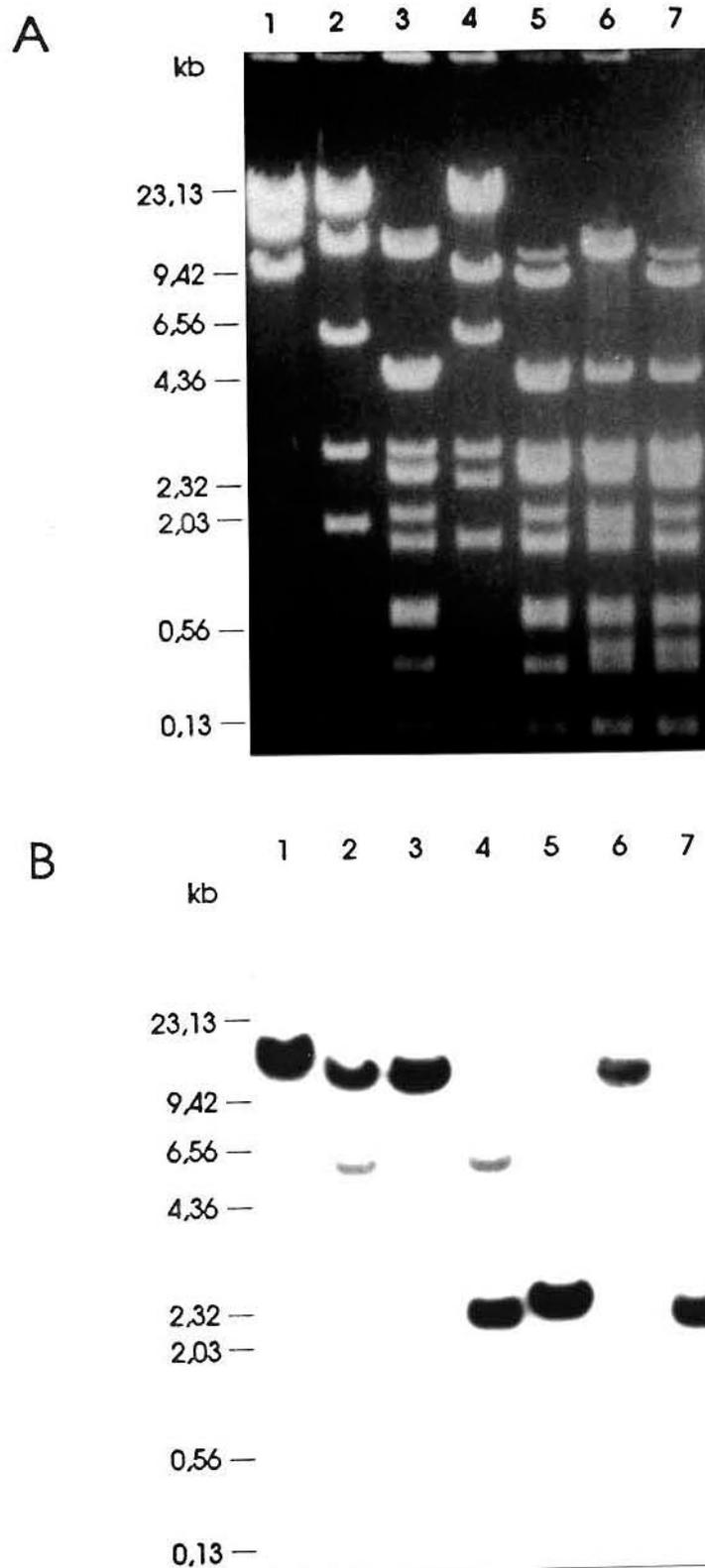


Fig. 4. (A) Análise de restrição do clone genômico de γ -kafirina 10S, onde as canaletas indicadas correspondem a 1) 10S EcoRI, 2) 10S BamHI, 3) 10S PstI, 4) 10S EcoRI/BamHI, 5) 10S EcoRI/PstI, 6) 10S BamHI/PstI e 7) 10S BamHI/PstI/EcoRI. B) As amostras de DNA do clone genômico 10S foram submetidas à eletroforese em gel de agarose 0,8% e transferidas para membranas de nylon, sendo hibridizadas com o inserto de PstI do clone pME 119 (cDNA de γ -zeína de 28 kDa)

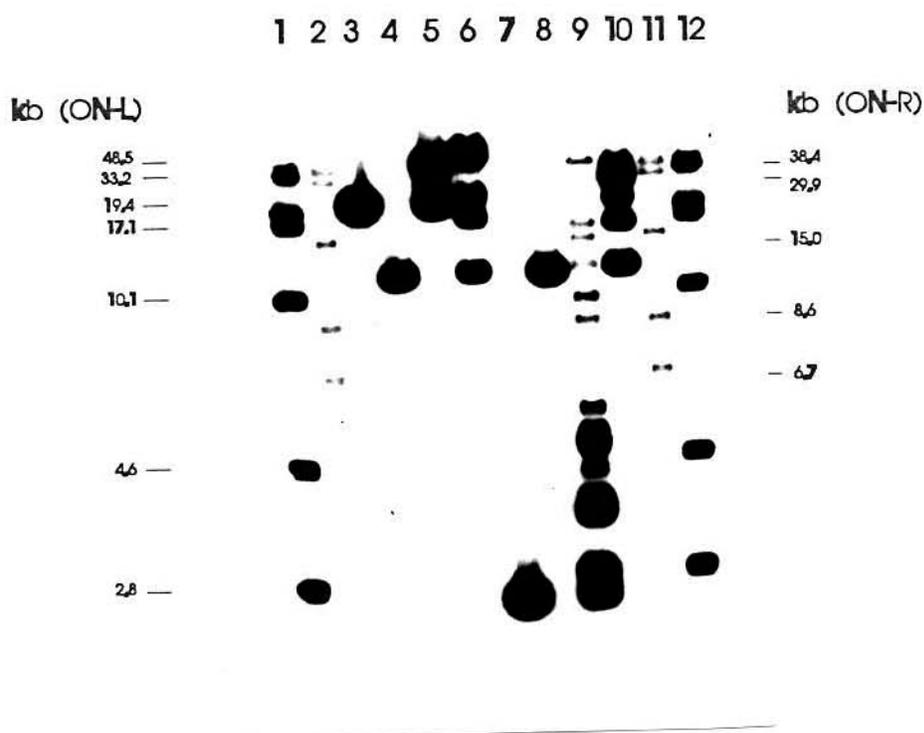


Fig. 5. Mapeamento do fago contendo o inserto de γ -kafirina, clone 10S. As enzimas de restrição utilizadas foram BamHI e PstI. Após as digestões, parciais ou totais, os DNAs foram hibridizados com os oligonucleotídeos correspondentes às extremidades *cos*, esquerda (ON-L) e direita (ON-R) do fago, conforme indicado em cada caso. O mesmo procedimento foi seguido para os marcadores de peso molecular (os pesos moleculares dos fragmentos de cada marcador estão indicados ao lado do gel). As canaletas indicadas correspondem a 1) padrão de peso molecular (ON-L), 2) padrão de peso molecular (ON-R), 3) digestão total com BamHI (ON-L), 4) digestão total com BamHI (ON-R), 5) digestão parcial com BamHI (ON-L), 6) digestão parcial com BamHI (ON-R), 7) digestão total com PstI (ON-L), 8) digestão total com PstI (ON-R), 9) digestão parcial com PstI (ON-L), 10) digestão parcial com PstI (ON-R), 11) padrão de peso molecular (ON-R) e 12) padrão de peso molecular (ON-L).

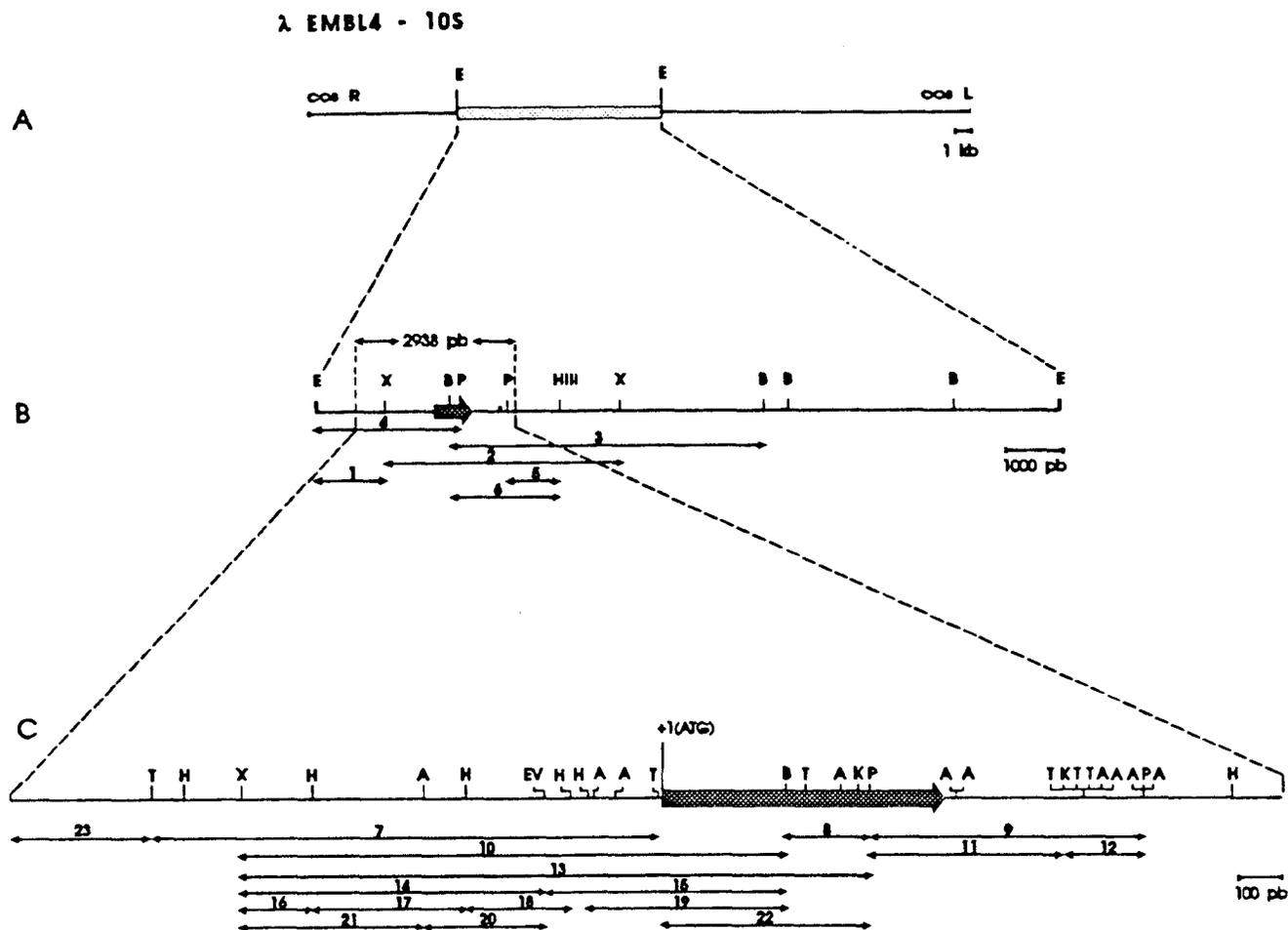


Fig. 6. Mapa de restrição do clone genômico de γ -kafirina. (A) Representação do bacteriófago λ EMBL4 contendo o clone genômico de 13,2 kb. Os braços R e L do bacteriófago λ possuem 8,8 kb e 19,9 kb. (B) Mapa de restrição obtido através do experimento de mapeamento de bacteriófago λ . Alguns sítios enzimáticos obtidos através de experimentos complementares também são indicados. (C) Região completamente sequenciada correspondente a 2938 pb, sendo indicados todos os subclones utilizados durante o sequenciamento. Somente os sítios enzimáticos relevantes ao sequenciamento são mostrados, sendo A, AluI; B, BamHI; EV, EcoRV; H, HinfI; HIII, HindIII; K, KpnI; P, PstI; T, TaqI e X, XbaI. Os clones numerados são mostrados na Tab. I. Em B) e C) as regiões sombreadas correspondem à fase de leitura do gene de γ -kafirina.

Tabela I. Relação dos subclones utilizados para o sequenciamento total do clone genômico de γ -kafirina

NOME	EXTREMIDADES DO INSERTO	TAMANHO DO INSERTO (pb)	VETOR
1) pKEX	EcoRI/XbaI	1262	rec pKEP/XbaI
2) pKXX	XbaI/XbaI	4166	KS (-)/XbaI
3) pKBB	BamHI/BamHI	5688	KS (-)/BamHI
4) pKEP	EcoRI/PstI	2542	KS (-)/EcoRI/PstI
5) pKPH3	BamHI/PstI	892	KS (+)/SmaI
6) pKBH3	BamHI/HindIII	1914	KS (+)/BamHI/HindIII
7) pKTT1	SaII/HindIII	1184	KS (+)/ClaI
8) pKBP	BamHI/PstI	195	KS (+)/SmaI
9) pKPP	BamHI/PstI	632	KS (+)/SmaI
10) pKXB	XbaI/BamHI	1280	KS (+)/XbaI/BamHI
11) pKPK	KpnI/BamHI	161	rec pKPP/KpnI
12) pKKP	KpnI	471	KS (-)/KpnI
13) pKXP	XbaI	1475	KS (+)/XbaI
14) pKXEV	XbaI/EcoRV	717	KS (+)/XbaI/EcoRV
15) pKEVB	EcoRV/BamHI	563	KS (+)/EcoRV/BamHI
16) pKXH	BamHI/PstI	179	KS (+)/SmaI
17) pKHH1	BamHI/PstI	348	KS (+)/SmaI
18) pKHH2	BamHI/PstI	242	KS (+)/SmaI
19) pKHB	BamHI/PstI	469	KS (+)/SmaI
20) pKAEV	PstI/BamHI	294	KS (+)/SmaI
21) pKXA	XbaI/HindIII	423	KS (+)/XbaI/EcoRV
22) pUCTP	PstI/HindIII	480	pUC19/PstI/HindIII
23) pKTT2	SaII/HindIII	320	KS (+)/ClaI

Na região flanqueadora 3' foram encontrados três possíveis sinais de poliadenilação, sendo que a seqüência AATAAT encontra-se repetida nas posições +658 e +715, e a seqüência AATGAA está presente na posição +784 (Fig. 7 e 8).

Outra característica importante deste gene, é a ausência de introns, verificada através da comparação da seqüência de nucleotídeos deste gene, com a seqüência descrita para um clone de cDNA de γ -kafirina (Barros et al. 1991). O alinhamento das seqüências dos clones genômico e de cDNA revelou a presença de 5 alterações de nucleotídeos na região estrutural, sendo 4 delas constituídas por transições silenciosas (2 G \leftrightarrow A e 3 C \leftrightarrow T) e uma transversão G \leftrightarrow C na posição +236, determinando a substituição Ala \leftrightarrow Pro. Foi encontrada ainda uma sétima diferença entre os clones genômicos e de cDNA representada por uma deleção de um nucleotídeo T na região flanqueadora 3' no clone genômico.

A seqüência de nucleotídeos do clone 10S foi depositada nos bancos de dados EMBL, GenBank e DDBJ com o número de acesso X62480.

4.4. Experimentos de reconstrução

O resultado do experimento de reconstrução do gene de γ -coixina, mostra que uma única banda no DNA genômico de *Coix* sofreu hibridização. Essa única banda apresentou intensidade correspondente a uma ou duas cópias do gene de γ -coixina no genoma haplóide do *Coix* (Fig. 9A). Enquanto que no experimento de reconstrução realizado com o gene de sorgo foram detectadas duas bandas com intensidade correspondente a aproximadamente 1 cópia cada uma (Fig. 9B).

4.5. Estudo comparativo das γ -prolaminas de *Coix*, sorgo e milho

4.5.1. Comparação das seqüências de aminoácidos das γ -prolaminas

A Fig. 10 mostra a comparação da seqüência deduzida de aminoácidos do gene de γ -kafirina, obtida através do seqüenciamento do clone 10S, e seqüências das γ -zeínas de 28 kDa e 16 kDa (Prat et al., 1985; Prat et al., 1987), bem como a seqüência da γ -coixina (descrita no item 4.2.2 e Leite et al., 1991). Prat et al. (1985) relataram que os 19 primeiros aminoácidos presentes na região N-terminal da seqüência deduzida de aminoácidos do gene da γ -zeína de 28 kDa, possuem propriedades de peptídeo sinal, presente em proteínas secretadas através de membranas. Através do seqüenciamento da região N-terminal da proteína madura de γ -kafirina (TLTTGG), realizada em colaboração com o Prof. Dr. José Camillo

Fig. 7. Seqüência de nucleotídeos do clone genômico de γ -kafirina. Na região 5' aparecem sublinhadas as seqüências correspondentes ao TATA box (1), CATC box (2), duas seqüências homólogas ao "endosperm box" (4 e 6) e duas seqüências GCN4 (3 e 5). Também são assinaladas duas seqüências repetidas (7 e 8). Na região 3' são sublinhados possíveis sinais de poliadenilação.

