

BC/22366

IB/81136

T/UNICAMP

V643_C

81150



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

**CLONAGEM E CARACTERIZAÇÃO DO FATOR DE
TRANSATIVAÇÃO OPACO2 DE *Coix lacryma-jobi***

ANDRÉ LUIZ VETTORE

ORIENTADOR: Dr. ADILSON LEITE

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

1994

ste exemplar corresponde à redação final
da tese defendida pelo(a) candidato(a)
André Luiz Vettore
e aprovada pela Comissão Julgadora.
03/08/94

UNIDADE	IB
N.º CHAMADA	
T/UNICAMI	
V643C	/
V.	
PREÇO	22366
PROD.	
C	☐
D	☐
RECEITA	R\$ 11,00
DATA	22/09/94
N.º CDS	

CM-00062155-0 /

**"Nada é ao acaso
Tudo é uma necessidade."**

**Ao meu Pai e à minha Mãe,
sem vosso apoio, eu jamais
chegaria até aqui.**

**À querida Berenice,
pela tranquilidade,
carinho e amor que
sempre encontrei no
brilho de seu olhar.**

Meu mais profundo agradecimento.....

Ao Dr. Adilson Leite, meu amigo e orientador, e sua estimada úlcera, pelo estímulo e contribuição para a minha formação científica.

Ao Prof. Dr. Paulo Arruda, por minha iniciação na carreira científica.

As Profas. Dra. Maricilda Palandi de Mello, Dra. Yoko Bomura Rosato e Dra. Solange Farah, pela participação no exame de qualificação e importantes contribuições para minha formação pedagógica.

A Dra. Laura M. M. Ottoboni, e aos Profs. Dra. Anete P. de Souza e Dr. Ladaslav Sodek, pela valiosa análise prévia deste trabalho e pelas importantes sugestões apresentadas.

Ao Dr. Robert J. Schmidt, por ceder o cDNA de opaco2 de milho.

A 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.

Ao "Fundão de Deus" (Xico de "Jesus", Andres "Loca", "Grande" Edson e Ricardinho "Jovem Cientista") e seus emissários terrenos (Germano, "Poderosa" Marymar, Tereza e Maria Elena), pela inestimável amizade e por nunca me deixarem esquecer da real força que nos move.

A amiga Silvia, pela agradável convivência nestes anos e pela correção dos inúmeros erros ortográficos presentes nos manuscritos deste trabalho.

Aos meio-irmãozinhos Fernando e Felipe, por sempre me auxiliarem na árdua e gratificante tarefa de manter sempre ativa a gloriosa úlcera de nosso amado Orientador.

Aos grandes amigos Gonçalo, Marília, Bia, Cláudia, Luciana, Maria do Carmo e Sibov, pelo agradável convívio e imensa amizade.

Aos amigos Ricardão Azevedo, David Moon, Isa (Maria Luísa P. N. Targon), Emi, Patrícia, Luzão e Agostinho, pela amizade e companheirismo.

Aos amigos do Laboratório de Genética de Bactérias (Regina, Susi, Geórgia, Sandra e Edmilson), pela amizade que cultivamos nos últimos anos.

Aos amigos do Laboratório de Genética Animal (Luzinho, Rosângela e Nair) e ao amigo Edson "Baiano", pelo carinho com que sempre me trataram.

Aos amigos do Laboratório de Genética Humana (Marcela, Raquel, Isa, Edi, Lee e Modalena), pela amizade e alegre convivência.

Ao grande Maurício, pela amizade e pela imprescindível ajuda no preparo da vidraria .

Aos amigos da secretaria, Sandra, Tânia, Ronaldo e Marciano, pelo carinho e prontidão que sempre empenharam na resolução de meus problemas burocráticos.

A todos aqueles que, de alguma forma, me incentivaram durante a execução deste trabalho.

À CAPES e FAPESP, pelo apoio financeiro.

Senhor, faz de mim um instrumento de Tua Paz

Onde há ódio, faça que eu leve o amor.

Onde há ofensa, faça que eu leve o perdão.

Onde há discórdia, faça que eu leve a união.

Onde há erro, faça que eu leve a verdade.

Onde há dúvida, faça que eu leve a esperança.

Onde há trevas, faça que eu leve a luz.

Onde há tristeza, faça que eu leve a alegria.

Concede, ó Mestre, que eu não procure tanto

Ser consolado, mas sim, consolar;

Ser compreendido, mas sim, compreender;

Ser amado, mas sim, amar.

Pois,

É dando, que se recebe.

É perdando, que se é perdoado,

*É morrendo na tua Paz, Senhor, que se nasce
para a vida eterna !*

FRANCISCO DE ASSIS

ÍNDICE

INTRODUÇÃO	1
Coix lacryma-jobi L.	1
Prolaminas	3
Regulação da expressão gênica das prolaminas	5
O gene o2	7
Leucine-Zippers	8
A função da proteína O2	10
O sítio de ligação da proteína O2 de milho	10
Localização nuclear da proteína O2	11
 OBJETIVOS	 13
 MATERIAIS E MÉTODOS	 15
Material vegetal	15
Linhagens de bactérias	15
Purificação de DNA de bacteriófago λ	16
Isolamento de DNA nuclear de planta	18
Isolamento em larga escala de DNA de plasmídio	18
Purificação de fragmentos de DNA a partir de géis de agarose	20
Biblioteca de cDNA de Coix	20
Seleção de clones de cDNA de opaco2	20
Excisão dos clones recombinantes de cDNA de opaco2	22
Biblioteca genômica de Coix	22
Sequenciamento de DNA	23
Deleção unidirecional com a enzima Exonuclease III	24
Construção de plasmídios utilizados nos testes de ligação seletiva	26
Teste de ligação seletiva	28
Etapa 1	28
Etapa 2	29
Etapa 3	29

Extração de RNA	30
Northern blot	31
Preparação das amostras	31
Preparação do gel desnaturante e eletroforese	31
Transferência	32
Hibridização	32
Métodos computacionais	32
RESULTADOS	33
O clone de cDNA	33
Isolamento de clones de cDNA de opaco2 de Coix	33
Sequenciamento do clone de opaco2	33
O clone genômico	35
Seleção do clone genômico de opaco2 de Coix	35
Sequenciamento do clone genômico de o2 de Coix	35
Descrição do clone genômico de o2 de Coix	36
Estudos comparativos	42
Comparação entre os genes opaco2	42
Comparação entre as proteínas Opaco2	46
RNA mensageiro de o2 de Coix	46
Tecido-especificidade da proteína O2 de Coix	46
Acúmulo de mRNA de opaco2 durante o desenvolvimento das sementes de Coix	50
O sítio de ligação da proteína O2 de Coix	52
DISCUSSÃO	56
CONCLUSÕES	65
SUMÁRIO	67
SUMMARY	69
LITERATURA CITADA	71
APÊNDICE I	79

APÊNDICE II	84
APÊNDICE III	86
APÊNDICE IV	91
APÊNDICE V	94
APÊNDICE VI	98

ÍNDICE DE FIGURAS E TABELAS

Figura 1 - Esquema de um ramo de Coix lacryma-jobi	1
Figura 2 - Esquema da dimerização de duas proteínas através de seus motivos bZIP	9
Figura 3 - Modelo "scissors grip" para interação de proteínas bZIP com o DNA	9
Figura 4- Representação esquemática do clone pCO2-1	34
Figura 5 - Mapa de restrição do clone genômico de opaco2 de Coix	39
Figura 6 - Sequência de nucleotídeos do clone genômico de opaco2 de Coix	40
Figura 7 - Sequência de nucleotídeos da região promotora do gene o2 de Coix	41
Figura 8 - Comparação das sequências de nucleotídeos dos genes o2, por "dot matrix"	43
Figura 9 - Representação esquemática dos genes o2 de Coix, milho e sorgo	44
Figura 10 - Alinhamento das regiões 5' dos genes o2 de Coix, milho e sorgo	45
Figura 11 - Alinhamento das sequências de aminoácidos das proteínas O2	47
Figura 12 - Alinhamento das sequências de aminoácidos de várias bZIP de plantas	48
Figura 13 - "Northern blot" com RNAs de diferentes tecidos de Coix	49
Figura 14 - "Northern blot" com RNAs de diferentes estágios de desenvolvimento de sementes de Coix	51
Figura 15 - Teste de ligação seletiva com as proteínas Opaco2 de Coix e milho	53
Figura 16 - Teste de ligação seletiva com a proteína Opaco2 de Coix	55
Figura 17 - Mapa do vetor λZAP	87
Figura 18 - Esquema do processo de exisão do "phagimid" pBluescript SK	89
Figura 19 - Mapa do vetor λDASH	90
Figura 20 - Mapa do vetor pBluescript KS	92
Figura 21 - Mapa do vetor pUR 278, 288 e 289	93
Figura 22 - Diagrama do método de deleção unidirecional	95
Figura 23 - Diagrama do método de ligação seletiva	97
 Tabela I - Relação dos subclones utilizados no sequenciamento do clone genômico	 37
Tabela II - Relação dos subclones obtidos com a deleção unidirecional	38
Tabela III - Relação dos sítios-alvo reconhecidos pela proteína O2 em diferentes genes	63
Tabela IV - Relação dos plasmídios utilizados neste trabalho	98

ABREVIACES

AMPc	adenosina 5'-mono-fosfato cclico	pb	pares de bases
ATP	adenosina 5'-trifosfato	PEG	polietileno glicol
DAP	dias aps a polinizao	pfu	unidade formadora de placa de lise
DDBJ	DNA Data Bank of Japan	q.s.p.	quantidade suficiente para
2-ME	2-mercaptoetanol	rpm	rotaes por minuto
DO ₆₀₀	absorbncia em 600 _{nm}	SDS	dodecil sulfato de sdio
DTT	ditiotreitol	Tris	tris-(hidroximetil)-aminometano
EDTA	cido etilenodiamino-tetractico	x g	fora centrfuga em relao  acelerao gravitacional
EMBL	European Molecular Biology Laboratories		
GenBank	Genetic Sequence Data Bank		
HEPES	cido N-2-hidroxitil-piperazina-N-2-etanosulfnico		
IPTG	isopropil tio--D-galactosdeo		
kb	kilobases		
MOPS	[3-(N-morfolino) cido propanosulfnico]		
o2	gene opaco2		
O2	protena Opaco2		

INTRODUÇÃO

***Coix lacryma-jobi* L.**

Coix lacryma-jobi L., juntamente com milho, sorgo e tripsacum, pertence a família Gramineae e a tribo Andropogoneae (Clayton, 1973; Jain & Benerjee, 1974; Clayton, 1983). Este cereal é nativo do sudeste asiático, onde foi utilizado, durante muito tempo, como fonte de alimento, na produção de bebidas alcoólicas, como planta medicinal e como ornamento (Venkateswarlu & Chaganti, 1973; Arora, 1977).

O Gênero *Coix* consiste de plantas monóicas, anuais ou perenes, com polinização efetuada pelo vento. As espiguetas carpeladas são em trios; sendo um fértil e dois estéreis. Cada um destes grupos encontra-se encerrado dentro de uma estrutura semelhante à contas, o invólucro. Este invólucro é formado por uma bainha modificada e endurecida, que apresenta, quando madura, coloração variável. Por um orifício, no ápice deste invólucro, emerge uma haste que sustenta a inflorescência masculina. As espiguetas masculinas são imbricadas em pares ou trios e uma delas, a central, é pedunculada. Cada

espiguetas apresenta duas floretas, ambas estaminadas, sendo a superior, às vezes, estéril (Fig. 1) (Venkateswarlu e Chaganti, 1973; Jain & Benerjee, 1974). *Coix lacryma-jobi* apresenta uma taxa média de fecundação cruzada de aproximadamente 40% (Mello, 1989).

Atualmente o *Coix* é encontrado em todas as regiões tropicais do mundo. No Brasil, por volta de 1938, foi introduzida a variedade Adlay com sementes trazidas das Filipinas (Torres & Bergamin, 1951; Shaafhawsen, 1952).

Apesar das sementes de *Coix* apresentarem conteúdo maior de proteínas que as sementes de milho (Ottoboni *et al.*, 1990), este cereal é considerado de pouca importância econômica devido a suas propriedades agronômicas desfavoráveis. Dentre estas propriedades, podemos destacar, a maturação não sincronizada das sementes, impedindo a mecanização da colheita deste cereal.

A



B

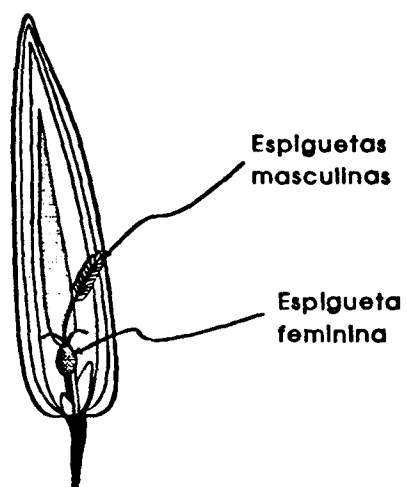


Figura 1: A. Esquema mostrando um ramo de *Coix lacryma-jobi*.
B. Detalhe de um fascículo indicando as espiguetas masculina e feminina de uma inflorescência.

As prolaminas

As proteínas de reserva acumulam-se nas sementes de inúmeras espécies de plantas e constituem uma fonte de nitrogênio para a germinação do embrião. Em cereais pertencentes a tribo

Andropogoneae, tais como milho, sorgo e *Coix*, as proteínas de reserva são constituídas por um grupo de proteínas solúveis em álcool, ricas em prolina e glutamina, e denominadas de prolaminas (Shewry & Tatham, 1990). Estas proteínas representam, aproximadamente, a metade da proteína presente nas sementes maduras, e, em milho, sorgo e *Coix*, recebem respectivamente os nomes de: zeínas, kafirinas e coixinas.

Baseado na solubilidade diferencial, as zeínas foram classificadas como: α -, β -, e γ -zeínas (Esen, 1986). Esta classificação foi posteriormente estendida para as kafirinas (De Rose *et al.*, 1989) e coixinas (Leite *et al.*, 1990). As α -, β - e γ -prolaminas, compreendem, respectivamente, as proteínas solúveis em 40-95%, 30-80% e 0-80% de isopropanol. Entretanto, esta classificação não reflete a relação estrutural entre as prolaminas, assim sendo o sequenciamento de um gene de uma zeína, previamente descrita como pertencente a classe das β -prolaminas,

revelou uma estrutura completamente diferente, determinando a criação de uma quarta classe de prolaminas denominada de δ -prolaminas (Kirihaara *et al.*, 1988). A coixina de 17 kDa constitui-se em outro exemplo, que apesar de apresentar solubilidade característica das α -prolaminas (Leite *et al.*, 1990), após o sequenciamento de um clone de cDNA, revelou alta similaridade com a β -prolina de milho (Leite *et al.*, 1992).

As α -prolaminas, a classe mais abundante, correspondem a cerca de 70% das prolaminas totais de milho, sorgo e *Coix*. Em milho, esta classe de prolaminas é constituída por polipeptídeos agrupados em duas classes de pesos moleculares aparentes de 19 e 22 kDa (Burr & Burr, 1982), o mesmo acontecendo com as α -prolaminas de *Coix*, que são constituídas pelos grupos de polipeptídeos de 25 e 27 kDa (Leite *et al.*, 1990).

Estudos de hibridização, bem como estudos imunológicos e análise de sequências de clones de cDNA e genômicos de α -prolaminas, têm revelado a existência de uma grande homologia entre os genes que codificam as α -zeínas de 22 kDa e aqueles que codificam as α -coixinas de 25 kDa (Leite *et al.*, 1990; Ottoboni *et al.*, 1993). Genes que codificam proteínas homólogas à α -zeína de

22 kDa foram também identificados em sorgo (De Rose *et al.*, 1989). Estudos de análise de reconstrução destes genes demonstram que os mesmos são codificados por uma família multigênica, constituída por, aproximadamente, 100 genes em milho (Viotti *et al.*, 1979), 20 genes em sorgo (De Rose *et al.* 1989) e 35 genes em *Coix* (Ottoboni *et al.*, 1993).

Em milho, as β -prolaminas correspondem à aproximadamente 15% do total das zeínas e são representadas por um único polipeptídeo de peso molecular aparente de 15 kDa (Pedersen *et al.*, 1986). Apesar da aparente ausência de β -prolaminas em endospermas de *Coix*, em 1992, Leite *et al.* descreveram a clonagem e a caracterização de um cDNA codificando uma coixina rica em metionina e cisteína, que apresentava 64% de similaridade com o gene de β -zeína.

As γ -prolaminas são solubilizadas em soluções alcoólicas (0-80%), na presença de agentes redutores, tais como, β -mercaptoetanol e ditioneitol. As γ -prolaminas em milho correspondem a 5-10% do total das zeínas, e são compostas por dois polipeptídeos, um de 28 kDa e outro de 16 kDa (Esen, 1986). A γ -zeína de 16 kDa apresenta

solubilidade característica de β -prolaminas (Esen, 1986), sendo codificada por um único gene, enquanto que a γ -zeína de 28 kDa, dependendo da linhagem de milho, é codificada por um ou dois genes (Gallardo *et al.*, 1988; Das & Messing, 1987; Das *et al.*, 1991a; Das *et al.*, 1991b). Em *Coix*, as γ -prolaminas representam cerca de 15% do total das prolaminas, e são constituídas por um único polipeptídeo de 22 kDa que apresenta alta similaridade com a γ -zeína de 28 kDa (Leite *et al.*, 1990; Leite *et al.*, 1991). O endosperma de sorgo apresenta menor concentração de γ -prolaminas do que milho e *Coix*. A γ -kafirina representa cerca de 2-5% do total de prolaminas e é composta por um único polipeptídeo de 27 kDa (Evans *et al.*, 1987). As caracterizações de clones de cDNA (Barros *et al.*, 1991), e genômico de γ -kafirina (Freitas *et al.*, 1994) mostraram que as γ -kafirinas, também, apresentam alta similaridade com a γ -zeína de 28 kDa.

Embora as sequências dos genes das diferentes classes de prolaminas não mostrem qualquer similaridade, estes genes apresentam algumas semelhanças estruturais. Primeiro: nenhum deles apresenta interrupções em suas regiões codificadoras, ou seja, não possuem

introns. Segundo: as sequências codificadoras das proteínas são organizadas em domínios estruturais típicos, que permitem às proteínas sintetizadas, assumirem uma estrutura secundária mais condizente com o rígido empacotamento dentro dos corpúsculos protéicos (Argos *et al.*, 1982; Garratt *et al.* 1993). Terceiro: todas as prolaminas são precedidas por um peptídeo sinal, responsável pelo direcionamento das prolaminas para o interior do retículo endoplasmático (Burr & Burr, 1976; Larkins & Hurkman, 1978; Burr *et al.*, 1978; Ludevid *et al.*, 1984; Lending & Larkins, 1989; Targon *et al.*, 1992).

Regulação da expressão gênica de prolaminas

As prolaminas são sintetizadas durante o desenvolvimento das sementes em polirribossomos ligados às membranas do retículo endoplasmático rugoso, e acumulam-se em organelas de armazenamento denominadas corpúsculos protéicos (Burr & Burr 1982; Larkins *et al.*, 1984; Targon *et al.*, 1992). A transcrição dos genes de prolaminas em milho e Coix é altamente regulada, sendo que os RNAs mensageiros acumulam-se em altas concentrações nos estágios precoces do desenvolvimento do

endosperma (Marks *et al.*, 1985; Targon *et al.*, 1992).

Tem sido demonstrado que, apesar dos genes que codificam as diferentes classes de prolaminas, nos cereais da tribo Andropogoneae, apresentarem baixa similaridade em suas sequências, estes genes são expressos de forma específica e coordenada no endosperma em desenvolvimento (Feix & Quayle, 1993). Até o momento, os mecanismos envolvidos na regulação da expressão coordenada são desconhecidos. Comparações dos promotores de genes de diversas classes de prolaminas têm indicado como uma das únicas regiões conservadas, o motivo conhecido pelos nomes de "prolamin box", "endosperm box", ou ainda "-300 box" (Feix & Quayle, 1993). Este motivo apresenta a sequência TGTAAG, altamente conservada, que é similar ao motivo conhecido como "core" do "enhancer" de SV40. Este motivo foi descrito em posições relativas similares nos genes das diferentes classes de prolaminas dos cereais da tribo Andropogoneae (De Rose *et al.*, 1989; Ottoboni *et al.*, 1993; Pedersen *et al.*, 1986; Freitas *et al.*, 1994; Brown *et al.*, 1986; Liu & Rubenstein, 1992), bem como em genes de outros cereais como cevada (Kreis *et al.*, 1986), trigo (Summer-

Smith *et al.*, 1985; Colot *et al.*, 1987), e centeio (Hull *et al.*, 1991). Apesar do alto grau de conservação da sequência e da posição deste motivo, e também de sua reconhecida capacidade de reconhecer proteínas nucleares (Maier *et al.*, 1987), a função deste motivo não foi ainda esclarecida.

Várias mutações afetando o acúmulo das zeínas têm sido descritas (Motto *et al.*, 1989). Alguns destes mutantes são conhecidos como *opaco*, pois, ao contrário das sementes normais, que apresentam o endosperma duro e translúcido, estes mutantes produzem sementes com endosperma macio e opaco.

A mutação *opaco2* (*o2*) foi primeiramente identificada em milho por Singleton & Jones na década de 30 (Emerson *et al.*, 1935). O efeito dessa mutação no acúmulo de proteínas na semente foi estabelecido somente 30 anos depois. Mertz *et al.* (1964) demonstraram que a mutação *o2* proporcionava um melhor balanço de aminoácidos nas sementes, devido, principalmente, ao aumento do conteúdo de lisina. Tal aumento resultava da redução do acúmulo da fração das prolaminas (zeínas), as quais são deficientes nos aminoácidos essenciais lisina e triptofano. O fato de uma simples mutação aumentar o

conteúdo de lisina, atraiu extensivas pesquisas sobre a mutação *o2* e outras mutações *opaco* de milho. No início, haviam perspectivas de gerar linhagens de milho com maior valor nutricional, entretanto, o milho *o2* não obteve sucesso comercial. Comparado com as linhagens normais, o milho *opaco2* apresenta um rendimento menor, suas sementes são mais susceptíveis a injúrias durante a colheita e armazenamento, apresentam ainda, uma grande susceptibilidade a doenças e pragas (Nelson, 1980).

Nos mutantes *o2* ocorre uma grande redução das zeínas de 22 kDa, em função de um decréscimo da taxa de transcrição gênica, gerando uma ausência parcial ou completa de RNAs mensageiros específicos da classe das zeínas de 22 kDa (Pedersen *et al.*, 1980; Langridge *et al.*, 1982; Marks *et al.*, 1985). Burr & Burr (1982) demonstraram, ainda, que esta redução não pode ser atribuída a alterações de síntese, transporte, ou processamento destas proteínas.

Mutações em outros genes afetam a expressão das α -zeínas de maneira diferente da mutação *o2*. Enquanto a mutação *o2* reduz significativamente a expressão das proteínas pertencentes a classe de

22 kDa, a mutação *opaco7* (*o7*) afeta principalmente a expressão da classe de 19 kDa, já a mutação *floury2* (*f12*) diminui a expressão de ambas as classes, de 19 e 22 kDa. A mutação *opaco6* (*o6*) causa uma redução generalizada das prolaminas, afetando todas as classes de zeínas. A mutação *De-B30* proporciona uma redução de, aproximadamente, 30% das zeínas, principalmente, da classe de 22 kDa (Motto *et al.*, 1989)

Apenas o efeito da mutação *o2* sobre a transcrição do gene das α -zeínas de 22 kDa foi confirmado (Kodrzycki *et al.*, 1989), entretanto, análises da expressão das zeínas, através de "Northern blot", também indicam que *o7* e *f12* alteram o nível de transcrição destes genes (Burr & Burr, 1982).

O gene *o2*

O gene *o2* de milho foi identificado e clonado através de estratégia de "transposon tagging", com o auxílio dos elementos móveis *Suppressor-mutator* (Schmidt *et al.*, 1987) e *Activator* (Motto *et al.*, 1988). O cDNA de *o2* de milho foi posteriormente clonado e sequenciado (Schmidt *et al.*, 1990; Hartings *et al.*, 1989). As sequências apresentadas pelos dois grupos mostram

algumas diferenças, principalmente em suas regiões 3'. Ainda não está claro se tais diferenças representam um polimorfismo entre alelos, uma vez que cada grupo trabalhou com linhagens diferentes de milho, ou ainda, se são devidas a um artefato de sequenciamento (Schmidt, 1993).

Pirovano *et al.* (1994) isolaram e caracterizaram o gene *opaco2* de sorgo. A sequência de aminoácidos deduzida a partir da sequência de cDNA apresenta grande similaridade com a sequência da proteína O2 de milho, com exceção da região carboxi-terminal que não apresenta qualquer semelhança com as duas sequências descritas para a proteína O2 de milho.

As sequências de aminoácidos deduzidas a partir dos cDNAs sugerem que o gene *o2* codifica um polipeptídeo que possui um motivo característico de proteínas que se ligam a DNA, tais como, reguladores transcricionais de fungos e oncoproteínas de mamíferos, denominados "leucine-zipper" (bZIP) (Landschulz *et al.*, 1988b; Vinson *et al.*, 1989).

Leucine-zippers

O domínio "leucine-zipper" (bZIP) foi primeiramente observado como uma sequência conservada de vários fatores de transcrição de eucariotos. Entre estas proteínas incluem-se o ativador transcricional de fungo GCN4, as oncoproteínas de mamíferos Jun, Fos e Myc, e a C/EBP ("enhancer binding protein"). A região de ligação ao DNA dessas proteínas geralmente possui 60-80 resíduos de aminoácidos e apresentam dois domínios distintos: o "zipper" de leucinas, responsável por interações com outras proteínas, promovendo dimerizações, e o domínio básico, uma região adjacente, rica em aminoácidos básicos, responsável pelo reconhecimento e ligação com o sítio alvo no DNA (Pabo & Sauer, 1992).

A região do "zipper", possui aproximadamente 30-40 resíduos, e é caracterizada por repetições de resíduos de leucina a cada sete aminoácidos (Landschulz *et al.*, 1988b), assumindo uma estrutura secundária de α -hélice, de modo que os resíduos de leucina fiquem alinhados em uma mesma face. Assim sendo, as leucinas de um polipeptídeo podem estabelecer interações hidrofóbicas com leucinas de uma α -hélice

análoga de um segundo polipeptídeo (Fig. 2; Pabo & Sauer, 1992).

O domínio básico está localizado em uma região adjacente anterior às repetições de leucinas. Este domínio contém cerca de 30 resíduos, sendo que a maioria dos aminoácidos apresenta carga positiva, representados, principalmente, por argininas e lisinas (Turner & Tjain, 1989; Neuberg *et al.*, 1989).

Experimentos de proteção sugerem que a região básica da bZIP liga-se ao sulco maior do DNA (Pabo & Sauer, 1992). Existem dois modelos para a interação entre as bZIP e o sítio-alvo no DNA (Vinson *et al.*, 1989; O'Neil *et al.*, 1990). Em ambos, as proteínas formam dímeros em forma de "Y". A região do "zipper" une duas proteínas formando a haste do "Y", e as regiões básicas ficam livres formando os braços do "Y", os quais estendem-se em direções opostas ao longo do sulco maior do DNA. O modelo "scissors grip" é o mais aceito e propõe a existência de uma quebra na α -hélice da região básica, permitindo que esta cadeia acompanhe a curvatura do sulco maior do DNA (Fig. 3; Vinson *et al.*, 1989). No modelo "induced helical fork", a hélice não apresenta dobras, e portanto, estende-se além do DNA,

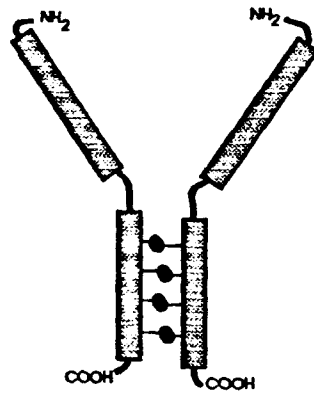


Figura 2: Representação esquemática da dimerização de duas proteínas através de seus domínios de "zipper" de leucinas.

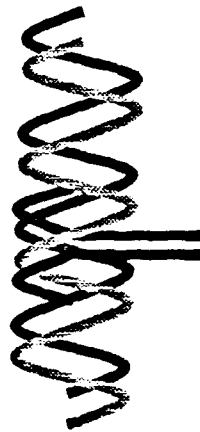


Figura 3: Modelo "scissors grip" (Vinson *et al.*, 1989) para a interação de proteínas bZIP com o sítio-alvo no DNA.

após contactar três ou quatro pares de base (O'Neil *et al.*, 1990).

Para as proteínas "leucine zipper" ligarem-se ao DNA, estas necessitam estar dimerizadas. A dimerização pode ocorrer entre duas proteínas idênticas, formando um homodímero, ou entre proteínas diferentes, formando heterodímeros. Algumas proteínas somente são ativas na forma de homodímeros, como C/EBP e Myc, enquanto outras, somente na forma de heterodímeros, como, Fos e Jun, e CREB e CREM (Johnson & McKnight, 1989).

Além de participar no reconhecimento do sítio-alvo, o domínio básico está envolvido, também, no transporte de proteínas do tipo bZIP para o núcleo de células vegetais (van der Krol & Chua, 1991).

