MAYRA AKEMI KUROKI

**Desenvolvimento de uma estratégia de
clonagem customizada de regiões
promotoras do genoma da cana-de-açúcar**

Dissertação apresentada ao
Programa de Pós-Graduação
Interunidades em Biotecnologia
USP/Instituto Butantan/IPT, para
obtenção do Título de Mestre em
Biotecnologia.

São Paulo
2012

MAYRA AKEMI KUROKI

**Desenvolvimento de uma estratégia de
clonagem customizada de regiões
promotoras do genoma da cana-de-açúcar**

Dissertação apresentada ao
Programa de Pós-Graduação
Interunidades em Biotecnologia
USP/Instituto Butantan/IPT, para
obtenção do Título de Mestre em
Biotecnologia.

Área de concentração: Biotecnologia

Orientador: Prof^a. Dra. Marie-Anne
Van Sluys

Versão original

São Paulo
2012

DADOS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
Serviço de Biblioteca e Informação Biomédica do
Instituto de Ciências Biomédicas da Universidade de São Paulo

reprodução não autorizada pelo autor

Kuroki, Mayra Akemi.

Desenvolvimento de uma estratégia de clonagem customizada de regiões promotoras do genoma da cana-de-açúcar / Mayra Akemi Kuroki. -- São Paulo, 2012.

Orientador: Profa. Dra. Marie-Anne Van Sluys.

Dissertação (Mestrado) – Universidade de São Paulo. Instituto de Ciências Biomédicas. Programa de Pós-Graduação Interunidades em Biotecnologia USP/IPT/Instituto Butantan. Área de concentração: Biotecnologia. Linha de pesquisa: Biologia Molecular de Plantas.

Versão do título para o inglês: Customized promoter cloning strategy.

1. Cana-de-açúcar 2. Protoplastos 3. Monocotiledôneas I. Van Sluys, Profa. Dra. Marie-Anne II. Universidade de São Paulo. Instituto de Ciências Biomédicas. Programa de Pós-Graduação Interunidades em Biotecnologia USP/IPT/Instituto Butantan III. Título.

ICB/SBIB0194/2012

UNIVERSIDADE DE SÃO PAULO
Programa de Pós-Graduação Interunidades em Biotecnologia
Universidade de São Paulo, Instituto Butantan, Instituto de Pesquisas Tecnológicas

Candidato(a): Mayra Akemi Kuroki.

Título da Dissertação: Desenvolvimento de uma estratégia de clonagem customizada
de regiões promotoras do genoma da cana-de-açúcar.

Orientador(a): Profa. Dra. Marie-Anne Van Sluys.

A Comissão Julgadora dos trabalhos de Defesa da **Dissertação de Mestrado**,
em sessão pública realizada a/...../....., considerou

☐ **Aprovado(a)**

☐ **Reprovado(a)**

Examinador(a): Assinatura:
Nome:
Instituição:

Examinador(a): Assinatura:
Nome:
Instituição:

Presidente: Assinatura:
Nome:
Instituição:

**Aos meus queridos pais.
Dedico**

AGRADECIMENTOS

À profa. Dra. Marie-Anne Van Sluys pela orientação, confiança e apoio ao longo de todos esses anos. O amadurecimento e crescimento profissional que tive durante este tempo se deve em grande parte pela sua postura firme nos momentos necessários e o devido reconhecimento quando merecido.

A todos os que permanecem ou passaram pelo GaTE lab: Alessandro, Andres, Andréia, Bruno, Cushla, Dani K., Dani Q., Douglas, Dudu, Edgar, Érika, Hana, Gabriel, Guilherme, Jonas, Juliana N., Juliane, Kléber, Léo, Luciene, Marcelo, Marisa, Miguel, Nathalia, Tati, Paula, Robson, Rosário e Verônica.

A todos do laboratório de Genômica e Biologia Molecular do Centro de Biotecnologia Agrícola Cebetec na ESALQ (USP), em especial à profª Dra. Helaine Carrer por possibilitar a realização e aprendizado da transformação permanente de cana-de-açúcar. Ao André Luis Barbosa e Daniela Silva por todo apoio e paciência no manuseio do material *in vitro*.

A todos do Núcleo de Genômica Funcional de Plantas, em especial à profª Dra. Marcia Margis e Fernanda Lazzarotto, ambas da UFRGS (Universidade Federal do Rio Grande do Sul) pela excelente colaboração realizada na geração de plantas transgênicas de arroz e aprendizado no isolamento de protoplastos.

Ao André e Amanda do laboratório de Biologia Celular de Plantas (BIOCEL) IB/USP pela grande disposição em ajudar no manuseio da microscopia de fluorescência.

À Juliana Nico que me ensinou os “primeiros passos” no cultivo de bactérias e sistema Gateway, além do excelente apoio na bancada foi a responsável por dar o “pontapé” inicial do presente projeto.

À Tati por ser para mim um grande exemplo profissional dentro e fora do lab, pelo excelente apoio técnico e amizade sincera. Por manter o ótimo funcionamento do lab assim como manter nossas “barrigas cheias” pelos bolos deliciosos e cafés da manhã recheados de risada.

Ao Bruno, Gui, Doug e Andres pela ótima disposição em participar de discussões acaloradas do cotidiano do lab, seja sobre a análise de um gel de agarose duvidoso ou sobre as situações que beiram o cômico/trágico. Pela excelente postura em ajudar, criticar e colaborar não só no meu projeto como tantos outros. Tudo com muita descontração!

Ao Jonas por ser peça fundamental em manter o funcionamento da bioinfo, assim como ser extremamente prestativo e paciente a todos que têm um conhecimento limitado de informática.

À Érika e Hana que deram a maior força na elaboração do documento da quali, com sugestões pertinentes e questionamentos sensatos ajudaram a amadurecer algumas idéias cruas.

Ao Gui pela grande ajuda na análise do sequenciamento das placas e dos clones isolados, assim como pelas idéias e sugestões criativas para o desenvolvimento da clonagem.

À Dani Q. pelo apoio na manutenção e funcionamento da casa de vegetação e sala de plantas.