-1511 AATATGCCTCAGTATTATATATAAACATGAGAAATTTGATGCTAGCATTATTTTTAATTTAGATTAAGATTGCTAT -1435
 -1434 GAAATATGAATCAAGCATCACATTAAATAGGCTACCATTAAATTTTTATAATAATTGGAGCAAGATAACTATAATTTT -1357
 -1356 AATTTTACAGTAAAAAGCATAGGCAAGCATTTCAAAAAAAAGTGTACTAAAATTTATGATATTTTTTGTACCAGCA -1279
 -1278 TGGATCTACATATGTGGAGCTGAACAAAATTTGTTTTATAATTTTTAGATTTTTGTATGAATTATTATGTATTTATTA -1201
 -1200 ATTTTCAGTCGATTTAAAAAATAAAAGAAAAGTAAATTTGGAGATTTTACATTGGGAACCTAGAAAATTTTTTATTTT -1123
 -1122 TTTCTCTCTCCATCGTAACGATTCTGGGCTGATTGGGCTACAGAGAGGAGAGCGAAAGCGAAGTAATATGAGATTT -1045
 -1044 TACATTAGGAACCCCTAGAATATTTTTTATTTTGTTTTTCCCTCTAGAAAAGTTGGATGGAGGAGGCCATCGGTAAA -967
 -966 CCGAACACGGACAAGATCAGGGGCAAAAGAAAATTTGGCAAATAAAATTTGGCTCTTATAGATAGGTATAGAT -889
 -888 TTGGGTTAAGAACAGTAGTCGGACATTCCACGCTTTGTTAAGTCTTTTTTTACATAACAAGAATAGAATCACGCTCT -811
 -810 GTAGGCTCGTCAGACCGTACCTTTCTAAGGAAGTCGCTTTGGGTAGTTCAGTTGGCGAGAAAAGCCTCCTACCTTTG -733
 -732 CAGGTCCATCGGGCCGACTACAACCCGTTGGCTCAATCCCGGTTCTGCGGTGTCTGGCAACATTCTGTTGGAAGAT -655
 -654 ACCAGAAGGTTGCTCCACGGGTAATCTTGACACGTATGTAAGTGTAGGGAACATTGAACGAACATTGGCATGTAAG -577
 -576 CTCTATAATTGGTGTATCCATAACAACGTCGAGAACATCACAAATGTCACGTCAAGGGATTGGGTGAGAAAACAAAT -499
 -498 CGTCTCCGTGTACAACGAAGTGGTGAAGTGTAGGCGATGAGCCATGTTGATCTGATATATACATAGCACACAGCATCACAAAC -421
 -420 AAGTCATACTACATTACAGAGTTAGTTTACCTTTCAAGTAAAAACAAAGTAGGCCGAGAGAGGACAATAATCCTTG -343
 -342 ACGTGTAAAGTGAATTTACAAGCCATATATCAATTTATATCTAATTCGTTTCATGTAGATATCAACAACCTGTAAAA -265
 -264 GGCAACAAATGAGCCACGCAAAATTACAAGTGAAGTCCAAATAAACCCCTCACATGTACATAAAAAGTGAATGATGAGT -187
 -186 CATGTATATCTGGCAAGAACTGTAGAAGTACAGTCAATCGGTAGCAAAGAAACACAAGAAAATGTGCTAATAAAAAGC -109
 -108 TATAAATAAACCCCTCGTACGCTATGCACATCTCCATCACCACCCTGGTCTTCATTCAGCCTATTAACCTTATATCTAT -31
 - 30 CTACTCCAGAGCAGACAAGAAGTCTGACACCATGAAGGTGTGCTCGTTGCCCTCGCTCTCCTGGCTCTCGCGGCGAGC 48
 1 M K V L L V A L A L L A L A A S 16
 49 GCCGCTCCACGCTTACAACCGGGGGCTGCGGCTGCCAGACACCTCATCTACCACCACCGCGGTTTCATCTGCCGCG 126
 17 A A S T L T T G G C G C Q T P H L P P P P V H L P P 42
 127 CCGGTGCATCTGCCACCGCGGTTGCACCTGCGCGCGCGGTTACAGTGCACACCGCCACCACAATGCCACCCACAC 204
 43 P V H L P P P V H L P P P V H V P P P P P Q C H P H 68
 205 CCTACTCTACCGCCACCACACATCCATGCGCTACATACCACCGCATCCAAGCCGTTGCCACCCAGGGCATCCCGGA 282
 69 P T L P P H P H P C A T Y P P H P S P C H P G H P G 94
 283 TCCTGCGGTGTTGGCGGGCGCCCGTCAACCCCGCGATCCTGGGCCAGTGCATCGAGTTCCTGAGGCATCAGTGCAGC 360
 95 S C G V G G G P V T P P I L G Q C I E F L R H Q C S 120
 361 CCGCGGGCAGCCCTACTGCTCGCCACAGTGCAGGCGTTGCGGCAGCAGTGTGTGTCAGCAGCTCAGGCAGGTGGAG 438
 121 P A A T P Y C S P Q C Q A L R Q Q C C Q Q L R Q V E 146
 439 CCGCTGCACCGGTACCAGGCGATCTCGGCGTGGTCTGCAGTCCATCCAGCAGCAGCAGCCGCAAGGCCAGTCTGCA 516
 147 P L H R Y Q A I F G V V L Q S I Q Q Q Q P Q G Q S S 172
 517 CCGTCCCAGCGCTGATGGCGGCGAAATAGCACAGCAACTGACGCGATGTGCGGTCTAGGAGTGGGCGAGCCAAAGT 594
 173 P L P A L M A A Q I A Q Q L T A M C G L G V G Q P S 198
 595 CCCTGCGCTTCTGAGCCCTTTTGGCGGTGGTGTCCACTATTAAGAAACTATCTATACTGTAATAATGTTGTATAG 672
 199 P C A S C S P F A G G V H Y 212
 673 CCGCGGATAGCTAGCTAGTTAGTCATTGACGCGGATGGGTAATAATAAAGTGTATCCATCCATCACCATGGGTGG 750
 751 CAACGTGAGCAATGACCTGATTGAACAAATGAAATGAAAAGAAGAAATATGTTATATGTCAACGAGATTTCTCATA 828
 829 ATGCCACTGACAACGTGTGTCGAAGAAATGATCAGTGATACGTATATTCACAATTTTTTATGACTTATACTCACAA 906
 907 TTTGTTTTTTACTACTTATACTCGAACATTTGTTGTGGGTACCATAACAATTTGATCGAATATATATCAGAAAGT 984
 985 TGACGAAAGTAAGCTCACTCAAAAAGTTAAATGGGCTGCGGAAGCTGCGTCAGGCCCAAGTTTTGGCTATTCTATCCG 1062
 1063 GTATCCACGATTTTGTGGCTGAGGACATATGTTGCGCTTAAGTGCAGCTTTGTAGTTAGTTTTGTTTTGATTAT 1140
 1141 ATTTAATACTCTATGCATGTGCACCAAGATTTTCTGGTGAATTAACAAGGCCTAATACGTGAGTAGCGTATCTAA 1218
 1219 CTGTGACCTATAAAGTAGAGCACCTTTTAGAGTAGGGCTCCTTTTTTGAAGTCTATTTATGCACCAACTTCA 1296
 1297 ATAAGGCTTTTTTATCCAAAATTAAGAGTCTTACATTACATCTAATCGTCTATTCATTGTCTATATTTAATAATA 1374
 1375 ATCTTACTGTATATCTGTAGCACACTAGTATGCCTCAAAGCCGACAATAAT 1427

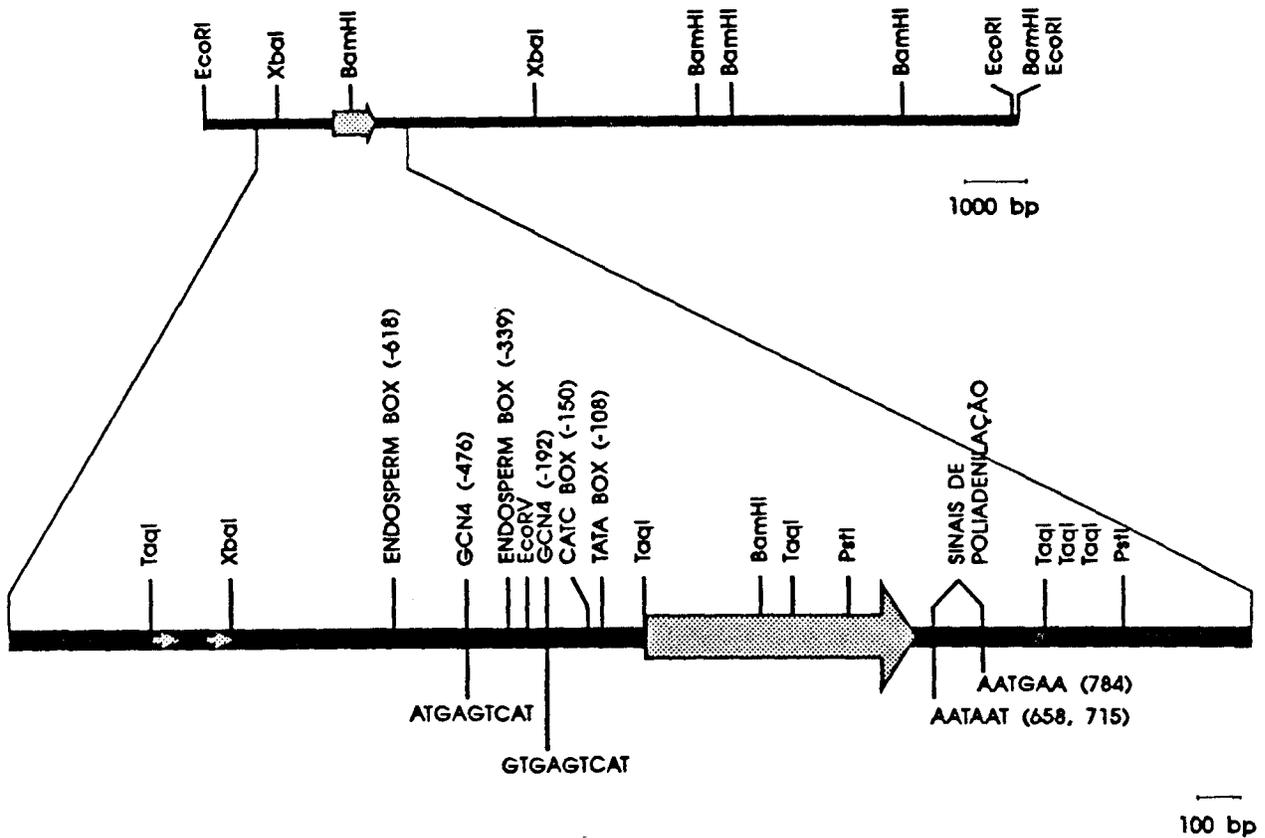


Fig. 8. Representação do clone genômico de γ -kafirina, onde é mostrado a localização das principais seqüências regulatórias encontradas na região 5', bem como sinais de poliadenilação presentes na região 3'.

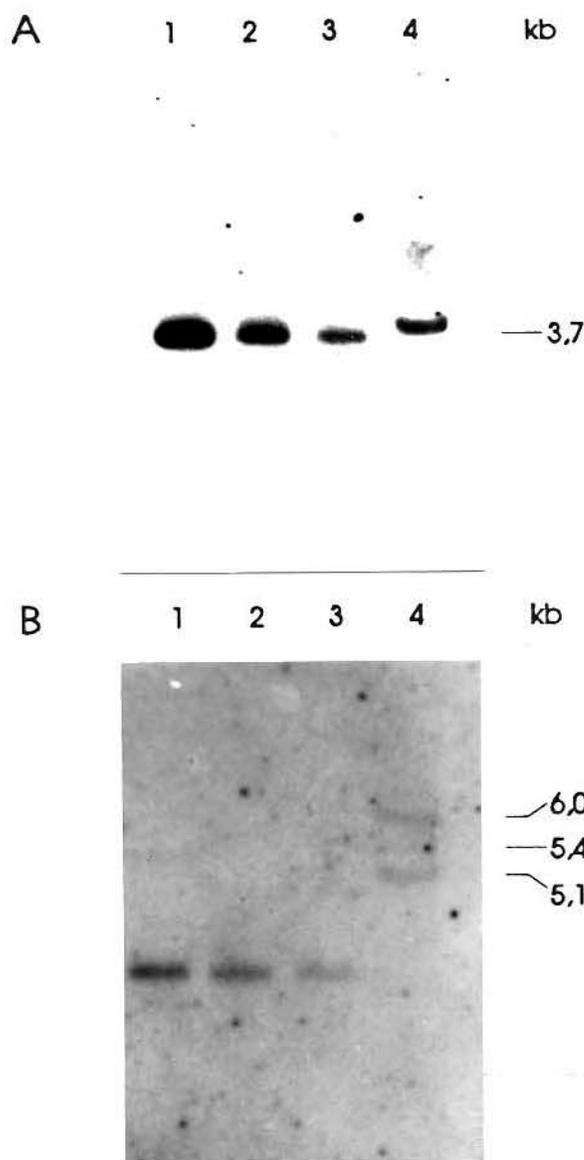


Fig. 9. A) Reconstrução para determinação do número de cópias do gene de γ -coixina no genoma haplóide de *Coix*. As linhas 1, 2 e 3 mostram o plasmídeo pBCX 22.5 linearizado com EcoRI, em quantidades equivalentes a 3, 2 e 1 cópia do gene no genoma haplóide do *Coix* e a linha 4 corresponde a 10 μ g de DNA genômico de *Coix* digerido com HindIII. As amostras de DNA foram submetidas à eletroforese em gel de agarose 0,8%. Em seguida o DNA foi transferido para membrana de nylon e hibridizado com o inserto proveniente da digestão do clone pBCX 22.5 com EcoRI/XhoI marcado radioativamente. O fragmento hibridizado, correspondente ao gene de γ -coixina (linha 4), possui aproximadamente 3700 pares de base. B) Auto-radiografia de um experimento similar realizado para o gene da γ -kafirina, onde o plasmídeo utilizado como controle do número de cópias foi o inserto liberado da digestão do clone pUCTP com PstI/HindIII. O DNA genômico de sorgo (10 μ g) foi digerido com HindIII.

Novello, do Departamento de Bioquímica do Instituto de Biologia da Unicamp, foi possível demonstrar que os primeiros 19 aminoácidos encontram-se ausentes na proteína madura, e possivelmente representam o peptídeo sinal. Da mesma forma, os 19 primeiros aminoácidos da γ -coixina e γ -zeína de 16 kDa podem representar o peptídeo sinal dessas proteínas, visto que possuem um alto grau de homologia com os peptídeos sinais da γ -zeína de 28 kDa e γ -kafirina. As seqüências das proteínas maduras das γ -prolaminas de milho, sorgo e *Coix*, apresentam regiões repetitivas na região N-terminal. A primeira região repetitiva é formada pelo hexapeptídeo PPPVHL. A γ -zeína de 28 kDa apresenta oito repetições desse hexapeptídeo, sendo que a sétima repetição apresenta uma modificação, onde o aminoácido Leu é trocado por Val (Esen et al., 1982). A γ -kafirina possui somente quatro repetições completas desse hexapeptídeo, sendo que a quarta repetição, da mesma forma que a sétima repetição da γ -zeína de 28 kDa, apresenta uma troca do aminoácido Leu por Val. Esta troca é resultado de uma única substituição, onde um nucleotídeo C foi trocado por G. A γ -coixina possui três repetições completas e uma incompleta desse hexapeptídeo. Porém, com relação a γ -zeína de 16 kDa, verifica-se que esta possui três repetições modificadas desse hexapeptídeo. A segunda seqüência repetitiva presente na região codificadora da γ -zeína de 28 kDa é composta por dois octapeptídeos, QPHPCPCQ e QPHSPCQ. A única diferença existente entre esses dois octapeptídeos, é a substituição do aminoácido Cys por Ser (Wang & Esen, 1986). As γ -prolaminas de sorgo e *Coix* apresentam uma versão menos conservada da seqüência duplicada, enquanto que a γ -zeína de 16 kDa apresenta uma única cópia do octapeptídeo.

4.5.2. Composição de aminoácidos

A Tabela II mostra a composição de aminoácidos das γ -prolaminas de sorgo, *Coix* e milho. Essas proteínas não possuem os aminoácidos essenciais, lisina e triptofano, com exceção da γ -zeína de 16 kDa, que apresenta 0,61% de triptofano. Entre os aminoácidos, destacam-se os altos teores de prolina e glutamina. As γ -prolaminas apresentam também um alto teor de cisteína, responsável pela formação de pontes dissulfeto intercadeias, o que determina a formação de complexos de alto peso molecular. Assim sendo, estas proteínas são solúveis somente em condições redutoras, isto é, na presença de 2-ME ou DTT. Devido a esta propriedade, as γ -prolaminas foram anteriormente classificadas em um grupo distinto de proteínas de reserva denominado RSP ("reduced soluble

protein") (Wilson et al., 1981). Uma outra classificação denominava a γ -zeína de 28 kDa como glutelinas-2 (Prat et al., 1985). A Tabela II mostra também os pesos moleculares das γ -prolaminas de *Coix*, sorgo e milho, estimados a partir de SDS-PAGE, bem como os pesos moleculares das proteínas maduras correspondentes, calculados a partir da seqüência deduzida de aminoácidos.