A função da proteína O2

O papel específico do *locus o2* na regulação da expressão dos genes de zeína foi demonstrado por Kodrzycki *et al.* (1989). Seus resultados mostraram claramente que o *locus o2* afeta o acúmulo de transcritos de genes da classe das α -zeínas de 22 kDa. Estes dados, aliados à presença de um domínio de ligação ao DNA (bZIP) na proteína O2 de milho, sugeriram que a proteína O2

regulava a expressão das zeínas através de uma interação direta com os promotores dos genes. O papel desempenhado pelo domínio básico da proteína O2, na regulação da expressão de prolaminas, foi comprovado através do estudo do mutante *o2-676* (Aukerman *et al.*, 1991). Este mutante apresenta uma substituição de uma arginina, do domínio básico, por uma lisina. Apesar de tratar-se de uma substituição conservativa, esta alteração proporciona uma acentuada queda na expressão das zeínas.

O *locus o2* também regula a expressão de uma albumina de 32 kDa, chamada B-32 (Soave *et al.*, 1976; Di Fonzo *et al.*, 1986; Lohmer *et al.*, 1991). Essa proteína é codificada por um pequeno número de genes (Hartings *et al.*, 1990). A proteína B-32 é descrita como uma proteína inativadora de ribossomos, abundante no endosperma de milho, e que, provavelmente, deve funcionar como agente de defesa contra o ataque de patógenos (Bass *et al.*, 1992).

O sítio de ligação da proteína O2 de milho

O palíndrome imperfeito 5'-TCCACGTAGA-3', localizado na região 5' não codificadora do gene de

α -zeína de 22 kDa, foi identificado através de análise por "footprinting", como sendo o sítio-alvo de ligação da proteína O2 de milho (Schmidt *et al.*, 1992). A presença da sequência ACGT, interna ao sítio de ligação, presente nos promotores de α -prolaminas de milho, foi descrita como condição essencial para a ligação com a proteína O2 *in vitro* (Schmidt *et al.*, 1992). Tal sequência tem sido identificada também na região central de sítios-alvo de diversos fatores da classe bZIP em plantas (Willians *et al.*, 1992). A especificidade de ligação destes fatores parece depender das sequências que flanqueiam a região ACGT e do estado de dimerização do efector (Willians *et al.*, 1992; Schindler *et al.*, 1992). Em plantas, as bZIPs também podem formar homodímeros ou heterodímeros (Armstrong *et al.*, 1992), sendo que, estes dímeros contribuem para aumentar a complexidade das possíveis interações com os sítios-alvo e outros fatores.

Para o promotor do gene *b-32* foram descritos cinco sítios de ligação da proteína O2 de milho, os quais apresentam a sequência consenso GATGAPyPuTGpu (Lohmer *et al.*, 1991).

Experimentos de hibridização, através de "Southern blot", revelaram

a presença de genes relacionados ao o2 de milho em endosperma de *Coix* e sorgo (Ottoboni *et al.*, 1993). Testes de transativação revelaram que a transcrição das α -coixinas de 25 kDa pode ser ativada pela proteína O2 de milho (Yunes *et al.*, 1994). A proteína O2 de milho reconhece uma sequência diferente daquela descrita como sítio-alvo nos promotores de α -zeínas. Este sítio-alvo, presente nos promotores de α -coixinas, é caracterizado pela ausência da sequência ACGT, e por apresentar maior similaridade com dois sítios (B1 e B4) descritos no promotor do gene *b-32* (Yunes *et al.*, 1994).

Localização nuclear da proteína O2

Proteínas sintetizadas no citoplasma funcionam em vários compartimentos diferentes dentro da célula. O maior tráfego de macromoléculas ocorre entre o citoplasma e o nucleoplasma. A importação e a exportação nuclear é um processo altamente específico, uma vez que o conteúdo nuclear é totalmente distinto do citoplasmático. Proteínas pequenas, com tamanho entre 40 e 60 kDa, devem penetrar no núcleo por difusão através do poro nuclear. Proteínas maiores necessitam de ATP e de um

sinal de localização nuclear (NLS), para atravessar o poro nuclear (Raikhel, 1992).

Os NLSs identificados até hoje podem ser classificados em três categorias: (1) NLS do tipo encontrado no vírus SV40: contém pequenas repetições seguidas de aminoácidos básicos, principalmente prolina e glicina, como PKKRKV (Kalderon *et al.*, 1984a; Kalderon *et al.*, 1984b). (2) NLS do tipo "mating α 2-like": consiste de pequenas regiões hidrofóbicas, as quais contém um ou mais aminoácidos básicos, como KIPK (Hall *et al.*, 1984). (3) NLSs bipartidos: constituídos, geralmente, pela combinação de duas regiões de aminoácidos básicos separados por aproximadamente dez aminoácidos (Dingwall & Laskey, 1991).

Proteínas regulatórias como a Opaco2, são requeridas apenas no núcleo de células de tecidos específicos e em determinados momentos, durante o desenvolvimento. É provável que estas proteínas regulatórias permaneçam no citoplasma até que sua função no núcleo seja necessária.

Varagona *et al.* (1991) demonstraram a localização nuclear da proteína O2 de milho em células de endosperma de milho e de plantas

transformadas de tabaco. Na proteína O2 de milho foram identificadas duas regiões NLSs: a NLS A e a NLS B (Varagona *et al.* 1992). A região NLS A está localizada na porção amino-terminal, e apresenta estrutura semelhante àquela descrita para SV40. A região NLS B está localizada no domínio básico, e apresenta estrutura do tipo bipartida.

OBJETIVOS

Os dados moleculares têm demonstrado um alto grau de conservação dos genes que codificam as prolaminas de milho, sorgo e *Coix*. Estes dados sugerem que eles provavelmente originaram-se de genes ancestrais comuns, e que, a divergência observada, ocorreu posteriormente à separação das espécies (Leite *et al.*, 1990; Leite *et al.*, 1991; Leite *et al.*, 1992; Garrat *et al.*, 1993; Ottoboni *et al.*, 1993).

A presença de sequências relacionadas a *opaco2* (gene que codifica um ativador transcricional de α -prolaminas em milho) nos genomas de *Coix* e sorgo (Ottoboni *et al.*, 1993) indica que o mecanismo de controle de expressão dos genes de prolaminas provavelmente também foi conservado nestes cereais.

Assim sendo, a comparação das sequências estruturais e regulatórias de genes relacionados, que participam da regulação da expressão gênica, pode permitir uma definição mais precisa dos domínios envolvidos nas diversas etapas dessa regulação.

Os objetivos deste trabalho foram o isolamento e sequenciamento

do gene *opaco2* de *Coix*, e a comparação com o gene *opaco2* de milho. Este estudo, além de permitir uma melhor caracterização de importantes domínios presentes nas proteínas do tipo Opaco2, pode também ajudar no esclarecimento da controvérsia existente entre as sequências descritas para as proteínas Opaco2 de milho e de sorgo.

Para alcançar tais objetivos foram realizadas as seguintes etapas:

- 1) Isolamento de clones recombinantes de *opaco2*, a partir de uma biblioteca de cDNA de endosperma imaturo de *Coix*.
- 2) Construção de uma biblioteca genômica de *Coix lacryma-jobi* e isolamento de clones genômicos recombinantes de *opaco2*.
- 3) Sequenciamento dos clones de cDNA e genômico.
- 4) Comparação das sequências de nucleotídeos e aminoácidos correspondentes, dos genes relacionados a *opaco2*.

- 5) Análise de ligação seletiva, visando a determinação da sequência reconhecida pela proteína Opaco2 de *Coix*, nas regiões promotoras das α -prolaminas de *Coix* e milho.
- 6) Detecção do mRNA de *opaco2* em diferentes tecidos de *Coix*.
- 7) Comparação do padrão de acúmulo dos RNAs mensageiros de *opaco2* e α -coixina durante o desenvolvimento da semente de *Coix*.

MATERIAIS E MÉTODOS

Metodologias básicas de rotina, tais como: a) Minipreparação de plasmídios, b) Digestões com enzimas de restrição, c) Clonagem em plasmídios, d) Transformação de *E. coli*, e) Marcação de sondas de DNA, entre outras, foram utilizadas de acordo com Sambrook *et al.*, (1989).

As soluções utilizadas estão descritas no Apêndice I.

Material Vegetal

As sementes de *Zea mays* L. var. Maya e *Coix lacryma-jobi* L. var. Adlay foram obtidas da coleção de germoplasma do Departamento de Genética da Universidade Estadual de Campinas.

Linhagens de Bactérias

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

DH5 α

Genótipo relevante: *F⁻ supE44 Δ (lacZYA - argF)U169 (ϕ 80 lacZ Δ M15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1* (Hanahan, 1983).

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

XL1-Blue

Genótipo relevante: *supE44 endA1 thi-1 lac Δ (lac-proAB) gyrA96 hsdR17 recA1 F' proAB⁺ lacI^q lacZ Δ M15 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.

DL538

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

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

Purificação de DNA de bacteriófago λ

A purificação de DNA do bacteriófago lambda foi realizada conforme a técnica Yamamoto *et al.* (1970), descrita em Sambrook *et al.* (1989), com algumas adaptações às condições do laboratório.

Inoculou-se 15 ml de meio LB contendo 10 mM de MgSO_4 com a bactéria *E. coli* (DL-538), incubando-se durante a noite a 37°C, sob agitação vigorosa (300 rpm). Em um tubo de cultivo, misturou-se 1×10^{10} células ($\text{DO}_{600} = 1$ corresponde a 8×10^8 células/ml) com 5×10^7 fagos, incubando-se durante 25 minutos a 37°C sob leve agitação (100 rpm). O conteúdo deste tubo foi inoculado em um Erlenmeyer com 500 ml de meio de cultura LB contendo 10 mM de MgSO_4 , previamente aquecido a 37°C, e mantido sob vigorosa agitação (300 rpm) nesta temperatura. A DO_{600} do cultivo foi acompanhada a partir de 3 horas após o inóculo, 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, e incubou-se a 37°C durante 10 minutos. O cultivo foi mantido em gelo durante 15 a 20 minutos, sendo em seguida, transferido para tubos 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 2 litros, ao qual adicionou-se RNase e DNase I para uma concentração final de 1 $\mu\text{g/ml}$, para eliminação de DNAs e RNAs bacterianos. O lisado foi incubado por 1 hora a 37°C, e adicionou-se NaCl para uma concentração final de 1M (29,2 g para 500 ml). O NaCl auxilia na precipitação de proteínas bacterianas. Após a dissolução do NaCl, o lisado foi mantido em gelo por 1 hora, em seguida, transferido para tubos 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 2 litros, e a precipitação das partículas fágicas foi realizada através da adição de polietilenoglicol (PEG-8000) para uma concentração final de 10% (m/v) (50 g para 500 ml de sobrenadante). Essa solução foi mantida em gelo durante a noite. O lisado foi transferido para tubos 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 sedimentos secos por 10 a 15 minutos à temperatura ambiente. O precipitado foi ressuspensão em 4 ml de tampão SM. A suspensão foi trans-

ferida para tubos Falcon de 15 ml. Adicionou-se 4 ml de clorofórmio, seguida de rápida agitação. O lisado foi centrifugado por 12 minutos a 1.500 rpm (centrifuga de mesa Beckman GPR). A fase superior foi transferida para um tubo e mantida a 4°C até o momento da ultracentrifugação. Para proceder a ultracentrifugação, adicionou-se 0,5 g de CsCl por mililitro de suspensão de bacteriófago. A suspensão foi depositada no topo de um gradiente descontínuo de CsCl em SM. O gradiente foi montado com as seguintes fases do topo para o fundo do tubo: 1 ml de CsCl 1,45 g/ml (índice de refração 1,3768), 1,5 ml de CsCl 1,50 g/ml (índice de refração 1,3815) e 1,5 ml de CsCl 1,7 g/ml (índice de refração 1,3990). Os tubos foram ultracentrifugados em rotor Beckman SW41 durante 4 horas, a 4°C, a 22.000 rpm. Após a centrifugação, a banda esbranquiçada contendo o vírus, depositada entre as fases contendo 1,45 e 1,50 g/ml de CsCl, foi retirada com o auxílio de uma pipeta automática P1000. A suspensão de bacteriófagos foi dialisada por duas vezes, durante 1 hora, em 1,5 litro (~1000 vezes o volume de fago recolhido) de solução contendo 50 mM de Tris-HCl (pH 8,0); 10 mM de NaCl; 10 mM de $MgCl_2$

Após a diálise, o conteúdo da membrana foi transferido para tubos de microcentrifuga, e adicionou-se EDTA, para uma concentração final de 20 mM, Pronase-E, para uma concentração final de 0,5 mg/ml e SDS, para uma concentração final de 0,5%. Nesta etapa os capsídios dos bacteriófagos serão rompidos e seu conteúdo, o DNA, será liberado. A amostra foi misturada por repetidas inversões e incubada por 1 hora a 37°C. Após o resfriamento, a amostra foi submetida a extração com fenol equilibrado com Tris-HCl 50 mM (pH 8,0), seguida de uma nova extração com fenol / clorofórmio / álcool isoamílico (25:24:1), e por último, uma extração com clorofórmio / álcool isoamílico (24:1). O DNA ficará na fase aquosa, e as proteínas na fase orgânica. Após a última extração, o DNA do bacteriófago foi precipitado com a adição de 0,1 volumes de acetato de sódio 3 M (pH 7,0) e 2 volumes de etanol gelado. A suspensão foi mantida à temperatura ambiente por 30 minutos e, em seguida, centrifugada por 30 minutos, a 4°C, em microcentrifuga. O sedimento obtido foi lavado com etanol 70%. Após a secagem, o sedimento foi solubilizado em 2 ml de TE contendo RNase (40 mg/ml).

Isolamento de DNA nuclear de planta

O DNA utilizado na construção da biblioteca genômica foi obtido utilizando-se a técnica descrita por Rivin *et al.* (1982).

Aproximadamente 1 a 2 g de folhas de plântulas obtidas em condições estéreis e armazenadas a -70°C foram maceradas em nitrogênio líquido até a obtenção de um pó fino. A amostra foi homogenizada em 10 a 20 ml de tampão Tris-HCl 50mM pH8.0; contendo 0,3 M de sacarose e 5 mM de MgCl_2 , em um homogenizador do tipo "polytron" por 30 segundos, na velocidade máxima. O homogenizado foi filtrado em 3 camadas de gaze estéril em funil de Buchner com ajuda de bomba de vácuo. Os núcleos foram sedimentados através de centrifugação a $350 \times g$ por 10 minutos, a 4°C (2.000 rpm em rotor Beckman JA-20). O sobrenadante foi descartado e o sedimento foi ressuspenso em tampão de lise (Tris-HCl 50mM; contendo 20 mM de EDTA). Adicionou-se sarcosyl (n-lauril-sarcosinato de sódio) suficiente para obter-se uma concentração final de 2%, incubou-se por 1 hora a 50°C . Após a incubação, adicionou-se 0,95 g de CsCl por mililitro de lisado. As proteínas

insolúveis e polissacarídeos foram removidos por centrifugação a $17.000 \times g$ (12.000 rpm em rotor Beckman JA-20), por 15 minutos, a 4°C . Ao sobrenadante, coletado adicionou-se brometo de etídio, para uma concentração final de $300 \mu\text{g/ml}$. A concentração de CsCl foi ajustada para o índice de refração da solução de 1,3895. As amostras foram transferidas para tubos de ultracentrífuga e submetidas a ultracentrifugação a 50.000 rpm, por 20 horas, a 20°C em rotor VTi-80 (Beckman). 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 por 2 vezes contra 1.000 volumes de TE, por 24 horas cada. Rendimento aproximado: 2 ml. A concentração de DNA foi determinada espectrofotometricamente.

Isolamento em larga escala de DNA de plasmídio

Para o isolamento em larga escala de DNA de plasmídio foi

utilizada a técnica de lise alcalina descrita em Sambrook *et al.* (1989), com algumas adaptações às condições do laboratório.

A etapa de amplificação descrita no método original foi dispensada para os plasmídios contendo origem de replicação derivada de Col E1, como no caso de pBluescript KS⁺-. A origem de replicação Col E1 apresenta um controle mais relaxado de replicação plasmidial, resultando em um alto número de cópias do plasmídio por célula, e portanto dispensa a amplificação.

A bactéria hospedeira (DH5 α), contendo o plasmídio, foi inoculada em 25 ml de meio líquido LB/ampicilina (60 μ g/ml) e incubada a 37°C durante a noite. Esse cultivo foi utilizado para inocular 500 ml de LB/ampicilina líquido, que foi incubado sob vigorosa agitação (300 rpm) durante a noite a 37°C. O cultivo foi centrifugado a 7.000 rpm em rotor Beckman JA-14, por 10 minutos, a 4°C e o sedimento foi ressuspenso em 4 ml de Solução I (Tris-HCl 25 mM pH 8,0, contendo 50 mM de glicose e 10 mM de EDTA). A suspensão foi transferida para um tubo de centrifuga próprio para rotor Beckman JA-20, adicionou-se 1 ml de solução de lisozima (25 mg/ml em Solução I). A suspensão foi misturada e mantida a

temperatura ambiente por 10 minutos. Em seguida, adicionou-se 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 o desaparecimento das diferentes fases. A suspensão foi mantida em gelo por 10 minutos e centrifugou-se a 12.000 rpm por 20 minutos, a 4°C em rotor Beckman JA-20. O sobrenadante foi filtrado em gaze e transferido para novo tubo de rotor Beckman JA-20. Adicionou-se 0,6 volumes de isopropanol e incubou-se a temperatura ambiente por 10 minutos. Centrifugou-se a 11.000 rpm em rotor Beckman JA-20 por 10 minutos, a temperatura ambiente, e o sedimento obtido foi lavado com etanol 70% a temperatura ambiente. Após a secagem, o sedimento foi ressuspenso em 4 ml de TE, e adicionou-se 4.4 g de CsCl e 0,4 ml de solução de brometo de etídio (10 mg/ml). A amostra foi centrifugada a 10.000 rpm, durante 5 minutos, a temperatura ambiente em rotor Beckman JA-20 para retirar proteínas precipitadas. A concentração de CsCl foi corrigida para um índice de refração da solução de 1,3860. As amostras foram submetidas a

ultracentrifugação em rotor Beckman VTi-80, 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 anteriormente descrito para o isolamento de DNA nuclear de plantas. A concentração de DNA foi determinada espectrofotometricamente.

Purificação de fragmentos de DNA a partir de géis de agarose

Após a digestão com enzimas de restrição apropriadas, o DNA foi submetido à eletroforese em gel de agarose contendo 0,5 µg/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 isolados dos géis utilizando-se papel DEAE (DE-81, Whatman), de acordo com a técnica descrita por Dretzen *et al.* (1981) e Danner (1982). Os fragmentos de DNA foram transferidos eletroforéticamente para papel DE-81 inserido na frente da banda de interesse. A eluição do DNA adsorvido no papel foi realizada incubando-se o papel DE-81 em 800 µl de tampão de eluição (Tris-HCl 20 mM pH 8,0, contendo 2 mM de EDTA e 1,5 M de NaCl), durante 1

hora em banho-maria a 45°C. Após este período, o papel DE-81 foi retirado, e a solução restante foi centrifugada por 5 minutos para a retirada dos fragmentos de papel remanescentes. O DNA foi precipitado com a adição de 2,5 volumes de etanol gelado, incubando-se a -20°C durante a noite. Para fragmentos menores que 100 pb, antes da adição do etanol, adicionou-se MgCl₂ para uma concentração final de 10 mM. Após centrifugação por 20 minutos, os sedimentos foram solubilizados em 10 µl de TE. A concentração de DNA foi determinada espectrofotometricamente.

Biblioteca de cDNA de Coix

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

Seleção de clones de cDNA de opaco2

A primeira busca de fagos recombinantes de *opaco2* foi realizada após a titulação da biblioteca (Sambrook *et al.*, 1989). Colocou-se

em um tubo de ensaio estéril, um volume da suspensão de bacteriófago contendo cerca de 30.000 pfu (unidades formadoras de placa), adicionou-se 0,1 ml de uma suspensão de bactéria XL1-Blue em MgSO_4 10 mM (Sambrook *et al*, 1989). Incubou-se sob leve agitação (100 rpm), a 37°C, durante 20 minutos. Em seguida, foram adicionados 8 ml de "top" agarose pré-aquecida a 60°C, vertendo rapidamente sobre uma placa de 135 mm de diâmetro contendo meio LB sólido. As placas assim preparadas foram incubadas a 37°C durante a noite. As placas foram mantidas em geladeira por 30 minutos, e em seguida, cobertas com membranas de nylon (Hybond-N, Amersham). As membranas foram marcadas em três posições assimétricas com uma agulha contendo tinta nanquim, e foram deixadas sobre as placas durante 1 minuto. As membranas foram então submetidas a incubação por 1 minuto em solução de desnaturação (NaOH 0,5 M; NaCl 1,5 M). Após a desnaturação, as membranas foram incubadas por 5 minutos em solução de neutralização (NaOH 0,25 M; NaCl 1,5 M) sendo em seguida, lavadas em 2X SSC (1X SSC contém 0,15 M de NaCl e 15 mM de citrato de sódio) por 15 minutos. Após a secagem, as

membranas foram incubadas por 2 horas em estufa a 80°C.

As membranas foram pré-hibridizadas a 42°C, por 4 horas, em uma solução contendo 5X SSC, 0,05X BLOTTO, 0,02% SDS, 0,1% sarcosyl (n-lauril-sarcosinato de sódio) e 50% de formamida deionizada. As hibridizações foram realizadas por 12 horas nas mesmas condições da pré-hibridização, utilizando-se como sonda o fragmento Eco RI de 1.050 pb do clone de cDNA de o2 de milho, o2 cDNA 1-4 (cedido pelo Dr. Robert Schmidt, Department of Biology and Center of Molecular Genetics, University of California, San Diego, U.S.A). Após dupla purificação em gel de agarose, o fragmento foi marcado pela incorporação de nucleotídeos radioativos ($[\alpha\text{-P}^{32}]$ dNTPs), através de reação de extensão catalisada pelo fragmento Klenow da DNA-polimerase I, utilizando-se como "primer" hexanucleotídeos aleatórios (Feinberg & Vogelstein, 1984).

Foram utilizadas sondas com atividade específica de 10^8 a 10^9 cpm/ μg , adicionadas às soluções de hibridização para uma atividade final de aproximadamente 10^6 cpm/ml. As sondas a serem utilizadas nas hibridizações foram previamente

desnaturadas em banho-maria fervente por 5 minutos.

Durante a primeira busca, foram analisados, cerca de, $2,5 \times 10^5$ bacteriófagos recombinantes, espalhados em 8 placas de Petri de 135 mm de diâmetro. Os recombinantes positivos foram isolados das placas (Sambrook *et al.*, 1989), e posteriormente, submetidos a mais duas análises sucessivas de hibridização em placas de 85 mm contendo de 1.000 a 3.000 pfu.

Excisão dos clones recombinantes de cDNA de *opaco2*

A técnica de excisão *in vivo* (Apêndice III) tem como objetivo recuperar o inserto (cDNA) clonado no vetor λ ZAP. A placa de lise de interesse, contendo o bacteriófago recombinante selecionado, foi transferida para um tubo de microcentrifuga contendo 0,5 ml de tampão SM e 20 μ l de clorofórmio. Esse tubo de microcentrifuga foi agitado em vortex e incubado por 2 horas à temperatura ambiente. Em um tubo Falcon de 50 ml, foram colocados 200 μ l do sobrenadante desse tubo de microcentrifuga (contendo $\sim 1 \times 10^5$ pfu), 200 μ l de bactéria XL1-Blue $DO_{600} = 1,0$ e mais 1 μ l de fago "helper" R408 (contendo $\sim 1 \times 10^6$

pfu/ml). Essa mistura foi incubada por 15 minutos a 37°C. Após esse período, adicionou-se 5 ml de meio 2X YT, incubando-se sob agitação (300 rpm) a 37°C durante 3 horas. Em seguida, foi feita outra incubação, agora a 70°C por 20 minutos (para inativação de fago "helper"), sendo posteriormente centrifugado por 5 minutos a 4.000 x g. O sobrenadante contendo o "phagemid" foi coletado em um tubo estéril. Para o plaqueamento do "phagemid", foram preparadas diluições seriadas do "phagemid" até 10^{-7} . Em 200 μ l de bactéria XL1-Blue ($DO_{600} = 1,0$) foram adicionados 10 μ l de cada diluição. A mistura foi incubada a 37°C, durante 15 minutos. Foram semeados 100 μ l de cada cultura em placas de LB/ampicilina (60 μ g/ml), as quais foram incubadas durante a noite a 42°C.

Biblioteca genômica de *Coix*

A biblioteca genômica de *Coix* foi construída em bacteriófago λ DASH (Stratagene). Maiores detalhes do vetor utilizado nesta etapa podem ser encontrados no Apêndice III.

O DNA genômico de *Coix* foi parcialmente digerido com a enzima de restrição Mbo I. Fragmentos de 9 a 23 kb foram isolados através de

ultracentrifugação em gradiente de sacarose por 16 horas a 30.000 rpm (Sambrook *et al.*, 1989).

Esses fragmentos foram então ligados aos braços do fago λ DASH (Stratagene), os quais foram previamente digeridos com a enzima de restrição Bam HI.

Para esta digestão parcial utilizou-se a enzima Mbo I, pois esta produz extremidades compatíveis com Bam HI. Além disto, esta enzima reconhece um sítio de apenas 4 pb produzindo um maior número de fragmentos de restrição de tamanho apropriado para a inserção no DNA do bacteriófago. A utilização do isoesquizômero Mbo I, cuja atividade é inibida pela metilação em adeninas, é justificada pelo fato de que a metilação em DNA de plantas ocorre preferencialmente em citosinas.

Após a ligação dos braços do bacteriófago λ DASH com os fragmentos de DNA genômico, os bacteriófagos recombinantes foram empacotados "*in vitro*" utilizando-se o kit GIGAPACK II - PACKAGING EXTRACT (Stratagene).

Foram analisados cerca de 1×10^6 bacteriófagos recombinantes, semeados em 35 placas de Petri de 135 mm de diâmetro. Esses fagos foram transferidos para membranas de nylon

(Hybond-N, Amersham), e hibridizados com a sonda radioativa de cDNA de o2 de Coix (pCO2-1). As semeaduras, transferências e hibridizações foram realizadas como descrito anteriormente para a biblioteca de cDNA. Os clones positivos foram isolados através de três seleções sucessivas.

Sequenciamento de DNA

Os sequenciamentos foram realizados com DNA dupla fita de plasmídios recombinantes de pBluescript KS⁺/. Amostras de 40 μ l, contendo 8 μ g de DNA plasmidial obtidos através de minipreparação, foram desnaturadas adicionando-se 40 μ l de uma solução contendo 0,4 M de NaOH e 0,4 mM de EDTA. As amostras foram incubadas por 5 minutos à temperatura ambiente, e a desnaturação foi interrompida pela adição simultânea de 8 μ l de acetato de amônio 2 M pH 4,5 e 176 μ l de etanol gelado. As amostras foram mantidas a -20°C por 1 hora e centrifugadas 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 25 μ l de água. Alíquotas de 10 μ l foram utilizadas nas reações de sequenciamento segundo a técnica de

dideoxi-nucleotídeos (Sanger *et al.*, 1977). As reações foram realizadas conforme as instruções fornecidas pelo fabricante do "T7 Sequencing Kit" (Pharmacia). Foram utilizados os "primers" direto, fornecido pelo kit, e o reverso (Boehringer-Mannheim). Os produtos da reação de sequenciamento, marcados radioativamente pela incorporação de (^{35}S)dATP α S (Amersham), foram analisados através de eletroforese em gel desnaturante de poliacrilamida. A eletroforese dos produtos de reação foi realizada carregando-se 2 μl de reação em géis desnaturantes 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 3 minutos a 80°C antes de serem aplicadas nos géis. Os dois géis foram submetidos a aproximadamente 2.000 volts, sendo que a potência máxima foi estabelecida para 50 watts para o conjunto dos géis, os quais foram montados no sistema de eletroforese MacroPhor (LKB). O gel, para leitura da sequência localizada próxima ao "primer" (gel de 6%) foi submetido à eletroforese até a saída

do corante azul de bromofenol. Enquanto, o gel para leitura da sequência mais afastada do "primer" (gel de 5%) foi submetido à eletroforese por 2 horas adicionais após a saída do corante xileno cianol. Após a eletroforese, os géis aderidos às placas foram incubados por 30 minutos sob leve agitação em uma solução contendo 10% de ácido acético e 10% de metanol (para retirada de uréia e fixação do DNA), e em seguida, lavados exaustivamente em água corrente. Os géis foram secos a 45°C durante a noite, e submetidos à auto-radiografia através de exposição de filmes sensíveis à radiação (Hyperfilm - Amersham) por 1 a 4 dias.