À Nathalia pelas críticas e sugestões oportunas às quais ajudaram a amadurecer o projeto, assim como o grande apoio na elaboração deste documento.

À Cushla pelo fornecimento de DNA genômico de excelente qualidade assim como a disposição em ajudar a resolver problemas na bancada.

Ao Homellhan pela manutenção das plantas na casa de vegetação e por manter o funcionamento adequado das instalações do jardim.

À todos os amigos queridos que me acompanharam nesta jornada, seja através das conversas descontraídas, festas, baladas, jogos de vôlei inesquecíveis ou até os que estão longe, via email, facebook. É a prova de que amizades verdadeiras superam o tempo e o espaço.

Ao Márcio Shibao por me apoiar em todas as decisões que tomei, por ser meu melhor amigo, meu companheiro de vida. Pelo carinho e amor que construímos juntos e por ser a minha força nos momentos mais difíceis.

Aos meus queridos pais, que mesmo sem conhecer os caminhos da área acadêmica sempre me apoiaram de forma incondicional. À minha batian por ser um exemplo de vida e fé presente no meu cotidiano.

A toda família Kuroki e Taira pelo exemplo incomum de serem famílias tão unidas nos dias de hoje, assim como no velho ditado “a união faz a força”. Pelo exemplo de garra e determinação, assim como por me lembrar da minha origem e dos meus antepassados.

Este trabalho foi realizado com o apoio financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP).

**“Pense profundamente.
Fale gentilmente.
Ame bastante.
Ria frequentemente.
Trabalhe com afinco.
Dê com generosidade.
Pague pontualmente.
Ore fervorosamente.
E seja bom.”**

Elmer Wheeler

RESUMO

KUROKI, M. A. **Desenvolvimento de uma estratégia de clonagem customizada de regiões promotoras do genoma da cana-de-açúcar.** 2012. 137 f. Dissertação (Mestrado em Biotecnologia) - Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2012.

O controle da expressão gênica é operado em diferentes níveis. A nível transcricional este controle é determinado pela interação de fatores de transcrição com sequências *cis*-regulatórias que flanqueiam as regiões a serem transcritas. Esta interação estabelece a especificidade espaço-temporal da expressão. A obtenção de OGMs (organismos geneticamente modificados) baseia-se na inserção de um ou mais genes de interesse sob o controle de regiões promotoras, de modo a direcionar a expressão de tais genes. Portanto, a identificação de promotores constitutivos, bem como a de promotores específicos a determinados tecidos ou fases do desenvolvimento do organismo, ou induzidos por fatores externos como estresses bióticos ou abióticos, são de extrema importância para o desenvolvimento de novos produtos biotecnológicos. O objetivo deste trabalho foi desenvolver uma metodologia para identificação de regiões promotoras funcionais a partir de segmentos de um genoma qualquer. A inovação desta abordagem encontra-se na busca, e identificação de regiões com potencial de promover a expressão gênica, sem a necessidade de conhecimento prévio de um gene específico, além de possibilitar o isolamento de promotores não canônicos. O genoma da cana-de-açúcar foi escolhido como modelo para o desenvolvimento desta estratégia. Este híbrido é de grande interesse econômico e biotecnológico devido a demandas nos setores alimentício e energético. No entanto, oferece desafio considerável na identificação de promotores funcionais, pois o alto nível de ploidia eleva o risco de isolamento de promotores provenientes de cópias inativas do gene. Neste trabalho, apresentamos o estabelecimento desta estratégia por meio da clonagem customizada de fragmentos provenientes do genoma dos cultivares R570 e SP80-3280, a qual consistiu no fracionamento do DNA genômico e seleção de fragmentos na faixa de 2 a 4kb. Estes fragmentos foram clonados a montante de uma fusão de genes repórteres GFP:GUS, gerando bibliotecas de vetores. A triagem destas bibliotecas foi realizada através de ensaios de expressão transiente por biobalística em epitélio de cebola na qual foram analisados 142 clones distintos. Como resultado desta triagem foram isolados quatro clones que abrigam módulos regulatórios funcionais, designados como F6, B4, F7 e C5. Esses foram submetidos a novos ensaios de expressão transiente para detecção do gene GFP em outro sistema: transformação de protoplastos de arroz. Os quatro clones apresentaram expressão de GFP, confirmando a funcionalidade das regiões isoladas como promotores de expressão gênica. Subsequentemente, foi realizado um ensaio de transformação permanente em arroz através de *Agrobacterium tumefaciens* em que foram obtidas 12 plantas. Foi detectada expressão do marcador GUS em calos, folhas e raízes das plantas de arroz as quais abrigam construções sob regulação dos promotores F6, B4 e F7, comprovando sua funcionalidade. Desta maneira, o desenvolvimento do presente trabalho permitiu estabelecer uma metodologia de recuperação de sequências regulatórias funcionais a partir de um genoma complexo como o da cana-de-açúcar e possibilitou a descoberta de quatro regiões genômicas distintas que abrigam

módulos regulatórios funcionais com ampla possibilidade de serem explorados biotecnologicamente e no campo da ciência básica.

Palavras-chave: Promotores. Monocotiledôneas. Cana-de-açúcar. Transformação genética. Protoplastos.

ABSTRACT

KUROKI, M. A. **Customized promoter cloning strategy**. 2012. 137 p. Masters thesis (Biotechnology) - Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2012.