4.6. Comparação das seqüências flanqueadoras 5' dos genes de γ -prolaminas de milho e sorgo

A Figura 11 mostra o alinhamento das seqüências flanqueadoras 5' dos genes de γ -prolaminas de milho e sorgo, indicando que a região mais conservada entre esses promotores compreende a região próxima ao ATG. Essa região contém o "TATA box", enquanto os possíveis "CATC box" encontram-se localizados em região menos conservada, sendo que o "CATC box" da γ -zeína de 28 kDa apresenta uma substituição A \leftrightarrow G e também a inserção do dinucleotídeo TC.

Os promotores de γ -kafirina e γ -zeína de 28 kDa apresentam duas cópias do "endosperm box" (TGTAAG). Uma das cópias deste elemento encontra-se localizado em posições equivalentes nos dois promotores (-339 no promotor de γ -kafirina e -353 no promotor de γ -zeína de 28 kDa). Conforme descrito anteriormente, estas posições correspondem aproximadamente às relatadas para outros genes de prolaminas. A segunda cópia encontra-se localizada em regiões distintas mais distais com relação ao ATG inicial dos dois promotores. O promotor de γ -zeína de 16 kDa apresenta uma extensa deleção em comparação com os outros dois promotores. Essa deleção inicia-se na região correspondente à cópia proximal do "endosperm box" dos promotores de γ -kafirina e γ -zeína de 28 kDa, sendo que o promotor de γ -zeína de 16 kDa apresenta apenas versões truncadas do "endosperm box", tais como as presentes nas posições -296 e -659, constituídas pelas seqüências TGAAAAG e TATAAAG, respectivamente.

A presença do "endosperm box" não é exclusividade dos genes de proteínas de reserva dos cereais da tribo Andropogoneae, sendo que já foram descritos também nos genes de prolaminas de cevada (Kreis et al., 1986), trigo (Summer-Smith et al., 1985; Colot et al., 1987) e centeio (Hull et al., 1991). Recentemente, Muller & Knudsen (1993) descreveram que o "endosperm box" e seqüências homólogas ao domínio reconhecido pelo fator de ativação transcrição em levedura, GCN4 (Hill et al., 1986), estão envolvidos na regulação de um gene de prolamina de cevada (hordeína-C) em resposta a níveis de nitrogênio. O

envolvimento conjunto do "endosperm box" e da seqüência GCN4 foi reforçado pelos estudos de "footprinting" *in vivo* realizados com uma prolamina de trigo (LMWG-1D1) (Hammond-Kosack et al., 1993). Tais resultados levaram-nos a realizar uma busca das seqüências similares ao domínio GCN4 (GG/ATGAG/CTCAT) nos promotores de γ -prolaminas. Duas cópias altamente conservadas foram encontradas nos promotores de cada gene. A cópia proximal em relação ao ATG inicial, localizada na posição -192 do promotor de γ -kafirina, encontra-se 100% conservada nos promotores de γ -kafirina e γ -zeína de 16 kDa, sendo que o promotor de γ -zeína de 28 kDa apresenta uma única substituição. A segunda cópia, localizada na posição mais distal, posição -476 do promotor de γ -kafirina, apresentou apenas uma substituição na seqüência do promotor de γ -zeína de 16 kDa.

4.7. Comparação das regiões flanqueadoras 3'

Com a finalidade de identificar regiões conservadas envolvidas no processamento da região 3' dos mRNAs de γ -prolaminas, foram alinhadas as seqüências flanqueadoras 3' dos genes de γ -kafirina (clone genômico 10S), γ -zeína de 28 kDa (clone genômico ZC2, Reina et al., 1990b), γ -zeína de 16 kDa (clone genômico ZC1, Reina et al., 1990a) e γ -coixina (clone de cDNA pBCX22.5). A Figura 12 mostra o alinhamento a partir do sinal de término de tradução até a região que contém os sítios de poliadenilação. Os sítios de poliadenilação das seqüências originárias de clones genômicos foram deduzidas a partir das seqüências de cDNA disponíveis para esses genes (Barros et al., 1991; Prat et al., 1987). Com base neste alinhamento foram reconhecidas 6 regiões contendo possíveis sinais de poliadenilação que foram denominadas de SP-1, SP-2, SP-3, SP-4 e SP-5 (Fig. 12).

O sinal de poliadenilação de mamíferos AATAAA (Proudfoot & Brownlee, 1976) está presente apenas na região SP-2 da seqüência flanqueadora do gene de γ -coixina. A seqüência AATAAT, anteriormente descrita como possível sinal de poliadenilação em genes de α -zeínas (Marks et al., 1985) e α -coixinas (Leite et al., 1992; Ottoboni et al., 1993), aparece nas regiões SP-1 de γ -kafirina, γ -zeína de 16 kDa e γ -coixina, e também na região SP-2 dos genes de γ -kafirina e γ -zeína de 28 kDa.

Tabela II. Composição de aminoácidos das γ -prolaminas de sorgo^a, *Coix*^b e milho^c

Aminoác.	γ -kafirina ^a (27 kDa)		γ -coixina ^b (22 kDa)		γ -zeína ^c (28 kDa)		γ -zeína ^c (16 kDa)	
	Quant	%	Quant	%	Quant	%	Quant	%
Ala	12	6,22	12	6,67	10	4,90	13	7,93
Arg	4	2,07	6	3,33	5	2,45	3	1,83
Asp	0	0,00	0	0,00	0	0,00	0	0,00
Asn	0	0,00	0	0,00	0	0,00	1	0,61
Cys	15	7,77	14	7,78	15	7,35	12	7,32
Gln	23	11,92	34	8,89	30	14,71	31	18,90
Glu	2	1,04	2	1,11	2	0,98	3	1,83
Gly	17	8,81	12	6,67	13	6,37	15	9,15
His	15	7,77	10	5,56	16	7,84	4	2,44
Ile	5	2,59	7	3,89	4	1,96	1	0,61
Leu	16	8,29	13	7,22	19	9,31	14	8,54
Lys	0	0,00	0	0,00	0	0,00	0	0,00
Met	2	1,04	3	1,67	1	0,49	3	1,83
Phe	3	1,55	4	2,22	2	0,98	7	4,27
Pro	44	22,80	36	20,00	51	25,00	25	15,24
Ser	10	5,18	6	3,33	8	3,92	9	5,49
Thr	9	4,66	8	4,44	9	4,41	6	3,66
Trp	0	0,00	0	0,00	0	0,00	1	0,61
Tyr	4	2,07	4	2,22	4	1,96	8	4,88
Val	12	6,22	9	5,00	15	7,35	8	4,88
Resíduos	193		180		204		164	
Mr deduzido	20253		19542		21824		17751	

Fig. 11. Alinhamento das regiões 5' dos genes de γ -kafirina, γ -zeína de 28 kDa e 16 kDa. Os nucleotídeos são numerados a partir do ATG inicial. Algumas seqüências conservadas envolvidas com a regulação desses genes são indicadas, tais como o TATA box, CATC box e as seqüências correspondentes ao "endosperm box" (end-b) e GCN4. A γ -kafirina possui 83% e 82% de homologia com a γ -zeína de 28 kDa e 16 kDa, respectivamente, enquanto que a γ -zeína de 28 kDa possui 79% de homologia com a γ -zeína de 16 kDa. O alinhamento mostra a existência de 388 nucleotídeos conservados.

Recentes estudos sobre o processamento da região 3' do mRNA de γ -zeína de 28 kDa (Wu et al., 1993) apontam a seqüência AATGAA como sinal de poliadenilação daquele gene. Cópias conservadas, bem como diferentes versões desta seqüência, encontram-se localizadas próximo às regiões que contém os sítios de poliadenilação, nas regiões identificadas como SP-3, SP-4 e SP-5, sendo que na região SP-3 todos os genes analisados apresentam a seqüência AATGAC. Na região SP-4 os genes de γ -kafirina e γ -coixina apresentam as seqüências ATTGAA e GATGAA, respectivamente, enquanto que os genes de γ -zeína de 28 e 16 kDa apresentam a seqüência AATGAA. Com excessão do gene de γ -zeína de 16 kDa que apresenta a seqüência AATGGA na região SP-5, os demais genes apresentam a seqüência AATGAA 100% conservada.

Dentre as regiões conservadas nas regiões flangeadoras 3' dos genes de γ -prolaminas, destacamos a região identificada como CATGG no alinhamento da Fig. 12. Essa região encontra-se 100% conservada nos genes de γ -prolaminas e inclui a seqüência CATGGGTGG, identificada como uma das regiões essenciais para o processamento 3' do mRNA de γ -zeína de 28 kDa (Wu et al., 1993). Esta seqüência foi também descrita na região 3' do transcrito 35S do vírus do mosaico de couve flor (CaMV35S), sendo que neste caso a deleção desta seqüência apresentou um menor efeito no processamento deste mRNA (Sanfaçon et al., 1991). A seqüência CAGTGTCAGTGT também é encontrada nas γ -zeínas de 28 kDa e 16 kDa na região denominada SR-1 (Wu et al., 1993), enquanto que a γ -kafirina e γ -coixina apresentam apenas versões modificadas desta seqüência.

4.8. Efeito do promotor da γ -kafirina na expressão do gene indicador *gus* em tecidos de milho

Para investigar a especificidade do promotor de γ -kafirina, endospermas imaturos de sorgo, *Coix* e milho foram bombardeados com micropartículas contendo DNA adsorvido de uma construção plasmidial possuindo 1190 pb do promotor de γ -kafirina dirigindo a expressão do gene *gus* (pPKGUS, Fig 13). A análise histoquímica de GUS (Figura 14) demonstrou que o fragmento do promotor de γ -kafirina foi capaz de dirigir a expressão nas sementes das diferentes espécies bombardeadas.

```

                                <SP-1>
γ-kafirin 637 TAAA----GAAACTATCTA-TACTG---TAATAATG---TTGTATAGCCGCCGGATAGCTAG-----CTA 690
γ-zein 28K 670 TGAA----GAAACTATG---TGCTGTAGTA-----TAGCCGCTGGCTAGCTAG-----CTA 713
γ-zein 16K 550 TGAG----GAAACTATG---TACTGTAGTAATAATG---TAATGGAGCCGCTGACTAGCTACCTATAGCTA 610
γ-coixin 598 TGAAGTGAAGAACTATATCCTACTGTAATAATAATGATGTAATATAGCCGCTGGCGAGCTAA-----GCTA 664
      * *      * * * * *      * * * * *      * * * * *      * * * * *      * * * * *
                                <SP-2>                                <---CATGG---> <---
γ-kafirin 691 GTT-AGTCATTCA--GCGGCGATG---GGTAATAATAAAGTGTCCATCCATCCATCACCATGGGTGGCAACG 755
γ-zein 28K 714 GTTGAGTCATTTA--GCGGCGATGATTGAGTAATAAT---GTGTCACG---CATCACCATGGGTGGCAGTG 776
γ-zein 16K 611 GTTCACTCGTTTA--GCGGCGATGAG-----TAACGGAGTGTCCACC---CATCACCATGGGTGGCAGTG 669
γ-coixin 665 GTTTAGTCATTTAGCGGCGATG---GGTAATAAAGT--GTCATC---CATCACCATGGGTGACAATA 726
      * * * * * * * * * *      * * * * *      * * * * *      * * * * *      * * * * *
-SR-1>      <SP-3>      <SP-4>      <SP-5>
γ-kafirin 756 TGAGC-----AATGACCTG--ATTGAACAAATTGAAATGAAAAGAAGAAA--TATGTTATATGTCAACGAGA 819
γ-zein 28K 777 TCAGTGTGAGCAATGACCTG--AATGAACAA--TTGAAATGAAAAGAAAAAGTATTGTTCCAAATTAACGT 846
γ-zein 16K 670 TGAGC-----AATGACCTG--AATGAACCA--TTGAAATGAAAAGGAATAA--TATATAGGAAAAAGGATGT 732
γ-coixin 727 TATGCGAA--CAAATGACCTGAAGATGAACAA--TTGAAATGAAAAGGAAAA--TATATTACAATTC 787
      * *      * * * * *      * * * * *      * * * * *      * * * * *

```

Fig. 12. Alinhamento das regiões flanqueadoras 3' dos genes de γ -prolaminas de milho, *Coix* e sorgo. Os sinais de poliadenilação estão distribuídos em 5 regiões denominadas de SP-1, SP-2, SP-3, SP-4 e SP-5. As seqüências CATGG e SR-1 também são indicadas.

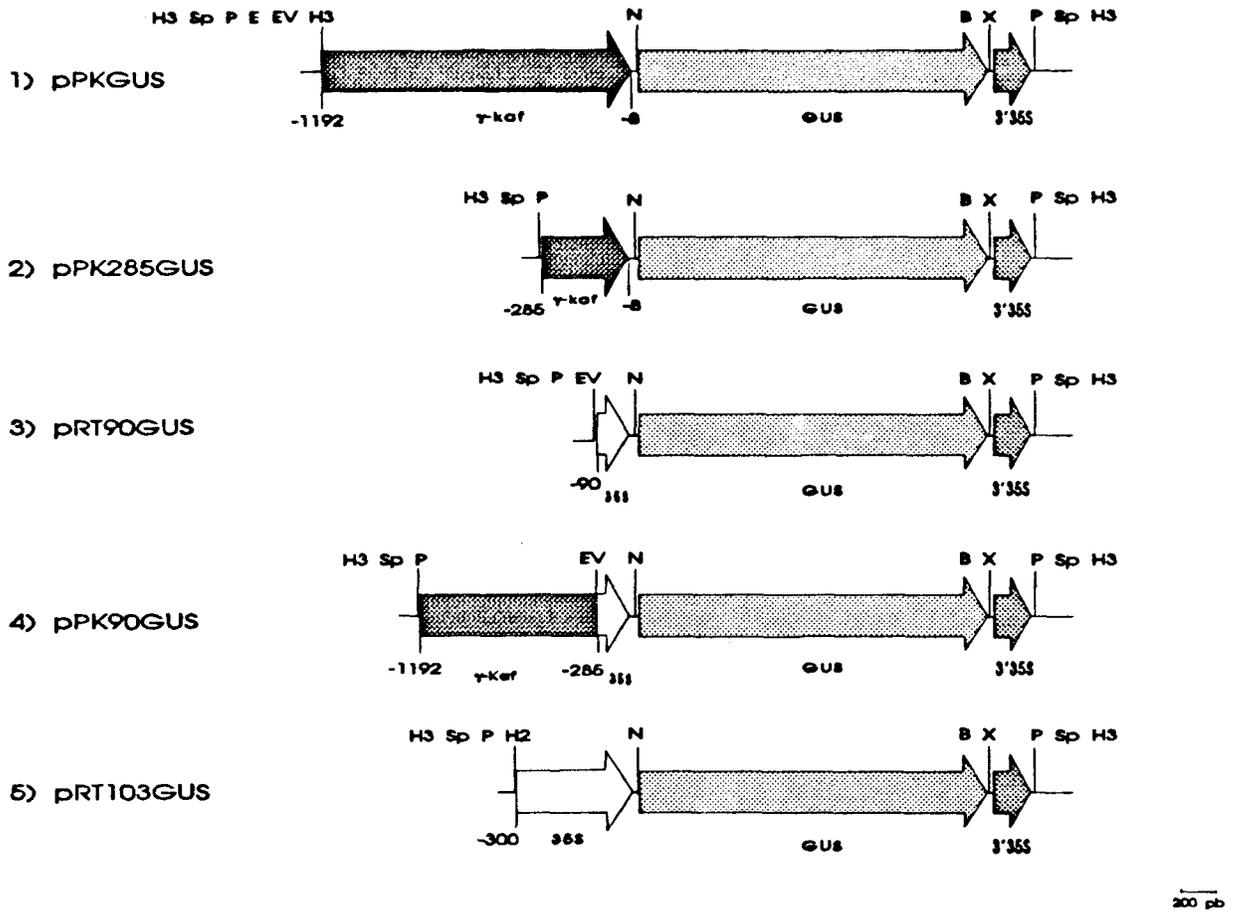


Fig. 13. Construções utilizadas nos experimentos de bombardeamento. 1) Construção pPKGUS: representa o gene *gus/3' 35S* sob o controle do promotor da γ -kafirina contendo aproximadamente 1200 pb; 2) construção pPK285GUS: possui o gene *GUS/3' 35S* sob o controle do promotor de γ -kafirina deletado, contendo 285 pb; 3) construção pRT90GUS: contém a região proximal do promotor CaMV35S, que possui 90 pares de base, dirigindo a expressão do gene *gus*; construção pPK90GUS: possui o promotor da γ -kafirina com a região -8/-285 deletada e substituída pelos primeiros 90 pares de base do promotor CaMV35S dirigindo a expressão do gene *gus*; construção pRT103GUS: possui o gene *gus/3' 35S* sob o controle do promotor constitutivo CaMV 35S. As distâncias indicadas em cada construção são com relação ao ATG. Essas construções foram feitas utilizando-se o plasmídeo pRT103. As enzimas indicadas são: B, BamHI; E, EcoRI; EV, EcoRV; H3, HindIII; N, NcoI; P, PstI; Sp, SphI; H2, HindII e X, XbaI, sendo que são indicados somente os sítios relevantes.