Deleção Unidirecional com a enzima Exonuclease III

A técnica utilizada foi a descrita por Henikoff (1987), com algumas adaptações às condições do laboratório. Os princípios básicos da técnica encontram-se descritos no Apêndice V

O fragmento HAVE1 (Tabela I) previamente clonado nos sítios Hind III e Eco RI do plasmídeo pBluescript KS⁺ foi digerido com as enzimas Xho I e Kpn I. Estas enzimas produzem extremidades 5' e 3' protuberantes, respectivamente. A

Exonuclease III age uniformemente sobre extremidades 5' protuberantes ou abruptas, não atacando extremidades 3' protuberantes. Portanto o tratamento do inserto Xho I-Kpn I com Exonuclease III permitiu a deleção unidirecional a partir da extremidade Xho I. Trinta microgramas de DNA foram digeridos com Xho I - Kpn I por 3 horas, a 37°C. O DNA digerido foi então extraído uma vez com fenol / clorofórmio / álcool isoamílico (25:24:1) e uma vez com clorofórmio / álcool isoamílico (24:1). Após a última extração, à fase aquosa foram adicionados 0,1 volumes de cloreto de sódio 2 M e 2 volumes de etanol gelado. A suspensão foi incubada a -70°C, por 30 minutos e, em seguida, centrifugada por 30 minutos, a 4°C. O sobrenadante foi descartado, e o precipitado foi lavado com etanol 70%. O precipitado liofilizado foi solubilizado em 65µl do Tampão de Exonuclease III (Tris-HCl 66 mM pH 8.0, contendo 66 mM de MgCl₂). Uma alíquota de 5 µl foi submetida a eletroforese em gel de agarose para confirmar a completa digestão do DNA. Aos 60µl restantes da amostra, foram adicionadas 300 U de Exonuclease III (GIBCO-BRL). Esta reação foi realizada a temperatura ambiente. Foram retiradas 25 alíquotas, contendo 2,5 µl, a cada 30

segundos. Estas alíquotas foram, imediatamente transferidas para tubos de microcentrifuga que continham 7,5 µl de Mistura de S1 Nuclease (Apêndice I). Esta reação foi incubada por 30 minutos, a temperatura ambiente. Adicionou-se então, 1 µl do Tampão "Stop" de S1 Nuclease (0,3 M de Trizma Base; 50 mM de EDTA), e incubou-se por 10 minutos a 70°C. Foram retiradas alíquotas de 2 µl, as quais foram submetidas à eletroforese em gel de agarose para conferir a extensão das deleções ocorridas. Enquanto que, ao restante das amostras, adicionou-se 1 µl de Mistura de Klenow (Tris-HCl 10 mM pH 8.0, contendo 20 mM de MgCl₂ e 3 U de polimerase Klenow) e 1 µl de Mistura de Nucleotídeos 0,4 mM (Apêndice I). Incubou-se as reações por 10 minutos, a 37°C. Após esta incubação, adicionou-se 40 µl da Mistura de Ligase (Apêndice I), incubando-se, por 2 horas, a 16°C.

Os plasmídios assim obtidos foram utilizados na transformação de bactérias DH5α. Após a seleção dos clones deletados, foram realizadas Minipreparações para obtenção do DNA utilizado no sequenciamento.

O fragmento EE/H3 (Tabela I) também foi previamente clonado nos sítios Hind III e Eco RI do plasmídio pBluescript KS⁺. Este plasmídeo foi

submetido ao mesmo procedimento para obtenção das deleções descrito anteriormente, exceto pelo fato de que as extremidades protuberantes 5' e 3' foram produzidas por Eco RI e Pst I, respectivamente.

Construção dos plasmídios utilizados nos testes de ligação seletiva

Todos os plasmídios utilizados estão listados no Apêndice VI. Todas as posições indicadas nas construções dos clones abaixo são com relação ao codon ATG inicial.

pO2CX

O clone de cDNA de o2 de *Coix*, pCO2-1, foi digerido com as enzimas de restrição Hind III e Bam HI, liberando um fragmento de 886 pb, contendo a região da bZIP. Este fragmento foi então inserido no plasmídio pUR289 (Apêndice V), previamente digerido com Hind III e Bam HI (Souza Filho, G.A., resultados não publicados). Esta construção permitiu a produção da proteína quimérica β -Gal::O2 de *Coix* utilizada nos experimentos de ligação seletiva.

P285 α -C

O fragmento Eco RI-Stu I de 0.3 kb do subclone do promotor do gene de α -coixina de 25 kDa, α -3B (Ottoboni *et al.*, 1993), foi amplificado utilizando-se o "primer" universal de sequenciamento e o "primer" Alfacoix-ATG (5'-TGCTGGATCCATGGCGACAAGATGCTGCT-3'; BioLabs), parcialmente complementar a região -18 a +8, devido a introdução dos sítios de restrição de Bam HI (5'-GGATCC-3') e Nco I (5'-CCATGG-3') próximo ao códon ATG inicial. O fragmento amplificado foi digerido com Eco RI e Bam HI e inserido no plasmídio pBluescript KS⁺. O fragmento resultante foi verificado através de sequenciamento (Yunes *et al.*, 1994).

P285 Δ O2

Os 285 pb do fragmento do promotor inserido no clone P285 α -C, foram amplificados utilizando-se o "primer" universal de sequenciamento e o "primer" Alfacoix-272 (5'-CTGCGAATTCAAGCGACACTTTTTTCTT-3'; BioLabs), complementar a região de -285 a -262 gerando a deleção da sequência CAATAGA, correspondente ao sítio O2b. O fragmento amplificado foi digerido com Eco RI e Bam HI e inserido no plasmídio pBluescript KS⁺.

O fragmento resultante foi verificado através de sequenciamento (Yunes *et al.*, 1994).

P285 CATG

O fragmento do promotor de α -coixina de 25 kDa, clonado no plasmídio P285 α -C, foi mutado através de reação assimétrica de PCR ("Polimerase Chain Reaction"), conforme descrito por Perrin & Gilliland (1990). Foram utilizados os "primers" de sequenciamento de DNA, universal e reverso, e um "primer" com uma mutação no sítio O2c, denominado O2c-ACGT (5'-TAGGA GACATATCATCTCTAG-3'; National Biosciences, Plymouth, USA). Este "primer" mutante é complementar a região -177 a -157 do promotor das α -coixina de 25 kDa, e gera a substituição de um "G" por "A" na sequência alvo da O2 de milho, GACATGTC (Yunes *et al.*, 1994). O fragmento amplificado foi digerido com Eco RI e Bam HI, e inserido no plasmídio pBluescript KS⁺. O fragmento resultante foi verificado através de sequenciamento.

P285 TGAC

Este plasmídio foi gerado de forma similar ao P285 CATG, exceto

que o "primer" mutante utilizado foi o oligonucleotídeo O2c-TGAC (5'-TGA CAAAATTAAGTAGGAGAC-3'; National Biosciences, Plymouth, USA). Este "primer" mutante é complementar a região -181 a -170 e gera a substituição de um "G" por "A" em uma sequência similar ao sítio de reconhecimento do transativador de fungos GCN4 (5'-ATTGACTA-3'). O fragmento amplificado foi digerido com Eco RI e Bam HI, e inserido no plasmídio pBluescript KS⁺. O fragmento resultante foi verificado através de sequenciamento.

P α -Cx103

O fragmento de -1084 até -285 do promotor de coixina, α -3C (Ottoboni *et al.*, 1993), foi retirado através da digestão com Hind III e Eco RI. Este fragmento de 0.8 kb foi inserido na construção P285 α -C, originando o clone P α -Cx. Este plasmídio contém o promotor inteiro do gene de α -coixina de 25 kDa, α -3C. Este promotor foi, então, retirado através da digestão com Hinc II e Nco I, e inserido nos mesmos sítios no plasmídio pRT103 (Töpfer *et al.*, 1987), gerando, assim, o plasmídio P α -Cx103. O fragmento resultante foi verificado através de sequenciamento (Yunes *et al.*, 1994).

P α -Z4103

Um fragmento do promotor de α -zeína de 22 kDa, cedido pelo Dr. Brian A. Larkins (Department of Plant Science, University of Arizona, Tucson, U.S.A), de -841 a +1, foi isolado a partir da digestão com Acc I, tratado com o fragmento Klenow da DNA polimerase I, e finalmente, digerido com Nco I. Esse fragmento de 842 pb foi, então, inserido no plasmídeo pRT103, nos sítios de Hinc II e Nco I (Yunes *et al.*, 1994).

Teste de ligação seletiva

Os teste de ligação seletiva de fragmentos de DNA e proteínas O2 de milho e Coix foram realizados conforme descrito por Schmidt *et al.* (1990). Os princípios da técnica estão descritos no Apêndice V

Etapa 1:

Aproximadamente, 1.5 μ g de DNA de cada plasmídeo foi digerido com as enzimas de restrição apropriadas, conforme mostrado nas Figuras 15 e 16. As digestões foram paralizadas com a adição de EDTA 0,5 M suficiente para obter-se uma concentração final de 10 mM, e incubadas a 65°C, por 15 minutos. No caso de enzimas de restrição termo-

estáveis foram realizadas extrações com fenol / clorofórmio / álcool isoamílico (25:24:1) e clorofórmio / álcool isoamílico (24:1), seguidas de precipitação dos fragmentos com a adição de 0,1 volumes de acetato de sódio 3 M pH 5,2 e 2 volumes de etanol. Os precipitados obtidos foram solubilizados em 40 μ l de TE.

Após a inativação, as digestões foram confirmadas através da eletroforese em géis de agarose 2%, com alíquotas correspondentes a 0,5 μ g de DNA.

Os fragmentos resultantes das digestões foram marcados através da incorporação, em suas extremidades, de α [P³²]dNTPs por catálise mediada pela polimerase Klenow (Sambrook *et al.*, 1989). As reações foram realizadas a temperatura ambiente, por 15 minutos. Após a incubação, foram adicionados 2 μ l EDTA 0,5 M, 50 μ l acetato de amônio 4 M pH 5,0 e 200 μ l etanol, e incubou-se a -70°C por 30 minutos. Centrifugou-se por 15 minutos, e o sedimento foi lavado com 500 μ l de acetato de amônio 0,67 M em etanol 67%, centrifugou-se por mais 5 minutos. O sedimento foi lavado novamente com 500 μ l de etanol 90%, e centrifugado por mais 10 minutos. O precipitado foi seco e ressuspenso em 100 μ l TE contendo 150 mM de cloreto de sódio.

Etapa 2:

As proteínas quiméricas recombinantes β -Gal::O2 de milho e Coix foram obtidas através de indução de culturas de bactérias XL1-blue contendo os plasmídios LC#37 (Schmidt *et al.*, 1992) e pO2CX, respectivamente.

Foram inoculados 25 ml de meio LB/ampicilina (60 μ g/ml) com uma alíquota de um pré-inóculo, obtido durante a noite, suficiente para se obter uma $DO_{600} = 0,1$. Incubou-se a 37°C, até atingir $DO_{600} = 0,4$, quando então adicionou-se 0,5 mM de IPTG (isopropil tio- β -D-galactosídeo), e, incubou-se a 37°C, por mais 4 horas. A cultura foi submetida a centrifugação por 20 minutos, a 4°C, o sedimento foi ressuspenso em 500 μ l de TE, congelado em nitrogênio líquido e armazenado a -70°C.

Após o descongelamento da suspensão de bactérias, adicionou-se 10 ml de TEP (Apêndice I). As bactérias foram lisadas através de sonicação em 5 ciclos de 10 segundos. O lisado foi centrifugado e distribuído, em alíquotas de 200 μ l, em tubos de microcentrifuga (um tubo para cada teste). As proteínas foram precipitadas adicionando-se 600 μ l de uma solução saturada de sulfato de amônio (35 g / 50 ml), Incubou-se em

gelo por 15 minutos e centrifugou-se por 10 minutos. O precipitado obtido foi dissolvido em 60 μ l de TEP, e adicionou-se 10,6 μ l de cloreto de sódio 1 M (para uma concentração final de 150 mM) e 3 μ g de anti-corpo monoclonal anti- β -Galactosidase (Promega). A mistura foi incubada em gelo por 30 minutos e, em seguida adicionou-se 15 μ l de Proteína A-Sepharose 50% (Pharmacia) em "Binding Buffer" (B.B.) (Apêndice I). A suspensão foi incubada em gelo, sob agitação constante. Após 2 horas, a suspensão foi submetida a uma breve centrifugação, 5 segundos, e o sedimento foi ressuspenso em 100 μ l de BB. Esta etapa foi repetida por mais três vezes para a completa retirada do excesso dos complexos anti- β -Gal::O2 não adsorvidos à resina.

Etapa 3:

O sedimento foi finalmente ressuspenso em 40 μ l de B.B. e recebeu a adição de 5 μ l de DNA de espermatozoides de salmão (1 μ g/ μ l) e 10 μ l dos fragmentos de DNA marcados anteriormente (~300.000 cpms). A suspensão foi incubada a temperatura ambiente, por 30 minutos, sob leve agitação. A resina foi então, lavada por no mínimo seis vezes com 200 μ l

de B.B., para a retirada dos fragmentos de DNA não adsorvidos ao complexo Prot. A-Seph.:Anti- β -Gal.: β -Gal-O2 e ressuspensa em 200 μ l de B.B..

Os fragmentos de DNA adsorvidos foram eluídos através de extração com fenol, e precipitados pela adição de 20 μ l de acetato de sódio 3 M pH 5,2 (0,1 volumes), 10 μ l de DNA de esperma de salmão (1 μ g/ μ l) e 580 μ l de etanol (2,5 volumes). As amostras foram incubadas por 30 minutos a -70°C ou durante a noite a -20°C. Os sedimentos obtidos após centrifugação por 30 minutos foram ressuspensos em 15 μ l de água. Alíquotas contendo atividades normalizadas para aproximadamente um mesmo número de cpms por fragmento marcado foram submetidas a eletroforese em gel de agarose 2%. Após a eletroforese, o gel foi seco ("Slab Gel Dryer", LKB) e submetido à auto-radiografia através de exposição de filmes sensíveis à radiações (Hyperfilm - Amersham) por 1 a 24 horas.

Extração de RNA

A metodologia empregada foi a descrita por Prescott & Martin (1987). Durante a preparação dos RNAs foi utilizado material de plástico descar-

tável, toda vidraria incubada por 24 horas em estufa a 180°C. Toda água e soluções utilizadas foram tratadas com DEPC (dietil-pirocarbonato) conforme descrito no Apêndice I. O DEPC é um agente que promove carboxi-metilação de proteínas, inativando possíveis RNases presentes.

Aproximadamente 2 g de tecido congelado, foram pulverizados e homogenizados em 10 ml de Tampão de extração (Tris-HCl 50 mM pH 9,0 contendo 150 mM de LiCl; 5 mM de EDTA; e 5% de SDS (m/v)). Acrescentou-se 10 ml de fenol equilibrado em Tampão citrato 0,1 M pH 4,5 / clorofórmio / álcool isoamílico (25:24:1). Misturou-se bem, e centrifugou-se por 10 minutos, a 3.000 rpm. A fase aquosa foi submetida a nova extração com fenol pH 4,5 / clorofórmio / álcool isoamílico. Após a centrifugação, 8 ml da fase aquosa foram transferidos para 8 tubos de microcentrifuga e adicionou-se à cada tubo 400 μ l de cloreto de lítio 8 M. Incubou-se, a -20°C, por 30 minutos e, em seguida, centrifugou-se por 20 minutos, a 10.000 rpm. O precipitado foi dissolvido em 600 μ l de água e adicionou-se 200 μ l de cloreto de lítio 8 M (para uma concentração final de 2 M). Incubou-se a -20°C por 60 minutos e, em seguida, centrifugou-se por 20 minutos, a

10.000 rpm. O precipitado foi dissolvido em 100 µl de água e, o conteúdo dos 8 tubos reunido em dois tubos. A estes tubos adicionou-se 40 µl de acetato de sódio 3 M pH 5,2 (0,1 volumes) e 1,1 ml de etanol (2,5 volumes). Incubou-se, a -20°C, durante a noite, e, centrifugou-se por 30 minutos, a 10.000 rpm. O precipitado foi lavado com 500 µl de etanol 70%, seco a temperatura ambiente e dissolvido em 100 µl de água. O RNA foi quantificado espectrofotometricamente.

Com a finalidade de verificar a integridade dos RNAs preparados, 2 µg de cada amostra foram desnaturados a 65°C por 15 minutos, e submetidos a eletroforese em gel de agarose 1% contendo 100 µg/ml de ácido iodo-acético (agente que promove carboxi-metilação de grupos sulfidríla de proteínas, inativando possíveis RNases presentes). O Tampão de eletroforese utilizado foi 1X TAE. A cuba de eletroforese bem como o pente foram previamente tratados com 0,4 M de NaOH por 2 horas (O NaOH promove a desnaturação de proteínas, inativando possíveis RNases presentes).

As amostras de RNA foram diluídas em água para uma concentração final de 1 µg/µl.

Northern blot

Os experimentos de Northern blot foram realizados através da transferência de amostras de RNA submetidas a eletroforese em géis desnaturantes de agarose, na presença de formaldeído 0,66 M, conforme descrito por Fourney *et al.* (1988).

Preparação das amostras

A 12 µl de amostra de RNA foram adicionados 38 µl de Tampão de Desnaturação de RNA (Apêndice I). Incubou-se a mistura a 65°C por 15 minutos. Após a adição de 10 µl de Tampão de Amostra de RNA (Apêndice I), as amostras de RNA foram submetidas a eletroforese em gel desnaturante de agarose 1% em 0,66 M de formaldeído, por cerca de 3 horas, 80Volts, Tampão TAE 1X.

Preparação do gel desnaturante e eletroforese

Três gramas de agarose foram dissolvidos em 150 ml de água tratada com DEPC, e em seguida, adicionou-se 30 ml de 10X MOPS, 15,3 ml de formaldeído 37% e 105 ml de água tratada com DEPC (concentração final do gel: 1%). O Tampão de eletro-

forese utilizado foi 1X MOPS. A cuba e o pente utilizados foram previamente tratados com NaOH 0,4 M por 2 horas.

Transferência

Após a eletroforese, o RNA foi transferido por capilaridade, para membrana de nylon (Hybond-N, Amersham), utilizando-se 10X SSC, conforme descrito por Sambrook *et al.* (1989). O RNA foi fixado a membrana através de incubação por 2 horas, a 80°C.

Hibridização

A metodologia empregada na hibridização das membranas foi a descrita por Siengel & Bresnik (1986). Durante a pré-hibridização, a membrana foi incubada a 42°C por 2 horas, em Solução de pré-hibridização (6X SSC; 0,2X BLOTTO; 50% formamida dionizada).

A hibridização foi realizada, durante a noite, em Solução de pré-hibridização acrescida da sonda radioativa desnaturada por 5 minutos em banho-maria fervente.

Após a hibridização, as membranas foram lavadas duas vezes em solução 2X SSC e 0,1% SDS (m/v)

a temperatura ambiente, por 20 minutos, e duas vezes por 20 minutos, em solução 0,3X SSC e 0,1% SDS (m/v) a 65°C.

Após a lavagem, as membranas foram envolvidas em filme de PVC e submetidas à auto-radiografia através de exposição de filmes sensíveis à radiação (Hyperfilm - Amersham) por 1 a 2 dias.

Métodos computacionais

Todos os programas utilizados foram executados em microcomputador do tipo PC 386.

A edição e análise das sequências de DNA foram realizadas com ajuda dos programas DNASIS (Pharmacia) e ESEE v. 1.04 (Cabot, 1988).

No alinhamento das sequências de DNA e proteínas, foi empregado o programa Clustal V (Higgins & Sharp, 1989).

Os alinhamentos através de "dot matrix" foram realizados utilizando-se o programa DIAGON (Staden, 1982).

RESULTADOS

O clone de cDNA

Isolamento de clones de cDNA de *opaco2* de *Coix*

Uma biblioteca de cDNA de endosperma imaturo de *Coix lacryma-jobi* em fago λ ZAP (Stratagene) previamente construída (Leite *et al.*, 1992), foi utilizada no isolamento de clones de sequências homólogas ao gene *opaco2* de milho. Desta biblioteca foram analisados aproximadamente $7,5 \times 10^5$ fagos recombinantes. Para esta análise utilizou-se como sonda o fragmento Eco RI de 1.050 pb do clone de cDNA de *opaco2* de milho, o cDNA 1-4, cedido pelo Dr. Robert Schmidt (Department of Biology and Center of Molecular Genetics, University of California, San Diego, U.S.A), marcado radioativamente.

Os clones positivos foram isolados após três ciclos de hibridizações sucessivas. Na 1ª seleção foram identificados dois possíveis clones positivos, os quais foram isolados e confirmados nas 2ª e 3ª seleções.

Após a seleção dos clones recombinantes positivos, os respecti-

vos plasmídios pBluescript SK⁻ foram excisados. Amostras de DNA desses plasmídios foram analisadas através de digestão com enzimas de restrição, seguidas de eletroforese em gel de agarose.

Os dois clones, denominados de pCO2-1 e pCO2-2, apresentaram mapas de restrição semelhantes. O mapa de restrição do clone pCO2-1 é mostrado na Figura 4.

Sequenciamento do clone de *opaco2*

Com o objetivo de sequenciar o clone pCO2-1, foram construídos três subclones: pCO2-1 / Eco B, pCO2-1 / Eco A, pCO2-1 / Eco-Hind B. Para a construção do subclone pCO2-1 / Eco B, o DNA do clone pCO2-1 foi digerido com a enzima de restrição Eco RI, e o inserto liberado foi isolado e ligado no sítio Eco RI do pBluescript KS⁺. O subclone pCO2-1 / Eco A foi obtido da recirculização, através da ligação do fragmento remanescente após a digestão do clone pCO2-1 com a enzima de restrição Eco RI. O subclone pCO2-1 / Eco-Hind B foi

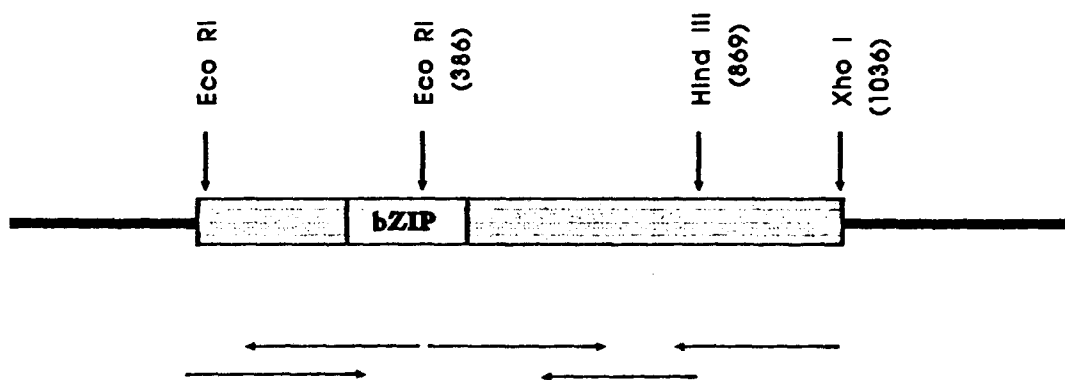
pCO2-1

Figura 4: Representação esquemática do clone de cDNA de *opaco2* de *Coix* (pCO2-1). O clone isolado apresenta 1.036 pb, contendo o motivo bZIP (retângulo branco), que inclui os domínios básico e "zipper" de leucina. Os sítios de restrição de algumas enzimas são indicados. As setas na parte inferior representam a estratégia adotada para o sequenciamento.

obtido pela inserção do fragmento Eco RI / Hind III, obtido a partir do subclone pCO2-1 / Eco A, em pBluescript KS⁺. O clone pCO2-1 e estes subclones foram então sequenciados, em ambas as direções, utilizando-se os "primers" universal e reverso.

A sequência de nucleotídeos do clone pCO2-1 mostrou alta similaridade com o gene *opaco2* de milho. Entretanto, o alinhamento das sequências de cDNA desses clones, indicou que o clone de cDNA de *Coix* obtido era incompleto. Este fato foi posteriormente confirmado com o isolamento e sequenciamento do clone genômico (a sequência de nucleotídeos do clone pCO2-1 está representada na Figura 6). A sequência de cDNA de *opaco2* de *Coix* foi depositada nos bancos de dados EMBL, GenBank e DDBJ com o número de acesso X78286, e código de identificação CLJOP2.

Na tentativa de se isolar um clone completo de cDNA de *opaco2* de *Coix*, a biblioteca de cDNA foi reanalisada por 2 vezes, utilizando-se como sonda o clone pCO2-1. Em nenhuma das tentativas foi obtido um clone de cDNA completo.

O clone genômico

Seleção do clone genômico de *opaco2* de *Coix*

De uma biblioteca genômica de *Coix*, recém construída em fago λ DASH, foram analisados cerca de 1×10^6 fagos recombinantes. Esses fagos foram transferidos para membranas de nylon (Hybond-N, Amersham), e hibridizados com uma sonda radioativa de cDNA de *opaco2* de *Coix* (pCO2-1). Os clones positivos foram isolados através de três seleções sucessivas. Na 1^a seleção foram isolados 7 clones positivos, dos quais apenas 3 foram confirmados nas demais seleções. Os clones recombinantes foram denominados gCXO2-1, gCXO2-3, e gCXO2-5.

Sequenciamento do clone genômico de *opaco2* de *Coix*

Foram realizadas preparações em grande escala dos DNAs dos clones recombinantes positivos. Esses DNAs foram então digeridos com enzimas de restrição e os padrões de restrição obtidos foram comparados com o padrão do clone genômico de *opaco2* de milho.

O padrão de restrição do clone gCXO2-1 apresentou uma grande semelhança com o do clone genômico

de o2 de milho (dados não apresentados). Esse clone, que contém uma inserção de aproximadamente 14 kb, foi digerido com algumas enzimas de restrição e, após a análise dos fragmentos, através de "Southern blot", utilizando-se o clone pCO2-1 como sonda, os fragmentos que hibridizaram com a sonda foram subclonados em pBluescript KS+. Todos os subclones obtidos são apresentados na Tabela I.

Com a finalidade de obter-se a superposição de sequências durante o processo de sequenciamento, foram também realizadas deleções unidirecionais dos subclones HAEI e EE/H3 I. Os subclones obtidos neste processo são apresentados na Tabela II. O sequenciamento dos subclones obtidos revelou uma sequência de 4.569 pb da região que contém o gene o2 de *Coix*. A partir desta sequência foi possível construir o mapa de restrição apresentado na Figura 5.

Descrição do clone genômico de o2 de *Coix*

A Figura 6 mostra a sequência de nucleotídeos da região sequenciada do clone gCXO2-1. A sequência inclui 987, 215, e 3.367 pb, respectivamente das regiões flanqueadoras 5' e 3', e da região estrutural do gene o2 de *Coix*. Foi considerada como

posição +1 a primeira base do codon ATG, que inicia a principal fase de leitura. Sequências correspondentes aos possíveis "TATA box" e "CCAAT box", conforme as sequências consenso descritas por Bucher (1990), foram identificadas nas posições -435 (5'-TATAAAA-3') e -511 (5'-CCAAT-3'), respectivamente.

Na região abaixo do possível "TATA box" foram identificadas 8 fases de leituras. A principal fase foi identificada através da comparação com as sequências genômica e de cDNA de o2 de milho (Hartings *et al.*, 1989; Schmidt *et al.*, 1990). Esta fase de leitura codifica uma proteína de 408 aminoácidos, correspondendo a um polipeptídeo de peso molecular 44.865. As demais fases de leitura localizadas entre o "TATA box" e a principal fase codificam polipeptídeos de tamanho variando entre 3 e 76 aminoácidos (Fig. 7).

A região codificadora do gene o2 de *Coix* é composta por seis exons, de tamanhos variando de 55 a 436 pb, e por cinco íntrons com tamanhos variando de 120 a 1094 pb. Todos cinco íntrons apresentam as terminações típicas de bordas de íntrons GT-AG (Fig. 6) (Breathnach & Chambon, 1981).

Um provável sinal de poliadenilação (5'-AAATAA-3') está presente

Tabela I: Relação dos subclones utilizados no sequenciamento do clone genômico de *opaco2* de *Coix*.

Nome	extremidades do inserto	vetor	posição 5'	posição 3'
HA	Hind III	KS ⁺ / Hind III	1	3363
HB	Hind III	KS ⁺ / Hind III	3363	4369
EE	Eco RI	KS ⁺ / Eco RI	2132	3716
EB	Eco RI	KS ⁺ / Eco RI	nd	3716
EE/H3	Eco RI/Hind III	KS ⁺ / Eco RI-Hind III	2132	3363
EE/H2 I	Eco RI	KS ⁺ / Eco RI	2132	3440
EE/H2 II	Hinc II	KS ⁺ / Hinc II	3440	3774
HB/H2 III	Hinc II	KS ⁺ / Hinc II	3440	3774
HB/H2 I	Hind III	KS ⁺ / Hind III	3774	4157
HB/H2 V	Hinc II	KS ⁺ / Hinc II	4157	4369
EB/H3	Eco RI/Hind III	KS ⁺ / Eco RI-Hind III	4369	nd
HA/EI	Eco RI/Hind III	KS ⁺ / Eco RI-Hind III	1	2132
HA/EI/Hinf I	Bam HI/Pst I	KS ⁺ / Sma I	1	1209
HA/EI/Hinf II	Bam HI/Pst I	KS ⁺ / Sma I	1209	2132
HA/EI/Hae II	Bam HI/Pst I	KS ⁺ / Sma I	1380	2132
HA/EI/Hae IIIc	Bam HI/Pst I	KS ⁺ / Sma I	1	503
HA/EI/Hae IIIb	Bam HI/Pst I	KS ⁺ / Sma I	503	730
HA/EI/Xba I	Eco RI/Hind III	KS ⁺ / Eco RI-Hind III	1	1597
HA/EI/Xba II	Xba I	KS ⁺ / Xba I	1555	nd

Tabela II: Relação dos subclones utilizados no sequenciamento do clone genômico de opaco2 de Coix, obtidos através de Deleção Unidirecional com Exonuclease III.