Gene expression is controlled at multiple levels. At the transcriptional level, control is determined by the interaction of transcription factors with *cis*-regulatory sequences flanking the regions to be transcribed. This interaction establishes the specificity of space-temporal expression. The basis of the construction of GMOs (Genetically Modified Organisms) is the insertion of one or more genes of interest under the control of promoter regions, which drive the expression of these genes. Therefore, the identification of constitutive promoters as well as specific, or inducible, promoters is extremely important to the development of new biotechnological products. The aim of this work is develop a strategy to identify functional promoter regions from any genome. The innovation of this approach relies on a screening method that identifies promoters without requiring previous knowledge of a specific gene, furthermore it allows the isolation of non canonical promoters. The modern sugarcane genome was chosen as a model for the development of this strategy. The modern sugarcane cultivar is a complex hybrid that is of major economic and biotechnological interest because of its importance to the agricultural and energy sectors. Notwithstanding, the sugarcane genome is a considerable challenge for identification functional promoters since the high level of ploidy increases the risk of isolating promoters from inactive copies of a gene. In this work we present the establishment of a strategy by means of customized cloning of fragments from the sugarcane cultivars R570 and SP80-3280. This strategy first involves fractionating genomic DNA and selecting fragments ranging from 2 to 4kb. Libraries were constructed by cloning these fragments upstream of a fusion of two gene reporters, GFP and GUS. The libraries were then screened by a transient expression assay using biolistic particle delivery of clones into epidermal onion cells. One hundred and forty two clones were analyzed and of these, four clones, F6, B4, F7 and C5, were isolated with putative functional regulatory modules. These clones were further analysed by another transient expression assay to detect GFP gene expression by transformation of rice protoplasts. All four clones showed GFP expression, confirming their activity in another monocot species. Three clones, F6, B4 and F7, were permanently transformed in rice, using *Agrobacterium tumefaciens*, and 12 plants were obtained. GUS expression was detected in the callus, leaves and roots of the rice plants with promoters from clones F6, B4 and F7, thus confirming the functionality of sequences in these clones. The present work has established a strategy to identify and extract functional regulatory sequences from a complex genome like sugarcane. It has also identified four genomic regions containing functional regulatory regions which show great potential of being useful in both the biotechnology field and in the field of basic science.

Key words: Promoter, monocot, sugarcane, genetic transformation, protoplast.

LISTA DE ILUSTRAÇÕES

Figura 1 - Estrutura de um cromossomo com detalhe em um segmento funcional..	25
Figura 2 - Estrutura de um promotor reconhecido pela RNA-polimerase II.....	27
Figura 3 - Estrutura de um típico promotor sintético de eucariotos o qual contém elementos <i>cis</i> -regulatórios e um promotor mínimo fusionados a um gene repórter à jusante seguido pela presença de um gene de seleção a antibiótico e uma sequência poliA.....	41
Figura 4 - Mapa do plasmídeo comercial pDONR TM 221. Plasmídeo utilizado para o primeiro passo da clonagem utilizando o sistema Gateway Invitrogen®.....	55
Figura 5 - Mapa do plasmídeo pHGWFS7. Plasmídeo utilizado para o segundo passo da clonagem utilizando o sistema Gateway Invitrogen®.....	57
Figura 6 - Esquematização do processo de obtenção de protoplastos viáveis de arroz utilizados na transformação transiente com PEG.....	63
Figura 7 - Mapa do vetor pART7 HA GFP (s65t) utilizado como controle positivo na transformação transiente de protoplastos de arroz.....	64
Figura 8 - Esquema representativo do método de transformação genética por biobalística utilizando como tecidos-alvo calos e discos de cana-de-açúcar.....	66
Figura 9 - Esquema representativo do plasmídeo pH9 o qual contém o cassete de expressão do gene <i>neo</i> que codifica a neomicina fosfotransferase (NPTII) sob controle do promotor da ubiquitina do milho (<i>Ubi-1</i>) e o terminador da nopalina sintetase T-Nos.....	68
Figura 10 - Esquema representativo do plasmídeo pFF19G o qual contém o cassete de expressão do gene repórter <i>uid-A</i> (GUS) sob o controle do promotor <i>CaMV35S</i> e o terminador 35S <i>polyA</i>	68
Figura 11 - A: Hidrólise do DNA genômico de cana-de-açúcar das cultivares R570 e SP80-3280 através da combinação de duas endonucleases de restrição: PstI/BamHI.....	77
Figura 12 - A: Análise dos fragmentos gerados na primeira reação de PCR os quais contém a parte inicial da região de recombinação, essencial para a reação de recombinação Gateway.....	78
Figura 13 - A: Análise dos fragmentos gerados na segunda reação de PCR contendo a região de recombinação completa.....	79

Figura 14 - Análise da integridade dos vetores pDONR221 recombinaados contendo fragmentos aleatórios de R570 (Placa 1) e SP80-3280 (Placa 2).....	81
Figura 15 - Placa 1 abriga clones de SP80-3280 gerados na primeira reação de recombinação BP.....	83
Figura 16 - Placa 2 abriga clones de R570 gerados na primeira reação de recombinação BP.....	84
Figura 17 - Representação esquemática da distribuição dos 33 clones da Placa 3 em oito mix.	88
Figura 18 - Análise histoquímica dos clones que apresentaram expressão transiente positiva 24 horas após o bombardeamento em epiderme de cebola.....	90
Figura 19 - Representação esquemática do processo de <i>primer walking</i> utilizado para o sequenciamento completo do clone F7.....	93
Figura 20 - Predição dos elementos TATA e CAAT na sequência completa do clone F6.	100
Figura 21 - Alinhamento entre o clone F6 (em rosa) e o cromossomo 1 de <i>Sorghum bicolor</i> através do <i>Phytozome</i>	102
Figura 22 - Filogenia das gramíneas analisadas quanto à presença do clone F6.....	104
Figura 23 - Alinhamento entre o clone C5 (em rosa) e o cromossomo 1 de <i>Sorghum bicolor</i> através do <i>Phytozome</i>	106
Figura 24 - Alinhamento entre o clone F7 (em rosa) e o cromossomo 9 de <i>Sorghum bicolor</i> através do <i>Phytozome</i>	107
Figura 25 - Alinhamento entre o clone B4 (em rosa) e o cromossomo 1 de <i>Sorghum bicolor</i> através do <i>Phytozome</i>	108
Figura 26 - Análise da expressão transiente em protoplastos de arroz através de visualização de GFP por microscopia de fluorescência.....	111
Figura 27 - Análise da presença do gene <i>neo</i> no possível transformante de cana-de-açúcar.....	116
Figura 28 - Expressão do gene GUS em tecidos de arroz transformados com as construções F6, B4 e F7.....	119