Baseado nas observações acima descritas e no fato de que a polinização em milho pode ser controlada, os demais experimentos de expressão transitória foram realizados utilizando-se apenas tecidos de milho.

A tecido-especificidade do promotor foi testada através do bombardeamento em endosperma imaturo, coleóptilo, embrião e folha de milho. Estes tecidos foram bombardeados com diferentes construções, as quais consistem do gene indicador *gus* sob controle do fragmento de 1190 pb do promotor de γ -kafirina (pPKGUS, Fig. 13), e também bombardeado com uma versão deletada do promotor (pPK285GUS, Fig. 13). Na construção pPK285GUS o gene *gus* é controlado pelos primeiros 285 pb do promotor de γ -kafirina, ficando portanto excluídas as cópias do "endosperm box", e também uma das cópias da sequência GCN4, previamente indicados como possíveis elementos de regulação da expressão gênica das γ -prolaminas (Fig. 8). A Figura 15 mostra que o bombardeamento com o fragmento de 1190 pb do promotor resultou em máxima atividade relativa nos endospermas imaturos seguida da atividade relativa em coleóptilo e em embrião, que apresentaram uma redução da atividade relativa de cerca de 5 e 50 vezes, respectivamente, em relação à atividade relativa em endosperma. Cabe ressaltar que o bombardeamento do clone pRT103GUS em endosperma, coleóptilo, embrião e folha resultou em 98 (± 16), 829 (± 151), 618 (± 182) e 131 (± 54) pontos azuis por placa, sendo que estes valores foram utilizados para os cálculos de atividade relativa. A deleção da região distal do promotor de γ -kafirina, representada pela construção pPK285GUS, resultou em uma redução de aproximadamente 80 e 60% da atividade relativa em endosperma e coleóptilo, respectivamente, quando comparada com a atividade relativa obtida com o fragmento de 1190 pb do promotor (construção pPKGUS). Nenhuma atividade foi detectada nos embriões bombardeados com a mesma construção.

Com a finalidade de verificar se a redução da atividade observada era devido à retirada de elementos ativadores *cis* presentes na região deletada do promotor, foram também realizados experimentos de expressão transitória com uma construção quimérica envolvendo os promotores de γ -kafirina e de CaMV35S. A região complementar do promotor de γ -kafirina, que se encontra ausente na construção contendo o promotor deletado (pPK285GUS), foi adicionada aos 90 pb da região proximal do promotor de CaMV35S, de forma que o plasmídeo obtido foi denominado de pPK90GUS (Fig. 13). O efeito da adição do fragmento do promotor de γ -kafirina foi avaliada através da comparação da

atividade da construção contendo o promotor quimérico (pPK90GUS) com a atividade da construção contendo apenas a região proximal do promotor de CaMV35S (pRT90GUS). Os 90 pb do promotor de CaMV35S, além incluir os elementos "TATA box" e "CAAT box", considerados essenciais para o início da transcrição e que se encontram ausentes no fragmento do promotor de γ -kafirina adicionado a construção do promotor quimérico, inclui também a seqüência as-1 (Benfey & Chua, 1990). Os resultados deste experimento, representados na Fig. 15 mostram que a adição da região complementar do promotor de γ -kafirina promoveu a triplicação da atividade de GUS em endospermas imaturos e a duplicação nos outros tecidos, quando comparada com a atividade da região proximal do promotor de CaMV35S. A Figura 16 mostra um resultado típico obtido após a revelação histoquímica da atividade de GUS de endospermas imaturos de milho bombardeados com as construções pPKGUS (Fig. 16A), pRT103GUS (Fig. 16B) e pPK285GUS (Fig. 16C).

A

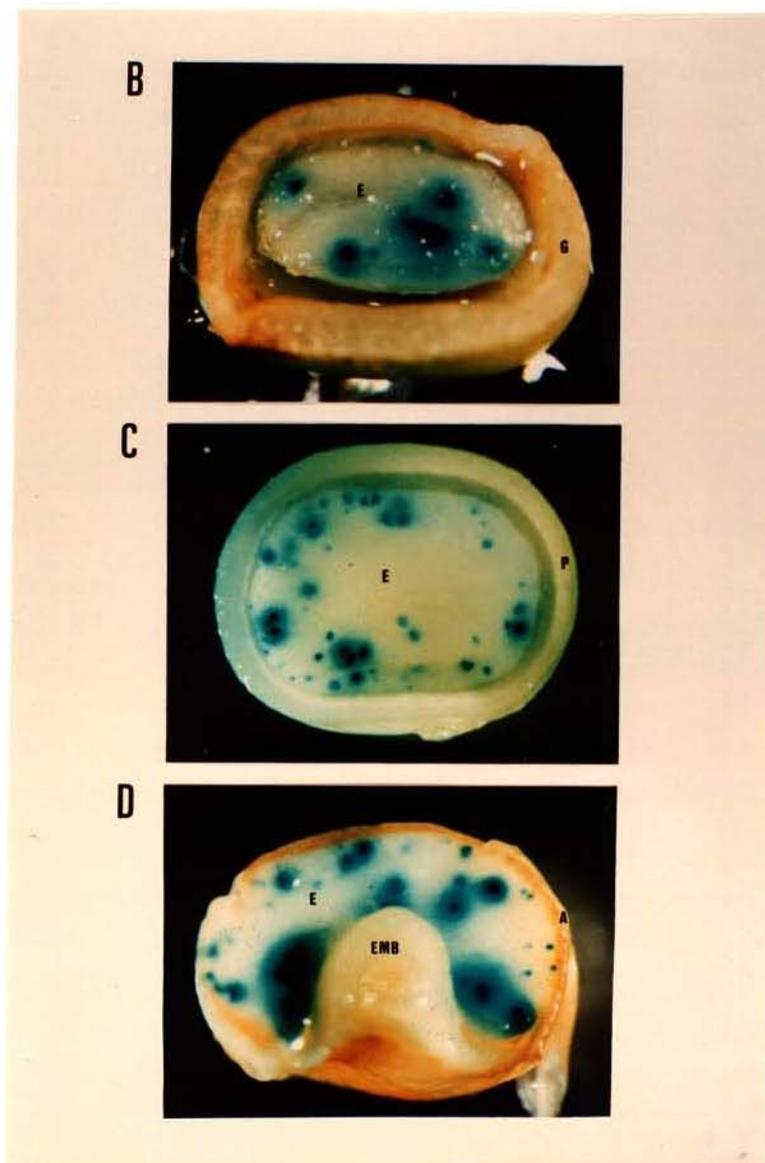
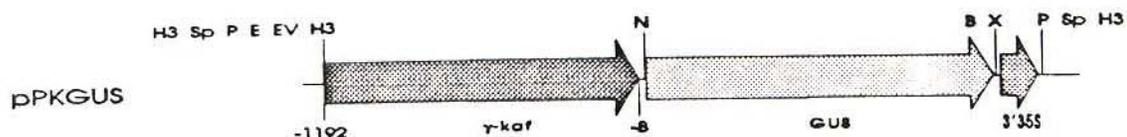


Fig. 14. A) Construção pPKGUS contendo a região do promotor de γ -kafirina, compreendida pela região -8/-1192, dirigindo a expressão do gene *gus*. Resultado do bombardeamento da construção pPKGUS em endosperma imaturo de B) sorgo C) milho e D) *Coix*.

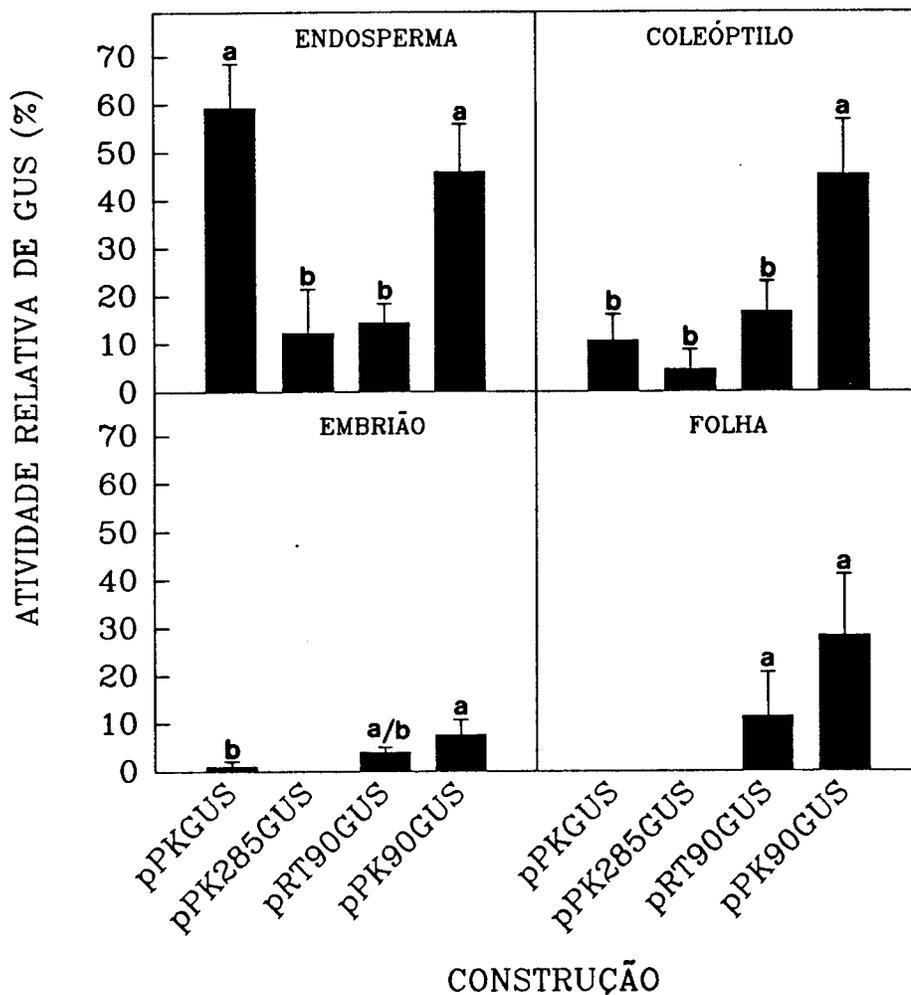


Fig. 15. Atividade relativa de cada construção após o bombardeamento em endosperma, embrião, coleóptilo e folha de milho. Os experimentos de bombardeamento foram feitos em triplicata (três placas por clone e dois disparos por placa). A atividade relativa representa a razão entre o número de pontos azuis resultantes do bombardeamento de cada construção, pelo número de pontos azuis resultantes do bombardeamento do clone pRT103GUS. As letras sobre as barras representam a separação de grupos obtida através do tratamento dos resultados pelo método SNK.

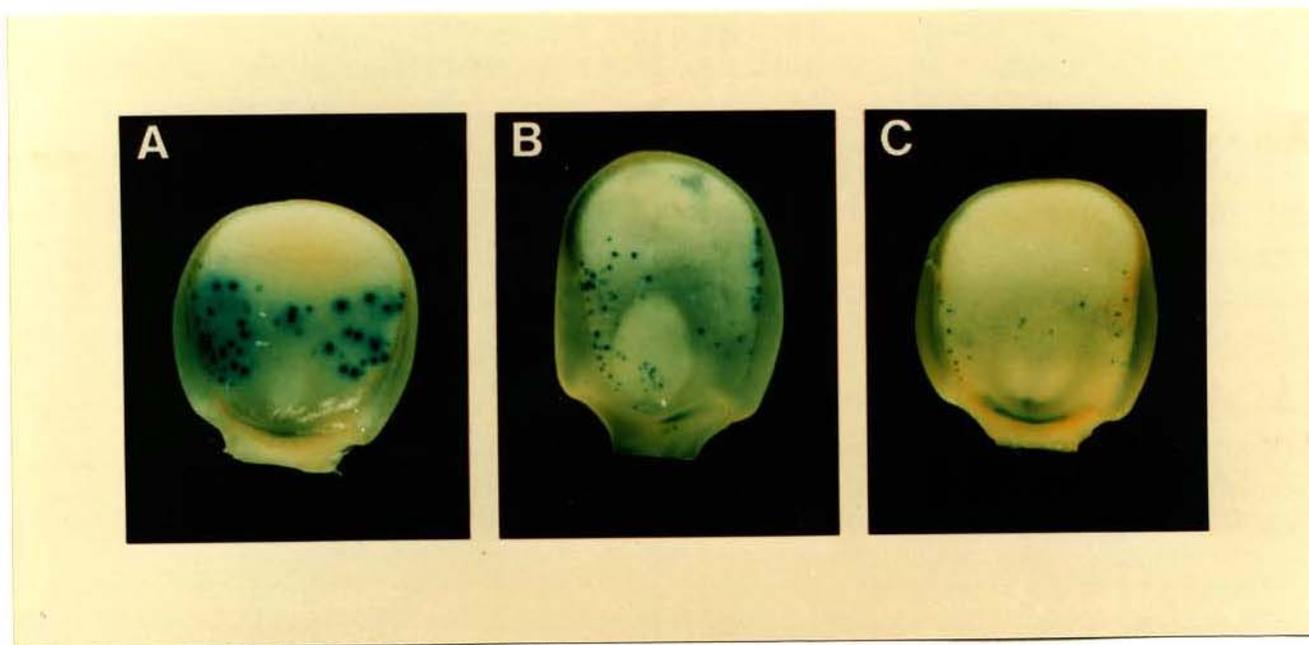


Fig. 16. Endosperma imaturo de milho (16DAP) bombardeado com as construções A) pPKGUS, B) pRT103GUS e C) pPK285GUS. A coloração azul dos pontos é resultado da degradação do substrato X-gluc pela enzima β -glucuronidase (GUS).

5. Discussão

As seqüências de aminoácidos das γ -prolaminas de milho, *Coix* e sorgo mostram características importantes na região N-terminal, apresentando seqüências repetidas formadas pelo hexapeptídeo PPPVHL e pelo octapeptídeo QPHPCPCQ, os quais parecem ter sido originados por eventos de duplicação. As diferenças existentes entre as γ -prolaminas de milho, *Coix* e sorgo, com relação ao número de repetições de cada polipeptídeo reflete-se na variação de peso molecular deduzido a partir da seqüência de aminoácidos para cada proteína. Os polipeptídeos deduzidos das seqüências de DNA das γ -zeínas de 28 kDa e 16 kDa, γ -kafirina e γ -coixina apresentam peso molecular igual a 21,8 kDa, 17,8 kDa, 20,3 kDa e 19,5 kDa, respectivamente. A γ -zeína de 16 kDa apresenta o menor peso molecular deduzido a partir da seqüência de aminoácidos, fato que pode ser explicado pelo reduzido número de repetições do hexapeptídeo PPPVHL e do octapeptídeo QPHPCPCQ, com relação as outras γ -prolaminas. A importância dessas seqüências repetidas na estrutura final de cada proteína e no seu empacotamento é desconhecida. Entretanto, essas seqüências podem ter algum papel importante na estrutura final e, juntamente com outros fatores como a presença de aminoácidos capazes de formar pontes dissulfeto, podem influenciar a distribuição dessas proteínas no interior do corpúsculo protéico. Essas seqüências parecem estar relacionadas com os epitopos reconhecidos pelos anticorpos utilizados no "Western blot", visto que a região compreendida por essas seqüências repetidas, é a região que apresenta menor similaridade entre a γ -zeína de 16 kDa e as outras γ -prolaminas. Experimentos de "Western blot", utilizando-se a fração total das prolaminas e anticorpos produzidos contra γ -zeína de 28 kDa e γ -coixina, mostraram que a γ -zeína de 16 kDa não foi reconhecida (resultados não apresentados). Este fato indica que as regiões correspondentes aos domínios repetidos podem ser responsáveis pelos epitopos antigênicos das proteínas γ -zeína de 28 kDa, γ -coixina e de γ -kafirina. A antigenicidade dessas proteínas poderia ocorrer pelo reconhecimento direto dos domínios repetidos pelos anticorpos ou ainda, pelas mudanças ocasionadas na estrutura final das proteínas, devido à presença desses domínios e a conseqüente exposição de regiões internas das outras γ -prolaminas, fato que não ocorreria na γ -zeína de 16 kDa.

O elevado teor de cisteínas apresentado pelas γ -prolaminas, aproximadamente 7%, aliado ao fato de que a solubilização destas proteínas depende da presença de agentes redutores, tais como 2-ME e ditioneitol (DTT),

indicam que as γ -prolaminas são capazes de promover a formação de um grande número de pontes dissulfeto. O alinhamento das γ -prolaminas de milho, *Coix* e sorgo (Fig. 10) mostra a conservação de resíduos de cisteínas em 10 posições diferentes, reforçando a importância destes resíduos na estrutura das proteínas. Em milho a localização preferencial das γ -prolaminas na periferia dos corpúsculos protéicos sugerem que esta classe de proteínas está envolvida na manutenção da integridade destes corpúsculos (Ludevid et al., 1984; Lending & Larkins, 1989). Porém em *Coix* constatou-se que as γ -prolaminas encontram-se dispersas, sem uma localização preferencial nos corpúsculos protéicos (Targon et al., 1992). Esses resultados sugerem que as γ -prolaminas podem apresentar algumas diferenças funcionais em diferentes espécies.