Nome	posição 5'	posição 3'
HA / EI - 1	69	2132
HA / EI - 2	201	2132
HA / EI - 3	171	2132
HA / EI - 5	201	2132
HA / EI - 8	823	2132
HA / EI - 15	630	2132
HA / EI - 18	740	2132
HA / EI - 21	829	2132
HA / EI - 24	848	2132
EE / H3 - 1	2260	3363
EE / H3 - 3	2815	3363
EE / H3 - 5	3026	3363
EE / H3 - 6	2240	3363
EE / H3 - 12	2405	3363
EE / H3 - 14	2380	3363
EE / H3 - 22	2492	3363

gCXO2-1

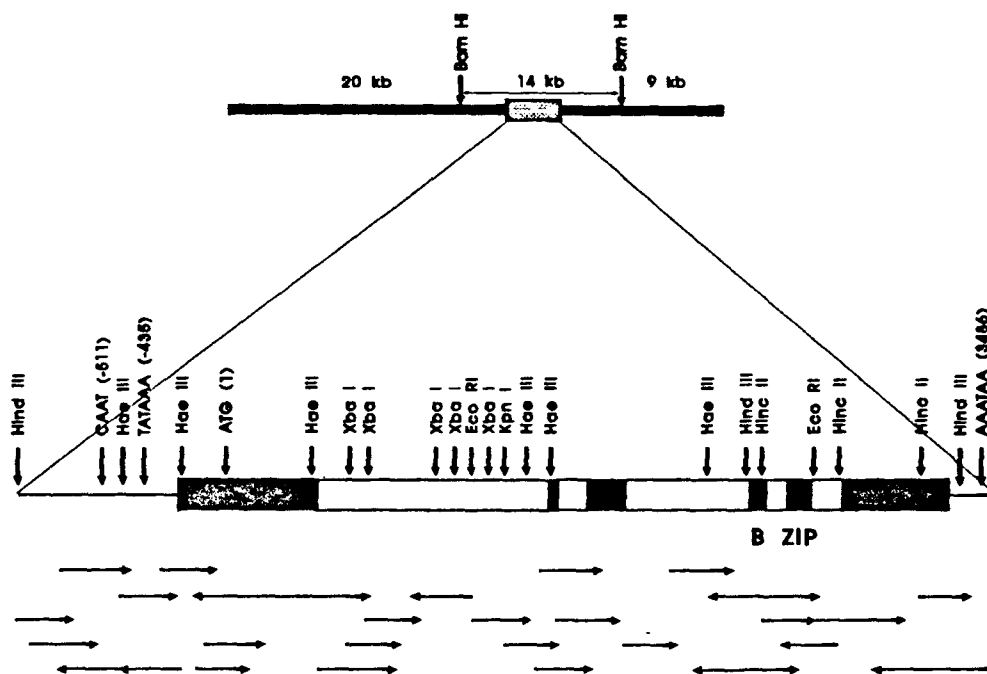


Figura 5: Mapa de restrição do clone genômico de *opaco2* de *Coix* (gCXO2-1). Na parte superior, estão representados os braços do bacteriófago λ DASH, de 20 Kb e 9 Kb, e o inserto genômico de 14 Kb. No centro, é apresentado o mapa de restrição do clone genômico de *Coix*, gCXO2-1. Os retângulos escuros representam os seis exons, e os brancos, os cinco íntrons. A região bZIP está indicada. As setas, na parte inferior, indicam a estratégia adotada no sequenciamento.


```

      | TATA|                               ①
TAGCACACGTGTTATAAAAGAGTTAAATTGAAACAAAATATGTATGTAGCGTGACGATAAAACCCATAATAA -376
      M Y V A * (4)

      ②
CCTTGTAGCATGCCTAACTTTGCTTTCTGTTAGCCTTTTCCAACCTTCCTTCCAAACAAAATTACTTGTCTTA -304
      M P N F A F C * (7)

      ③                               ④
TTTGCCATGGTCACATAGAACAACAGGCTACTTGGTTCACGTTGAGTGCAAGGGCAATGCACTACTCCTT -232
      M V T * (3)                               M H Y S F

TTGCCTCTCCAGAAATTTCTGCACGCACAAGCCCTCACTAGATAAATAGAGGGGAGAACACCCTTGAAGCCT -160
      C L S R N F C T H K P S L D K * (20)

      ③ ⑤                               ⑦
ACTCACCCAAGCATAAACTCAATCATGGATGCATAGAAATAAATACCAACCATATCTATGGGTCCTTCCCTT -88
      M D A * (3)                               M G P S L
      M H R N K Y Q P Y L W V L P F

TCTCCATCTCCTCCAAAGCCAACCTGATTTAGTGCTTGTTCCTTCTTCTTGAACCTTCCCTTGAAAACATT -16
      S P S P P K P T * (13)
      L H L L Q S Q P D L V L V S F F L E L P L K T L

      ⑧
GATAGTTTCTTGGACATGGAGCATGTCATCTCAATGGAGGAGATCCTCGGACCCTTTTGGGACCTGCCACCA +57
      M E H V I S M E E I L G P F W D L P P
      I V S W T W S M S S Q W R R S S D P F G T C H H

TCACCACCACCCTACCCTACCAGAGCAGCAGCCTCTGGTGACCGACACCGGCAGTGTGTCATAG..... +124
      S P P P L P L P E Q Q P L V T D T G S V V I . . .
      H H H H Y H Y Q S S S L W * (76)

```

Figura 7: Sequência de nucleotídeos da região 5' não traduzida do gene *opaco2* de *Coix*. São indicadas as oito fases de leitura existentes abaixo do provável "TATA box". O início de cada fase de leitura está indicado por números. O número de aminoácidos de cada polipeptídeo é dado entre parenteses, com exceção da fase principal (⑧). Os codons de terminação estão indicados por asteriscos (*).

117 pb abaixo do códon de terminação TAG.

A sequência de nucleotídeos do clone genômico foi depositada nos bancos de dados EMBL, GenBank e DDBJ com o número de acesso X78287, e código de identificação CLJOP2D.

Estudos comparativos

Comparação entre os genes *opaco2*

A comparação das sequência de nucleotídeos dos genes *o2*, através da análise por "dot matrix" (Fig. 8), revelou alta similaridade entre os exons do gene nas três espécies comparadas. Com exceção do exon 3 do gene de sorgo, e do exon 5 do gene de *Coix*, que apresentam respectivamente deleções de 84 e 71 pb, em relação aos mesmos exons do gene de milho, os demais exons dos genes *o2* de *Coix*, milho e sorgo são altamente conservados tanto em sequência quanto em posição.

Os íntrons de *Coix* apresentaram maior similaridade com os de sorgo do que com os íntrons de milho (Fig. 8). Sendo que o primeiro íntron apresentou maior grau de conservação nos três genes. A Figura 9 representa,

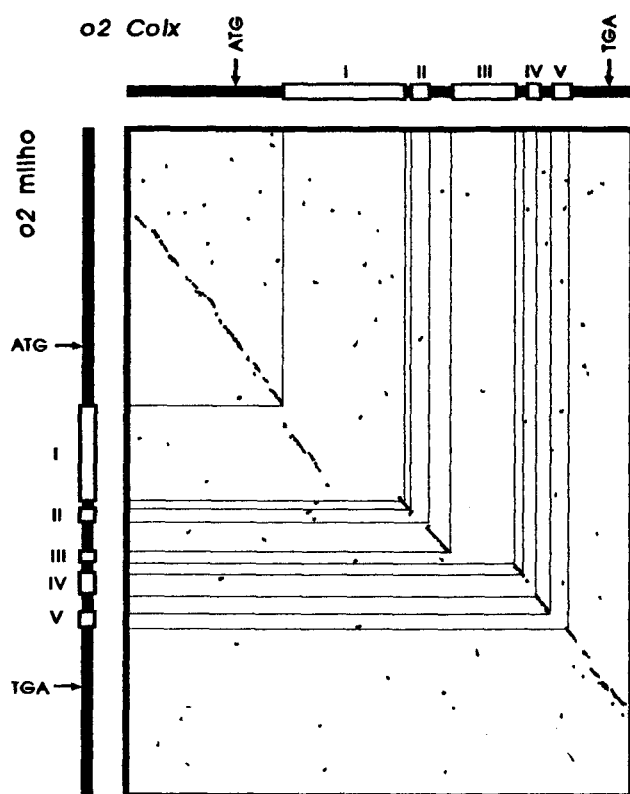
esquemáticamente, os genes *o2* de *Coix*, milho e sorgo.

O alinhamento das regiões flangeadoras 5' dos genes *o2* está representada na Figura 10. A sequência canônica do provável "TATA box" (TATA I, Fig. 10), identificado no gene de *Coix*, encontra-se ausente na sequência do gene *o2* de milho, e modificada na de sorgo. A sequência CTATTTG identificada como possível "TATA box" no gene de milho (Hartings *et al.*, 1989), foi também localizada nos genes de sorgo (Pirovano *et al.*, 1994) e de *Coix* (TATA II, Fig. 10). As sequências TATA II de milho e sorgo são completamente conservadas, enquanto que em *Coix* a mesma sequência apresenta uma única substituição.

A sequência CCAATC, correspondente a um provável "CCAAT box", é conservada nos três genes analisados (Fig. 10).

Nas regiões flangeadoras 5' dos genes *o2* analisados foi encontrada ainda uma sequência palindrômica similar ao "G box", definido por Izawa *et al.* (1993). Esta sequência têm sido descrita na região central do sítio de ligação de um grande número de transativadores de plantas, tais como TAF-1, EmBP-1, HBP-1a, CPRF-1, OCSBF-1, TGA1b (Izawa *et al.*, 1993).

A



B

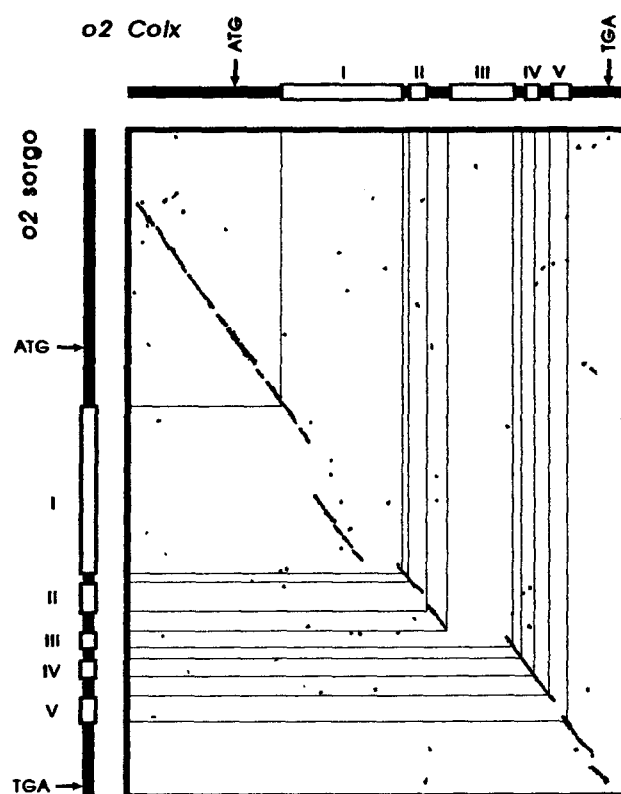


Figura 8: Comparação das seqüências de nucleotídeos dos genes *opaco2* através de "dot matrix", com o programa DIAGON (Staden, 1982), utilizando-se janela de 15 nucleotídeos com identidade de 80%. Em A foram comparadas as seqüências dos genes *opaco2* de *Coix* (horizontal) e milho (vertical). Em B foram comparadas as seqüências dos genes *opaco2* de *Coix* (horizontal) e sorgo (vertical). Os genes encontram-se representados esquematicamente nas laterais, sendo que as caixas brancas indicam as posições dos íntrons. As posições dos códons de iniciação (ATG) e terminação (TGA) estão indicadas.

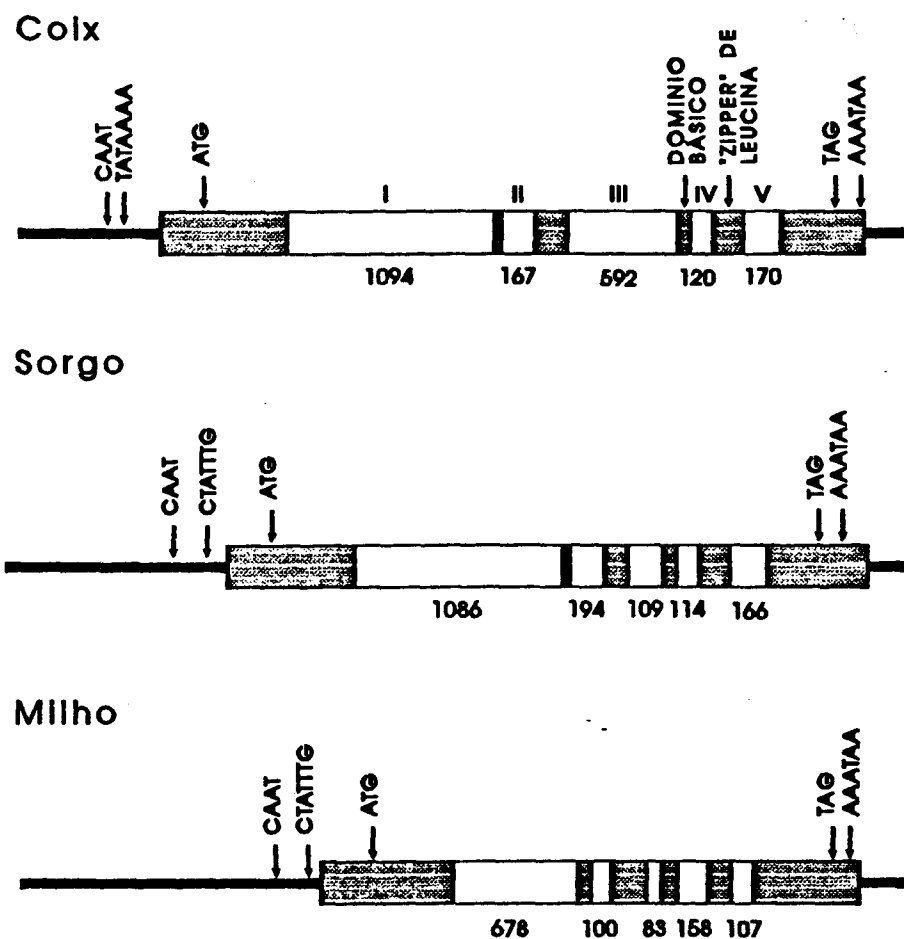


Figura 9: Comparação esquemática dos genes *opaco2* de *Coix*, sorgo e milho. Os retângulos escuros representam os exons, e os claros, os íntrons. Abaixo de cada íntron é indicado o tamanho em pb. O codon inicial (ATG), o codon de terminação (TAG), a região do Domínio Básico e do "Zipper" de leucinas, os prováveis "CAAT box" e "TATA box", e o sinal de poliadenilação estão indicados.

Coix	ATCCTCATATA--TAGCATATA-GTCGAGATTCTGCTTA-GCTTTCCATTCAATTG-----GAACAT	-691
Milho	ATCCTCATATATATAGCATGTAAGCTGAGATTCTGTTTATGCTTTCCCTTTTCGTTTAAATTTGGACAGCAG	-734
Sorgo	ATCATCGTATATA--GCATGTA-GTTGAGATTCTCTTTA-GCTTTCCCTCCATTG--CTGGA--CAG	-702
	*** ** *	
Coix	ATTACTTGTACGCATTTTCATTAT-TATCATATT----AAAAGTGAAGATTGTGTACATTTTTTGTGTTTC	-626
Milho	ATTACTTGTACGCATAT-ATTGTCTATATCACGATAAAAAAGTAAAGATCGTGTAACCTTTTGTGTTTC	-665
Sorgo	ATTACTTGTACGCATTT--TCAT-TAT--CACATAAA--AGTGAAGATTGTGTACATTTTTTGTGTTTC	-639
	***** * * * *	
Coix	TTGTACTTCCTTCCAC-TGCACATGAACCTTAAAGGTCTAGT--AGGTAGAACAGAAAGTGCAA----AT	-563
Milho	TTGCACGTAGTTTCGTCGTCCACATGAACCTCAAAGTTGT-GT--AGGTAGAACAGAAAGTGCAATACAAT	-598
Sorgo	TTGTAGATCCTCC-----ACATGAACCTCAAAGATATGATCTAGGTAGAACAGAAAGTGCA----AT	-581
	*** * * *	
	<- O2 ->	CAAT
Coix	ACATG-----TTGGCAAGGCATGGTATCTAGCACA--AAGTCAGTGCA--GCTTGTATCCAATCCA--	-505
Milho	ACATGTTGACGTTGGTAAGGCATGGTGTCTAGCACACAAAGCCAGTGCAAACTCGTTATCCAATCCAGT	-528
Sorgo	ACATGT-----TTGGTAAGGCGTGGTATCCAGCACA--AA----GTACAA--CATGTTATCCAATCCAAG	-524
	***** * * * *	
Coix	-----AGTT--AGAACACAC-----GTACATGGCCAACAGTTTGATCTATTGACACGC--TTGGCTT	-453
Milho	GGTTTAAAGGTAGTAGAACACACACGGGCATGCATGGGTA-----CACGC--TTGGCTT	-476
Sorgo	GATATAAAGTT--AGAACACACGG-----ACATGGCCAGCAGTTTGATCCACTGCACACGTTTGATT	-463
	* * * *	
	G BOX TATA I	
Coix	ATTACTAGCACACGTGTTATAAAGAGTTAAAT-TGAAACAAAATATGTATGTAGCGTGACGATAAAACC	-384
Milho	ATTGCTA--ACACGTGTG-----GTTAAAA-CAAAACACACATGCATGTAGCGTCACGATAAAACC	-417
Sorgo	GTCATAACACACGTGTTATTGAAGAGTTAAAAATGAAGCAAAATATATACATAGCGTCACGATAAAACC	-393
	* * * * *	
Coix	CATA---ATAACCTTGTAGC-----AT--GCCTAACTTTGCTTTCTGTTAGCCTTTTCCAACCTTCCTTC	-325
Milho	CATGGTACGTAACTTGGTAGCGCACCATCTAGCTAACTTTGCTTTCTGTTAGCCT-----TTC	-359
Sorgo	CATA---ATAACCTTGTAGCGCAC-----CTAACTTTGTTTTCTGTTAGCCTTTTCCA-TTCCCTTC	-335
	*** * * * *	
	TATA II	
Coix	CAAAACAAATTACTTGTCTTATTTGCCATGGTCACATAGAACACACAGGCTACTTGGTTCACGTTGAGT	-255
Milho	CAAAACAAATCACTTGTCTTATTTGCTGCCCTGCA-----GGTTCACATTGAGT	-310
Sorgo	CAAAACAAATTACTTGTCTTATTTGCCATGGTCACATTGAACACAGCAGGCTACTTGGTTCACATTGAGT	-265
	***** * * * *	
Coix	GCAAGGGCAATGCACTAC-----TCCTTTTGCCCTCCTCAGAAATTTCTGCAC----GCACA-----	-203
Milho	GCAAGGCCGATGCATTTTTTTCTCTTTTTTTTGCCCTCCTCAGAAATTTCTGCAGGGAGGCACAGCAAG	-240
Sorgo	GCAAGGGCAGTGCACTAC-----TTTTGCCTCTCCAGAAATTTCTGCA----GACAC-----	-217
	***** * * * *	
Coix	-----AGCCCT-CACTAGATAAATAGAGGGG--AGAACACCCTGAAG-----CCT---ACTCACCC	-152
Milho	AGAGAGAGCCCAGCACTAGATAAGTAGGGGGGAAAGAGCATCCAAGCGTACTCCTCGATCTCCTTC	-170
Sorgo	-----AGCCCAACACTAGATAAATAGGGGGG--AAGAACACCCTTCAAG-----CCCAACTCACCC	-162
	***** * * * *	
Coix	AAG-----CATAAATCAATCATGGATGCATAGAAATA-----AATACCAA-CCATATCTATGGG	-98
Milho	AAGCAGCTAACATAA-TTAACTCATGGGTGCATAGATATGCCGCCAGCACCAT-CATATCTATGGG	-102
Sorgo	AAGCA---ACATAA-CTCAATCATGGGTGCATAGAAATATG-----CCAACCATATCTATGGG	-107
	*** * * * *	
Coix	-TCCTTCCCTTT---CTCCATCTC-----CTCCAAAGCCAACCTGATTTAGTGCTTGTTCCTTCTTCC	-38
Milho	-TCCCTCCCCTTGCACTGCATCTCCT-----TCAATTCCTAGTGTTTGTCTTCTCCCTTCC	-48
Sorgo	GTCCTTCCCTT-GAAGTCTTCTCCATCTCTCCAAAGCCACCTTGACCTAGTGCTTGT---TCTTCC	-31
	*** * * * *	
Coix	TTGAACCTTCCCTTGAAACATT---GATAGT-----TTCTTGGACATG +3	
Milho	TTGACCTTTGCTTGGAAACATTGATTGATAGTTACTTATTATGGGCATG +3	
Sorgo	TTGATCTTCCCTTGAAATCATT---GATAGTTAC---TTGTTGGACATG +3	
	***** * * * *	

Figura 10: Alinhamento das regiões flangeadoras 5' dos genes *opaco2* de Coix, milho e sorgo. Os nucleotídeos conservados nos três genes estão indicados com asteriscos (*). Os espaços para maximizar o alinhamento estão representados por hífens. As sequências foram numeradas a partir do códon de iniciação. O sítio de ligação da proteína O2 no promotor de milho é indicado (O2; Lohmer *et al.*, 1991). Os prováveis "TATA box" (TATA I e TATA II), "CAAT box" e "G box" (Izawa *et al.*, 1993) também estão indicados. O alinhamento foi realizado utilizando-se o programa CLUSTAL V (Higgins and Sharp, 1989) e otimizado manualmente.

Comparação entre as proteínas Opaco2

A proteína O2 de milho apresenta um peso molecular deduzido da sequência de nucleotídeos de 47.074, enquanto que as proteínas de O2 de Coix e sorgo apresentam pesos moleculares de 44.865 e 43.234, respectivamente. As diferenças nos pesos moleculares devem-se principalmente às deleções de 26 e 27 aminoácidos presentes, respectivamente, nas regiões amino e carboxi-terminal das proteínas de Coix e sorgo (Fig. 11).

A região correspondente aos domínios básico e "zipper" de leucina (bZIP) é altamente conservada em Coix, milho e sorgo. Nesta região ocorrem apenas sete substituições de aminoácidos, sendo que destas, cinco são conservativas (Fig. 11).

As proteínas O2 apresentam dois sinais de localização nuclear (NLS), necessários para o seu transporte do citoplasma, local de sua síntese, para o núcleo, onde exercem a função de ativadores transcricionais. O primeiro, NLS A (KRK) está localizado na região amino-terminal da proteína O2 de Coix, entre os aminoácidos 123 e 125, e tem estrutura similar aos NLS do tipo encontrado em SV40 (Raikhel, 1992). Conforme indicado na Figura 11, as

sequências dos NLS A não se encontram completamente conservadas nas três proteínas O2. O segundo, NLS B, completamente conservado nas três proteínas, está localizado na região básica do motivo bZIP (Fig. 11) e apresenta uma estrutura bipartida (Raikhel, 1992).

A proteína O2 de Coix apresenta ainda os resíduos de aminoácidos, arginina e asparagina, que encontram-se presentes em todas as proteínas do tipo bZIP sequenciadas (Vison *et al.*, 1989) (Fig. 12).

RNA mensageiro de o2 de Coix

Tecido-especificidade da proteína O2 de Coix

A tecido-especificidade da proteína O2 de Coix foi analisada através de hibridização de RNAs totais ("Northern blot") extraídos de diferentes tecidos de Coix. Conforme mostra a Figura 13, os transcritos do gene o2 encontram-se presentes exclusivamente no endosperma, não sendo detectados em coleóptilo, embrião, folha ou raiz.

```

Coix  MEHVISMEEILGPFWDLPSPPLPLPE-QQPLVTDTGSSVIDGVVTQGGGD  51
Milho MEHVISMEEILGPFWELL--PPPAPEPEREQPPV--TGIVVGSVIDVAAAGH  48
Sorgo MEPVFSMEEILGPFWDL-PSPPP--PE-QQPLVIGTSSVIDGVVTH-GGN  46
      ** * .*****. *   ** .*** .** * . * . * . * .

Coix  GEGGDMMDQ--NTTEWTFERLLEEEILINKTTL----VTNSSCS-TLNID-P  95
Milho GD-GDMMDQ--HATEWTFERLLEEEALTSTPPPVVVPNSCCSGALNADRP  98
Sorgo GEGSNMMDQIQNTTEWTFERLLEEEELLTDTP-----VANSSCPA-LNVDP-  91
      * . .*** ..*****. *   *   *   *   *   *   *   *

      ###

Coix  VVEVDQGTMASGAVS-AV-GDPMYNAILKRKLEVDLVAFKMWR-ASS-VVN  143
Milho VMEEAVTMAPAAVSSAVVGDPMEYNAILRRKLEEDLEAFKMWRASS-VVT  149
Sorgo VVEVDQGAPEAVS-AV-GDPMYNAILKRNVEDLMFAFKMWR-ASTSGVN  140
      * . .** .*** ** *****. * . * * ***** ** . *

Coix  SE-RSQDSNNHNGGS---KNVVQNKI-NGEDPI-NNHAQNVDLRVRLATSSS  189
Milho SDQRSQGSNNHTGGSSIRNPNVQNKLMGEDPINNHAQTAGLGVRLATSSS  201
Sorgo SE-----GSNNENG-----DLHVRLTTSSS  161
      * . .***. ** .*****.

      |<--MOTIVO BÁSICO->|
      #####  ##  ##

Coix  SRDPSPSDEMDGEVEILGFKMPTEERVVRKRESNRESARRSRYRKAHLKE  241
Milho SRDPSPSDEMDGEVEILGFKMPTEERVVRKRESNRESARRSRYRKAHLKE  253
Sorgo SRDPSPSDEMDGEVEILGFKMPTEERVVRKRESNRESARRSRYRKAHLKD  213
      *****.
      |<----- REPETIÇÕES DE LEUCINA ----->|
      1       2       3       4       5       6       7
Coix  LEDQVEQLKAENSCLLRRLAALNQKYNEANVDNRVLRADMETLRKVKGED  293
Milho LEDQVAQLKAENSCLLRRIAALNQKYNDANVDNRVLRADMETLRKVKGED  305
Sorgo LEDQVDKLKAENSCLLRRLAALNQKYNHATVDNRVLRADMETLRKVKGED  265
      *****.

Coix  SLKRVEMSSSLPP-SMIPALPSSSDASVP---IQDDIINYFSTTPAEDDA  341
Milho SLKRVIEMSSSVSPSSMIPISAPTPSSDAPVPPPIRDSIVGYFSATAADD-DA  356
Sorgo SLKRRIEMTSLT--SIPIELPSSSDVPVH---IQDNIVNYFTTTPAG--DA  310
      *****.

Coix  PVDNNSFIIMPADP--LQLVQAEQPTM-----  368
Milho SVGN-GF-----LRL-QA-HQE--PASMVVGGLSATEMNRVAAATH  393
Sorgo LADN-SF--MPMPDPLPLQL-QAEE-PTI-----NGALNATEMNQI--ATH  347
      * . *   * . *   * .

Coix  ---GA---MELIQKTMGAMP-TSP-GSAL--QESQLLGL-GPDETINMDMY  408
Milho CA-GA---MELIQTAMGSMPPTSASGSTPPQDYELLGPN---AIHMDMY  437
Sorgo CAAGSQPSMQLIQETMGAMPPTS-SGSTL--QESSELLGPN---ETINMRY  394
      * . * .*** .*** * * * * . * .*** .***.

```

Figura 11: Alinhamento das sequências de aminoácidos das proteínas Opaco2, deduzidas das sequências de nucleotídeos dos genes *opaco2* de Coix, milho e sorgo, respectivamente a partir dos clones gCXO2-1, o2 cDNA 1-4 (Schmidt *et al.*, 1990) e da sequência de nucleotídeos do clone genômico de sorgo depositada nos bancos de dados do EMBL, GenBank, DDBJ Nucleotide Sequence Databases, com o número de acesso X71636. Os resíduos idênticos e as substituições conservativas são indicadas por asteriscos e pontos, respectivamente. Estão indicadas as posições dos domínios Básico e "Zipper" de leucinas. Os aminoácidos relacionados com o transporte nuclear (NLS) da proteína Opaco2 estão indicados pelo símbolo (#). Os espaços para maximizar o alinhamento estão representados por hífen. O alinhamento foi realizado utilizando-se o programa CLUSTAL V (Higgins and Sharp, 1989).