LISTA DE TABELAS

Tabela 1 - Condições das reações de dupla-hidrólise do DNA genômico dos cultivares SP80-3280 e R570.....	51
Tabela 2 - Condições das reações de ligação dos fragmentos de DNA digeridos das cultivares SP80-3280 e R570.....	52
Tabela 3 - Condições da primeira reação de amplificação dos fragmentos de R570 e SP880-3280.....	53
Tabela 4 - Condições da segunda reação de amplificação dos fragmentos de R570 e SP80-3280.....	54
Tabela 5 - Relação de placas de miniprep geradas pelo Sistema Gateway de Clonagem.....	85
Tabela 6 - Listagem das placas de miniprep 3 a 6 as quais contêm os clones de pHGWFS7 recombinados com fragmentos provenientes das placas de miniprep de clones de pDONR221 recombinados. A eficiência da reação de recombinação LR das placas 3 a 6 é dada em porcentagem.....	86
Tabela 7 - Relação das placas de miniprep submetidas a bombardeio em epitélio de cebola e quantidade de mix bombardeados por placa seguido pelo número de clones identificados com expressão transiente positiva.....	88
Tabela 8 - Estimativa do tamanho dos fragmentos liberados pela hidrólise enzimática com SacII dos clones identificados por apresentar expressão transiente positiva em epitélio de cebola.	92
Tabela 9 - Relação de oligonucleotídeos utilizados para sequenciar por completo os clones F7, C5 e B4.....	94
Tabela 10 - Proteínas homólogas à proteína Sb01g048200.1 de <i>Sorghum bicolor</i>	103

LISTA DE ABREVIATURAS E SIGLAS

DNA – ácido desoxirribonucleico

dNTP – deoxirribonucleotídeos trifosfato

EST – do inglês *expressed sequence tag*

pb – pares de base

Kb – kilo pares de base

μl, μg – microlitro, micrograma

ng - nanograma

ORF – do inglês *open reading frame*

PCR – do inglês *polymerase chain reaction*

RNA – ácido ribonucléico

mRNA – RNA mensageiro

SDS – dodecil sulfato de sódio

TBE – Tris-borato-EDTA

NCBI – National Center for Biotechnology Information

SUMÁRIO

1	INTRODUÇÃO	20
2	REVISÃO BIBLIOGRÁFICA	22
2.1	O que é um gene? Uma visão pós-projeto ENCODE	22
2.2	Regulação da transcrição gênica em eucariontes	23
2.2.1	Elementos <i>cis</i> -regulatórios	26
2.2.2	Elementos <i>trans</i> -regulatórios	28
2.3	Importância da identificação de regiões promotoras em plantas	30
2.4	Promotores constitutivos, modificados, regulados e sintéticos	33
2.4.1	Promotores constitutivos	33
2.4.2	Promotores constitutivos modificados	35
2.4.3	Promotores regulados	36
2.4.3.1	<i>Promotores tecido-específicos</i>	36
2.4.3.2	<i>Promotores induzidos</i>	38
2.4.4	Promotores sintéticos	40
2.5	Métodos de identificação e isolamento de regiões promotoras em plantas	43
2.6	Identificação de regiões promotoras em cana-de-açúcar	46
3	OBJETIVOS	49
3.1	Objetivo geral	49
3.2	Objetivos específicos	49
4	MATERIAL E MÉTODOS	50
4.1	Processamento do genoma para a produção de fragmentos de DNA compatíveis ao sistema de clonagem	50
4.1.1	Extração do DNA genômico e hidrólise com endonucleases de restrição	50
4.1.2	Excisão e purificação de fragmentos de DNA de tamanho desejado e ligação de adaptadores	51
4.1.3	Primeira amplificação dos fragmentos	52
4.1.4	Segunda amplificação dos fragmentos	53
4.2	Geração de bibliotecas de vetores de expressão: associação dos fragmentos de DNA preparados na seção 4.1 a genes repórteres GUS e GFP	54
4.2.1	Reações de Recombinação BP e LR do Sistema Gateway	54
4.2.1.1	<i>Reação de recombinação BP</i>	55
4.2.1.2	<i>Transformação da reação de recombinação BP em E.coli seguido por miniprep de placa 96 poços e confirmação da obtenção dos clones recombinantes</i>	55
4.2.1.3	<i>Reação de recombinação LR</i>	56

4.2.1.4 Transformação de <i>E.coli</i> com as reações de recombinação LR seguido por extração de DNA plasmidial em placas de 96 poços e sequenciamento.....	57
4.3 Busca por regiões regulatórias funcionais através da triagem da biblioteca de vetores por biobalística e análise da expressão do gene repórter GUS	58
4.3.1 Ensaios de expressão transiente em epitélio de cebola	58
4.3.2 Ensaio histoquímico	59
4.4 Identificação e análise das regiões regulatórias funcionais	59
4.4.1 Hidrólise enzimática dos clones identificados na análise de expressão transiente	59
4.4.2 Sequenciamento completo dos clones com expressão do gene GUS	60
4.4.3 Análise <i>in silico</i> das putativas regiões promotoras identificadas.....	60
4.4.4 Extração de DNA plasmidial em pequena escala (miniprep) dos clones com expressão transiente positiva	61
4.4.5 Ensaios de expressão transiente em arroz	62
4.4.5.1 Obtenção de protoplasto de arroz (<i>Oryza sativa</i>).....	62
4.4.5.2 Transformação transiente de protoplastos de arroz por PEG.....	63
4.4.5.3 Análise da expressão transiente através da visualização da expressão de GFP.....	64
4.5 Validação da funcionalidade dos promotores através de transformação permanente em sistema nativo	65
4.5.1 Ensaios de expressão permanente em cana-de-açúcar.....	65
4.5.1.1 Transformação genética de calos e discos de cana-de-açúcar via biobalística.....	65
4.5.1.2 Co-transformação com o vetor pHA9.....	67
4.5.1.3 Seleção, regeneração e obtenção de plantas transgênicas.....	69
4.5.1.4 Extração de DNA total e análises de PCR.....	69
4.6 Material vegetal, soluções estoque, meios de cultura e procedimentos de biologia molecular	70
4.6.1 Material vegetal	70
4.6.2 Soluções estoque	70
4.6.3 Antibióticos	71
4.6.4 Soluções para isolamento de protoplastos.....	72
4.6.5 Meios de cultura.....	72
4.6.5.1 Meio para crescimento de <i>Escherichia coli</i>	72
4.6.5.2 Meio de cultura para germinação de sementes de <i>Oryza sativa</i> (arroz).....	72
4.6.5.3 Meio osmótico para calos de cana-de-açúcar.....	73
4.6.5.4 Meio para seleção de calos e discos transformantes de cana-de-açúcar.....	73
4.6.6 Bactérias eletrocompetentes	73
4.6.7 Transformação de bactérias eletrocompetentes	74
4.6.8 Extração de DNA plasmidial em placa de 96 poços	74
5 RESULTADOS E DISCUSSÃO	76