As γ -prolaminas de milho, *Coix* e sorgo apresentam considerável diferença entre os pesos moleculares calculados através do comportamento em gel de poliacrilamida na presença de SDS, e os obtidos através da seqüência deduzida de aminoácidos para a proteína madura. Wang & Esen (1986) deduziram que a grande diferença encontrada entre estes valores nos casos das γ -zeínas de 28 kDa, poderia ser explicada pelo alto teor de resíduos de aminoácidos carregados positivamente (histidina e arginina), que reduziriam a carga total negativa daquela proteína, reduzindo a sua velocidade de migração em um campo elétrico. Isto explicaria porque após o processamento do peptídeo sinal, a proteína madura da γ -kafirina possui peso molecular igual a 20,3 kDa, cerca de 7 kDa menor do que o peso molecular estimado por SDS-PAGE, o mesmo ocorrendo com a γ -coixina, que apresenta uma diferença de 2,5 kDa. Este fato não é verificado no caso da γ -zeína de 16 kDa, onde os valores de pesos moleculares obtidos através de SDS-PAGE e da seqüência deduzida de aminoácidos são muito próximos. Este comportamento pode estar relacionado com o baixo teor de aminoácidos com carga positiva.

Os experimentos de reconstrução realizados para a determinação do número de cópias dos genes de γ -kafirina e de γ -coixina, mostram, a exemplo do que ocorre para os genes de γ -zeína, que estes genes em sorgo e *Coix* encontram-se representados por uma ou duas cópias. Em algumas linhagens de milho foi demonstrada no *locus* da γ -zeína de 28 kDa, localizado no cromossomo 7, a presença de um alelo contendo uma duplicação deste gene (Das & Messing, 1987). Ao contrário, outras linhagens apresentam alelos contendo uma única cópia do gene, resultante da recombinação entre as cópias duplicadas (Das et al, 1991a; Das et al., 1991b). A variação encontrada na determinação do número de cópias por

genoma haplóide das γ -prolaminas citadas, oscilando entre uma a duas cópias, é consequência da dependência desse experimento na precisão da determinação da concentração dos DNAs utilizados. Portanto o resultado obtido no experimento de reconstrução realizado com o DNA genômico de *Coix*, onde se observa a presença de uma banda, pode indicar a presença de uma única cópia ou ainda duas cópias do gene apresentando sítios conservados para a enzima de restrição HindIII. No caso da γ -kafirina, o experimento de reconstrução resultou na obtenção de 3 bandas de tamanhos diferentes. Duas destas bandas apresentaram intensidade correspondente à aproximadamente uma cópia do gene, enquanto que a terceira banda de tamanho intermediário entre as outras duas hibridizou menos intensamente do que o controle que representa uma cópia. Possivelmente as bandas que hibridizaram mais intensamente correspondem a cópias do gene de γ -kafirina, enquanto que a terceira banda corresponde à uma seqüência que apresenta baixa similaridade com o gene de γ -kafirina, podendo corresponder aos clones recombinantes que apresentaram hibridização mais fraca durante o processo de seleção do gene de γ -kafirina do banco genômico de sorgo.

A seqüência de nucleotídeos do gene de γ -kafirina, clone 10S, é praticamente a mesma descrita para o cDNA correspondente, clone sorgH (Barros et al., 1991). As duas seqüências podem ser perfeitamente alinhadas, com exceção de seis alterações de base na região codificadora, sendo que cinco alterações são silenciosas (sinônimas) e uma produz a modificação do resíduo de prolina do cDNA para alanina no clone genômico. Essa modificação é resultante de uma transversão, onde o triplete GCT do clone genômico 10S encontra-se modificado para CCT no clone de cDNA sorgH. Cabe destacar que as 5 alterações silenciosas correspondem à transições, sendo que 3 envolvem mudanças de T para C ou vice-versa. A maior freqüência observada para esta transição pode ser explicada pela deaminação de 5-metilcitosinas, produzindo timidina, sendo que esta idéia é corroborada pelo fato de que a grande maioria das citosinas em plantas encontram-se metiladas. Outra diferença observada entre o clone genômico de γ -kafirina e o correspondente cDNA, é uma deleção de um T na região 3' do clone genômico.

A expressão das diversas classes de prolaminas é controlada de maneira coordenada temporalmente e espacialmente nos endospermas imaturos de milho. Resultados recentemente publicados sobre a acumulação das proteínas e mRNAs de prolaminas em endospermas imaturos de milho mantidos ou não em cultura, descrevem diferentes padrões de regulação para as diversas classes de prolaminas (Balconi et al., 1993). As α -zeínas de 22 kDa e β -zeínas apresentam um controle

de expressão bifásico. A primeira fase, correspondendo ao estágio até 14 DAP, quando a expressão destes genes é controlada através de um mecanismo relacionado aos níveis de nitrogênio. E uma segunda fase correspondente ao estágio após os 14-19 DAP, quando ocorre um grande aumento da transcrição dependente do ativador transcricional Opaco-2 (O2). Quanto às α -zeínas de 19 kDa, o nível de expressão parece manter-se constante durante as duas fases de desenvolvimento do endosperma, enquanto que a expressão de γ -zeínas apresenta um padrão de regulação completamente diferente dos anteriores, já que essas proteínas são produzidas somente na segunda fase, isto é, produzidas a partir dos 14 DAP em endospermas imaturos de milho. Estudos realizados com endospermas de *Coix* indicam também um acúmulo diferencial de proteínas e mRNAs das diferentes classes de prolaminas neste cereal (Targon et al., 1992). A γ -coixina, a exemplo da γ -zeína de 28 kDa, é produzida em estágios mais tardios do desenvolvimento do endosperma, quando comparados com as demais coixinas. Portanto, apesar das γ -prolaminas apresentarem um padrão de controle temporal ligeiramente diferente, as mesmas são expressas de uma forma coordenada com as demais prolaminas.

A expressão coordenada das diferentes classes de prolaminas indica a existência de fatores de ativação transcricionais comuns. Uma das seqüências reconhecidamente conservada nos promotores de prolaminas de diversos cereais, corresponde à seqüência do "endosperm box" (TGTAAG). Apesar deste fato ter sido descrito por Kreis et al. (1986), apenas recentemente a sua presença nos promotores de prolaminas de cereais foi correlacionada com a regulação gênica. Müller & Knudsen (1993) demonstraram através de experimentos de expressão transitória, que em cevada a atividade do promotor de hordeína-C em resposta aos níveis de nitrogênio, é mediada por seqüências de "endosperm box" em cooperação com seqüências similares à seqüência reconhecida pelo fator de ativação transcricional GCN4 de leveduras. A interação de fatores que reconhecem os elementos "endosperm box" e seqüências similares ao GCN4, foi descrita para uma prolamina de trigo (Hammond-Kosack et al., 1993).

Porém, a existência de um mecanismo que vincule a síntese das diferentes classes de prolaminas de milho, *Coix* e sorgo, ao suprimento de nitrogênio, bem como o papel desempenhado pelo "endosperm box", ainda não são conhecidos. Entretanto, a localização das seqüências de GCN4 em posições conservadas nos promotores dos genes da γ -prolaminas de milho e sorgo, bem como a presença de seqüências "endosperm box" (Fig. 8), constituem fortes indícios da existência de

um mecanismo de controle envolvendo tais seqüências. Os promotores dos genes de γ -zeína de 28 kDa e γ -kafirina possuem seqüências homólogas ao "endosperm box" presentes em duplicata e completamente homólogas, enquanto que a γ -zeína de 16 kDa apresenta apenas 2 versões modificadas. Uma das cópias presentes nos promotores de γ -zeína de 28 kDa e γ -kafirina encontra-se localizada em região equivalente à descrita para inúmeros genes de prolaminas de diversos cereais (Summer-Smith et al., 1985; Kreis et al., 1986; Boronat et al., 1986; Brown et al., 1986; Pedersen et al., 1986; Colot et al., 1987; DeRose et al., 1989; Hull et al., 1991; Liu & Rubenstein, 1992; Ottoboni et al., 1993).

Os experimentos preliminares de bombardeamento em sementes de sorgo, *Coix* e milho, demonstraram que o promotor de γ -kafirina foi capaz de dirigir a expressão do gene *gus* em endosperma das três espécies testadas (Fig. 14). Porém, a dificuldade em determinar, com precisão, o estágio de maturação das sementes de sorgo e *Coix*, bem como o fato de toda a metodologia de bombardeamento ter sido estabelecida para milho, levou-nos a utilizar esse cereal para a continuidade do trabalho. Dessa forma, os experimentos de bombardeamento foram conduzidos em tecidos de milho, sendo endosperma, embrião, coleóptilo e folha.

As atividades do promotor de γ -kafirina nos diversos tecidos foram comparadas através da atividade relativa, que foi calculada através da razão entre o número de pontos azuis obtidos pelo bombardeamento de uma construção contendo o gene *gus* dirigido pelo promotor constitutivo CaMV35S (pRT103GUS) e o número de pontos azuis obtidos pelo bombardeamento das construções restantes (Fig. 13A). A utilização da atividade relativa tem por objetivo eliminar os efeitos de área, viabilidade do tecido, dureza do tecido e competência das células para a transformação. Porém, os valores obtidos através desta razão incorpora a variação da atividade do promotor de CaMV35S nos diferentes tecidos. Os resultados de atividade relativa obtidos nos experimentos de expressão transitória do promotor de γ -kafirina nos diferentes tecidos, indicam que o promotor de γ -kafirina apresenta maior atividade em endosperma, seguido de coleóptilo e embrião.

A expressão do gene *gus* em tecidos diferentes do endosperma, quando este gene está sob controle do promotor da γ -kafirina, não era esperada, visto que muitos autores têm descrito a ausência de mRNA de zeínas em tecidos como embrião, folha e raiz (Sánchez-Mártinez et al., 1987; Bianchi & Viotti, 1988; Ueda & Messing, 1991). Uma primeira explicação para este comportamento está relacionada às variáveis da metodologia de bombardeamento com micropartículas.

Primeiramente, a quantidade de DNA utilizado para a precipitação com as micropartículas consiste em uma massa muito elevada, fazendo com que um grande número de cópias do gene de interesse seja adsorvido em cada partícula. Um elevado número de cópias do gene, pode fazer com que ocorra falta de algum fator de regulação, como um repressor, por exemplo, fato que permitiria que o promotor bombardeado ativasse a expressão do gene indicador de forma não específica. A tecido especificidade do promotor de γ -kafirina pode ter sido ainda afetada pela ausência de elementos regulatórios presentes em regiões mais distais do promotor. A presença de seqüências importantes para a transcrição em regiões distantes do ATG inicial foi demonstrada por Zheng et al. (1993), verificando-se que seqüências presentes entre as posições -1800 e -5000 do promotor de um gene de glutelina (proteína de reserva de arroz), clone Gt1, é responsável por um nível de atividade aproximadamente 20 vezes superior à observada para a região compreendida entre o ATG inicial e a posição -1800.

Um outro motivo pode ser o estado de metilação do DNA bombardeado, pois já foi demonstrado que as seqüências promotoras dos genes de zeínas apresentam um nível inferior de metilação no endosperma, ao contrário do que ocorre em outros tecidos. Esse padrão de metilação particular das zeínas em endosperma, concorda com a idéia de que a metilação seja mais um mecanismo de regulação gênica (Bianchi & Viotti, 1988; Gallardo et al., 1988). Até o momento, não se tem informações sobre o mecanismo de metilação do DNA introduzido através de bombardeamento com micropartículas, no sentido de esclarecer o quão rápido e eficiente seria este processo. Assim, o promotor introduzido, apresentando um diferente estado de metilação, poderia apresentar respostas inespecíficas, como por exemplo, dirigir a expressão em tecidos onde normalmente isto não ocorreria. Finalmente, as diferentes formas topológicas existentes entre os plasmídios bombardeados e o DNA cromossômico do milho, pode ser responsável por diferentes níveis de associação DNA/histonas. Sabendo-se que as histonas participam do processo de regulação gênica, este poderia ser mais um fator para explicar a atividade do promotor da γ -kafirina em tecidos diferentes do endosperma, pois talvez a associação das histonas com o plasmídio seja diferente do que ocorre com o DNA cromossômico do milho. Porém, o elevado nível de expressão do clone pPKGUS em coleóptilo não pode ser completamente explicado sem que sejam realizados experimentos com a finalidade de detectar mRNAs de γ -prolaminas neste tecido.

Os resultados de redução da atividade relativa obtido a partir da deleção

do promotor de γ -kafirina, e de ativação obtida através da complementação da região proximal do promotor de CaMV35S, indicam que a região distal do promotor de γ -kafirina, compreendendo a região de -285 até -1190 em relação ao ATG inicial, deve incluir elementos envolvidos na ativação da transcrição do gene de γ -kafirina. Esta hipótese é corroborada pela diferença de intensidade dos pontos azuis observada nos experimentos de bombardeamento em endospermas imaturos de milho (Fig. 15). O bombardeamento com a construção apresentando a versão deletada do promotor de γ -kafirina (pPK285GUS) produziu pontos azuis de menor intensidade do que o bombardeamento com a construção contendo a versão mais completa do promotor (pPKGUS). A região ausente na construção pPK285GUS inclui as duas cópias do "endosperm box" e uma das cópias da seqüência GCN4, bem como a região de 39 pb duplicada nas posições -1052 e -1162 do promotor de γ -kafirina. A ausência das repetições de 39 pb nos promotores das γ -zeínas, permite-nos deduzir que os efeitos observados nos experimentos de expressão transitória com o promotor de γ -kafirina, estão provavelmente relacionados com as seqüências "endosperm box" e GCN4. A redução no número de pontos azuis observada no endosperma bombardeado com a construção pPK285GUS, indica que somente as células mais viáveis apresentaram atividade de GUS.

Tem sido demonstrado que a expressão gênica das prolaminas de cereais é regulada primeiramente a nível transcricional, através de seqüências regulatórias presentes nas regiões flanqueadoras 5' e 3'. As seqüências regulatórias da região flanqueadora 5' respondem diretamente pela atividade do promotor, enquanto que as seqüências regulatórias da região 3' estão geralmente relacionadas com o processamento e estabilidade do mRNA. O alinhamento das regiões flanqueadoras 3' dos genes de γ -prolaminas de sorgo, *Coix* e milho, e a comparação com seqüências regulatórias já descritas, permitiram-nos identificar possíveis seqüências envolvidas no processamento da região 3' dos mRNAs transcritos por estes genes.

Assim como ocorre na maioria dos mRNAs de eucariotos, grande parte dos mRNAs de plantas apresentam adição de cauda poli(A) na extremidade 3' (Proudfoot, 1991). A adição da cauda de poli(A) em mRNAs de mamíferos parece apresentar uma grande dependência na conservação da seqüência AATAAA, enquanto que o sinal de poliadenilação em plantas pode apresentar maiores variações (Joshi, 1987). Assim sendo, seqüências apresentando similaridade com o sinal de poliadenilação de mamíferos foram encontrados nas regiões identificadas como SP-1, SP-2, SP-3, SP-4 e SP-5 na Fig. 12. Wu et al. (1993) demonstraram

que a seqüência AATGAA (região SP-5) funciona como sinal de poliadenilação no mRNA de γ -zeína de 28 kDa, ainda que a seqüência AATGAA localizada 13 nucleotídeos acima (região SP-4), também possa determinar o posicionamento do sítio de adição da cauda poli-A. Evidências para este fato foram obtidas através do sequenciamento dos clones de cDNA de γ -coixina, clones pBCX22.4, pBCX22.5 e pBCX22.9 (Fig. 3). O clone pBCX22.4 apresentou o sítio de poliadenilação localizado aproximadamente 10 nucleotídeos abaixo da seqüência AATGAC (região SP-3), indicando que esta seqüência provavelmente esteja funcionando como sinal de poliadenilação. A conservação desta seqüência nos genes de γ -prolaminas de milho, *Coix* e sorgo é um forte indício de que a mesma possa estar relacionada com o processamento do mRNA. Os clones pBCX22.5 e pBCX22.9 apresentam sítios de poliadenilação localizados em posições mais distais com relação ao término de tradução, sugerindo que as seqüências GATGAA e AATGAA (regiões SP-4 e SP-5, respectivamente, Fig. 13) podem ser responsáveis pela determinação do sítio de adição da cauda poli-A, ainda que as seqüências vizinhas também possam estar relacionadas com este processo. Essas variações sugerem a existência de um mecanismo flexível no processamento do mRNA dos genes considerados, o que é confirmado pela presença da seqüência SR-1 nos genes de γ -zeínas de 28 kDa (CAGTGTCAGTGT) e 16 kDa (CAGTGTGAGC), ao passo que os genes de γ -kafirina e γ -coixina apresentam apenas versões modificadas da mesma.

O presente trabalho representa a primeira etapa no estudo do controle de expressão dos genes de γ -prolaminas. Muito do que se conhece hoje, com relação ao controle de expressão gênica de prolaminas em endosperma de cereais da tribo Andropogoneae, está relacionado com o isolamento do fator de transativação codificado pelo gene de opaco-2 de milho (Schmidt et al., 1992; Lohmer S. et al., 1991). A proteína O-2 apresenta domínio característico da família das "leucine zipper" e constitui-se em um ativador transcricional para as α -prolaminas. No caso das γ -prolaminas, a indisponibilidade de mutantes mais específicos para esta classe de proteínas, determina que o estudo do controle de expressão seja realizado, em um primeiro estágio, através da comparação das seqüências das possíveis regiões regulatórias. Seqüências conservadas nos vários genes isolados das diferentes espécies, podem representar possíveis elementos de regulação *cis*, tais como sítios de ligação para fatores de transativação.