	<-- MOTIVO BÁSICO-->	<-ZIPPER DE LEUCINA-->
	##### *	1 2 3 4
Opaco2 de Coix	...RVRKRKESNRESARRSRYRKA [#] AHLKELEDQVEQLKAENSCILLRRLAAL...	
Opaco2 de milho	...RVRKRKESNRESARRSRYRKA [#] AHLKELEDQVAQLKAENSCILLRRIAAL...	
Opaco2 de sorgo	...RVRKRKESNRESARRSRYRKA [#] AHLKDLEDQVDKLKAENSCILLRRLAAL...	
EMBP-1 (trigo)	...KRERRKQSNRESARRSRLRKQ [#] EECELAQKVSELTAANGTLRSELDQL...	
HBP-1 (trigo)	...KKQKRKLSNRESARRSRLRKQ [#] AECEELGQRAEALKSENSSLRIELDRL...	
OCSBF-1 (milho)	...RREKRRLSNRESARRSRLRKQ [#] HLDELVQEVARLQADNARVAARAATS...	
OCSBF-2 (milho)	...KKMRQIRNRDSAMKS [#] RERKKS [#] YIKDLET [#] SKSHLEAECRRLLTYALQSY...	
TGA1A (tabaco)	...KVLRRLAQNREAA [#] RKSRLRKAYVQQL [#] ENSKLKL [#] IQLEQE [#] ELERARKQG...	
TGA1B (tabaco)	...KKRRLVRNRESAQLSRQ [#] RKKHYVEELEDKVRIMHSTIQDLNAKVAYI...	
GBF1 (<i>Arabidopsis</i>)	...KRQKRKQSNRESARRSRLRKQ [#] AECEQLQQRVESLSNENQSLRDELQRL...	
GBF2 (<i>Arabidopsis</i>)	...KRQKRKQSNRESARRSRLRKQ [#] AECEQLQQRVESLSNENQSLRDELQRL...	
GCN4 (levedura)	...PAALKRARNT [#] EARRSRARLQ [#] RMKQLEDKVEELLSKNYHLENEVARL...	
CPC-1 (<i>Neurospora</i>)	...VVAMKRARNTLAARKS [#] RERKAQRLEEELEAKIEELIAERDRWKNLALAH...	
C/EBP (rato)	...EYRVRRERN [#] NI [#] AVRKS [#] RD [#] KAKQ [#] RVETQ [#] QVLELTSDNDRLRKRVEQL...	
C-JUN (humano)	...KAERKRN [#] RN [#] RIAASKCRK [#] RKL [#] ER [#] IA [#] LEEKV [#] KT [#] LKAQ [#] NSELASTANML...	
C-FOS (humano)	...KRRIRRERN [#] KMAA [#] ACRN [#] RRRELTD [#] TLQAETD [#] QLEDEKSALQTEIANL...	
CRE-BP1 (humano)	...KRRKFLE [#] RN [#] RAA [#] SR [#] CQ [#] RK [#] VWVQSLEKKAEDLSS [#] LNQ [#] LQSEVTIL...	
XBP-1 (humano)	...KALRRKLKN [#] RVAAQTARD [#] RK [#] KARMSELEQQVVDLEEENQKLLLENQILL...	

Figura 12: Alinhamento das seqüências de aminoácidos dos domínios básicos de proteínas bZIP de plantas, fungos e mamíferos. Os resíduos em negrito indicam similaridade em relação a proteína Opaco2 de Coix. Os asteriscos indicam a posição dos resíduos de asparagina e arginina conservados nos domínios básicos das bZIP. O provável domínio de fosforilação está indicado pelo símbolo (#). As seqüências deduzidas de aminoácidos foram obtidas a partir das seguintes referências; O2 de milho (Schmidt *et al.*, 1990), O2 de sorgo (Pirovano *et al.*, 1994), EmBP-1 (Guiltinam *et al.*, 1990), HBP-1 (Tabata *et al.*, 1989), OCSBF-1 e OCSBF-2 (Singh *et al.*, 1990), TGA1a e TGA1b (Katagiri *et al.*, 1989), GBF1 e GBF2 (Schindler *et al.*, 1992), GCN4 (Hinnebusch, 1984), CPC-1 (Paluh *et al.*, 1988), C/EBP (Landschulz *et al.*, 1988a), c-Jun (Angel *et al.*, 1988), c-Fos (Van Straaten *et al.*, 1983), CRE-BP1 (Maekawa *et al.*, 1989), e XBP-1 (Liou *et al.*, 1990).

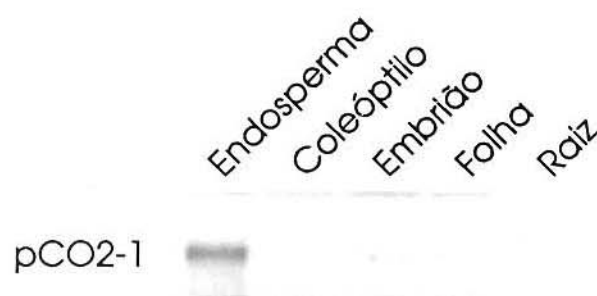


Figura 13: Análise através de "Northern blot" da expressão do gene *opaco2* em diferentes tecidos de *Coix*. Foram extraídas amostras de RNA total de endosperma, coleóptilo, embrião, folha e raiz. Foram utilizados 10 µg de RNA total de cada amostra. Na hibridização foi utilizado, como sonda, o clone cDNA de *o2* de *Coix*, pCO2-1.

Acúmulo de mRNA de *opaco2* durante o desenvolvimento das sementes de *Coix*

De acordo com a coloração e aspecto do pericarpo, as sementes de *Coix* foram classificadas em 5 estágios durante o seu desenvolvimento (Targon *et al.*, 1992). A quantidade de RNA total é mínima no Estágio 1, atinge o máximo no Estágio 3, e decai para níveis não detectáveis no Estágio 5. O acúmulo das α -coixinas inicia-se no Estágio 1, e atinge o nível máximo nos estágios 4 e 5 (Targon *et al.*, 1992).

Os níveis de RNA mensageiros (mRNAs) do gene *opaco2* em *Coix* foram analisados através de hibridizações de amostras de RNA total, extraídos de sementes correspondentes as diferentes fases de desenvolvimento do endosperma. Os RNAs foram submetidos a eletroforese em gel desnaturante de agarose na presença de formaldeído, transferidos para membranas de nylon (Hybond-N, Amersham), e hibridizados com a sonda de cDNA de *o2* de *Coix*, clone pCO2-1.

O padrão de acúmulo dos mRNAs de *opaco2* é similar ao descrito para os mRNAs de α -coixinas (Targon *et al.*, 1992). Os mRNAs de *opaco2* começam a ser detectados no Estágio 1, atingem o nível máximo no

Estágio 3, e decaem para níveis não detectáveis no Estágio 5 (Fig. 14).

Como controle das quantidades de RNA neste experimento foi realizada uma hibridização utilizando-se como sonda o clone de actina de *Trichoderma reesei*, pKpn 1.9, cedido pelo Prof. Dr. Hamza El-Dorry (Instituto de Química, USP, S.P.). A actina de *T. reesei*, assim como as demais actinas citoplasmáticas, além de apresentarem alta similaridade, parecem ser um constituinte universal do citoesqueleto celular. Por esta razão, e devido ao carácter constitutivo da expressão deste gene, decidiu-se usá-lo como controle no experimento. Conforme demonstra a Figura 14, os Estágios 2 e 3 apresentaram os maiores níveis de expressão de actina, coincidindo com os estágios de maior atividade de divisão celular. Enquanto que no Estágio 5, correspondente a semente madura e que apresenta baixa atividade celular, não foi detectado RNA de actina através da hibridização. Embora o gel corado com brometo de etídio (resultado não apresentado) indique quantidades equivalentes de RNA total de cada estágio, a hibridização com o gene de actina demonstrou que os níveis de RNA mensageiro deste gene em endosperma dependem da fase de desenvolvimento do tecido.

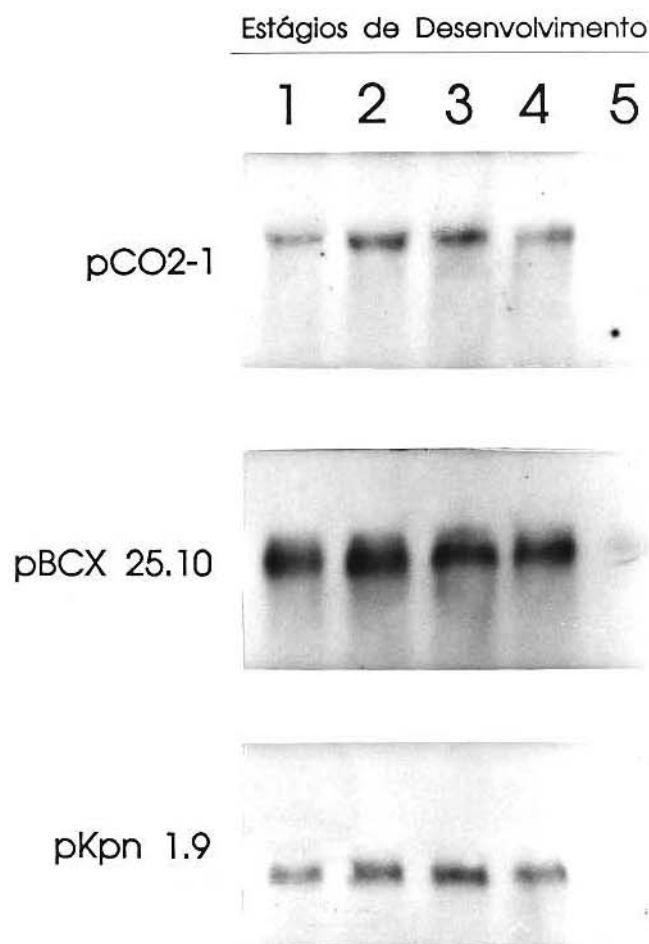


Figura 14: Análise através de "Northern blot" da expressão do gene *opaco2* durante os cinco estágios de desenvolvimento da semente de *Coix*. Foram extraídas amostras de RNA total do endosperma de *Coix* nos cinco diferentes estágios de desenvolvimento descritos por Targón *et al.* (1992). De cada amostra foram utilizados 10 µg de RNA. Após a transferência os filtros foram hibridizados contra pCO2-1 (o2 de *Coix*), pBCX 25.10 (α -coixina) e pKpn 1.9 (actina de *Trichoderma reesei* cedido pelo Prof. Dr. Hamza El-Dorry, Inst. de Química, USP, S.P). A hibridização com o clone de actina de fungo foi realizada como controle das quantidades de RNA.

O sítio de ligação da proteína O2 de Coix

A proteína O2 de milho reconhece nos promotores de α -zeínas de 22 kDa o sítio 5'-TCCACGTAGA-3' (Schmidt *et al.*, 1992). Esta sequência foi posteriormente denominada de sítio O2a (Yunes *et al.*, 1994). Nas α -coixinas de 25 kDa, a proteína O2 de milho reconhece a sequência denominada sítio O2c, 5'-GACATGTC-3' (Yunes *et al.*, 1994). O promotor das α -coixinas apresenta ainda, outros dois sítios similares ao sítio O2a de milho (Ottoboni *et al.*, 1993), denominados de O2a e O2b. Testes de ligação seletiva *in vitro* demonstraram que os sítios O2a e O2b de Coix são incapazes de interagir com a proteína Opaco2 de milho (Yunes *et al.*, 1994).

Para determinar se a proteína Opaco2 de Coix apresenta a mesma especificidade de ligação da proteína de milho, foram realizados experimentos de ligação seletiva, conforme descrito em Materiais e Métodos. Nestes experimentos foram utilizadas proteínas quiméricas obtidas através de fusão da região amino-terminal da β -galactosidase na região amino-terminal das proteínas O2 de Coix e de milho (β -Gal::O2).

Para a obtenção da proteína quimérica β -Gal::O2 de Coix foi

construído o plasmídeo pO2CX, através da inserção do fragmento Hind III / Bam HI do clone pCO2-1 no plasmídeo de expressão pUR289 (Sambrook *et al.*, 1989). Um plasmídeo semelhante expressando a proteína quimérica β -Gal::O2 de milho foi cedido pelo Prof. Dr. Robert Schmidt (plasmídeo LC#37).

A Figura 15 A mostra que a proteína quimérica β -Gal::O2 de Coix (coluna 4) reconhece o fragmento Sau 3AI de 260 pb do promotor de α -coixina de 25 kDa. Esse fragmento contém o sítio O2c, que é reconhecido pela proteína Opaco2 de milho (Yunes *et al.*, 1994), e também os sítios O2a e O2b. O plasmídeo P285 α -C, que contém o inserto Eco RI / Nco I de 285 pb do promotor α -coixina de 25 kDa, e que exclui o sítio O2a (Fig. 15 B), foi retido pela β -Gal::O2 de Coix (Fig. 15 A, coluna 6). A proteína O2 de Coix manteve, ainda, uma forte capacidade de ligação com o plasmídeo P285 Δ O2b (Fig. 15 B), que apresenta deleção da sequência CAATAGA, correspondente ao sítio O2b (Fig. 15 A, coluna 8). Esses resultados indicaram que o sítio de ligação da proteína Opaco2 de Coix, nos promotores, das α -coixinas de 25 kDa, corresponde ao sítio O2c.

O teste de ligação seletiva do promotor de α -zeína de 22 kDa, digerido com Sau 3AI (plasmídeo

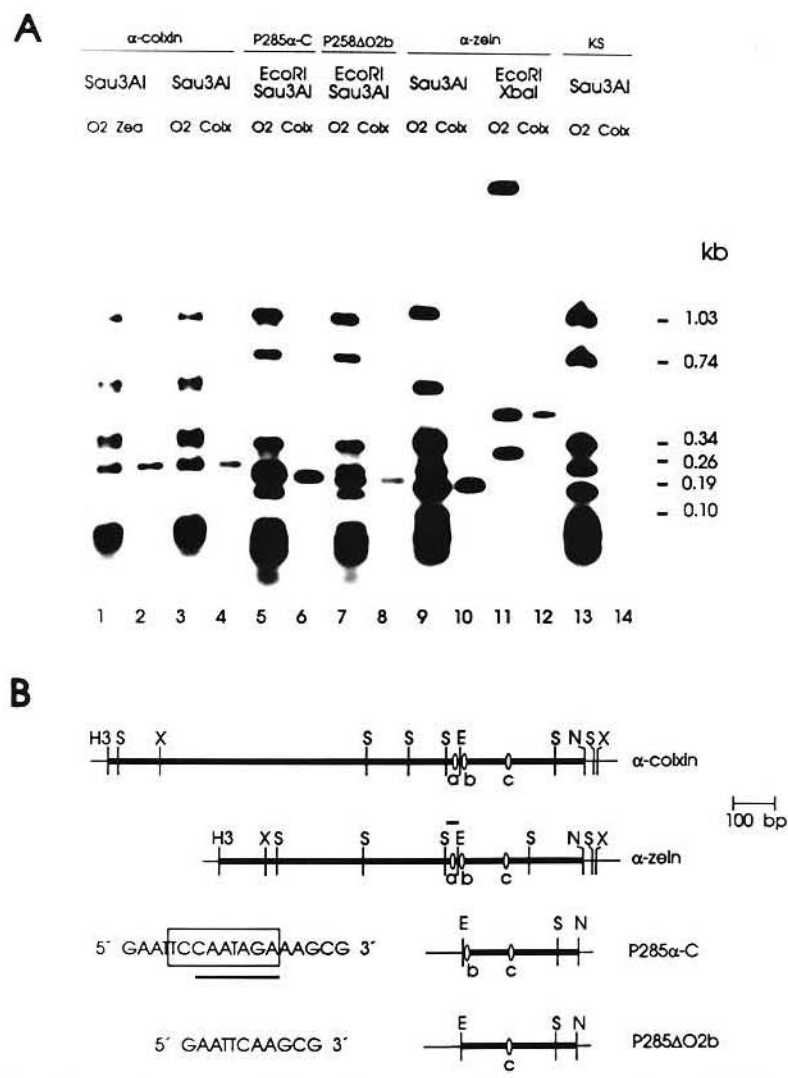


Figura 15: Ligação seletiva da proteína O2 de Coix e de milho em fragmentos de restrição de α -coixina, de clones mutantes do promotor de α -coixina e de promotores de α -zeína de 22 kDa. A, plasmídios contendo a sequência de α -coixina (P α -Cx103), um fragmento de 285 pb do promotor de α -coixina (P285 α -C), uma versão mutante desse fragmento com o sítio O2b deletado (P258 Δ O2b), o clone 22 Z-4 da α -zeína de 22 kDa (P α -Z4103) e o vetor pBluescript KS⁺ foram digeridos com Sau 3AI, Eco RI / Sau 3AI ou Eco RI / Xba I. Os fragmentos obtidos foram marcados em suas extremidade, e incubados com o imunocomplexo Proteína A-Sepharose- β -gal::O2. As linhas pares mostram os fragmentos que foram especificamente retidos pelo imunocomplexo. O tamanho dos fragmentos do KS⁺ digerido com Sau 3AI é indicado no lado direito. B, mapa de restrição do promotor da α -coixina (P α -Cx103), da α -zeína de 22 KDa (P α -Z4103), e dos subclones da região promotora da α -coixina. Os prováveis sítios alvo da proteína Opaco2, O2a, O2b e O2c estão indicados. A sequência normal e a versão mutante do sítio O2b estão detalhadas. Foram utilizadas as seguintes abreviações para os sítios das enzimas de restrição: E, Eco RI; N, Nco I; H3, Hind III; S, Sau 3AI; X, Xba I.

P α Z4103, Fig. 15 B), mostra que a proteína O2 de *Coix* reconhece o fragmento de 200 pb que contém os sítios O2a, O2b, e O2c (Fig. 15 A, coluna 10). O mesmo ocorre com o fragmento Eco RI / Xba I de 455 pb (Fig. 15 A, coluna 12), entretanto este fragmento contém apenas o sítio O2a.

Como controles positivo e negativo foram utilizados, respectivamente, o promotor da α -coixina de 25 kDa digerido com Sau 3AI (Fig. 15 A, coluna 2), e o plasmídeo pBluescript KS⁺ digerido com Sau 3AI, selecionados com a proteína O2 de milho.

Através de experimento de "footprinting", Yunes *et al.* (1994) mostraram que a proteína O2 de milho protege, no promotor das α -coixinas, uma região de 23 nucleotídeos no sítio O2c, incluindo o motivo CATG, considerado essencial para a interação proteína-DNA (Yunes *et al.*, 1994). Visando a confirmação do sítio O2c, do promotor da α -coixina de 25 kDa, também como local de ligação da proteína Opaco2 de *Coix*, foi realizado um segundo teste de ligação seletiva. Com esta finalidade foram introduzidas, através de PCR conforme descrito em Materiais e Métodos, duas mutações pontuais e independentes na região protegida no experimento de "footprinting". A primeira mutação foi introduzida no

motivo CATG do sítio O2c, através da substituição do nucleotídeo "G" por um "A", gerando o plasmídeo p285CATG. A segunda mutação foi introduzida na sequência TGAC, localizada em uma região adjacente a descrita anteriormente, e que apresenta alta similaridade com o sítio-alvo da proteína GCN-4, um fator transcricional de fungos. Esta mutação foi introduzida através da substituição do nucleotídeo "G" por um "A", gerando o plasmídeo p285TGAC (Fig. 16 B). Os plasmídios foram, então, digeridos com Eco RI / Sau 3AI e submetidos a testes de ligação seletiva com a proteína β -Gal::O2 de *Coix*. Como mostra a Figura 16 A, as duas mutações impediram a ligação ao DNA (Fig. 16 A, colunas 2 e 4).

Estes resultados confirmam, definitivamente, o sítio O2c como sítio de ligação da proteína Opaco2 de *Coix* nos promotores das α -coixinas de 25 kDa.

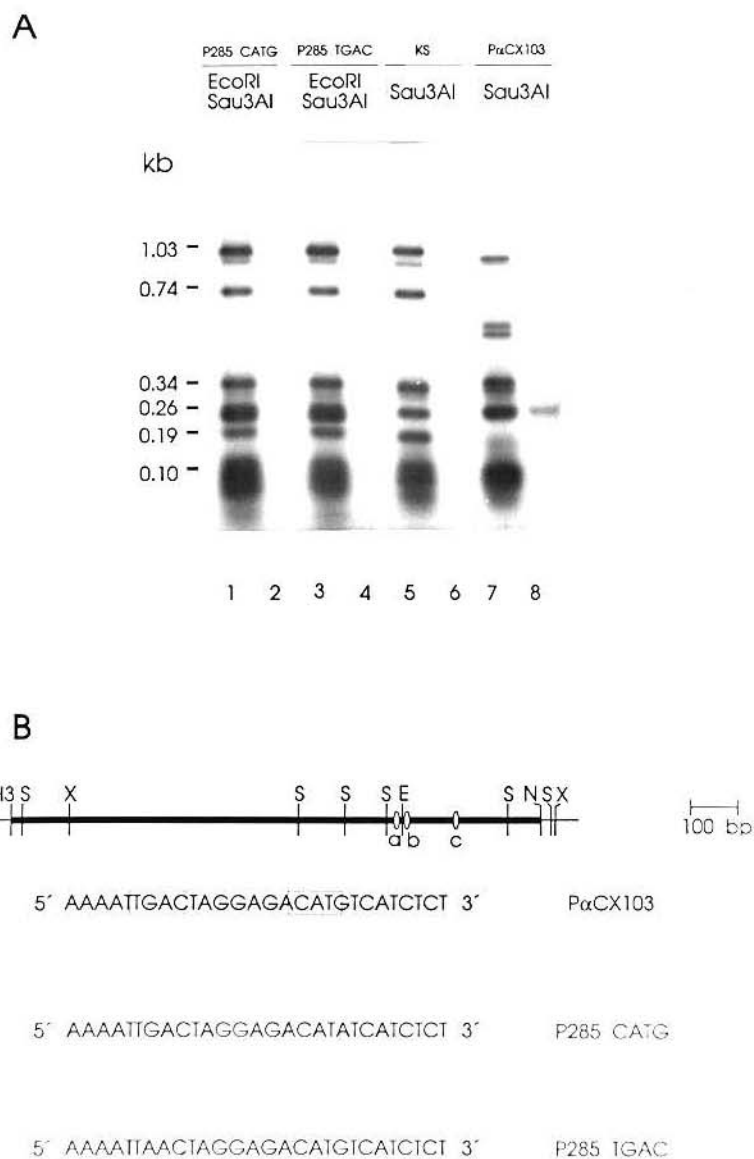


Figura 16: Ligação da proteína Opaco2 de *Coix* em fragmentos do promotor da α -coixina e em versões mutantes de 285 pb. A, plasmídios contendo a sequência do promotor da α -coixina (P α -C103), duas versões mutantes de 285 pb do promotor de α -coixina, nas quais o sítio O2c foi mutado (P285 CATG and P285 TGAC), e o vetor pBluescript KS⁺ foram digeridos com Eco RI / Sau 3AI ou Sau 3AI. Os fragmentos obtidos foram marcados em suas extremidades, e incubados com o imunocomplexo Proteína A-Sepharose- β -gal::O2. As linhas ímpares mostram os fragmentos de restrição que foram incubados com o imunocomplexo. As linhas pares mostram os fragmentos de restrição que foram especificamente retidos pelo imunocomplexo. O tamanho dos fragmentos do KS⁺ obtidos após a digestão com Sau 3AI é indicada no lado esquerdo. B, sequência do sítio O2c do promotor de α -coixina (P α -Cx103), e das duas versões mutantes (P285 CATG and P285 TGAC). O sítio O2a, O2b e O2c estão indicadas. As abreviações dos sítios das enzimas de restrição seguem: E, Eco RI; N, Nco I; S, Sau 3AI; H3, Hind III; X, Xba I.

DISCUSSÃO

Duas sequências diferentes foram descritas para o gene *opaco2* de milho. A sequência de aminoácidos da proteína O2 de milho apresentada por Hartings *et al.* (1989) mostra uma proteína de 460 aminoácidos, enquanto que Schmidt *et al.* (1990) apresentaram uma sequência de cDNA que codifica uma proteína de 437 aminoácidos. A maior discrepância entre as duas sequências ocorre na região codificadora 3' dos cDNAs. A sequência apresentada por Hartings apresenta a deleção de um nucleotídeo "G" na posição 1.282, quando comparada com a sequência descrita por Schmidt. Esta diferença resulta na alteração da fase de leitura, de modo que, os últimos 33 aminoácidos apresentados por Hartings diferem dos últimos 16 aminoácidos apresentados, por Schmidt, para a proteína O2.

Na sequência adicional de aminoácidos da proteína O2, Hartings *et al.* (1989), identificaram um motivo semelhante aos domínios de ligação a DNA denominados "zinc fingers".

A sequência de aminoácidos descrita por Pirovano *et al.* (1994) para a proteína O2 de sorgo

apresenta uma sequência carboxi-terminal diferente das encontradas em ambas as proteínas O2 descritas em milho. Entretanto, a sequência de aminoácidos deduzida a partir da sequência genômica de o2 de sorgo depositada por estes autores no banco de dados EMBL, com o número X71636, apresenta-se diferente daquela descrita no artigo (Pirovano *et al.*, 1994), revelando uma região carboxi-terminal semelhante à encontrada na proteína O2 de milho descrita por Schmidt *et al.* (1990).

A sequência da proteína O2 de Coix é similar as sequências da proteína O2 de milho apresentada por Schmidt e O2 de sorgo deduzida a partir da sequência genômica, depositada no banco de dados EMBL, sobretudo na região carboxi-terminal. O alto grau de conservação entre as proteínas O2 em espécies diferentes sugere, que nestas espécies, não apenas os genes das prolaminas, mas também, os mecanismos de regulação da expressão destes genes devem ter sido conservados.

As extensas deleções observadas nas proteínas O2 de Coix e sorgo (Fig. 11), principais responsáveis pela

redução do peso molecular em relação à proteína de milho, parecem não interferir na função destas proteínas. Este fato foi confirmado através de experimentos de transativação realizados com as proteínas O2 de sorgo (Pirovano *et al.*, 1994) e de *Coix* (Yunes, J.A., comunicação pessoal).

A proteína O2, assim como as demais proteínas do grupo das bZIPs, são caracterizadas pela presença de uma região bifuncional, contendo dois diferentes domínios. O domínio básico, rico em aminoácidos básicos, é responsável pelo reconhecimento específico do sítio-alvo no DNA (Turner & Tijain, 1989; Neuberg *et al.*, 1989), enquanto que a região contendo repetições de leucinas a cada sete aminoácidos ("zipper" de leucinas) é responsável por interações com outras moléculas de proteínas do tipo bZIP, determinando a formação de homodímeros ou heterodímeros (Kouzarides & Ziff, 1988).

O alinhamento das sequências de aminoácidos das três proteínas relacionadas indicam uma completa conservação do domínio básico, e apenas sete substituições de aminoácidos, sendo cinco delas conservativas, na região do "zipper" de leucinas (Fig. 11). Este alto grau de conservação indica uma estreita

relação entre a estrutura e função destes domínios.

O domínio básico da proteína O2 de *Coix* apresenta os resíduos de asparagina e arginina conservados como as demais bZIP descritas até o momento (Fig. 12). Vinson *et al.* (1989) propuseram um modelo para o reconhecimento da sequência do DNA pelas proteínas do grupo bZIP ("scissors-grip model"), no qual a asparagina conservada apresenta a função de permitir a angulação da região básica, facilitando a interação da proteína com o sulco maior da dupla hélice do DNA. Outro aminoácido, que não a asparagina, nesta mesma posição, provavelmente não permitiria a angulação correta da região básica, e, conseqüentemente, inibiria a ligação com o DNA. O efeito da contribuição da arginina no reconhecimento e ligação não foi ainda completamente esclarecido. Aukerman *et al.* (1991) descreveram um mutante *opaco2* (o2-676), que apresenta uma lisina substituindo a arginina no domínio básico. Apesar da substituição conservativa, este mutante produz uma proteína que não apresenta afinidade, *in vitro*, pelo promotor de α -zeínas de 22 kDa.

O domínio básico tem sido apontado, também, como um importante mediador da importação das

proteínas bZIP para o interior do núcleo de células vegetais (van der Krol & Chua, 1991). O transporte de proteínas de alto peso molecular ou de complexos oligoméricos, para o interior do núcleo, depende de sinais de localização nucleares (NLS), que têm sido encontrados num grande número de proteínas de mamíferos, anfíbios, leveduras e plantas (Silver, 1991; Raikher, 1992). A comparação de várias sequências NLS não foi capaz de indicar uma sequência consenso de aminoácidos. A sequência parece depender apenas de um pequeno número de resíduos básicos contínuos. Na proteína O2 de milho foram descritas duas regiões com atividade de NLS, o NLS A e o NLS B (Raikhel, 1992). Conforme indicado na Figura 11, a sequência correspondente ao NLS A, apesar de estar localizada em uma região altamente conservada, apresenta substituições conservativas nos resíduos apontados como essenciais para permitir a importação para o interior do núcleo. A sequência do NLS B encontra-se localizada no domínio básico e apresenta completa conservação dos resíduos funcionais. Na proteína O2 de milho o NLS B é tido como o NLS efetivamente ativo (Raikhel, 1992).

A busca de possíveis sítios de fosforilação na proteína O2 de *Coix*,

utilizando-se o banco de dados PROSITE (EMBL, Heidelberg, Alemanha), revelou, na região inicial do domínio básico, um motivo similar ao sítio de fosforilação da proteína-quinase dependente de AMPc e GMPc, com a sequência consenso de [RKRK]-X-[ST] (Glass *et al.*, 1983). O motivo RKRKESN (Fig. 12) é completamente conservado nas proteínas O2 de *Coix*, milho e sorgo. O resíduo de serina, provável alvo da fosforilação, é conservado em grande parte das proteínas bZIP. Devido ao caráter bifuncional desempenhado pelo domínio básico do motivo bZIP, a fosforilação da serina pode estar relacionada com o reconhecimento do sítio-alvo no DNA, ou com o controle do transporte da proteína O2 para o interior do núcleo. Até o momento, entretanto, não existem evidências comprovadas da presença de proteínas-quinase dependentes de AMPc no endosperma de sementes.