5.1 Processamento do genoma para a produção de fragmentos de DNA compatíveis ao sistema de clonagem	76
5.2 Geração de bibliotecas de vetores de expressão: associação dos fragmentos de DNA preparados no item 5.1 a genes repórteres GUS e GFP	80
5.2.1 Clonagem dos fragmentos de DNA em vetores intermediários pDONR221 através da reação de recombinação BP.....	80
5.2.2 Clonagem dos fragmentos de DNA genômico em vetores de expressão os quais abrigam os genes repórteres GUS e GFP	82
5.2.3 Identificação das sequências obtidas das cultivares R570 e SP80-3280. ...	85
5.3 Busca por regiões regulatórias funcionais por meio da triagem da biblioteca de vetores por biobalística e análise da expressão do gene repórter GUS	87
5.4 Identificação e análise das regiões regulatórias funcionais.....	91
5.4.1 Identificação de regiões regulatórias funcionais por meio de hidrólise enzimática e sequenciamento	91
<i>5.4.1.1 Hidrólise enzimática dos clones com expressão transiente positiva.....</i>	<i>91</i>
<i>5.4.1.2 Sequenciamento completo dos clones com expressão positiva em epitélio de cebola: B4, C5 e F7.....</i>	<i>92</i>
5.4.2 Análise <i>in silico</i> dos clones F6, B4, F7 e C5.....	94
<i>5.4.2.1 Análise da sequência completa dos clones F6, C5, B4, F7 no banco de dados PLACE e PlantCare.....</i>	<i>94</i>
<i>5.4.2.2 Análise de elementos regulatórios flanqueantes aos clones isolados: F6, B4, C5 e F7.....</i>	<i>98</i>
<i>5.4.2.3 Predição do TATA e CAAT para os clones F6, C5, F7 e B4.....</i>	<i>99</i>
<i>5.4.2.4 Análise do clone F6.....</i>	<i>101</i>
<i>5.4.2.5 Análise do clone C5.....</i>	<i>105</i>
<i>5.4.2.6 Análise do clone F7.....</i>	<i>107</i>
<i>5.4.2.7 Análise do clone B4.....</i>	<i>108</i>
5.4.3 Ensaios de expressão transiente em arroz	110
<i>5.4.3.1 Análise da expressão transiente de protoplastos através da visualização da expressão de GFP.....</i>	<i>110</i>
5.5 Validação dos promotores através da transformação permanente em sistema heterólogo e nativo	112
5.5.1 Ensaios de expressão permanente em cana-de-açúcar.....	112
<i>5.5.1.1 Transformação genética de calos e discos de cana-de-açúcar via biobalística, seleção e regeneração de plantas transformadas.....</i>	<i>112</i>
<i>5.5.1.2 Análises Moleculares.....</i>	<i>114</i>
<i>5.5.1.3 Extração de DNA genômico das plântulas transformadas com o clone F6 em cana-de-açúcar.....</i>	<i>115</i>
<i>5.5.1.4 Análise por PCR com os iniciadores nptII.....</i>	<i>115</i>
<i>5.5.1.5 Análise por PCR com os iniciadores eGFP.....</i>	<i>116</i>
5.5.2 Ensaios de expressão permanente em arroz realizado em colaboração com o Núcleo de Genômica Funcional de Plantas (UFRGS)	118
6 CONCLUSÕES.....	121
7 PERSPECTIVAS	122

REFERÊNCIAS.....123

1 INTRODUÇÃO

A biotecnologia tem se destacado nos últimos anos como uma inovação tecnológica capaz de agregar riquezas a diferentes áreas do setor produtivo, entre elas, o setor da saúde e da produção de alimentos. No setor agrícola e florestal, a engenharia genética de plantas tem contribuído substancialmente para o maior entendimento da regulação gênica e desenvolvimento de plantas por meio da geração de organismos transgênicos para amplo uso na agricultura e fins industriais e farmacêuticos. A produção de organismos transgênicos de interesse comercial se baseia no desenvolvimento de métodos capazes de controlar a expressão do transgene de maneira precisa.

Desta maneira existe um grande empenho para se identificar novos promotores, já que estes são conhecidos por serem os processadores centrais da regulação da expressão de um gene e, portanto, são considerados ferramentas moleculares na pesquisa e desenvolvimento da biotecnologia. O objetivo central deste trabalho é o estabelecimento de uma metodologia de identificação rápida de promotores em média escala, a qual poderá ser aplicada a qualquer genoma de interesse. A inovação contida nesta proposta se baseia em adotar uma estratégia de clonagem customizada de sequências regulatórias capazes de regular a expressão gênica, sem a necessidade do conhecimento prévio do gene de interesse. Trata-se, portanto, de um trabalho com enfoque acadêmico e aplicado, o qual está vinculado ao programa FAPESP de Pesquisa em Bioenergia (BIOEN). Esse programa objetiva estimular e articular atividades exploratórias que possam gerar novos conhecimentos sobre a cana-de-açúcar, com o intuito de aprimorar sua capacidade na indústria em tecnologias dirigidas ao bioetanol, assim como aumentar sua competitividade interna e externa. Para o desenvolvimento deste projeto, foi escolhido como modelo biológico, a cana-de-açúcar, um híbrido de elevado interesse econômico e de alta complexidade genética. Poucos promotores foram isolados para a cana-de-açúcar, principalmente devido ao elevado nível de ploidia e o alto risco de se identificar regiões regulatórias provenientes de cópias inativas dos genes. Desta maneira, o desenvolvimento desta estratégia inovadora em cultivares de alto interesse permitirá resgatar novos motivos regulatórios capazes de dirigir a expressão gênica. Essas sequências irão permitir a construção de cassetes de expressão destinados a

programas de biotecnologia cujo objetivo é o aumento da biomassa desta cultivar e de outras, assim como possibilitará estudos de genômica funcional os quais permitirão desvendar módulos regulatórios inéditos.