A importância destes sítios putativos de regulação pode ser ainda testada *in vivo*, através de experimentos de expressão transitória utilizando seqüências

modificadas através de deleções ou mutações. O segundo estágio no estudo da regulação gênica, compreende a utilização das seqüências dos sítios comprovadamente envolvidos nos processos moleculares de regulação e no isolamento de fatores capazes de reconhecer especificamente tais sítios. Para tanto, seqüências destes sítios podem ser utilizadas como sondas na seleção de clones recombinantes em bibliotecas de expressão construídas a partir de mRNA do endosperma em desenvolvimento.

O estudo da regulação deverá ter continuidade com a identificação de seqüências regulatórias, envolvidas com a tecido especificidade, através da obtenção de um número maior de modificações, e abrangendo uma maior extensão do promotor a ser utilizado em experimentos de expressão transitória.

Cabe destacar que a utilização do promotor de γ -kafirina na construção de vetores para expressão de proteínas heterólogas, especificamente em endosperma, bem como a expressão de γ -kafirina em milho, vem sendo realizada no Laboratório de Biologia Molecular de Plantas do CBMEG.

7. CONCLUSÕES

1) O alto grau de homologia existente entre as cadeias de polipeptídeos das γ -prolaminas de milho, *Coix* e sorgo fornecem evidências moleculares sobre a existência de um ancestral comum entre esses cereais.

2) A homologia entre os genes de γ -prolaminas de milho e sorgo estende-se às regiões regulatórias dos mesmos, indicando que estes possuem mecanismos de regulação semelhantes, os quais foram conservados após a divergência das espécies. A existência de mecanismos de regulação semelhantes é corroborada pela presença de seqüências regulatórias conhecidas, tais como as seqüências correspondentes ao "endosperm box" e GCN4, e também pelos resultados de expressão transitória que confirmam a atividade do promotor de γ -kafirina nos endospermas de milho, *Coix* e sorgo.

3) Os resultados obtidos através dos experimentos de expressão transitória sugerem a existência seqüências presentes acima da posição -285, com relação ao ATG, envolvidas com a otimização da expressão do gene da γ -kafirina.

4) A atividade do promotor de γ -kafirina em tecidos diferentes do endosperma, tais como coleóptilo e embrião pode ser explicada da seguinte maneira:

a) Inespecificidade relativa do promotor de γ -kafirina, que pode ser uma característica inerente do mesmo, ou devido à ausência de seqüências regulatórias normalmente encontradas em regiões mais distais não incluídas na seqüência testada do promotor.

b) Diminuição da tecido-especificidade do promotor de γ -kafirina em experimentos de expressão transitória, quando utilizou-se tecidos de milho

c) A quantidade de DNA utilizado para a precipitação com as micropartículas, consiste em uma massa muito elevada, fazendo que um grande número de cópias do gene de interesse seja precipitado em cada micropartícula.

d) O estado de metilação do DNA introduzido pode ser diferente daquele necessário para uma expressão temporal e/ou tecido-específica correta, fazendo com que o promotor apresente atividade em tecidos onde normalmente não apresentaria.

e) As diferentes formas topológicas existentes entre os plasmídios bombardeados e o DNA cromossômico do milho, pode ser responsável por diferentes níveis de associação DNA/histonas. Dessa forma, sabendo-se que as histonas participam da regulação gênica, este pode ser mais um fator para explicar a atividade do promotor da γ -kafirina em tecidos diferentes do endosperma.

4) A partir do alinhamento da região 3' dos genes de γ -prolaminas de milho, Coix e sorgo, foi possível verificar a conservação na distribuição de possíveis sinais de poliadenilação entre esses genes, sendo reconhecidas 5 regiões distintas denominadas de SP-1, SP-2, SP-3, SP-4, SP-5 e SR-1. Através do seqüenciamento de clones de cDNA de Coix, foi possível confirmar a existência de um mecanismo variável no processamento do mRNA deste gene, sendo que os sinais de poliadenilação mais distais, com relação ao término de tradução, parecem ser os responsáveis pela adição da cauda poli-A.

8. BIBLIOGRAFIA

- Arora, R.K. - Job's Tears (*Coix lacryma-jobi*) - A minor food and fodder crop from North Eastern India. *Econ. Bot.* **31**:358-366 (1977).
- Balconi, C; Rizzi, E.; Motto, M.; Salamini, F.; Thompson, R. - The accumulation of zein polypeptides and zein mRNA in cultured endosperms of maize is modulated by nitrogen supply. *Plant Mol. Biol.* **3**:325-334 (1993).
- Barros, E.G; Takasaki, K.; Kirleis, A.W.; Larkins, B.A. - Nucleotide sequence of a cDNA clone encoding γ -kafirin protein from *Sorghum bicolor*. *Plant Physiol.* **97**:1606-1607 (1991).
- Benfey, P.N. & Chua, N.-H. - The cauliflower mosaic virus 35S promoter: combinatorial regulation of transcription in plants. *Science* **250**:959-966 (1990).
- Bianchi, M.W. & Viotti, A. - DNA methylation and tissue-specific transcription of the storage protein genes of maize. *Plant Mol. Biol.* **11**:203-214 (1988).
- Boronat, A.; Martínez, M.C.; Reina, M.; Puigdomènech, P.; Palau, J. - Isolation and sequencing of a 28kD glutelin-2 gene from maize common elements in the 5' flanking regions among zein and glutelin. *Plant Science* **47**:95-102 (1986).
- Brown, J.W.S.; Wandelt, C.; Feix, G.; Neuhaus, G.; Schweiger, H.G. The upstream regions of zein genes - Sequence analysis and expression in the unicellular alga *Acetabularia*. *Eur. J. Cell. Biol.* **42**:161-170 (1986).
- Bullock, W.O.; Fernandez, J.M.; Short, J.M. - XL1-Blue: A high efficiency plasmid transforming recA *Escherichia coli* strain with beta-galactosidase selection. *Biotechniques* **5**:376-379 (1987).
- Burr, B. & Burr, F.A. - Three mutations in *Zea mays* affecting zein accumulation: A comparison of zein polypeptides, *in vitro* synthesis and processing, mRNA levels and genomic organization *J. Cell Biol.* **94**:201-206 (1982).
- Cabot, E. - The Eyeball Sequence Editor. (programa obtido a partir do banco de dados de seqüências EMBL). Email: NETSERV@EMBL-Heidelberg.DE (1988)

- Colot, V.; Robert, L.S.; Kavanagh, T.A.; Bevan, M.V.; Thompson, R.D. - Localization of sequences in wheat endosperm protein genes which confer tissue-specific expression in tobacco. *EMBO J.* 6:3559-3564 (1987).
- Danner, D.B. - Recovery of DNA fragments from gels by transfer to DEAE-paper in electrophoresis chamber. *Anal. Biochem.* 125:139-142 (1982).
- Das, O.P. & Messing, J.W. - Allelic variation and differential expression at the 27-kilodalton zein locus in maize. *Mol. Cell. Biol.* 7:4490-4497 (1987).
- Das, O.P.; Ward, K.; Ray, S.; Messing, J. - Sequence variation between alleles reveals two types of copy correction at the 27-kDa zein locus of maize. *Genomics* 11:849-856 (1991a).
- Das, O.P.; Poliak, E.; Ward, K.; Messing, J. - A new allele of the duplicated 27kD zein locus of maize generated by homologous recombination. *Nucl. Acids Res.* 19:3325-3330 (1991b).
- DeRose, R. T.; Ma, D-P.; Kwon, E-S.; Hasnain, S.E.; Klassy, R.C.; Hall, T.C. - Characterization of the kafirin gene family from sorghum reveals extensive homology with zein from maize. *Plant Mol. Biol.* 12:245-256 (1989).
- Diaz, N.M. - Caracterização e utilização do *Coix* em panificação. *Alimentos e Tecnologia* 4:17-18 (1985).
- Dretzen, G.; Bellard, M.; Sassone-Corsi, P.; Chambon, P. - A reliable method for the recovery of DNA fragments from agarose and acrylamide gels. *Anal. Biochem.* 112:295-298 (1981).
- Esen, A.; Bietz, J.A.; Paulis, J.W.; Wall, J.S. - Tandem repeats in the N-terminal sequence of a proline-rich protein from corn endosperm. *Nature* 296:678-679 (1982).
- Esen, A. - Separation of alcohol-soluble proteins (zeins) from maize into three fractions by differential solubility. *Plant Physiol.* 80:623-627 (1986).
- Esen, A. & Stetler, D.A. - Immunocytochemical localization of δ -zein in the protein bodies of maize endosperm cells. *Am. J. Bot.* 79:243-248 (1992).
- Feinberg, A.P. & Vogelstein, B. - A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132:6-13 (1983).

- Feinberg, A.P. & Vogelstein, B. - A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Addendum Anal. Biochem.* 137:266-267 (1984).
- Frischauf, A-M.; Lehrach, H.; Poustka, A.; Murray, N. - Lambda replacement vectors carrying polylinker sequences. *J. Mol. Biol.* 170:827-842 (1983).
- Gallardo, D.; Reina, M.; Rigua, J.; Boronat, A.; Palau, J. - Genomic organization of the 28 kDa glutelin-2 gene from maize. *Plant Sci.* 54:211-218 (1988).
- Garrat, R.; Olliva, G.; Caracelli, I.; Leite, A.; Arruda, P. - Studies of the zein-like α -prolamins based on an analyses of amino acid sequences: implications for their evolution and three-dimensional structure. (submetido à publicação)
- Gianazza, E.; Righetti, P. G.; Pioli, F.; Galante, E.; Soave, C. - Size and charge heterogeneity of zein in normal and *opaque-2* maize endosperms. *Maydica* 21:1-17 (1976).
- Geetha, K.B.; Lending, C. R.; Lopes, M.A.; Wallace, J.C.; Larkins, B.A. - *Opaque-2* modifiers increase γ -zein synthesis and alter its spatial distribution in maize endosperm. *Plant Cell* 3:1207-1219 (1991).
- Goldsbrough, P.B.; Gelvin, S.B.; Larkins, B.A. - Expression of maize zein genes in transformed sunflower cells. *Mol. Gen. Genet.* 202:374-381 (1986).
- Hagen, G. & Rubenstein, I. - Complex organization of zein genes in maize. *Gene* 13:239-249 (1981).
- Hammond-Kosack, M. C. U.; Holdsworth, M. J.; Bevan, M. W. - *In vivo* footprinting of a low molecular weight glutenin gene (LMWG-1D1) in wheat endosperm. *EMBO J.* 12:545-554 (1993).
- Hanaham, D. - Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 166:557 (1983).
- Harlan, J.R. - Indigenous african agriculture. In: "The origins of agriculture - an international perspective", pp. 59-70, ed. C. Wesley Cowan & Patty Jo Watson, Smithsonian Institution Press, Washington, 1992.
- Higgins, D.G & Sharp, P.M. - Fast and sensitive multiple sequence alignments on a microcomputer. *CABIOS*, 5:151-153 (1989) (programa obtido a partir do banco de dados de seqüências EMBL). Email: NETSERV@EMBL-Heidelberg.DE

- Higgins, T. J. V. - Synthesis and regulation of major proteins in seeds. *Ann. Rev. Plant Physiol.* **35**:191-221 (1984).
- Hill, D.E.; Hope, I.A.; Macke, J.P.; Struhl, K. - Saturation mutagenesis of the yeast *his3* regulatory site: Requirements for transcriptional induction and for binding by GCN4 activator protein. *Science* **234**:451-457 (1986).
- Hull, G.A.; Halford, N.G.; Kreis, M.; Shewry, P.R. - Isolation and characterization of genes encoding rye prolamins containing a highly repetitive sequence motif. *Plant Mol. Biol.* **17**:1111-1115 (1991).
- Jefferson, R.A. - Assaying chimeric genes in plants: The *gus* gene fusion system. *Plant Mol. Biol.* **5**:387-405 (1987).
- Joshi, C.P. - Putative polyadenylation signals in nuclear of higher plants: a compilation and analysis. *Nucl. Acids Res.* **15**: 9627-9640 (1987).
- Kirihara, J.A.; Husperger, J.P.; Mahoney, J.W. - Differential expression of a gene for a methionine-rich storage protein in maize. *Mol. Gen. Genet.* **211**:477-484 (1988).
- Kodrzycki, R.; Boston, R.S.; Larkins, B.A. - The *opaque-2* mutation of maize differentially reduces zein gene transcription. *Plant Cell* **1**:105-114 (1989).
- Kreis, M.; Williamson, M.S.; Forde, J.; Schmutz, D.; Clark, J.; Buxton, B.; Pywell, J.; Marris, C.; Henderson, J.; Harris, N.; Shewry, P.R.; Forde, B.G.; Miflin, B.J. - Differential gene expression in the developing barley endosperm. *Phil. Trans. R. Soc. Lond.* **314**:355-356 (1986).
- Laemmeli, U.K. - Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**:680-684 (1970).
- Landry, J. & Moureaux, T. H. - Distribution and amino acid composition of protein groups located in different histological parts of maize grain. *J. Agric. and Food Chem.* **28**:1186 (1980).
- Langridge, P.; Eibel, H.; Brown, J.W.S.; Feix, G. - Transcription from maize storage protein gene promoters in yeast. *EMBO J.* **3**:2467-2471 (1984).
- Larkins, B. A. - In: *The Biochemistry of Plants*, Stumpf, P. K.; Conn, E. E. (eds), Academic Press, New York, 1981, pp. 449-489.
- Leite, A.; Ottoboni, L.M.M.; Targon, M.L.P.N.; Silva, M.J.; Arruda, P. - Phylogenetic relationship of zeins and coixins as determined by

- immunological cross-reactivity and Southern blot analysis. *Plant Mol. Biol.* 14:743-751 (1990).
- Leite, A.; Freitas, F.A.; Yunes, J.A.; Arruda, P. - Nucleotide sequence of a cDNA clone encoding γ -coixin from *Coix lacryma-jobi* seeds. *Plant Physiol.* 97:1604-1605 (1991).
- Leite, A.; Yunes, J.A.; Turcinelli, S.R.; Arruda, P. - Cloning and characterization of a cDNA encoding a sulfur-rich coixin. *Plant Mol. Biol.* 18:171-174 (1992).
- Lending, C.R.; Kriz, A.L.; Larkins, B.A.; Bracker, C.E. - Structure of maize protein bodies and immunocytochemical localization of zeins. *Protoplasm* 143:51-62 (1988).
- Lending, C.R. & Larkins, B.A. - Changes in the zein composition of protein bodies during maize endosperm development. *The Plant Cell* 1:1011-1023 (1989).
- Liu, C-N. & Rubenstein, I. - Molecular characterization of two types of 22 kilodalton α -zein genes in a gene cluster in maize. *Mol. Gen. Genet.* 234:244-253 (1992).
- Lohmer, S.; Maddaloni, M.; Motto, M.; Di Fonzo, N.; Hartings, H.; Salamini, F.; Thompson, R.D. - The maize regulatory locus *opaque-2* encodes a DNA-binding protein which activates the transcription of the *b-32* gene. *EMBO J.* 10:617-624 (1991).
- Ludevid, M. D.; Torrent, M.; Martinez-Izquierdo, J. A.; Puigdomènech, P.; Palau, J. - Subcellular localization of glutelin-2 in maize (*Zea mays* L.) endosperm. *Plant Mol. Biol.* 3:227-234 (1984).
- Matzke, M.A.; Susani, M.; Binns, A.N.; Lewis, E.D.; Rubenstein, I.; Matzke, A.J.M. - Transcription of a zein gene introduced into sunflower using a Ti plasmid vector. *EMBO J.* 3:1525-1531 (1984).
- Marks, M.D.; Lindell, J.S.; Larkins, B.A. - Nucleotide sequence analysis of mRNAs from maize endosperm. *J. Biol. Chem.* 260:16451-16459 (1985).
- Messing, J. - New M13 vectors for cloning. *Methods Enzimol.* 101:20 (1983).
- Moreno, M.R.; Smith, J.F.; Smith R.V. - Silver staining of proteins in polyacrylamide gels: Increased sensibility through a combined Comassie Blue-silver stain procedure. *Anal. Biochem.* 151:466-470 (1985).