Assim como em milho e sorgo, o domínio básico e o "zipper" de leucinas da proteína O2 de *Coix*, estão codificados em dois pequenos exons distintos. A região básica e a primeira das sete leucinas do "zipper" estão codificadas no exon 4, enquanto o restante do "zipper" encontra-se codificado no exon 5. A separação da região de ligação ao DNA e a região

de dimerização em dois exons distintos está de acordo com a hipótese do embaralhamento dos exons (Gilbert, 1978; Gilbert, 1985). Segundo esta hipótese, os exons devem codificar porções específicas da estrutura de uma proteína: regiões funcionais, estruturais, domínios ou sudomínios - segmentos que podem ser distribuídos de forma independente durante a evolução.

O alinhamento das proteínas O2 (Fig. 11) aponta, ainda, duas outras regiões conservadas. A primeira corresponde à região amino-terminal rica em resíduos de prolina, e a segunda, rica em resíduos ácidos, localizada entre os aminoácidos 57 e 76 da proteína de *Coix*. Regiões ricas em prolina, bem com as ricas em resíduos ácidos têm sido apontadas como importantes domínios de transativação (Mermoud *et al.*, 1989; Bohmann & Tjian, 1989).

A conservação das proteínas O2 nos cereais, a exemplo da conservação das prolaminas, corrobora a hipótese de uma origem comum para *Coix*, milho e sorgo.

Os mRNAs de o2, assim como os de vários transativadores, apresentam uma longa região não traduzida na extremidade 5', denominada sequência líder (Hartings *et al.*, 1989; Schmidt *et al.*, 1990; Pirovano *et al.*,

1994). A presença de códons de iniciação nesta região, em geral determinam a redução da expressão gênica (Kosak, 1984). Segundo o modelo proposto por Kosak (1989), denominado "scanning model of translation", a subunidade ribossômica 40S liga-se a extremidade 5' do mRNA e desloca-se através da sequência até encontrar o primeiro códon de iniciação, AUG. Tem início, então, a tradução a partir deste AUG. O processo de reiniciação da tradução, a partir de outros códons de iniciação localizados após o primeiro, é ineficiente. Assim, a presença de múltiplos códons de iniciação AUG na sequência líder deve resultar na redução da tradução a partir do códon que abre a fase de leitura principal. A região localizada abaixo do provável "TATA box", possivelmente corresponde a sequência líder do gene o2 de *Coix*. Esta região apresenta sete codons AUGs de iniciação de tradução anteriores ao AUG da principal fase de leitura (Fig. 7). Entre os polipeptídeos deduzidos a partir destes códons de iniciação, apenas o sexto, apresenta o códon de terminação localizado na região codificadora do gene. Este fato não ocorre em milho e sorgo, onde todos os códons de terminação dos polipeptídeos da sequência líder encontram-se antes do início da sequência codificadora.

Lohmer *et al.* (1993) demonstraram o envolvimento dos códons de iniciação, presentes na sequência líder do cDNA de *o2* de milho, no controle traducional deste gene.

Usando anticorpos contra a proteína O2 de milho, Schmidt (1993) relatou resultados indicando a manutenção da concentração da proteína O2, apesar do sensível aumento de mRNA acumulado durante o desenvolvimento do endosperma nas sementes de milho. As análises de mRNA de *opaco2* em *Coix* indicam, que a exemplo de milho, a transcrição do gene *o2* ocorre especificamente no endosperma acumulando-se durante o desenvolvimento das sementes.

A presença de múltiplos códons de iniciação na sequência líder não traduzida dos genes *opaco2* de *Coix*, milho e sorgo, e o fato da concentração da proteína O2 em milho permanecer constante durante o desenvolvimento das sementes, indicam que o gene *opaco2*, além de controle transcricional deve apresentar algum tipo de controle traducional.

O alinhamento das regiões flangeadoras 5' dos genes *opaco2* é de grande importância, visto que esta região pode conter elementos *cis* relacionados ao controle da expres-

são no nível transcricional. O gene *o2* de *Coix* foi o único dos genes sequenciados a apresentar um "TATA box" com motivo similar à sequência consenso descrita para este elemento (Bucher, 1990) (Fig. 10, TATA I). Nos genes *o2* de sorgo e milho a sequência CCTATTTGC (Fig. 10, TATA II) foi apontada como provável "TATA box" (Hartings *et al.*, 1989; Pirovano *et al.*, 1994). Apesar de ser considerada uma sequência incomum para "TATA box", sua funcionalidade não pode ser completamente descartada, conforme a mais recente descrição de elementos *cis* envolvidos na interação com a RNA polimerase II de eucariotos (Bucher, 1990). A provável função desta sequência é reforçada ainda pela presença de prováveis sinais de iniciação de transcrição ("Cap signal") localizados à distâncias esperadas em relação ao "TATA box". Corrobora com este fato o alto grau de conservação desta região nos três genes.

O elemento *cis* conhecido como "CCAAT box" tem sido descrito como um elemento que atua no controle transcricional, funcionando em qualquer orientação em distâncias variáveis em relação ao início de transcrição (Maniatis *et al.*, 1987). Um motivo altamente conservado nos genes *o2* apresentando similaridade

com a sequência consenso do "CCAAT box" (Bucher, 1990) foi identificado no alinhamento apresentado na Figura 10.

Os promotores de *opaco2* apresentam, ainda, uma cópia conservada do motivo CACGTG, definida como região central do elemento *cis* denominado "G box" (Izawa *et al.*, 1993). Este elemento tem sido identificado em transativadores de plantas, relacionados, principalmente, com fatores de resposta à efetores como a luz e o hormônio vegetal ácido abscísico (Maniatis *et al.*, 1987). O motivo encontrado nos promotores de *o2* apresentam sequências flanqueadoras que determinam uma baixa afinidade pelos transativadores descritos até o momento (Isawa *et al.*, 1993). Porém esta constatação não descarta a possibilidade deste elemento interagir eficientemente com fatores ainda não descritos.

O mecanismo de transcrição é, reconhecidamente, conservado nos eucariotos. Várias proteínas ativadoras de fungos funcionam, perfeitamente bem, em uma série de organismos eucariotos, e fatores de transcrição de vertebrados são ativos em células de fungos (Guarente, 1988; Buratowski *et al.*, 1988).

Nos procariotos cada fator envolvido na ativação da transcrição (transativador), geralmente, reconhece uma sequência alvo específica no DNA. Em contraste, um dos aspectos fundamentais na regulação da expressão gênica nos eucariotos é o de que eles possuem famílias de proteínas que interagem com sequências similares de DNA (Struhl, 1989). Como exemplo podemos citar a série de proteínas bZIP de mamíferos Jun, Jun B, Jun D, Fos, Fra, entre outras, que interagem com uma sequência comum, usualmente conhecida como sítio AP-1 (Nakabeppu *et al.*, 1988; Halazonetis *et al.*, 1988; Kouzarides & Ziff, 1988). Muitos dos transativadores são conservados nas diferentes classes de eucariotos e reconhecem sequências alvos similares, embora possam responder a diferentes efetores. Como exemplo podemos citar a proteína GCN4 de levedura e a oncoproteína Jun de vertebrados.

A proteína *Opaco2* possui estrutura similar e desempenha a mesma função em milho, sorgo e *Coix*, ativando a transcrição dos genes de α -prolaminas. Entretanto, diferentemente de outros transativadores interrelacionados, que reconhecem sequências similares, a proteína *O2* reconhece sítios distintos nos

promotores de α -prolaminas de milho e Coix. O sítio reconhecido pela proteína O2 no promotor do gene de α -coixina de 25 kDa apresenta maior similaridade com os sítios descritos no promotor de *b-32* (Lohmer *et al.*, 1991), do que com aquele encontrado no promotor de α -zeína.

A proteína O2 foi descrita originalmente como um transativador específico de α -prolaminas, entretanto, Cord *et al.* (1994) demonstraram que esta proteína é também capaz de ativar a transcrição dos genes de β -prolaminas em milho e Coix. Sendo que os promotores de β -prolaminas apresentam um sítio de ligação similar ao sítio O2a descrito no promotor do gene de α -zeína de 22 kDa (Yunes *et al.*, 1994).

Baseado na similaridade, os diversos sítios reconhecidos pela proteína Opaco2 descritos até o momento, podem ser agrupados em duas classes distintas: Tipo α -zeína e Tipo α -coixina (Tabela III).

Devido ao caráter multigênico das α -prolaminas, e ao pequeno número de genes sequenciados não se pode descartar a presença de sítios de reconhecimento do Tipo α -coixina em promotores de α -zeínas, e vice-versa.

Através de experimentos de retardamento em gel e "footprinting", Lohmer *et al.* (1991) identificaram no gene *opaco2* de milho um sítio reconhecido pela própria proteína O2. Estes mesmos autores demonstraram, através de experimentos de transativação, a auto-ativação do gene *opaco2* de milho. O motivo GTTGACGTTG, identificado como sítio de ligação da O2 em seu próprio promotor, encontra-se truncado nos promotores dos genes *opaco2* de Coix e sorgo (Fig. 10). A inexistência de sítios de ligação para proteína O2 de Coix, nos promotores de *opaco2* de Coix, foi confirmada através de testes de ligação seletiva (dados não apresentados).

Durante a busca de possíveis sítios de ligação da proteína O2 no gene *opaco2* de milho, foi identificada uma sequência similar ao sítio-alvo, descrito para o gene de α -coixina de 25 kDa, na região codificadora do gene *opaco2*, cerca de 150 pb abaixo do códon de iniciação (Cord Neto, G. comunicação pessoal). O sítio-alvo foi confirmado através de experimentos de "footprinting", e inclui o motivo CATG, considerado essencial para o reconhecimento das sequências de ligação (Yunes *et al.*, 1994). Experimentos de ligação seletiva realizados com a região codificadora

Tabela III: Sítios-alvo da proteína Opaco2 em diversos genes.

A) Tipo zeína

Sítio	origem	gene	sequência ACGT	ligação	referência
O2a	α -zeína	22Z-4	GTCATTCCACGTAGATGAAAA	+	Schmidt <i>et al.</i> , 1992
O2	β -zeína	gZ15A	ATGTTTCCACGTCATGCAACG	+	Cord Neto <i>et al.</i> , 1994
O2	β -coixina	β -55.1	GTATTTCCACGTCATGCAGTG	+	Cord Neto <i>et al.</i> , 1994
O2a	α -coixina	α -3B	GTCATTCCCTCATGAATTAGAA	-	Yunes <i>et al.</i> , 1994

B) Tipo coixina

Sítio	origem	gene	sequência CATG	ligação	referência
O2c	α -coixina	α -3B	AATTGACTAGGAGACATGTCA	+	Yunes <i>et al.</i> , 1994
O2c	α -coixina	α -3A	AATCGACTAGGAGCCATGTCA	+	Yunes <i>et al.</i> , 1994
O2c	α -zeína	22Z-4	AATAGAGTAGATGCCATATCA	-	Yunes <i>et al.</i> , 1994
B1	B-32		GGTTAGTATGATGACATGGCT	+	Lohmer <i>et al.</i> , 1991
B4	B-32		TGTTAGTGGGATGACATGACT	+	Lohmer <i>et al.</i> , 1991

do gene *opaco2* de *Coix* apresentaram resultados negativos (Cord Neto, G. comunicação pessoal). O alinhamento da região contendo o sítio encontrado na região estrutural do gene de *opaco2* de milho e a região correspondente no gene *opaco2* de *Coix* revela a ruptura do motivo CATG na sequência da região codificadora do gene *opaco2* de *Coix*. Estes resultados indicam que a ligação da proteína O2 na região codificadora do gene *opaco2* de milho deve constituir um mecanismo alternativo de auto-regulação da expressão desse gene em milho. Uma hipótese para a auto-regulação, supõe que a ligação do transativador no início da região codificadora pode interferir na etapa de elongação da transcrição, e neste caso provocando uma possível inibição da transcrição (Cord Neto, G., resultados não publicados). O gene *opaco2* de *Coix* não apresenta as sequências descritas como sítios de auto-ativação (região promotora) e possível auto-inibição (região codificadora) encontrados no gene *opaco2* de milho. Tal fato corrobora com a hipótese de auto-regulação proposta, visto que, a ausência de um dos sítios desobriga a presença do outro, não sendo possível a presença de um sítio de auto-ativação na

ausência de um de sítio de auto-inibição e vice-versa.

Os resultados descritos neste trabalho representam uma primeira fase na elucidação dos mecanismos moleculares envolvidos na função de transativação da proteína O2. Assim sendo, a comparação das sequências de aminoácidos das proteínas O2 de milho, sorgo e *Coix*, nos proporcionaram o reconhecimento de regiões conservadas, possivelmente envolvidas no reconhecimento de sítios no DNA, transporte para o núcleo e ativação de transcrição. Baseados nestes resultados, as funções dos motivos conservados serão testadas com a obtenção de proteínas recombinantes mutadas.

A comparação das regiões promotoras dos genes *opaco2* nos permitiu também identificar sequências conservadas nesta região. Este resultado deverá direcionar os estudos de análise de deleção a serem realizados com estes promotores, visando o reconhecimento de sequências envolvidas na regulação da transcrição do gene *opaco2*.

CONCLUSÕES

- 1) A sequência de aminoácidos da proteína Opaco2 de *Coix* é similar às sequências das proteínas O2 de milho apresentada por Schmidt *et al.* (1990) e de sorgo deduzida a partir da sequência genômica depositada nos bancos de dados EMBL, GenBank e DDBJ.
- 2) As sequências de nucleotídeos dos exons do gene *opaco2* de *Coix*, milho e sorgo apresentam alta similaridade. As sequências de nucleotídeos dos íntrons de *Coix* apresentam maior similaridade com as sequências de íntrons de sorgo do que com as de milho.
- 3) Na região regulatória do gene *opaco2* de *Coix* foram identificadas possíveis sequências "TATA BOX" (-435) e "CAAT BOX" (-511). Um provável sinal de poliadenilação está localizado 117 pb abaixo do códon de terminação TAG.
- 4) Abaixo da sequência do provável "TATA BOX" existem sete fases de leitura além da fase principal, as quais, possivelmente, estão envolvidas com a regulação gênica.
- 5) A região do domínio bZIP apresenta-se extremamente conservada nas proteínas Opaco2.
- 6) A transcrição do gene *opaco2* ocorre, especificamente, no endosperma das sementes de *Coix*. Sendo que, seus transcritos acumulam-se durante o desenvolvimento das sementes.

- 7) Assim como a proteína Opaco2 de milho, a proteína O2 de Coix também reconhece como sítio-alvo de ligação a sequência 5'-GACATGTC-3', presente nos promotores das α -coixinas de 25 kD, e a sequência 5'-TCCACGTAGA-3', presente nos promotores das α -zeínas de 22 kD.

Sumário

Coix lacryma-jobi L., juntamente com milho, sorgo e tripsacum, pertence a família *Gramineae* e a tribo *Andropogoneae* (Clayton, 1973; Jain & Benerjee, 1974; Clayton, 1983).

Nos membros desta tribo, as proteínas de reserva mais abundantes são representadas pelas prolaminas, as quais chegam a atingir 50% a 70% da quantidade total de proteína no endosperma (Otto et al., 1990; Leite et al., 1990). As α -prolaminas correspondem a cerca de 70% das prolaminas totais de milho, sorgo e *Coix*.

Muitas mutações afetando o acúmulo das prolaminas de milho (zeínas) têm sido descritas (Motto et al., 1989). Nos mutantes *opaco2* ocorre a quase completa eliminação das zeínas de 22 kDa (Pedersen et al., 1980; Langridge et al., 1982). O gene *opaco2* codifica um transativador transcricional da classe das bZIPs, o qual se liga a região regulatória dos genes das α -zeínas de 22 kD. O sítio de ligação da proteína O2 nos genes das α -zeínas contém a sequência 5'-TCCACGTAGA-3' (O2a, Schmidt et al., 1992). Nos promotores das α -coixinas de 25 kD, a proteína O2 de

milho reconhece a sequência 5'-GACATGTC-3' (O2c, Yunes. et al., 1994).

Pirovano et al. (1994) isolaram e caracterizaram o gene *opaco2* de sorgo.

Neste trabalho, foram isolados clones de cDNA e genômico do gene *opaco2* de *Coix lacryma-jobi*. Para o isolamento do clone de cDNA, denominado pCO2-1, foram analisados aproximadamente $7,5 \times 10^5$ fagos recombinantes de uma biblioteca de cDNA construída em fago λ ZAP. Para o isolamento do clone genômico foram analisados cerca de 1×10^6 fagos recombinantes.

O clone de cDNA está incompleto e apresenta 1.050 pb. O clone genômico possui 4.569 pb, sendo 987, 215, e 3.367 pb, respectivamente das regiões flancadoras 5' e 3', e da região estrutural do gene. Sequências correspondentes aos possíveis "TATA box" e "CCAAT box" foram identificadas nas posições -435 (5'-TATAAAA-3') e -511 (5'-CCAAT-3'). Na região abaixo do possível "TATA box" foram identificadas sete fases de leituras além da fase principal, a qual

codifica uma proteína de 408 aminoácidos, correspondendo a um polipeptídeo de peso molecular 44.865. Os seis exons dos genes *opaco2* de *Coix*, milho e sorgo são altamente conservados tanto em sequência quanto em posição. Os cinco íntrons do gene *o2* de *Coix* apresentam similaridade com os íntrons do mesmo gene em sorgo, o mesmo não ocorrendo com os íntrons do gene *opaco2* de milho. Nas regiões flangeadoras 5' dos genes *opaco2* foi encontrada ainda uma sequência palindrômica similar a um "G box", descrito por Izawa *et al.* (1993).

Um provável sinal de poliadenilação (5'-AAATAA-3') está presente 117 pb abaixo do códon de terminação TAG.

A região correspondente aos domínios básico e "zipper" de leucina (bZIP) é altamente conservada nas proteínas *Opaco2* de *Coix*, milho e sorgo.

Análises de RNA mostraram que, transcritos do gene *opaco2* encontram-se, especificamente, no endosperma das sementes de *Coix*, e que, estes transcritos são acumulados durante o desenvolvimento destas sementes.

Através de experimentos de ligação seletiva, demonstrou-se que

as proteínas *Opaco2* de *Coix* e de milho, reconhecem, como sítio-alvo, o sítio O2a nas α -zeínas e o sítio O2c nas α -coixinas.

Summary

Coix lacryma-jobi L., together with maize, sorghum and tripsacum belongs to the grass tribe Andropogoneae (Clayton, 1973; Jain and Benerjee, 1974; Clayton, 1983).

As for other members of the tribe, prolamins represent 50-70% of the total seed storage proteins of *Coix* (Ottoboni *et al.*, 1990; Leite *et al.*, 1990). The α -prolamins fraction accounts for more than 70% of the total prolamins of maize, sorghum, and *Coix*.

Several mutations are known to affect the accumulation of prolamins in maize (Motto *et al.*, 1989). The maize *opaque2* mutant shows almost complete elimination of the 22 kD α -zein (Pedersen *et al.*, 1980; Langridge *et al.*, 1982). The *opaque2* gene from maize encodes a trans-acting factor, belonging to the bZip class, that activates the transcription of 22 kD α -zein genes.

The maize Opaque2 protein recognizes the sequence 5'-TCCACGTAGA-3' (O2a) in the promoter region of 22 kDa α -zein gene (Schmidt *et al.*, 1992) and the sequence 5'-GACATGTC-3' (O2c) in

the 25 kDa α -coixin promoter (Yunes *et al.*, 1994).

A *opaque2*-related gene from sorghum has been recently reported by Pirovano *et al.* (1994).

We have isolated a cDNA and a genomic clone corresponding to the *opaque2* gene in *Coix lacryma-jobi*. In order to isolate the cDNA clone, named pCO2-1, 7,5 x 10⁵ recombinant phages were screened. The genomic clone, named gCXO2-1, was isolated by screening 1 x 10⁶ recombinant phages of a genomic library.

The cDNA clone of 1050 bp is incomplete. The genomic clone possesses 4569 bp. The 5' and 3' flanking regions possess 987 and 215 bp respectively. The coding region is 3367 bp long. The regions with the strongest similarity to the TATA and CAAT boxes were found at positions -435 (5'-TATAAAA-3') and -511 (5'-CCAAT-3'), respectively. Downstream to the putative TATA box we have found seven open reading frames (ORF) in addition to the main ORF. This ORF encodes a polypeptide of 408 amino acids.

The six exons of *Coix*, maize, and sorghum *opaque2* genes exhibit a

high similarity in their sequence and positions. The five introns of the *opaque2* gene from *Coix* present some similarity to the introns of sorghum *opaque2* gene. However there is no similarity when compared to those present in maize *opaque2* gene. At the 5' flanking region we have found a G-box like sequence (Izawa *et al.*, 1993).

A putative polyadenylation signal (5'-AAATAA-3') is present 117-bp downstream to the TAG stop codon

The amino acid sequence in the bZip region is very similar in the three *opaque2* proteins.

RNA blot analysis revealed that *Coix opaque2* transcripts are present specifically in endosperm tissue, and that they accumulate during seed development.

Using DNA binding assays, with both *Coix* and maize *Opaque2* proteins, we have demonstrated that both recognize the O2a sequence present in the 22 kDa α -zein promoter and the O2c present in the 25 kDa α -coixin promoter.

LITERATURA CITADA

- Angel, P., Allegretto, E.A., Okino, S.R., Hattori, K., Boyle, W.J., Hunter, T., Karim, M. (1988). Oncogene *jun* encodes a sequence-specific *trans*-activator similar to AP-1. *Nature* **332**, 166-171.
- Argos, P., Pedersen, K., Marks, M.D., Larkins, B.A. (1982). A structural model for maize zein proteins. *J. Biol. Chem.* **257**, 9984-9990.
- Armstrong, G.A., Weisshaar, B., Hahlbrock, K. (1992). Homodimeric and heterodimeric leucine zipper proteins and nuclear factors from parsley recognize diverse promoter elements with ACGT cores. *Plant Cell* **4**, 525-537.
- Arora, R.K. (1977). Job's tears (*Coix lacrym-jobi*) - a minor food and fodder crop from North Eastern India. *Econ. Bot.* **31**, 358-366.
- Aukerman, M.J., Schmidt, R.J., Burr, B., Burr F.A. (1991). An arginine to lysine substitution in the bZip domain of an *opaque-2* mutant in maize abolishes specific DNA binding. *Genes Dev.* **5**, 310-320.
- Barros, E.G., Takasaki, K., Kirleis, A.W., Larkins, B.A. (1991). Nucleotide sequence of a cDNA clone encoding γ -kafirin protein from *Sorghum bicolor*. *Plant Physiol.* **97**, 1606-1607.
- Bass, H.W., Webster, C., Obian, G.R., Roberts, J.K.M., Boston, R.S. (1992). A maize ribosome-inactivating proteins controlled by the transcriptional activator *opaque-2*. *Plant Cell* **4**, 225-234.
- Bohmann, D. & Tjian, R. (1989). Biochemical analysis of transcriptional activation by Jun: differential activity of c- and v-Jun. *Cell* **59**, 709-717.
- Breathnach, R. & Chambon, P. (1981). Organization and expression of Eukaryotic split genes coding for proteins. *Annu. Rev. Biochem.* **50**, 349-383.
- Brown, J.W.S., Wandelt, C., Feix, G., Neuhaus, G., Schweiger, H.G. (1986). The upstream regions of zein genes - Sequence analysis and expression in the unicellular alga *Acetabularia*. *Eur. J. Cell. Biol.* **42**, 161-170.
- Bucher, P. (1990). Weight matrix descriptions or four eukaryotic RNA Polymerase II promoter elements derived from 502 unrelated promoter sequences. *J. Mol. Biol.* **212**, 563-578.
- Bullock, W.O., Fernandez, J.M., Short, J.M. (1987). XL1-Blue: A high efficiency plasmid transforming *recA Escherichia coli* strain with beta-galactosidase selection. *Biotechniques* **5**, 376-379.
- Buratowski, S., Hahn, S., Sharp, P.A., Guarente, L. (1988). Function of a yeast TATA element-binding protein in a mammalian transcription system. *Nature* **334**, 37-42.
- Burr, B. & Burr, F.A. (1976). Zein synthesis in maize endosperm by polyribosomes attached to proteins bodies. *Proc. Natl. Acad. Sci. USA* **70**, 1767-1769.
- Burr, B. & Burr, F.A. (1978). In Vitro uptake and processing of prezein and others maize preproteins by maize membranes. *J. Cell Biol.* **90**, 427-434.
- Burr, B. & Burr, F.A. (1982). 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.

- Cabot, E. (1988).** The Eyeball Sequence Editor. (programa obtido a partir do banco de dados de seqüências EMBL). Email: NETSERV@EMBL-Heidelberg.DE
- Clayton, W.D. (1973).** The awnless species of Andropogoneae. *Kew Bull.* **28**, 49-58.
- Clayton, W.D. (1983).** Notes on tribe Andropogoneae (Gramineae). *Kew Bull.* **35**, 813-818.
- Colot, V., Robert, L.S., Kavanagh, T.A., Bevan, M.V., Thompson, R.D. (1987).** Localization of sequences in wheat endosperm protein genes which confer tissue-specific expression in tobacco. *EMBO J.* **6**, 3559-3564.
- Cord-Neto, G., Yunes, J.A., Vettore, A.L., Souza, G.A., Silva, M.J., Ottoboni, L.M.M., Arruda, P., Leite, A. (1994).** The regulation of β -prolamin gene expression by the Opaque2 protein in maize and *Coix lacryma-jobi*. *in preparation*
- Danner, D.B. (1982).** Recovery of DNA fragments from gels by transfer to DEAE-paper in eletrophoresis chamber. *Anal. Biochem.* **125**, 139-142.
- Das, O.P. & Missing, J.W. (1987).** Allelic variation and differential expression at the 27-kilodalton zein locus in maize. *Mol. Cell. Biol.* **7**, 4490-4497.
- Das, O.P., Ward, K., Ray, S., Messing, J. (1991a).** Sequence variation between alleles reveals two types of copy correction at the 27-kDa zein locus of maize. *Genomics* **11**, 849-856 .
- Das, O.P., Poliak, E., Ward, K., Messing, J.W. (1991b).** A new allele of the duplicated 27 kDa zein locus of maize generated by homologous recombination. *Nucl. Acids Res.* **19**, 3325-3330.
- De Rose, R.T., Ma, D.P., Kwon, I.S., Hasnain, S.E., Klassy, R.C., Hall, T.C. (1989).** Characterization of the kafirin gene family from sorghum reveals extensive homology with zein from maize. *Plant Mol. Biol.* **12**, 245-256.
- Di Fonzo, N., Mandocchi, L., Salamini, F., Soave, C. (1986).** Purification and properties of and endospermic protein of maize associated with the *opaque-2* and *opaque-6* genes. *Planta* **167**, 587-594.
- Dingwall, C. & Laskey, R.A. (1991).** Nuclear targeting sequences - A consensus ? *Trends Biol. Sci.* **16**, 478-481.
- Dretzen, G., Bellard, M., Sassone-Corsi, P., Chambon, P. (1981).** A reliable method for the recovery of DNA fragments from agarose and acrylamide gels. *Anal. Biochem.* **112**, 295-298.
- Emerson, R.A., Beadle, G.W., Fraser, A.C. (1935).** A summary of linkage group in maize. *Cornell Univ. Agric. Exp. Sta. Men.* **180**.
- Esen, A. (1986).** Separation of alcohol-soluble proteins (zeins) from maize into three fractions by differential solubility. *Plant Phys.* **80**, 623-627.
- Evans, D.J., Schüssler, L., Taylor, J.R.N. (1987).** Isolation of a reduced-soluble protein from sorghum starchy endosperm. *J. Cereal Sci.* **5**, 61-65.
- Feinberg, A.P. & Vogelstein, B. (1984).** A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Addendum Anal. Biochem.* **137**, 266-267.
- Feix, G. & Quayle, T. (1993).** Structure and expression of zein genes of maize. *Crit. Rev. Plant Sci.* **12**, 111-127.