2 REVISÃO BIBLIOGRÁFICA

2.1 O que é um gene? Uma visão pós-projeto ENCODE

A definição do termo gene evoluiu e se tornou mais complexa desde a primeira vez em que foi proposta, em 1909 (GERSTEIN et al., 2007). Atualmente, o termo gene é definido como uma “região localizável de sequência genômica, a qual corresponde a uma unidade de herança e que está associada a regiões regulatórias, regiões transcritas e/ou outras regiões funcionais” (PEARSON, 2006). No entanto, existem aspectos problemáticos relacionados à definição do gene trazidos a tona pela iniciativa do projeto ENCODE:

- uma grande quantidade de sequência de DNA, não anotada em genes conhecidos, é na realidade transcrita em RNA. Estas novas regiões transcritas são geralmente chamadas de TARs (do Inglês, transcriptionally active regions);
- existe um grande número de TSSs (do Inglês, transcription start site) não anotados, sendo que muitos genes que codificam proteínas conhecidas possuem TSSs alternativos que podem estar a mais de 100 kb à montante do TSS anotado. Portanto, muitas das isoformas alternativas codificam para a mesma proteína, diferindo apenas nas 5' UTRs (do Inglês, 5' UnTranslated Regions).;
- a regulação é dispersa ao longo do genoma, sendo que muitos elementos regulatórios na realidade estão situados dentro do primeiro exon, introns ou no corpo inteiro do gene (CAWLEY et al., 2004; EUSKIRCHEN et al., 2004; GERSTEIN et al., 2007; KIM et al., 2005; ZHANG et al., 2007).

Em função do exposto acima, uma menor distinção pode ser feita entre regiões gênicas e intergênicas no genoma. Agora, os genes parecem se estender no que uma vez foi chamado de espaço intergênico e a descoberta crescente de novos transcritos que possuem origem a partir de sítios regulatórios adicionais apontam para esta nova realidade (GERSTEIN et al., 2007). Desta maneira, a clássica visão do gene como uma unidade de informação hereditária alinhada ao longo do cromossomo, no qual cada gene é responsável por codificar uma proteína, têm dramaticamente mudado no último século. O que não têm mudado é o conceito de que o genótipo determina o fenótipo e a nível molecular isso significa que as sequências de DNA determinam sequências de moléculas funcionais (GERSTEIN et

al., 2007). É importante a perspectiva de que a proteína ou RNA codificados sejam produtos funcionais, já que a nível molecular assume-se que o fenótipo está relacionado à função bioquímica. Essa ênfase dada na geração de produtos funcionais claramente implica no conceito do que realmente é função biológica. Indo além, outra questão levantada é: O que é um gene? Qual é sua função?

O projeto ENCODE permitiu a identificação de elementos funcionais pertencentes a 1% do genoma humano através do emprego de técnicas computacionais e de experimentos moleculares de alta performance (THE ENCODE PROJECT CONSORTIUM, 2007). As descobertas baseadas naquele projeto expõem uma visão surpreendente relacionada à complexa e extensa atividade transcricional produzida em um genoma. Neste sentido, o ENCODE fortalece o que outros estudos têm apontado: a complexidade e quantidade de transcritos produzidos em um genoma deve ser levado em conta quando se pensa na definição do que é um gene.

A definição da função de uma dada região transcrita dependerá de análises bioquímicas e de mutagênese em larga escala (LAN et al., 2002, 2003). Em muitos casos é apenas questão de tempo para que sejam obtidas evidências experimentais que irão estabelecer o que muitos RNAs transcritos e proteínas fazem. Até esse momento chegar, iremos utilizar termos como TAR (regiões transcricionalmente ativas) ou indicar o grau de confiança em assumir função para um determinado produto gênico (GERSTEIN et al., 2007).

É baseado neste contexto que o presente trabalho se insere, numa fase em que o conceito do que é um gene já está em discussão, juntamente com a comprovação de que as regiões regulatórias compõem um mecanismo muito mais complexo na regulação da expressão gênica e geração de produtos funcionais do que era considerado há alguns anos atrás.

2.2 Regulação da transcrição gênica em eucariontes

Apenas alguns genes na célula eucariótica são expressos em certo momento do desenvolvimento celular. A proporção e composição dos genes transcritos variam consideravelmente durante o ciclo de vida, entre os tipos celulares e em resposta a condições ambientais e mudanças fisiológicas (ARBEITMAN et al., 2002; IYER et al.,

2001; KAYO et al., 2001; MODY et al., 2001; WHITE et al., 1999). Considerando que o genoma eucariótico contém cerca de 0,5 a 5×10^4 genes, a regulação diferencial da expressão gênica requer uma complexa ostentação de interações físicas específicas entre macromoléculas (WRAY et al., 2003).

As células eucarióticas apresentam o processamento da expressão gênica organizada em dois compartimentos: núcleo e citoplasma. O DNA dos eucariotos está armazenado no núcleo, onde ocorre a transcrição, enquanto a tradução do RNA originado a partir do DNA tem lugar no citoplasma, mais precisamente nos ribossomos. Desta forma, a regulação da expressão gênica ocorre nestes dois compartimentos, o que promove um significativo aumento em número e complexidade das etapas dos processos envolvidos na ação gênica (JOAQUIM; EL-HANI, 2010).

Para que ocorra a transcrição de um gene, uma série de estágios amplamente regulados deve suceder de forma concatenada. Primeiramente, há a ativação da estrutura gênica através de mudanças na estrutura local da cromatina a qual expõe sítios regulatórios (Figura 1) e permite que ocorra a interação do promotor a fatores basais de transcrição e co-ativadores. A montagem do aparato basal de transcrição na região do promotor permite sua interação com a RNA polimerase e desencadeia o início da transcrição. O transcrito primário de mRNA dos eucariontes é amplamente processado através da aquisição de revestimento *cap* (7 metil-guanosina) em sua extremidade 5', adição da cauda poli-A em sua extremidade 3' e a remoção de introns (*splicing*), gerando mRNAs maduros que serão exportados para o citoplasma e poderão ser traduzidos em proteínas (SNUSTAD; SIMMONS, 2008).