- Motto, M.; Maddaloni, M.; Ponziani, G.; Brembilla, M.; Marotta, R.; Di Fonzo, N.; Soave, C.; Thompson, R.; Salamini, F. - Molecular cloning of the o2-m5 allele of *Zea mays* using transposon marking. *Mol. Gen. Genet.* 212:488-494 (1988).
- Murashige, T. & Skoog, F. - A revised medium for rapid growth and bioassays with tobacco tissue culture. *Physiol. Plant.* 15: 473-497 (1962).
- Müller, M. & Knudsen, S. - The nitrogen response of a barley C-hordein promoter is controlled by positive and negative regulation of the GCN4 and endosperm box. *The Plant Journal* 4:343-355 (1993).
- Nkonge, C. & Ballance, G.M. - A sensitive colorimetric procedure for nitrogen determination in micro-Kjeldal digests. *J. Agric. Food Chem.* 30:416-420 (1982).
- Norlander, J.; Kempe, T.; Messing, J. - Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. *Gene* 26:101 (1983).
- Osborne, T. B. & Mendel, L. B. - Nutritive properties of proteins of maize kernel. *J. Biol. Chem.* 18:1-6 (1914).
- Ottoboni, L. M. M.; Leite, A.; Targon, M.L.N.; Crozier, A.; Arruda, P. - Characterization of the storage protein in seed of *Coix lacryma-jobi* var. Adlay. *J. Agric. Food Chem.* 38:631-635 (1990a).
- Ottoboni, L.M.M.; Leite, A.; Targon, M.L.P.N.; Silva, M.J.; Arruda, P. - Heterogeneity of *Coix*, maize and teosinte prolamins detected by isoelectric focusing. *Rev. Bras. Genet.* 13:313-22 (1990b).
- Ottoboni, L.M.M.; Leite, A.; Yunes, J.A.; Targon, M.L.P.N.; Souza Filho, G.A.; Arruda, P. - Sequence analysis of 22 kDa-like α -coixin genes and their comparison with homologous zein and kafirin genes reveals highly conserved protein structure and regulatory elements. *Plant Mol. Biol.* 21:765-778 (1993).
- Pedersen, K.; Argos, P.; Naravana, S.L.V.; Larkins, B.A. - Sequence analysis and characterization of a maize gene encoding a high-sulfur zein protein of M_r 15000. *J. Biol. Chem.* 14:6279-6284 (1986).
- Plikaytis, B.D.; Carlone, G.M.; Edwards, P.; Mayer, W. - Robust estimation of standard curves for protein molecular weight and linear-duplex DNA based pair number after gel electrophoresis. *Anal. Biochem.* 152:346-364 (1986).

- Pernollet, J.C. & Mossé, J. - Structure and location of legume and cereal seed proteins. In: Seed Proteins, Daussant, J.; Mossé, J.; Vaughan, J. (eds), Academic Press, London, pp. 155-191 (1983).
- Prat, S.; Cortadas, J.; Puigdomènech, P.; Palau, J. - Nucleic Acid (cDNA) and amino acid sequences of the maize endosperm protein glutelin-2. *Nucleic Acids Res.* 13:1493-1504 (1985).
- Prat, S.; Pérez-Grau, L.; Puigdomènech, P. - Multiple variability in the sequence of a family of maize endosperm proteins. *Gene* 52:41-49 (1987).
- Proudfoot, N. - Poly (A) signals. *Cell* 64:671-674 (1991).
- Proudfoot, N.J. & Brownlee, G.G. - 3' Non-coding region sequences in eukaryotic messenger RNA. *Nature* 263:211-214 (1976).
- Reina, M.; Guillén, P.; Ponte, I.; Boronat, A.; Palau, J. - DNA sequence of the gene encoding the Zc1 protein from *Zea mays* W64 A. *Nucl. Acids Res.* 18:6425 (1990a).
- Reina, M.; Ponte, I.; Guillén, P.; Boronat, A.; Palau, J. - Sequence analysis of a genomic clone encoding a Zc2 protein from *Zea mays* W64 A. *Nucl. Acids Res.* 18:6426 (1990b)
- Rivin, C.J.; Zinner, E.A.; Walbot, V. - Isolation of DNA and DNA recombinants from maize. In: Sheridan, W.F.(ed) *Maize for Biological Research*, pp. 161-164. University Press, University of North Dakota, Grand Forks, (1982).
- Sambrook, J.; Fritsch, E.F.; Maniatis, T.- *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory (2nd ed.), Cold Spring Harbor, New York, 1989.
- Sánchez--Mártinez, D.; Gomes, J.; Ludevid, M.D.; Torrent, M.; Puigdomènech, P.; Pages, M. - Absence of storage protein synthesis in the embryo of *Zea mays*. *Plant Sci.* 53:215-222 (1987).
- Sanfaçon, H.; Brodmann, P.; Hohn, T. - A dissection of the cauliflower mosaic virus polyadenylation signal. *Genes Devel.* 5:141-149 (1991).
- Schernthaner, J.P.; Matzke, M.A.; Matzke, A.J.M. - Endosperm-specific activity of a zein gene promoter in transgenic tobacco plants. *EMBO J.* 7:1249-1255 (1988).

- Schmidt, R.J.; Burr, F.A.; Burr, B. - Transposon tagging and molecular analysis of the maize regulatory locus *opaque-2*. *Science* 238:960-63 (1987).
- Schmidt, R.J.; Ketudat, M.; Aukerman, M.J.; Hoschek, G. - *Opaque-2* is a transcriptional activator that recognizes a specific target site in 22-kD zein genes. *Plant Cell* 4:689-700 (1992).
- Schwall, M. & Feix, G. - Zein promoter activity in transiently transformed protoplasts from maize. *Plant Sci.* 56:161-166 (1988).
- Shaaafhawsen, R.V. - Adlay or Job's tears - a cereal of potentially greater economic importance. *Econ. Bot.* 6:216-227, 1952.
- Shewry, P.R. & Tatham, A.S. - The prolamin storage of cereal seeds: structure and evolution. *Biochem. J.* 267:1-12 (1990).
- Short, J.M.; Fernandez, J.M.; Sorge, J.A.; Huse, W.D. - 1 ZAP : A bacteriophage 1 expression vector with *in vivo* excision properties. *Nucl Acids Res.* 16:7583 (1988).
- Shull, J.M. & Kirleis, A.W. - Localization of alfa and beta-kafirins in *Sorghum bicolor*. *Cereal Foods World* 34:784 (1989).
- Small E. - Hybridization in the domesticated-weed-wild-complex. In: Plant Biosystematics, Grant, W. F. (ed), Academic Press, London, pp. 195-210 (1984).
- So, J-S. & Larkins, B.A. - Binding of an endosperm-specific nuclear protein to a maize beta-zein gene correlates with zein transcriptional activity. *Plant Mol. Biol.* 17:309-319 (1991).
- Steel, R.G.D. & Torrie, J.H. - Principles and procedures of statistics, a biometrical approach (2nd ed), McGraw-Hill, New York, 1980, pp.186-187.
- Su, J. C. & Preiss, J. - Purification and properties of structure of sucrose synthase from maize kernels. *Plant Physiol.* 61:389 (1978).
- Summer-Smith, M.; Rafalski, J.A.; Sugyama, T.; Stoll, M.; Söll, D. - Conservation and variability of wheat α/β -gliadin genes. *Nucl. Acids Res.* 13:3905-3916 (1985).
- Targon, M.L.N.; Ottoboni, L.M.M.; Leite, A.; Ludevid, D.; Puigdomènech, P.; Arruda, P. Synthesis and deposition of coixin in seeds of *Coix lacryma-jobi*. *Plant Sci.* 83:169-180 (1992).

- Töpfer, R.; Matzeit, V.; Gronenborn, B.; Schell, J.; Steinbiss, H-H. - A set of plant expression vectors for transcriptional and translational fusions. *Nucl. Acids Res.* 15:5890 (1987)
- Tosello, G.A. - Milhos Especiais e seu valor nutritivo. Em: Paterniani, E. & Viégas, G. P. (eds), *Melhoramento e Produção do Milho*, Fundação Cargill, Campinas, pp. 375 - 402 (1987).
- Ueda T. & Messing, J. - A homologous expression system for cloned zein genes. *Theor. Appl. Genet.* 82:93-100 (1991).
- Venkateswarlu, J. & Chaganti, R.S.K. -Job's Tears (*Coix lacryma-jobi* L.) ICAR Tech. Bull. (Agric.) 43-44: pp. 1-54 (1973).
- Wandelt, C. & Feix, G. - Sequence of a 21 kDa zein gene from maize containing an in frame stop codon. *Nucl. Acids Res.* 17:2354 (1989).
- Wang, S-Z & Esen, A.- Primary structure of a prolamine-rich zein and its cDNA. *Plant Physiol.* 81:70-74 (1986).
- Whittaker, P.A.; Campbell, Southern, E.M.; Murray, N.E. - Enhanced recovery and restriction mapping of DNA fragments cloned in a new λ vector. *Nucl. Acids Res.* 16:6725-6736 (1988).
- Williamson, J.D.; Galili, G.; Larkins, B.A.; Gelvin, S.B. - The synthesis of a 19 kilodalton zein protein in transgenic *Petunia* plants. *Plant Physiol.* 88:1002-1007 (1988).
- Wilson, C. M.; Shewry, P. R.; Mifflin, B. J. - Maize endosperm proteins compared by SDS-gel electrophoresis and isoelectric focusing. *Cereal Chem.* 58:275-281 (1981).
- Wu, L.; Ueda, T.; Messing, J. - 3'-end processing of the maize 27 kDa zein mRNA. *The Plant Journal* 4:535-544 (1993).
- Yanisch-Perron, C.; Vieira, J.; Messing, J. - Improved M13 phage cloning vectors and hosts strains: Nucleotide sequences of the M13mp18 and pUC 19 vectors. *Gene* 33:103 (1985).
- Ye, G-N; Daniell, H. & Sanford, J.C. - Optimization of delivery of foreign DNA into higher plant chloroplasts. *Plant Mol. Biol.* 15:809-819 (1990).
- Yunes, J.A.; Cord Neto, G.; Silva, M.J.; Leite, A.; Ottoboni, L.M.M. & Arruda, P.- The transcriptional activator Opaque-2 recognizes two completely

different target sequences in the 22 kDa-like α -prolamin genes (submetido à publicação).

Zheng, Z.; Kawagoe, Y.; Xiao, S.; Li, Z.; Okita, T.; Hau, T. L.; Lin, A.; Murai, N.
- 5' distal and proximal *cis*-acting regulator elements are required for developmental control of a rice seed storage protein *glutelin* gene. *The Plant Journal* 4:357-366 (1993).

APÊNDICE I

A) XL1-Blue:

A linhagem lac^- , supressora de mutação ambar (*supE44*), defectiva em recombinação (*recA1*), deficiente em restrição e proficiente em metilação no sistema EcoK (*hsdR17*), foi utilizada como hospedeira do bacteriófago λ -ZAP durante a seleção de clones recombinantes da biblioteca de cDNA de *Coix* e nos experimentos de excisão dos "phagemids" dos clones positivos. Bactérias apresentando o fenótipo F^+ (contendo um epissomo do tipo F' , tal como XL1-Blue) codificam para proteínas de pilus, permitindo a infecção por fagos fita simples macho específico, tais como fagos filamentosos do tipo f1. O epissomo F' desta linhagem contém o gene *lacZ* deletado na região N-terminal (*lacZ Δ M15*), o que permite a α -complementação (vide descrição do plasmídio pUC, Apêndice IV) do segmento do gene presente em determinados plasmídios como pBluescript e pUC. O epissomo F' contém também o gene para super produção de repressor do operon *lac* (*lacI^q*), bem como genes marcadores tais como para resistência a tetraciclina (*Tn10*) e os genes envolvidos no metabolismo de prolina (*proAB*). A linhagem apresenta as marcas cromossômicas para resistência ao ácido nalidíxico (*gyrA*) e requerimento de tiamina para crescimento em meio mínimo. A mutação *endA1* no gene que codifica para endonuclease I, aumenta o rendimento e a qualidade do DNA plasmidial preparado através de técnicas rápidas de minipreparação (Bullock et al., 1987).

b) DH5 α :

Esta linhagem, defectiva em recombinação (*recA1*), foi utilizada na preparação de células competentes para os experimentos de transformação realizados em todas as subclonagens. A deleção dos genes do operon *lac* (Δ *lacU169*) e a presença do fragmento *lacZ Δ M15* no profago ϕ 80 permite a utilização de α -complementação na seleção de recombinantes. A ausência de superprodução do repressor do operon *lac* (*lacI^q*) possibilita a realização do teste de complementação na ausência de IPTG (Hanahan, 1983). Os demais marcadores foram descritos anteriormente para XL1-Blue.

C) DL538:

Esta linhagem foi utilizada como hospedeira do bacteriófago λ EMBL4 durante a seleção de clones recombinantes da biblioteca genômica de sorgo. A mutação do gene *recD* que extingue a atividade da exonuclease V, bem como a mutação *sbcC*, supressora de mutantes do sistema *recBC* agem diminuindo a perda de fagos recombinantes que carregam fragmentos de DNA genômicos com seqüências invertidas repetidas. As mutações *mcrA* e *mcrB* evitam a restrição de DNAs contendo metilcitosinas nas seqüências 5'G^mCGC3' e 5'AG^mCT3', o que favorece a construção de bibliotecas genômicas de plantas superiores, cujo DNA é rico em citosinas metiladas (Whittaker et al., 1988).

APÊNDICE II

A) Vetor λ -Zap:

A grande vantagem na utilização do vetor λ -Zap, consiste na possibilidade da excisão de um "phagemid" contendo o inserto de cDNA clonado. A seguir, são descritas as características do vetor λ -ZAP envolvidas no processo de excisão. Os vetores derivados de λ -ZAP apresentam origem e término de replicação de bacteriófago f1, sendo o sítio de iniciação (I) e o sítio de terminação (T) (Fig. 17). Proteínas derivadas de fagos do tipo f1 (fago "helper") em *E. coli* submetidas a infecção simultânea com o vetor e o fago simples fita, reconhecem o sítio de iniciação produzindo um corte em uma das duas fitas do DNA. A partir deste corte, uma nova fita de DNA é sintetizada, ocorrendo a duplicação de toda a seqüência que encontra-se no sentido 5'→3'. A nova fita sintetizada inclui o inserto clonado no vetor, sendo interrompida no sítio de terminação. O DNA fita simples é recircularizado pelo produto do gene II do fago f1, formando uma molécula de DNA circular contendo toda a seqüência presente entre os sítios I e T (o vetor λ -ZAP II inclui a seqüência do "phagemid" pBluescript SK-). O "phagemid" apresenta sinais necessários ao empacotamento (terminador f1) permitindo que o DNA recircularizado possa ser empacotado e secretado da *E. coli* para o meio de cultura. As células são mortas incubando-se a 70°C e após a centrifugação, o "phagemid" empacotado é isolado no sobrenadante. Para a obtenção de DNA fita dupla, células de *E. coli* são infectadas com o "phagemid" e plaqueadas em meio contendo ampicilina. As colônias resistentes à ampicilina, que crescem lentamente (colônias pequenas), ou ainda minipreparações de DNA mostrando múltiplas formas de plasmídeo, podem resultar da presença do fago "helper" nas células. O fago "helper" pode ser removido aumentando-se a razão de bactéria em relação ao fago durante o plaqueamento. A coinfeção com fago "helper" R408, que apresenta o produto do gene II termolábil, permite a seleção de transformantes livres do fago "helper", quando incubadas a 42°C. As colônias isoladas podem ser utilizadas na obtenção de DNA para análise do inserto.

APÊNDICE III

Mapeamento do bacteriófago:

O sistema consiste em hibridizar o DNA do bacteriófago contendo o inserto a ser mapeado, parcialmente digerido, com os oligonucleotídeos complementares às extremidades *cos* presentes no fago, marcados radioativamente com [γ -³²P]ATP. A digestão parcial do DNA é feita de modo a escolher as condições onde ocorre o aparecimento do maior número de fragmentos possíveis. Isso indica que todos os possíveis fragmentos estão representados na digestão parcial. As reações de hibridização do DNA, direita (ON-R) e esquerda (ON-L), são feitas separadamente com os respectivos oligonucleotídeos. Em seguida, cada reação é submetida à eletroforese em gel de agarose 0,5%, sendo que o tamanho de cada fragmento reproduz a distância do respectivo sítio enzimático em relação à extremidade marcada pela hibridização.

APÊNDICE IV

A) Vetor pUC19:

Os vetores pUC 18/19 possuem 2690 pares de bases e diferenciam-se somente pela orientação do sítio de policlonagem. Estes vetores possibilitam a identificação rápida do clone recombinante, através da identificação histoquímica das colônias obtidas pelo plaqueamento das células transformadas. Carregam um segmento de DNA derivado do operon *lac* de *E. coli* que codifica a região N-terminal da enzima β -galactosidase. Este fragmento, cuja síntese pode ser induzida por IPTG, é complementado por uma forma defectiva da β -galactosidase codificada pela célula hospedeira. Este processo é conhecido como α -complementação. Dada a síntese de ambos os fragmentos da enzima, a forma ativa degrada o substrato X-gal, previamente espalhado sobre o meio de cultura, fazendo com que as colônias adquiram a coloração azul. Entretanto, como os sítios de policlonagem do plasmídio estão na região N-terminal do fragmento de β -galactosidase, ocorrerá a inativação desta, caso um inserto seja clonado nesta região. Assim, o polipeptídeo produzido não será hábil para participar da α -complementação e a célula contendo este plasmídio não tomará a coloração azul. Os vetores pUC 18 e 19 apresentam o gene que confere resistência à ampicilina (β -lactamase), e origem de replicação derivada de ColE1 (Fig. 19) (Messing, 1983; Norrander et al., 1983; Yanisch-Perron et al., 1985).