- Fourney, R.M., Miyakoshi, J., Day III, R.S., Peterson, M.C. (1988). Northern blotting: Efficient RNA staining and transfer. Bethesda Res. Lab. Focus 10 (5), 5-7.
- Freitas, F.A., Yunes, J.A., Silva, M.J., Arruda, P., Leite, A. (1994). Structural characterization and promoter activity of the γ -kafirin gene from sorghum. Mol. Gen. Genet. *in press*
- Gallardo, D., Reina, M., Rigua, J., Boronat, A., Palau, J. (1988) Genomic organization of the 28 kDa glutelin-2 gene from maize. Plant Sci. 54, 211-218.
- Garret, R., Olliva, G., Caracelli, I., Leite, A., Arruda, P. (1993) Studies of the zein-like α -prolamins based on an analyses of amino acid sequences: implications for their evolution and three-dimensional structure. Prot. Stru. Func. Genet. 15, 88-99.
- Gilbert, W. (1978). Why genes in pice? Nature 271, 501.
- Gilbert, W. (1985). Genes-in-pice revisited. Science 228, 823-824.
- Glass, D.B. & Smith, S.B. (1983). Phosphorlation by cyclic GMP-dependent protein kinase of a synthetic peptide corresponding to a autophosphorylation site in the enzyme. J. Biol. Chem. 258, 14797-14803.
- Guarente, L. (1988). UASs and enhances: Common mechanism of transcriptional activation in yeast and mammals. Cell 52, 303-305.
- Guiltinan, M.J., Marcotte, W.R., Quatrano, R.S. (1990). A plant leucine zipper that recognizes an abscisic acid response element. Science 250, 267-271.
- Halazonetis, T., Georgopoulos, K., Greenberg, M.E., and Leder, P. (1988). Transcriptional activation by the SV40-AP-1. Recognition Element in yeast is mediated by a factor simule to AP-1 that is different from GCN-4. Cell 53, 917-942.
- Hall, M.N., Hereford, L., Herskowitz, I. (1984). Targeting of *E. coli* β -galactosidase to the nucleus in yeast. Cell 36, 1057-1065.
- Hanaham, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. 166, 557.
- Hartings, H., Maddaloni, M., Lazzaroni, N., Di Fonzo, N., Motto, M., Salamini, F., Thompson, R.D. (1989). The O2 gene which regulates zein deposition in maize endosperm encodes a protein with structural homologies to transcriptional activators. EMBO J. 8, 2795-2801.
- Hartings, H., Lazzaroni, N., Aljmone Marsan, P., Aragay, A., Thompson, R.D., Salamini, F., Di Fonzo, N., Palau, G., Motto, M. (1990). The b-32 protein from maize endosperm: characterization of a genomic sequences encoding two alternatives central domains. Plant Mol. Biol. 14, 1031-1040.
- Henikoff, S. (1987). Unidirectional digestion with exonuclease III in DNA sequence analysis. Method. Enzymol. 155, 156-165.
- Higgins, D.G. & Sharp, P.M. (1989). Fast and sensitive multiple sequence alignments on a microcomputer. CABIOS, 5, 151-153. (programa obtido a partir do banco de dados de seqüências EMBL). Email: NETSERV@EMBL-Heidelberg. DE.
- Hinnebush, A.G. (1984). Evidence for translational regulation of the activator of general amino acid control in yeast. Proc. Natl. Acad. Sci. USA. 81, 6442-6446.
- Hull, G.A., Halford, N.G., Kreis, M., Shewry, P.R. (1991). Isolation and characterization of genes encoding rye prolamins containing a highly repetitive sequence motif. Plant Mol. Biol. 17, 1111-1115
- Izawa, T., Foster, R., and Chua, N.-H. (1993). Plant bZIP protein DNA binding specificity. J. Mol. Biol. 230, 1131-1144.

- Jain, S.K., & Benerjee, D.K. (1974). Preliminary observations on the ethnobotany of the genus *Coix*. *Econ. Bot.* **28**, 38-42.
- Johnson, P.F. & McKnight, S.L. (1989). Eukaryotic transcriptional regulatory proteins. *Ann. Rev. Biochem.* **58**, 799-839.
- Kalderon, D., Richardson, W.D., Markham, A.F., Smith, A.E. (1984a). Sequences requiriments for nuclear localization of simian virus 40 large-T antigen. *Nature* **331**, 33-38.
- Kalderon, D., Roberts, B.L., Richardson, W.D., Smith, A.E. (1984b). A short aminoacid sequence able to sprcify nuclear localization. *Cell* **39**, 499-509.
- Katagari, F., Yamazaki, K., Horikoshi, M., Roeder, R.G., Chua, N.-H. (1989). A palnt DNA-binding protein increases the number of preinitiation complexes in a human *in vitro* transcription system. *Genes Dev.* **4**, 1899-1909.
- Kirihara, J.A., Husperger, J.P., Mahoney, J.W. (1988). Differential expression of a gene for a methionine-rich storage protein in maize. *Mol. Gen. Genet.* **211**, 477-484.
- Kodrzycki, R., Boston, R.S., Larkins, B.A. (1989). The *opaque-2* mutation of maize differentially reduces zein gene transcription. *Plant Cell* **1**, 105-114.
- Kouzarides, T., & Ziff, E. (1988). The role of the leucine zipper in the fos-jun interaction. *Nature* **336**, 646-651.
- Kozak, M. (1984). Selection of initiation sites by eukaryotics ribossomes: Effect of inserting AUG triplets upstream from the coding sequence for preproinsulin. *Nucl. Acids res.* **12**, 3873-3893.
- Kozak, M. (1989). A scanning model for translation: an update. *J. Cell Biol.* **108**, 229-241.
- 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. (1986). Differential gene expression in the developing barley endosperm. *Phil. Trans. R. Soc. Lond.* **314**, 355-356.
- Landschulz, W.H., Johnson, P.F., Adashi, E.Y., Graves, B.J., McKnight, S.L. (1988a). Isolation of a recombinant copy of a gene encoding C/EBP. *Genes Dev.* **2**, 786-800.
- Landschulz, W.H., Johnson, P.F., McKnight, S.L. (1988b). The leucine zipper: A hypothetical structure common a new class of DNA binding protein. *Science* **240**, 1759-1764.
- Langridge, P., Pintor-toro, J.A., Feix, G. (1982). Zeins precursor mRNAs from Maize endosperm. *Mol. Gen. Genet.* **187**, 432-438.
- Larkins, B.A. & Hurkman, W.J. (1978). Synthesis and deposition of zeins in proteins bodies of maize endosperm. *Plant Physiol.* **62**, 256-263.
- Larkins, B.A., Pedersen, K., Marks, M.D., Wilson, D.R. (1984). The zein proteins of maize endosperm. *Trends Biochem. Sci.* **9**, 306-308.
- Leite, A., Ottoboni, L.M.M., Targon, M.L.P.N, Silva, M.J., Turcinelli, S.R., Arruda, P. (1990). Phylogenetic relationship of zeins and coixins as determined by immunological cross-reactivity and Southern blot analysis. *Plant Mol. Biol.* **14**, 743-751.
- Leite, A., Freitas, F.A., Yunes, J.A., Arruda, P. (1991). Nucleotide sequence of a cDNA clone encoding γ -coixin from *Coix lacryma-jobi* seeds. *Plant Physiol.* **97**, 1604-1605.

- Leite, A., Yunes, J.A., Turcinelli, S.R., Arruda, P. (1992). Cloning and characterization of a cDNA encoding a sulfur-rich coixin. *Plant Mol. Biol.* **18**, 171-174.
- Lending, C.R. & Larkins, B.A. (1989). Changes in the zein composition of protein bodies during maize endosperm development. *Plant Cell* **1**, 1011-1023.
- Liou, H.C., Boothby, M.R., Finn, P.W., Davidon, R., Nabavi, N., Zeleznik-Le, N.J., Ting, J.P.Y., Glincher, L.H. (1990). A new member of the leucine zipper class of the proteins that binds to the HLA DR-alpha promoter. *Science* **247**, 1581-1584.
- Liu, C.-N. & Rubenstein, I. (1992). Molecular characterization of two types of 22 kilodalton α -zein genes in a gene cluster in maize. *Mol. Gen. Genet.* **234**, 244-253.
- Lohmer, S., Maddaloni, M., Motto, M., Di Fonzo, N., Hartings, H., Salamini, F., Thompson, R.D. (1991). 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.
- Lohmer, S., Maddaloni, M., Motto, M., Salamini, F., Thompson, R.D. (1993). Translation of the mRNA of the maize transcriptional activator *opaque-2* is inhibited by upstream open reading frames present in the leader sequence. *Plant Cell* **5**, 65-73.
- Ludevid, M.D., Torrent, M., Martinez-Izquierdo, J.A., Puigdomenech, P., Palau, J. (1984). Subcellular localization of glutelin-2 in maize (*Zea mays* L.). *Plant Mol. Biol.* **3**, 227-234.
- Maekawa, T., Sakura, H., Kanei-Ishii, C., Sudo, T., Yoshimura, T., Fugisawa, J.I., Yoshida, M., Ishii, S. (1989). Leucine zipper structure of the protein CRE-BP1 binding to the cyclic AMP response element in brain. *EMBO J.* **8**, 2023-2028.
- Maier, U.G., Brown, J.W.S., Toloczkyk, C., Feix, G. (1987). Binding of nuclear factor to a consensus sequence in the 5' flanking region of zein genes from maize. *EMBO J.* **6**, 17-22.
- Maniatis, T., Goodbourn, S., Ficher, J.A. (1987). Regulation of inducible and tissue-specific gene expression. *Science* **236**, 1237-1245.
- Marks, M.D., Lindell, J.S., Larkins, B.A. (1985). Quantitative analyses of the accumulation of zein mRNA during maize endosperm development. *J. Biol. Chem.* **260**, 16451-16459.
- Mello, L.V. (1989). Estudo da biologia da reprodução em *Coix* ssp. Tese apresentada ao Instituto de Biologia da Universidade Estadual de Campinas para obtenção do Título de Mestre em Ciência.
- Mermod, N., O'Neill, E.A., Kelly, T.J., Tjian, R. (1989). The proline-rich transcriptional activator of CTF/NF-1 is distinct from the replication and DNA binding domain. *Cell* **58**, 741-753.
- Mertz, E.T., Bates, L.S., Nelson, O.E. (1964). Mutant gene that changes protein composition and increases lysine content of maize endosperm. *Science* **145**, 279-280.
- Motto, M., Maddaloni, M., Ponziani, G., Brembilla, M., Marotta, R., Di Fonzo, N., Soave, C., Thompson, R., Salamini, F. (1988). Molecular cloning of the *o2-m5* allele of *Zea mays* using transposon marking. *Mol. Gen. Genet.* **212**, 488-494.
- Motto, M., Di Fonzo, N., Hartings, H., Maddaloni, M., Salamini, F., Soave C., Thompson, R.D. (1989). Regulatory genes affecting maize storage protein synthesis. In *Oxford Surveys of Plant Molecular and Cell Biology*, Vol. 6, B.J. Mifflin, ed. (Oxford: Oxford University Press), pp. 87-114.

- Nakabeppu, Y., Ryder, K., Nathans, D. (1988). DNA binding activates activities of three murine Jun proteins: stimulation by Fos. *Cell* **55**, 907-915.
- Nelson, O.E. (1980). Genetic control of polysaccharide and storage protein synthesis in endosperm of barley, maize, and sorghum. In *Advances in Cereal Science and Technology*, Vol. 3, Y. Pomares (Ed.) American Association of Cereal Chemists, St. Paul, MN, pp. 41-71.
- Neuberg, M., Schuermam, M., Hunter, J.B., Müller, R. (1989). Two functionally different regions in Fos are required for sequence-specific DNA interaction of the Fos/Jun protein complex. *Nature* **338**, 589-590.
- O'Neil, K.T., Hoess, R.H., DeGrado, W.F. (1990). Design of DNA-binding peptides on the Leucine Zipper motif. *Science* **249**, 774-778.
- Ottoboni, L. M. M., Leite, A., Targon, M.L.N., Crozier, A., Arruda, P. (1990). Characterization of the storage protein in seed of *Coix lacryma-jobi* var. Adlay. *J. Agric. Food Chem.* **38**, 631-635.
- Ottoboni, L.M.M., Leite, A., Yunes, J.A., Targon, M.L.P.N., Souza Filho, G.A., Arruda, P. (1993). 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.
- Pabo, C.O. & Sauer, R.T. (1992). Transcriptional Factors: Structural families and principles of DNA recognition. *Ann. Rev. Biochem.* **61**, 1053-1095.
- Paluh, J.L., Orbach, M.J., Legerton, T.L., Yanofsky, C. (1988). The cross-pathway control gene of *Neurospora crassa*, *cpc-1*, encodes a protein similar to GCN4 of yeast and DNA-binding domain of the oncogene *v-jun*-encoded protein. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 3728-3732.
- Pedersen, K., Bloom, K.S., Anderson, J.N., Glover, D.V., Larkins, B.A. (1980). Analysis of the complexity and frequency of zein genes in the maize genome. *Biochemistry* **19**, 1644-1650.
- Pedersen, K., Argos, P., Naravana, S.L.V., Larkins, B.A. (1986). Sequence analysis and characterization of a maize gene encoding a high-sulfur zein protein of M_r 15 000. *J. Biol. Chem.* **261**, 6279-6284.
- Perrin, S., & Gilliland, G. (1990). Site-specific mutagenesis using asymmetric polymerase chain reaction and a single mutant primer. *Nucleic Acid Res.* **18**, 7433-7438.
- Pirovano, L., Lanzini, S., Hartings, H., Lazzarini, N., Rossi, V., Joshi, R., Thompson, R.D., Salamini, F., Motto, M. (1994). Structural and functional analysis of an *Opaque-2*-related gene from sorghum. *Plant. Mol. Biol.* **24**, 515-523.
- Preccott, A. & Martim, C. (1987). Rapid method for quantitative assessment of levels of specific mRNAs. *Plant Mol. Biol. Report* **4**, 219-224.
- Raikhel, N. (1992). Nuclear targeting in plants. *Plant Physiol.* **100**, 1627-1632.
- Rivin, C.J., Zimmer, E.A., Walbot, V. (1982). Isolation of DNA and DNA recombinants from maize. In: Sheridan, W.F. (ed.). *Maize for Biological Research*, University Press, University of North Dakota, Grand Forks. pp.161-164.
- Sanger, F., Nicklen, S., Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci.* **74**, 5463-5467.
- Sambrook, J., Fritsch, E.F., Maniatis, T. (1989). *Molecular Cloning: A laboratory Manual*. 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

- Schindler, U., Beckmann, H., Kadesch, T., Cashmore, A.R. (1992). DNA binding site preferences and transcriptional activation properties of the *Arabidopsis* transcription factor GBF1. *EMBO J.* **11**, 1275-1289.
- Schmidt, R.J. (1993). Opaque-2 and zein gene expression. In *Control of plant gene expression*, Desch Pal. S. Verna, ed. (Boca Raton: CRC Press), pp. 337-355.
- Schmidt, R.J., Burr, F.A., Burr, B. (1987). Transposon tagging and molecular analysis of the maize regulatory locus *opaque-2*. *Science* **238**, 960-963.
- Schmidt, R.J., Burr F.A., Aukerman, M.J., Burr, B. (1990). Maize regulatory gene *opaque-2* encodes a protein with a "leucine zipper" motif that binds to zein DNA. *Proc. Natl. Acad. Sci. USA* **87**, 46-50.
- Schmidt, R.J., Ketudat, M., Aukerman, M.J., Hoschek, G. (1992). Opaque-2 is a transcriptional activator that recognizes a specific target site in 22-kD zein genes. *Plant Cell* **4**, 689-700.
- Shaaffhausen, R.V. (1952). Adlay or Job's tears - a cereal of potentially greater economic importance. *Econ. Bot.* **6**, 216-227.
- Shewry, P.R. & Tatham, A.S. (1990). The prolamin storage of cereal seeds: structure and evolution. *Biochem. J.* **267**, 1-12.
- Siegel, L.I. & Bresnick, E. (1986). Northern hybridization analysis of RNA using diethylpyrocarbonate-treated nonfat milk. *Anal. Biochem.* **159**, 82-87.
- Silver, P.A. (1991). How proteins enter the nucleus. *Cell* **64**, 489-497.
- Singh, K., Dennis, E.S., Ellis, J.G., Llewellyn, D.J., Tokuhisa, J.G., Wahleithner, J.A., Peacock, W.J. (1990). OCSBF-1, a maize ocs enhancer binding factor: isolation and expression during development. *Plant Cell* **2**, 981-903.
- Soave, C., Righetti, P.G., Lorenzoni, C., Gentinetta, E., Salamini, F. (1976). Expressivity of *opaque-2* gene at the level of zein molecular components. *Maydica* **21**, 61-75.
- Stadem, R. (1982) An interactive graphics program for comparing and aligning nucleic acid and amino acid sequences. *Nucleic Acid Res.* **10**, 2951-2961.
- Struhl, K. (1989). Helix-turn-helix, zinc finger, and leucine zipper motifs for eukaryotic transcriptional regulatory proteins. *Trend Biochem. Sci.* **14**, 137-140.
- Summer-Smith, M., Rafalski, J.A., Sugyama, T., Stoll, M., Söll, D. (1985). Conservation and variability of wheat a/b-gliadin genes. *Nucl. Acids Res.* **13**, 3905-3916.
- Tabata, R., Takase, H., Takayama, S., Mikami, K., Nakatsuka, A., Kawata, T., Nakayama, T., Iwabushi, M. (1989). A protein that binds a cis-acting element of wheat histone genes has a leucine zipper motif. *Science* **245**, 965-967.
- Targon, M.L.N., Ottoboni, L.M.M., Leite, A., Ludevid, D., Puigdomènech, P., Arruda, P. (1992). Synthesis and deposition of coixin in seeds of *Coix lacryma-jobi*. *Plant Sci.* **83**, 169-180.
- Töpfer, R., Matzeit, V., Gronenborn, B., Schell, J., Steinbiss, H.H. (1987). A set of plant expression vectors for transcriptional fusions. *Nucl. Acids Res.* **15**, 5890.
- Torres, P. & Bergamin, A. (1951). O cereal Adlay. Estudo do seu valor agrícola e nutricional para aves. *Anais da Escola Superior de Agricultura "Luiz de Quiroz"* **8**, 669-685.
- Turner, R. & Tjain, R. (1989) Leucine repeats and an adjacent DNA binding domain mediate the formation of functional cFos-cJun heterodimers. *Science* **243**, 1689-1694.

- van der Krol, A.R. & Chua, N.H. (1991). The basic domain of plant B-ZIP proteins facilitates import of a reporter protein into plant nuclei. *Plant Cell* **3**, 667-675.
- Van Straaten, F., Müller, R., Curran, T., van Beveren, C., Verna, I.M. (1983). Complete nucleotide sequence of human c-onc gene: deduced amino acid sequence of human c-fos protein. *Proc. Natl. Acad. Sci. U.S.A.* **80**, 3183-3187.
- Varagona, M.J., Schmidt, R.J., Raikhel, V. (1991). Monocot regulatory protein Opaque-2 is localized in the nucleus of maize endosperm and transformed tobacco plants. *Plant Cell* **3**, 105-113.
- Varagona, M.J., Schmidt, R.J., Raikhel, V. (1992). Nuclear localization signal(s) required for nuclear targeting of maize regulatory protein, Opaque-2. *Plant Cell* **4**, 1213-1227.
- Venkateswarlu, J., & Chaganti, R.S.K. (1973). Job's Tears (*Coix lacryma-jobi* L.). ICAR Tech. Bull. (Agric.) **43-44**, 1-54.
- Vinson, C.R., Singler, P.B., McKnight, S.L. (1989). Scissors-grip model for DNA recognition by a family of leucine zipper proteins. *Science*, **246**, 911-916.
- Viotti, A., Sala, E., Marotta, R., Alberti, P., Paducci, C., Soave, C. (1979) Genes and mRNA coding for zein polypeptides in *Zea mays*. *Eur. J. Biochem.* **102**, 211-222.
- Whittaker, P.A., Campbell, Southern, E.M., Murray, N.E. (1988). Enhanced recovery and restriction mapping of DNA fragments cloned in a new λ vector. *Nucl. Acids Res.* **16**, 6725-6736.
- Willians, M.E., Foster, R., Chua, N.H. (1992). Sequences flanking the hexameric G-box core CACGTG affect the specificity of protein binding. *Plant Cell* **4**, 485-496.
- Yamamoto, K.R., Alberts, B.M., Benzinger, R., Lawhorne, L., Treiber, G. (1970). Rapid bacteriophage sedimentation in the presence of polyethylene glycol and its application to large-scale virus purification. *Virology* **40**, 734.
- Yunes, J.A., Cord Neto, G., Silva, J.M., Leite, A., Ottoboni, L.M.M., Arruda, P. (1994). The transcriptional activator Opaque2 recognize two different target sequences in the 22-kD-Like α -prolamin genes. *Plant Cell* **6**, 237-249.

APÊNDICE I

Soluções

Acetato de Amônio 2 M

Acetato de Amônio	30,8 g
H ₂ O q.s.p.	200 ml

-Ajustar para pH 4,5 ou 5,0 com ácido acético.

-Esterilizar por filtração.

Acetato de Potássio 3 M

Acetato de Potássio	58,9 g
H ₂ O q.s.p.	200 ml

-Ajustar para pH 4,8 com ácido acético

Acetato de Sódio 3 M

Acetato de Sódio (anidro)	49,2 g
H ₂ O q.s.p.	200 ml

-Ajustar para pH 5,2 ou 7,0 com ácido acético

Binding Buffer (BB)

50 mM Tris-HCl pH7,8, contendo
150 mM de NaCl, 1 mM de EDTA, e
1 mM de DTT

Tris-HCl 1 M pH 7,8	0,5 ml
NaCl 5 M	0,3 ml
EDTA 0,5 M	20 µl
DTT 1 M	10 µl
H ₂ O	9,17 ml

BLOTTO 1X

Leite em pó desnatado (Molico)	10 g
H ₂ O q.s.p.	200 ml

Brometo de Etídio 10 mg/ml

Brometo de Etídio	0,2 g
H ₂ O q.s.p.	20 ml

Cloreto de Lítio 8 M

Cloreto de Lítio	13,57 g
H ₂ O q.s.p.	40 ml

-Esterilizar em filtro de 0,2 µm

EDTA 0,5 M

EDTA (dissódico, 2·H ₂ O)	37,2 g
H ₂ O q.s.p.	200 ml

-Ajustar para pH 8,0 com NaOH

Formamida Deonizada

Formamida	200ml
Resina AG 501 - X8	5g

-Agitar lentamente por 1 hora.

-Filtrar em papel Whatman 1MM.

-Armazenar a -70°C

Mistura de Ligase

Tampão de ligase 5X	200 µl
T4 DNA Ligase	5 U
H ₂ O	800 µl

Mistura de Nucleotídeos (0,4 mM)

dATP 100 mM pH 7,5	4 µl
dTTP 100 mM pH 7,5	4 µl
dGTP 100 mM pH 7,5	4 µl
dCTP 100 mM pH 7,5	4 µl
H ₂ O q.s.p.	1,0 ml

Mistura de S1 Nuclease

Tampão da S1 (10X)	27 µl
S1 Nuclease (1 U/µl)	60 µl
H ₂ O	183 µl

MOPS/EDTA 10X

O,2 M MOPS pH 7,0, contendo 50 mM de acetato de sódio, 10 mM de EDTA

MOPS	10,46 g
Acetato de Sódio	1,02 g
EDTA	0,93 g
H ₂ O q.s.p.	250 ml

-Ajustar para pH 7,0 com NaOH

PMSF 100 mM

Fluoreto de Fenilmetilsufonil	17 mg
etanol gelado	1 ml

Poliacrilamida 40%

Acrilamida	38 g
N,N'-metileno diacrilamida	2 g
H ₂ O q.s.p.	100 ml

-Aquecer a 40°C, em banho-maria.

-Adicionar 1 g de resina AG 501 -XB

-Agitar por 1 hora.

-Filtrar em papel Whatman 1MM.

SDS 10% (m/v)

Dodecil Sulfato de Sódio	20 g
H ₂ O q.s.p.	200 ml

SM

50 mM Tris-HCl pH 7,5, contendo 100 mM de NaCl, e 10 mM de MgSO₄

NaCl	1,16 g
MgSO ₄ 6·H ₂ O	0,46 g
Tris-HCl 1 M pH 7,5	10 ml
H ₂ O q.s.p.	200 ml

-Esterilizar por autoclavagem.

SSC 20X

0,3 M de citrato de sódio, 3 M de NaCl

NaCl	175,3 g
Citrato de sódio 2·H ₂ O	88,2 g
H ₂ O q.s.p.	800 ml
Ajustar para pH 7,0 com NaOH	
H ₂ O q.s.p.	1000 ml

-Esterilizar por autoclavagem

Tampão de ligase 5X

0,25 M Tris-HCl pH 7,6, contendo 50 mM de MgCl₂, 5 mM de ATP, 5 mM de DTT, 25 % (m/v) de polietileno glicol-8.000 (PEG)

Tris-HCl 1M pH 7,6	250 µl
MgCl ₂ 1M	50 µl
ATP 100 mM	50 µl
DTT 1M	5 µl
PEG 50%	500 µl
H ₂ O q.s.p.	1 ml

TAE 50X

2 M Tris-base, 50 mM de EDTA

Tris-Base	242,2 g
EDTA dissódico	18,6 g
H ₂ O q.s.p.	600 ml
Ácido Acético glacial	51 ml

-Após a dissolução, acertar pH7,4 com NaOH e completar o volume para 1000 ml

TBE 5X

1 M Tris-Base, 0,5 M de ácido bórico, 10 mM de EDTA

Ácido bórico	30,9 g
Tris-Base	21,1 g
EDTA dissódico	3,73 g
H ₂ O q.s.p.	600 ml

-Após a dissolução, completar o volume para 1000 ml

TE

10 mM Tris-HCl pH 7,5, 1 mM de EDTA

Tris-HCl 1M pH 7,5	100 µl
EDTA 0,5 M pH 8,0	20 µl
H ₂ O q.s.p.	10 ml

TEP

100 mM Tris-HCl pH 7,4, contendo 5 mM de EDTA, e 1 mM de PMSF

Tris-HCl 1 M pH 7,4	1 ml
EDTA 0,5 M	200 µl
PMSF 100 mM	100 µl
H ₂ O	8,7 ml

Tris-HCl 1 M

Tris-Base	121,1 g
H ₂ O	800 ml

-Ajustar o pH com HCl concentrado

pH	HCl
7,4	~70 ml
8,0	~42 ml
9,0	~20 ml
H ₂ O q.s.p	1000 ml

Soluções para RNA

Toda a vidraria foi incubada a 180°C por 24 horas.

Todas as soluções estoques foram preparadas conforme descrito acima.

Às seguintes soluções foram adicionados 0,1 volumes de dietil-pirocarbonato (DEPC), incubando-as a 37°C durante a noite e autoclavando-as: acetato de sódio 3 M pH 5,2; EDTA 0,5 M pH 8,0; SSC 20X; BLOTTO 1X; água.

As seguintes soluções foram preparada utilizando-se água, previamente tratada com DEPC: Tris-HCl pH 9,0; SDS 10% (m/v); cloreto de lítio 8 M; brometo de etídio 10 mg/ml; MOPS/EDTA 10X pH 7,0.

Tampão de Desnaturação de RNA

Formamida deionizada	250 µl
Formaldeído 37%	80 µl
MOPS/EDTA 10X	50 µl
Brometo de Etídio 10 mg/ml	10 µl

Tampão de Amostra de RNA

Glicerol	3 ml
Tris-HCl 1 M pH 8.0	0,5 ml
H ₂ O	6,5 ml
Azul de Bromofenol	
10% (m/v)	50 µl

Meios de cultura

LB (Meio Luria-Bertani)

Bacto-tryptone	10 g
Bacto-yeast extract	5 g
NaCl	10 g
H ₂ O q.s.p.	1000 ml

-Ajustar o pH para 7,0 com NaOH 2 M e esterilizar por autoclavagem por 20 minutos

2X YT

NaCl	10 g
Bacto-yeast extract	10 g
Bacto-tryptone	16 g
H ₂ O q.s.p.	1000 ml

-Ajustar para pH 7.5 com NaOH

Top agarose

agarose	1,4 g
MgSO ₄ ·6·H ₂ O	0,457 g
H ₂ O q.s.p.	200 ml

-Esterilizar por autoclavagem

APÊNDICE II

XL1-Blue

Genótipo relevante:

supE44 endA1 thi-1 lac⁻Δ(lac-proAB) gyrA96 hsdR17 recA1 F' proAB⁺ lacI^q lacZΔM15 Tn10(tet^r)

A linhagem *lac⁻*, supressora de mutação âmbar (*supE44*), defectiva em recombinação (*recA1*), deficiente em restrição e proficiente em metilação no sistema Eco K (*hsdR17*), foi utilizada como hospedeira do bacteriófago λZAP 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 *lac Z* deletado na região N-terminal (*lacZΔM15*), o que permite a α-complementação do segmento do gene presente em determinados plasmídios como pBluescript. O

epissomo F' contém também o gene para super produção de repressor do operon *lac* (*lac I^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).

DL538

Genótipo relevante:

hsdR mcrA mcrB recD sbcC

Esta linhagem foi utilizada como hospedeira do bacteriófago λ DASH durante a seleção de clones recombinantes da biblioteca genômica de *Coix*. A mutação do gene *recD* que elimina a atividade de 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^mCGC-3' e 5'-AG^mCT-3', o que favorece a construção de bibliotecas genômicas de plantas, cujo DNA é rico em citosinas metiladas (Whittaker et al., 1988).

DH5 α

Genótipo relevante:

F- supE44 Δ (lacZYA - argF) U169
(ϕ 80 lacZ Δ M15) hsdR17 recA1
endA1 gyrA96 thi-1 relA1.

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* (*lac^r*) possibilita a realização do teste de complementação na ausência de IPTG (Hanahan, 1983). Os demais marcadores foram descritos anteriormente para XL1-Blue.