Figura 1 - Estrutura de um cromossomo com detalhe em um segmento funcional.

São representados os elementos *cis*-regulatórios (promotores, sítios de ligação a fatores de transcrição) e elementos de extenso alcance (*enhancers*, repressores/*silencers*, *insulators*) intercalados por regiões não codificantes, os quais permitem a regulação do gene e produção do transcrito desejado.

Fonte: Adaptado de Illumina (2012).

De acordo com a Figura 1, o promotor (bloco em vermelho) é o processador central da regulação da expressão de um gene, uma vez que contém os sítios de ligação para fatores de transcrição e para a RNA polimerase II, a qual é responsável pela efetiva transcrição gênica. Para iniciar a transcrição, a RNA polimerase deve reconhecer um local específico onde começará a síntese de RNA. As combinações de interações entre os elementos *cis*-regulatórios (promotores e sítios de ligação a fatores de transcrição), elementos regulatórios de extenso alcance (*enhancers*, repressores/*silencers*, *insulators*) e fatores de transcrição, os quais se ligam a estes elementos regulatórios, ativam a RNA polimerase II a qual se liga a este complexo de iniciação e inicia a síntese de RNA a partir de um único ponto (Figura 1). Posteriormente, os elementos *cis*-regulatórios, elementos de extenso alcance e fatores de transcrição serão melhor caracterizados.

É importante enfatizar que algumas sequências de DNA podem ser transcritas em mais do que de um RNA primário e, conseqüentemente, em diferentes mRNA chamados de isoformas de transcritos, as quais podem ser produzidas de várias maneiras, por exemplo, mediante o uso de diferentes TSS, ou usando regiões promotoras de locos gênicos completamente diferentes.

O aprofundamento no estudo destes componentes é de fundamental importância para um maior entendimento sobre os mecanismos moleculares envolvidos no processo de regulação da expressão gênica. Do ponto de vista biotecnológico, existe um grande interesse no emprego destas sequências regulatórias para a construção de cassetes que permitam a expressão temporal e espacial controlada de genes desejados no organismo de interesse por meio da transformação genética de plantas.

2.2.1 Elementos *cis*-regulatórios

Elementos *cis*-regulatórios estão distribuídos ao longo do DNA em basicamente quatro regiões: região promotora mínima (região mínima de DNA necessária para dirigir *in vitro* um baixo nível de início de transcrição), *enhancers*, *silencers* e *insulators* (LEVINE; TIJAN, 2003). Região promotora ou promotor são termos frequentemente usados neste trabalho e sua referência abrangerá os elementos *cis*-regulatórios presentes próximos ao promotor mínimo e ao longo da sequência de DNA localizados à montante do ponto de início da transcrição, que se localiza na posição +1 (Figura 2).

Existe uma série de elementos que garantem o acoplamento da maquinaria de transcrição ao início da sequência funcional. O elemento conservado mais próximo do ponto de início de transcrição é chamado motivo TATA, também conhecido como região promotora mínima (Figura 2). (BUTLER; KADONAGA, 2002). Essa região possui a sequência consenso TATAAAA e está localizada em torno da posição -30, ou seja, à montante da posição de início da transcrição +1. Esse motivo tem um papel importante em posicionar o início da transcrição, já que o complexo de proteínas de ligação ao TATA (*TATA binding protein* - TBP) se liga a esse motivo em grande parte dos genes transcritos pela RNA polimerase II (SMALE; KADONAGA, 2003). No entanto, muitos genes não possuem o motivo TATA, mas abrigam outro

tipo de sequência que permite a ancoragem do complexo de proteína de ligação ao TATA (TBP) denominado: *initiator element* (INR) e *downstream promoter element* (DPE) (LEVINE; TIJAN, 2003).

Outro elemento conservado é o motivo CAAT, localizado perto da posição -80 e tem a sequência de consenso GGCCAATCT (Figura 2). Dois outros elementos conservados, o motivo GC, consenso GGGCGG localizado a -100, e o motivo octâmero, consenso ATTTGCAT, a -130, geralmente estão presentes nos promotores da RNA polimerase II (SNUSTAD; SIMMONS 2008).

Figura 2 - Estrutura de um promotor reconhecido pela RNA-polimerase II.

Os motivos TATA e CAAT frequentemente estão situados nas mesmas posições nos promotores da maioria dos genes nucleares que codificam proteínas, mas existem variações. Os motivos GC e octâmeros podem estar presentes ou ausentes. Quando presentes, eles ocorrem em muitos locais diferentes, seja isoladamente ou em várias cópias. As sequências mostradas aqui são as sequências consenso para cada um dos elementos promotores.

Fonte: Snustad e Simmons (2008).

A eficiência do início da transcrição é modulada por regiões chamadas de *enhancers* e *silencers*. Os *enhancers* são formados por curtas sequências de DNA que apresentam um sítio de ligação em que vários tipos de fatores de transcrição se ligam. Os *silencers* são elementos reguladores equivalentes aos *enhancers*, porém, inibem a atividade de transcrição de genes específicos (ARMELIN-CORREA, 2007).

Os *enhancers* podem atuar a distâncias relativamente grandes – até vários milhares de pares de bases de seus genes regulados. Sua influência na expressão gênica é independente de orientação, podendo ativar a transcrição tanto na orientação senso como antisenso do gene a ser regulado. Seus efeitos são independentes de posição, ou seja, eles podem estar situados à montante ou à

jusante de um intron em um gene, ou mesmo dentro dele (SNUSTAD; SIMMONS, 2008). Existem estratégias que utilizam *enhancers* para aumentar a atividade de promotores e uma delas é a adição de cópias-extra de elementos *enhancer* a jusante de um promotor (CHEN et al., 2002).