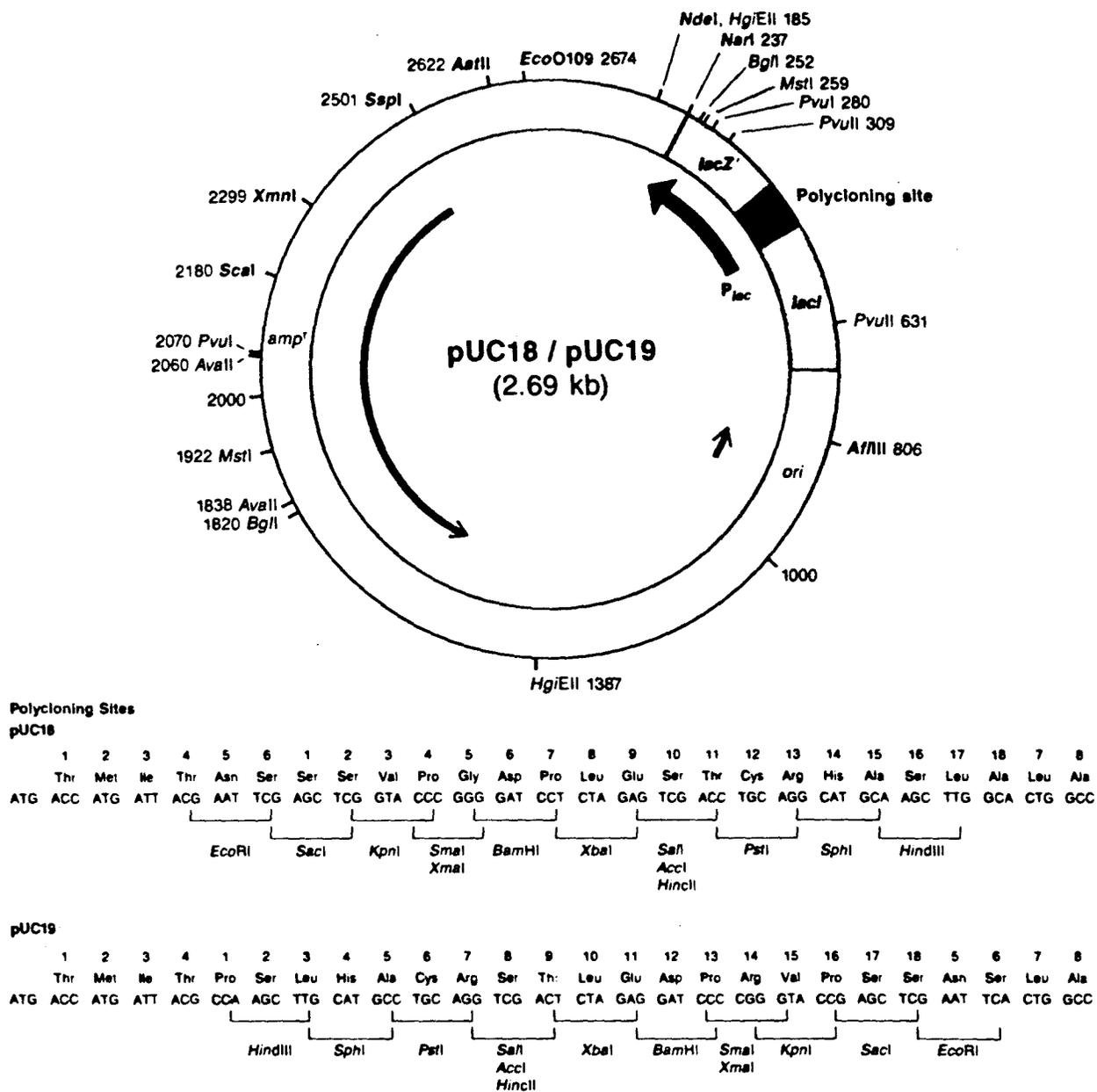


Fig. 19. Mapa de restrição do vetor pUC19.

B) Vetor pBluescript KS +/-

O vetor pBluescript KS +/- é um "phagemid" que contém a região intergênica (IG) do fago filamentososo do tipo f1, podendo ser secretado com um DNA fita simples na presença do fago "helper" do tipo f1 (Stratagene Cloning System). A região IG codifica para as funções *cis* requeridas para o empacotamento e replicação. Existem "phagemids" pBluescript contendo uma das duas possíveis orientações de região IG. O pBluescript KS+ replica-se de forma que a fita codificadora do gene β -galactosidase seja secretada na partícula do fago, enquanto que o pBluescript KS- secreta a fita não codificante do mesmo gene. O vetor pBluescript KS +/- possui 2961 pb e é derivado dos plasmídios da série pUC. Portanto, esse vetor também apresenta α -complementação, além de permitir a utilização dos mesmos primers para sequenciamento de insertos clonados em pUC. A denominação KS refere-se apenas à orientação da região de policlonagem, com relação às enzimas que ficam na extremidade (KpnI \rightarrow SacI). Este tipo de plasmídio é denominado "phagemid" já que apresenta origem de replicação de fago simples fita e também origem de replicação plasmidial (derivada de ColE1), a qual é utilizada na ausência do bacteriófago "helper". Da mesma forma que os plasmídios da série pUC, a resistência a antibiótico dá-se pela presença do gene que confere resistência à ampicilina (Fig. 20).

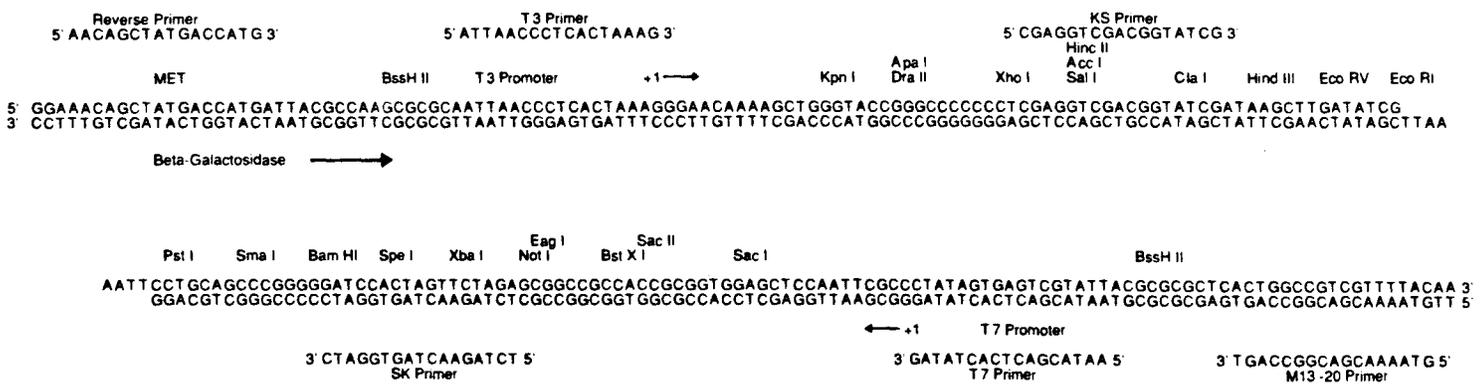
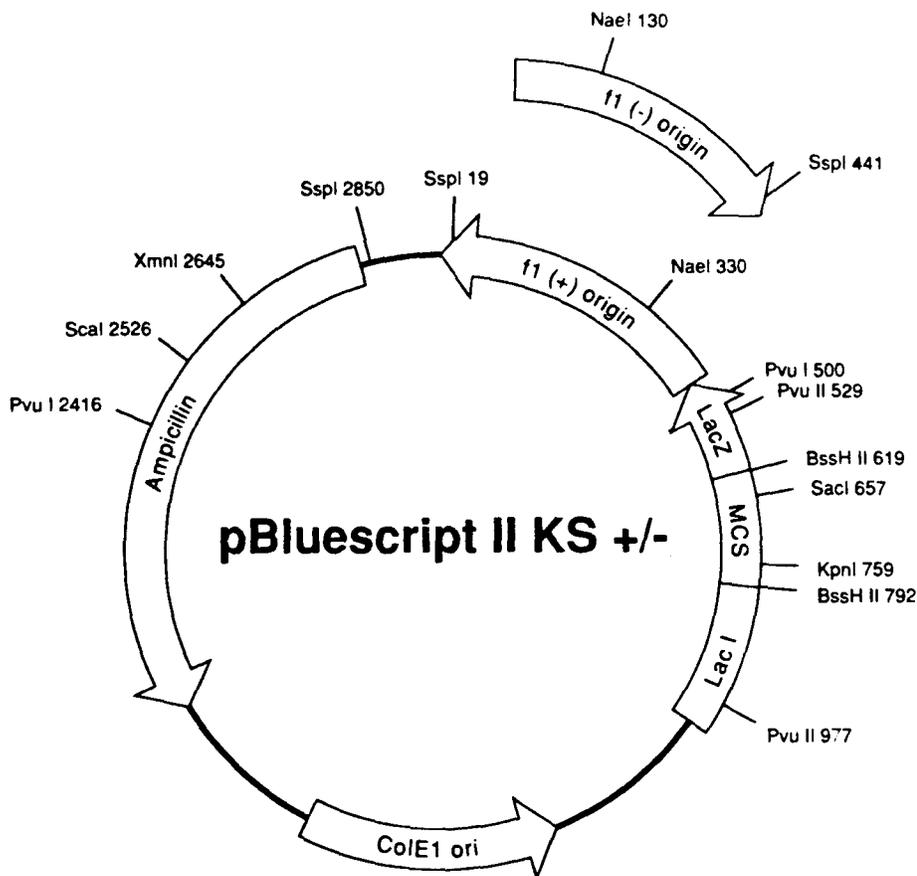


Fig. 20. Mapa de restrição do vetor pBluescript KS +/-.

C) Vetor de expressão pRT103.

O vetor pRT103 é derivado dos vetores pUC18/19 e possui 3340 pares de base (Töpfer et al, 1987). Este plasmídeo, bem como os outros da série pRT, são adequados para expressão em plantas, pois carregam o promotor do gene do transcrito 35S do vírus do mosaico da couve flor (CaMV35S), apresentando expressão constitutiva. Este promotor está representado por aproximadamente 360 pares de base, sendo seguido pela região contendo os sinais de poliadenilação, também do CaMV. Entre essas duas regiões, encontram-se os sítios enzimáticos para clonagem do gene de interesse. É importante ressaltar que o sítio de *Nco*I, presente nesta região, coincide com o codon de iniciação de transcrição ATG, sendo que este codon está contido em uma seqüência consenso, ideal para a iniciação da transcrição pelos ribossomos em eucariotos. Este plasmídeo apresenta marca de resistência ao antibiótico ampicilina (Fig. 21.).

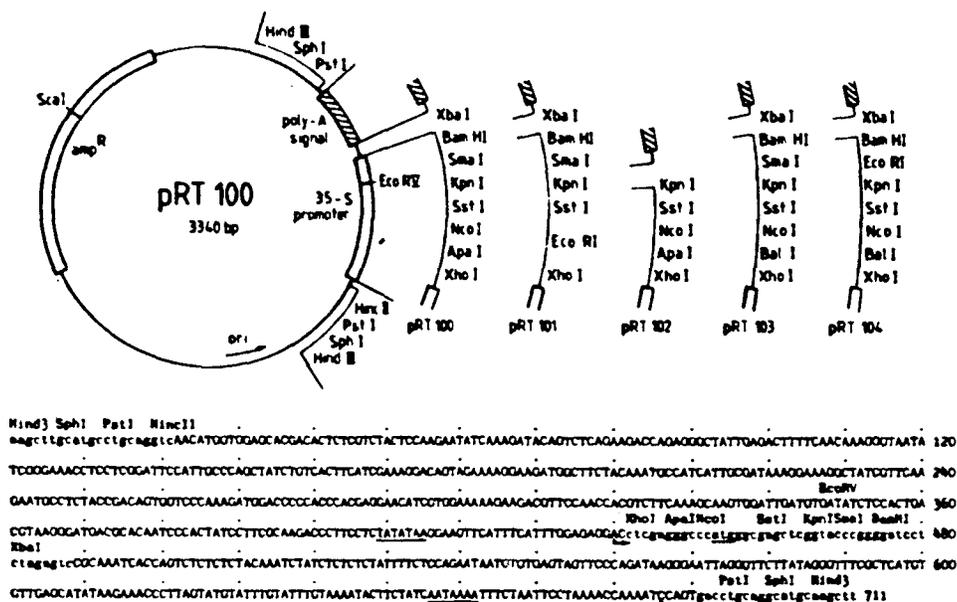


Fig. 21. Mapa de restrição do vetor de expressão pRT103.

APÊNDICE V

Experimento de reconstrução:

Os cálculos relacionados com a determinação da quantidade de DNA a ser utilizada como controle do número de cópias no experimento de reconstrução, são mostrados a seguir:

A) Determinação da massa do plasmídeo pBCX22,5 correspondente a uma cópia do gene de γ -coixina por genoma haplóide do *Coix* (o tamanho do genoma haplóide do *Coix* foi considerado similar ao do milho, isto é, $5,6 \times 10^6$ kb (Laurie & Bennett, 1985)):

tamanho do genoma do *Coix*: $5,7 \cdot 10^6$ kb

tamanho do plasmídeo pBCX22,5: 3,8 kb

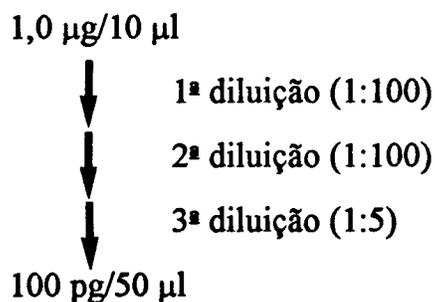
massa de DNA genômico digerida: 10 μ g

$$\begin{array}{l} 5,7 \cdot 10^6 \text{ kb} \longrightarrow 10 \mu\text{g} \\ 3,8 \text{ kb} \longrightarrow x \mu\text{g} \end{array} ,$$

onde $x = 6,7 \cdot 10^{-6}$ μ g (6,7 pg) corresponde à massa de DNA plasmidial que contém o mesmo número de cópias do gene de γ -coixina que está presente em 10 μ g de DNA genômico, considerando-se a presença de uma única cópia do gene por genoma haplóide de *Coix*.

B) Diluições do DNA plasmidial:

Concentração do clone pBCX22,5 após a digestão: 1,0 μ g/10 μ l



Portanto, temos:

1 cópia: 6,7 pg (3,4 μ l)

2 cópias: 13,4 pg (6,7 μ l)

3 cópias: 33,5 pg (16,8 μ l)

APÊNDICE VI

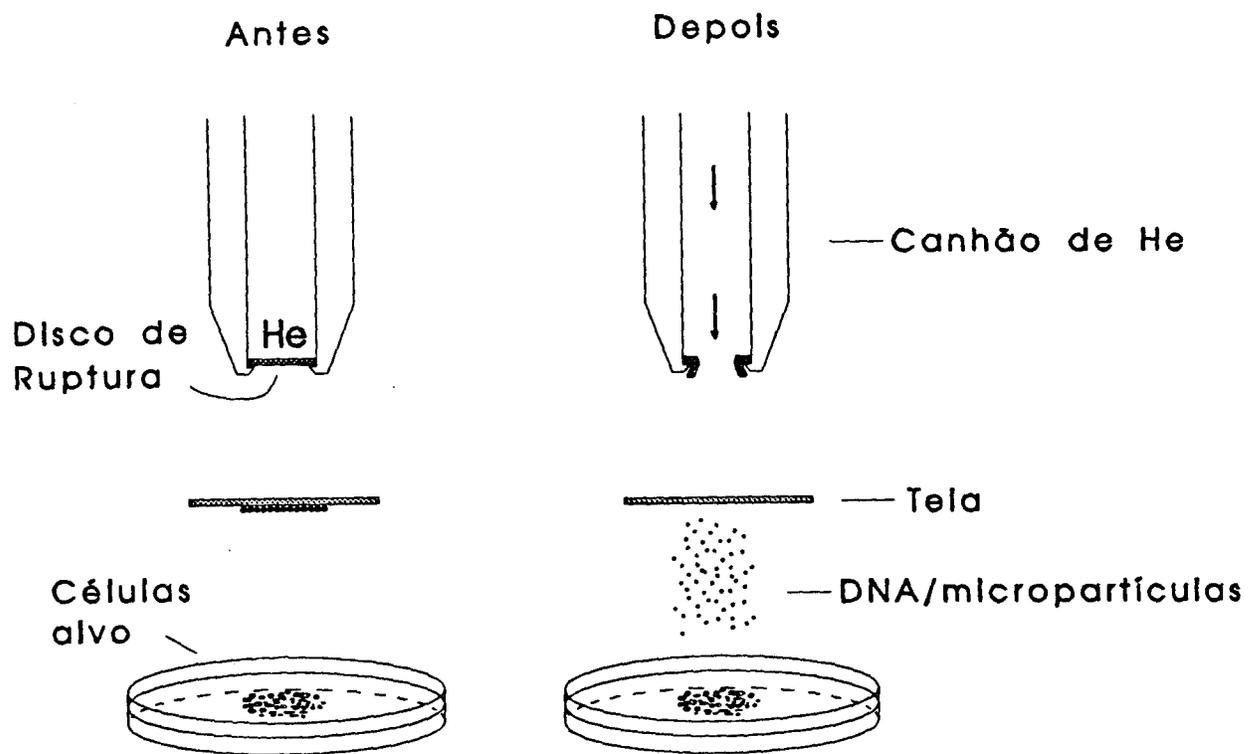


Fig. 22. Esquema do aparelho utilizado nos experimentos de bombardeamento com micropartículas.