APÊNDICE III

Vetor λ Zap

A grande vantagem na utilização do vetor λ ZAP, consiste na possibilidade da excisão de um "phagemid" tipo pBluescript, 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"), presentes 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 sequência 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 sequência presente entre os sítios I e T (o vetor λ ZAP II inclui a sequência 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 inativadas incubando-se a 70°C e, após 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 semeadas em meio contendo ampicilina. Colônias resistentes à ampicilina, que crescem lentamente (colônias pequenas), 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 a semeadura. A infecção simultânea com fago "helper" R408, o qual apresenta o produto do *gene II* termolábil,

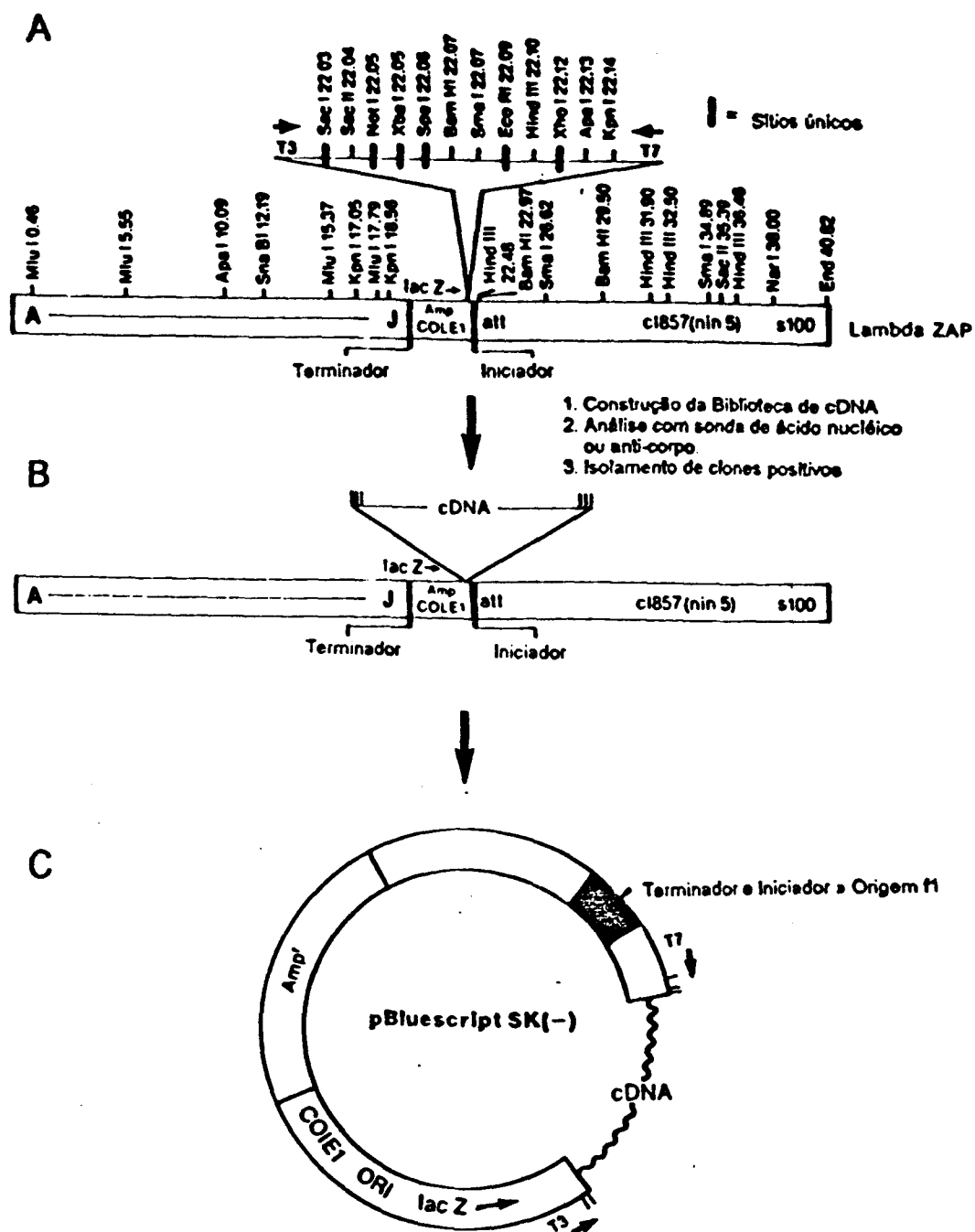


Figura 17: Diagrama esquemático mostrando os mapas do vetor λ ZAP original (A), após inserção do cDNA (B), e após a excisão do "phagemid" pBluescript (C).

permite a obtenção de transformantes livres deste fago através da incubação a 42°C.

As colônias isoladas podem ser utilizadas na obtenção de DNA para análise do inserto através de digestão com enzimas de restrição e eletroforese em geis de agarose (Fig. 18).

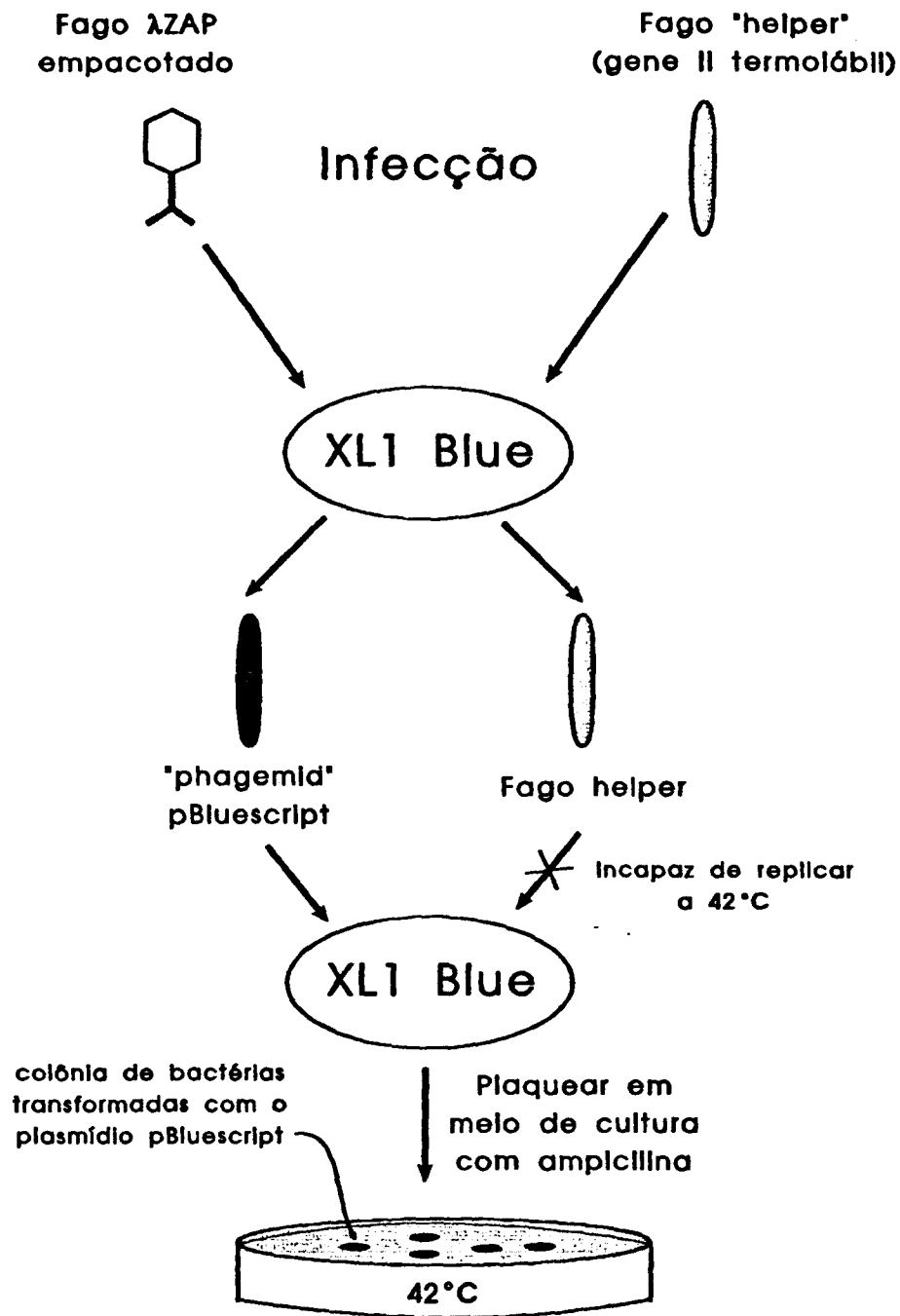


Figura 18: Esquema do processo de excisão do "phagemid" pBluescript do bacteriófago λZAP.

Vetor λ DASH

O vetor λ DASH é um vetor de substituição, onde o fragmento de 14 kB, contendo os genes das funções *red* e *gam*, pode ser substituído por fragmentos de 9 a 23 kB. Os bacteriófagos recombinantes, assim obtidos, apresentam o genótipo *red⁻* e *gam⁻*, permitindo sua seleção em bactérias *Rec⁻*, lisogênicas para o bacteriófago P2 (fenótipo: *spi*).

Por permitir a inserção de fragmentos de DNA de 9 a 23 kB,

este vetor é utilizado na construção de bibliotecas genômicas.

O vetor possui um sítio de policlonagem que possibilita a inserção de fragmentos provenientes de digestão com várias enzimas de restrição. A região de clonagem desse vetor é flanqueada pelos promotores de bacteriófagos T3 e T7, permitindo a transcrição *in vitro* do fragmento de DNA clonado (Fig. 19).

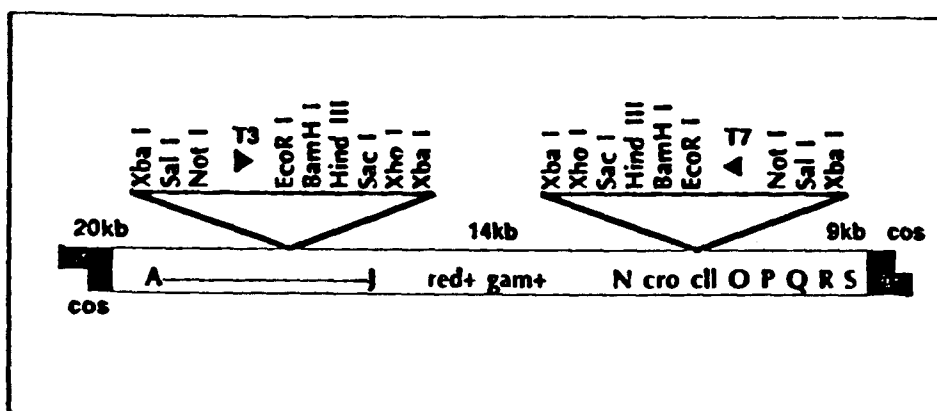


Figura 19: Mapa do bacteriófago λ DASH.

APÊNDICE IV

Vetor pBluescript KS⁺/⁻

O vetor pBluescript KS⁺/⁻ é um "phagemid" que contém a região intergênica do fago filamentosso do tipo f1, podendo ser secretado com um DNA fita simples na presença de um fago "helper" do tipo f1 (Stratagene Cloning System). Esta região intergênica codifica para as funções *cis* requeridas para empacotamento e replicação do fago. Existem "phagemids" pBluescript contendo uma das duas possíveis orientações de região intergênica, sendo denominados pBluescript KS⁺ e pBluescript KS⁻. O pBluescript KS⁺ replica-se de forma que a fita codificadora do gene β -galactosidase seja secreta-da na partícula do fago, enquanto que o pBluescript KS⁻ secreta a fita não codificante do mesmo gene. O vetor pBluescript KS⁺/⁻ possui 2.961 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 sequencia-

mento de insertos clonados em pUC.

A denominação KS refere-se à orientação da região de policlonação, com relação às enzimas de restrição localizadas nas extremidades (Kpn I $\rightarrow$ Sac I). 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).

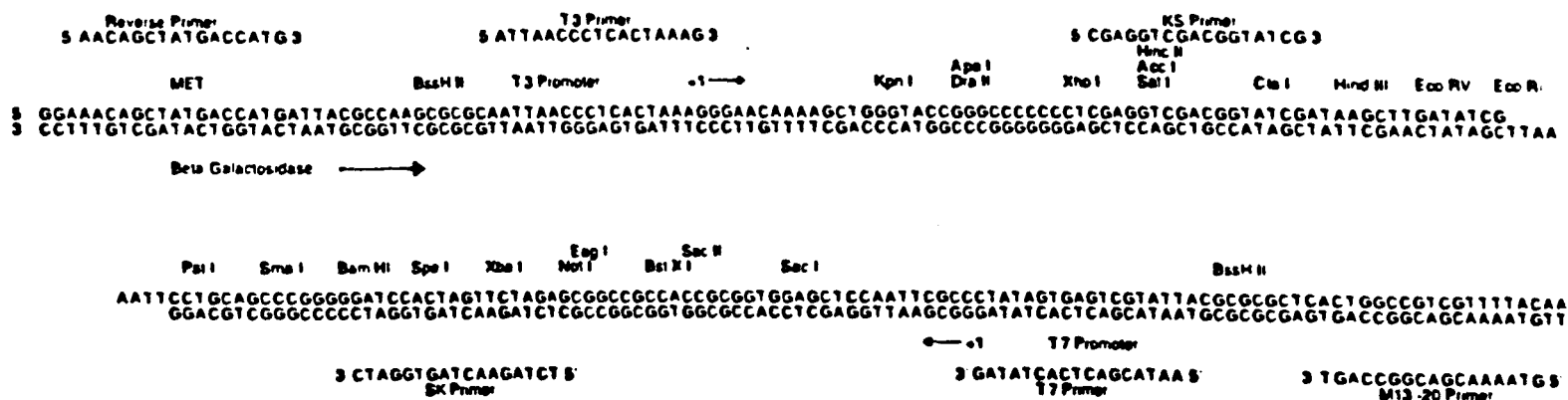
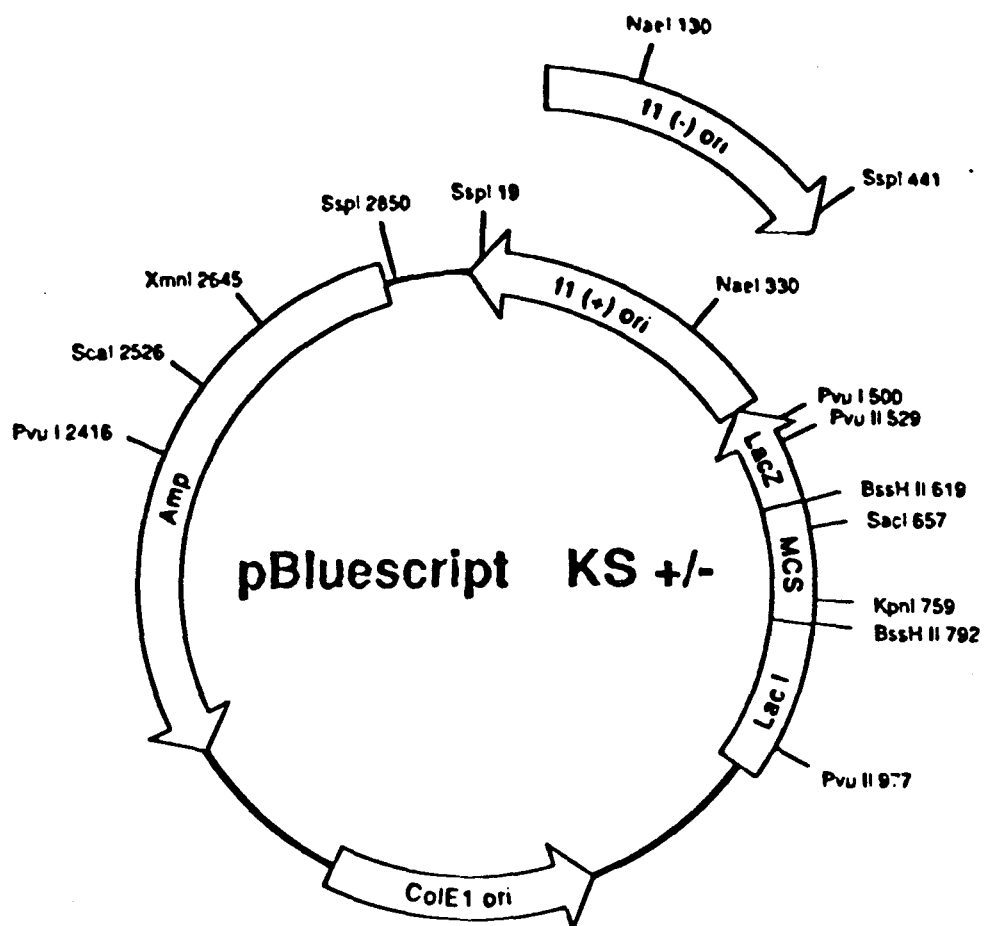


Figura 20: Mapa do plasmídeo pBluescript KS[±].

pUR

Os vetores da série pUR (278, 288 e 289) permitem a produção de uma proteína quimérica, resultante da fusão da proteína de interesse na região carboxi-terminal da β -Galactosidase (Fig. 21). Estes vetores possuem sítios de clonagem, Bam HI, Sal I, Pst I, Xba I, Hind III e Cla I, nas três fases de leitura possíveis.

Inserções de sequências de cDNA nos sítios de poli-clonagem apropriados permitem a produção de proteínas de fusão entre a β -Gal e o peptídeo codificado pelo cDNA.

Estes plasmídios contêm o gene *amp^r* que confere resistência ao antibiótico ampicilina.

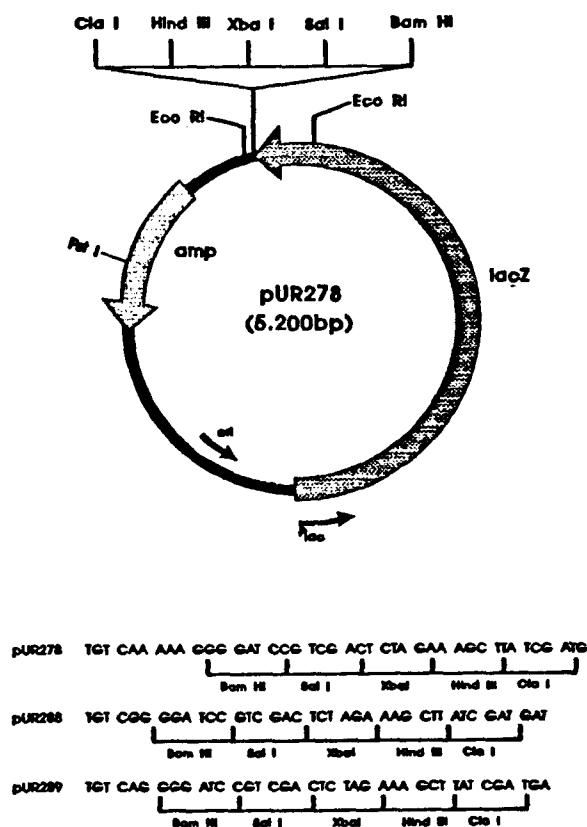


Figura 21: Mapa dos vetores pUR 278, pUR 288 e pUR 289.

APÊNDICE V

Deleção Unidirecional com Exonuclease III

O método é baseado na digestão com Exonuclease III do fragmento de DNA a partir de uma de suas extremidades.

A Exonuclease III ataca apenas as extremidades protuberantes 5' e abruptas, e não, as protuberantes 3'. Assim sendo, o fragmento de interesse deve ser clonado em um sítio de policlonação, de forma que, permaneçam, pelo menos, dois sítios únicos entre o inserto e o sítio de ligação do "primer" de sequenciamento. A enzima de restrição que corta próximo ao inserto de interesse deve deixar uma extremidade com protuberância 5' ou abrupta, a partir da qual serão realizadas as deleções. A

digestão com a outra enzima, a qual possui seu sítio no "polylinker" do plasmídeo posicionado após o sítio da primeira enzima, deve deixar uma extremidade protuberante 3' de, no mínimo, 4 bases. Por vezes, o ataque da Exonuclease III na extremidade 5', cria extremidades protuberantes 3', esta fita remanescente de DNA é removida pela ação da S1 nuclease, na presença de cátions Zn^{2+} , em pH ácido. Após inativação térmica da S1 nuclease, adiciona-se polimerase Klenow para reparar as extremidades. Finalmente os fragmentos são recircularizados através da ligação de suas extremidades (Fig. 22).

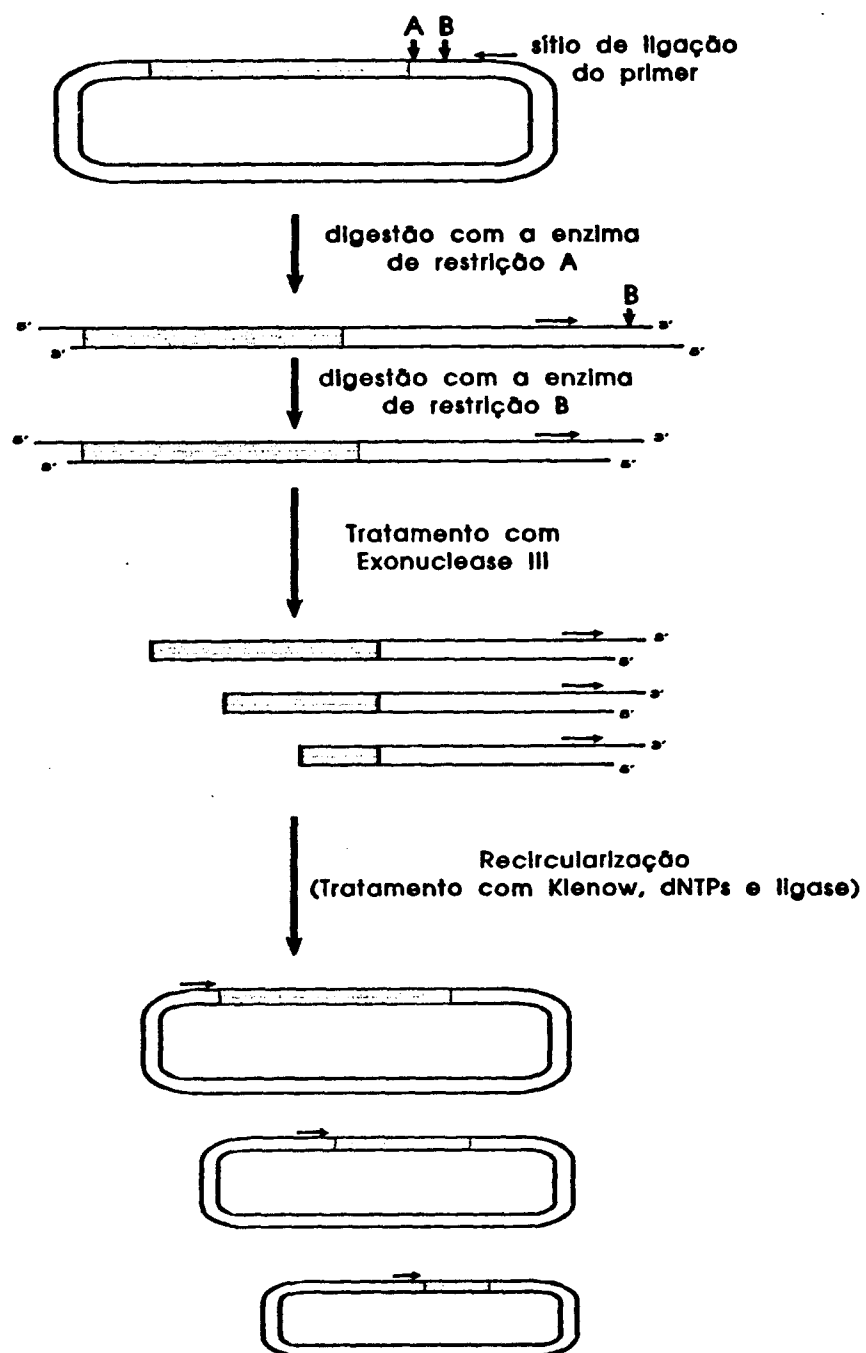


Figura 22: Diagrama esquemático do método de obtenção de deleções unidirecionais com Exonuclease III.

Ligação Seletiva

Conforme mostra a Figura 23, esta técnica permite a seleção de fragmentos de DNA, que contenham o sítio de ligação de uma determinada proteína de interesse. Extrato de bactéria, expressando a proteína de interesse fusionada na região carboxi-terminal da β -galactosidase, é incubado com anti-corpo monoclonal contra β -galactosidase. O conjugado obtido é então incubado com a Proteína A-Sepharose, obtendo-se o imuno-complexo Proteína A-Sepharose::anti- β -gal:: β -gal-O2.

O imuno-complexo é, então, incubado com fragmentos de DNA

que foram marcados radiativamente nas extremidades. Após centrifugação, e posterior lavagem do sedimento obtido para remoção dos fragmentos de DNA não selecionados, os imuno-complexos são denaturados com fenol, liberando o fragmento retido.

Os fragmentos selecionados são então analisados através de eletroforese em gel de agarose. Como controle de tamanho de fragmento, utiliza-se fragmentos marcados, não submetidos a ligação seletiva (Fig.23. Após a secagem, o gel é submetido a autoradiografia.

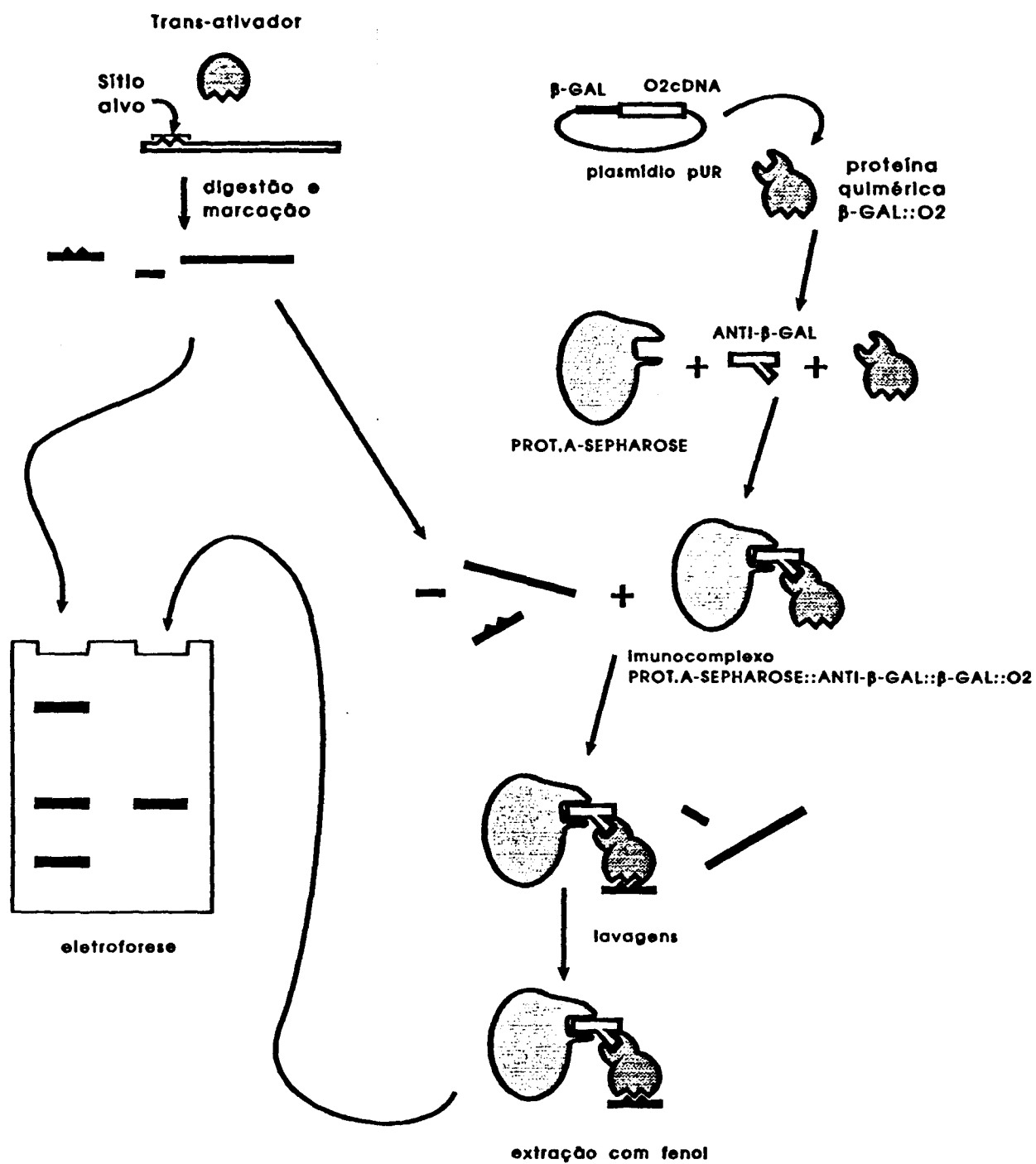


Figura 23: Diagrama esquemático do método de Ligação Seletiva.

APÊNDICE VI

Tabela VI: Relação dos plasmídios utilizados neste trabalho.

Plasmídio	Vetor	Inserto	Referências
P α -Cx103	pRT103	promotor de α -coixina inteiro	Yunes <i>et al.</i> , 1994
P α -Z4103	pRT103	promotor de α -zeína inteiro	Yunes <i>et al.</i> , 1994
P285 α -C	pBluescript KS ⁺	promotor de α -coixina (O2b, O2c)	Yunes <i>et al.</i> , 1994
P285 Δ O2	pBluescript KS ⁺	promotor de α -coixina (O2c)	Yunes <i>et al.</i> , 1994
P285 CATG	pBluescript KS ⁺	promotor de α -coixina (O2c mutado)	Yunes, J.A. (não publicado)
P285 TGAC	pBluescript KS ⁺	promotor de α -coixina (O2c mutado)	Yunes, J.A. (não publicado)
pO2CX	pUR289	o2 Coix + β -Gal	Souza Filho, G.A. (não publicado)
LC#37	pUR288	o2 milho + β -Gal	Schmidt <i>et al.</i> , 1992
pBCX 25.10	pBluescript KS ⁺	porção estrutural do gene de α -coixina	Ottoboni <i>et al.</i> , 1993
pKpn 1.9	pBluescript KS ⁺	porção estrutural do gene de actina de <i>T. reesei</i>	El-Dorry, H. (não publicado)
pCO2-1	pBluescript SK ⁺	cDNA o2 Coix	Este Trabalho