Mesmo presentes a muitos kilobases de distância, os *enhancers* podem interagir diretamente com as proteínas ligadas nas redondezas do início da transcrição, sugerindo que a organização do DNA deve ser flexível o suficiente para permitir que ele e promotor permaneçam proximamente localizados. A forma como acontece essa “ação à distância” é por meio da formação de alças no DNA, que permite acomodar a interação entre as proteínas que exercem controle da transcrição, caso observado em *enhancers* bacterianos (SNUSTAD; SIMMONS, 2008).

Elementos denominados *insulators* podem bloquear a ação de elementos regulatórios. Esses são sequências de DNA presentes nos eucariotos e são definidos a partir de duas atividades funcionais: a primeira é uma atividade bloqueadora do *enhancer* que é dependente da posição. O *insulator* bloqueia a ação do *enhancer* somente quando está situado entre um *enhancer* e o promotor, mas não quando está à montante ou à jusante da dupla promotor-*enhancer* (CHUNG et al., 1993). A segunda atividade é a propriedade do *insulator* em participar da organização da estrutura da cromatina (ISHII et al., 2002) e a possível formação de alças. Se um *enhancer* estiver posicionado dentro da alça ele não poderia se comunicar com o promotor posicionado dentro de outro domínio de cromatina.

Para um melhor entendimento da regulação da expressão gênica é importante estudar tanto os elementos *cis*-regulatórios como outros segmentos do genoma responsáveis por codificar fatores de transcrição específicos os quais irão se ligar aos elementos *cis*-regulatórios.

2.2.2 Elementos *trans*-regulatórios

A região promotora contém sequências regulatórias que permitem a ligação de elementos *trans*-regulatórios conhecidos como fatores de transcrição. Esses elementos são capazes de reconhecer sequências de DNA específicas que se encontram nas regiões regulatórias dos genes por eles controlados. Além disso, os

fatores de transcrição atuam em sinergia e a existência de vários sítios de ligação próximos uns aos outros levam ao agrupamento desses fatores em uma determinada região regulatória, potencializando a atividade do complexo transcricional (LEVINE; TIJAN, 2003; SCHREM et al., 2002).

Os fatores de transcrição podem ser classificados em famílias de acordo com a presença de estruturas específicas em seu domínio de ligação ao DNA ou em seus domínios de interação proteína-proteína. In vitro, os fatores de transcrição geralmente reconhecem sequências de DNA de 6 a 8 pb de comprimento, o que claramente é insuficiente para a especificidade regulatória refinada que eles demonstram in vivo. Isso sugere que um grande número de fatores de transcrição formam os complexos regulatórios ativos e fornecem as bases para a regulação gênica por meio de ação combinatória (YILMAZ; GROTEWOLD, 2010). Um típico fator de transcrição possui múltiplos domínios funcionais, não somente para reconhecer e se ligar de forma apropriada ao elemento cis-regulatório alvo, como também para interações com outros fatores de transcrição, co-ativadores, com a RNA polimerase II, com complexos de remodelamento da cromatina e com pequenos RNAs não codificantes (PHILLIPS; HOOPES, 2008).

No entanto, existem alguns fatores de transcrição que atuam de maneira indireta, ou seja, não se ligam diretamente ao DNA, porém interagem com proteínas as quais vão se ligar à região promotora e dar início à transcrição. Esses fatores são chamados de fatores basais de transcrição e participam na geração do complexo de iniciação da transcrição, o que inclui a RNA polimerase II e muitos outros fatores de transcrição. A RNA polimerase II não é capaz de se ligar ao DNA sozinha, porém a presença desses fatores basais de transcrição (TFIIA, B,D,E,F) é capaz de formar um complexo multiprotéico, sendo que um dos componentes (TFIID) reconhece o motivo TATA e a ele se liga, o que cria uma estrutura capaz de ser reconhecida pela RNA polimerase II (BUCHANAN et al., 2001).

Desta maneira, fatores basais de transcrição são necessários para a iniciação da transcrição em todos os genes, porém não são capazes de aumentar ou diminuir a taxa de transcrição. Esta tarefa é realizada por elementos *trans*-regulatórios chamados ativadores e repressores, os quais são responsáveis por transportar mensagens de estímulo ou repressão para a RNA polimerase, que pode acelerar ou diminuir a taxa de transcrição de determinados genes. Os ativadores se ligam aos

enhancers e contribuem para determinar quais genes serão transcritos e em que velocidade a taxa de transcrição irá ocorrer. Já os repressores são proteínas que se ligam a *silencers*, interferindo no funcionamento dos ativadores e, portanto, diminuindo a taxa de transcrição. Os co-ativadores são moléculas “adaptadoras”, responsáveis por integrar sinais dos ativadores e repressores e transmiti-los a fatores basais de transcrição (TJIAN, 1995).

Deste modo, o controle da transcrição gênica é uma das formas de regulação da expressão gênica. É governado em nível de transcrição pela complexa interação entre os elementos *cis*-regulatórios e os fatores de transcrição, co-ativadores, ativadores e repressores, que atuam de maneira coordenada para que a RNA polimerase seja capaz de se acoplar ao aparato molecular e realizar a transcrição no núcleo de células eucarióticas.

2.3 Importância da identificação de regiões promotoras em plantas

A diversidade de promotores identificados com o decorrer dos últimos 40 anos e os benefícios da sua utilização, tanto para dirigir com sucesso a expressão gênica adequada no organismo desejado, como para fornecer um maior entendimento da complexa regulação de genes na planta têm sido inestimável. Estes avanços no estudo da expressão gênica ocorreram principalmente por meio da transformação genética de plantas com cassetes que possuíam um promotor fusionado a genes de interesse. Esta abordagem permitiu a manipulação e análise de processos bioquímicos, além da integração e estudo de genes que não poderiam ser manipulados pelo melhoramento clássico (POTENZA et al., 2004). A consequência foi uma explosão de cultivares geneticamente modificadas de alto valor agregado e com grande potencial para aumentar a produção de alimentos e criar produtos farmacêuticos que beneficiariam a saúde humana.

No entanto, a preocupação pública com relação às questões de segurança no consumo de plantas geneticamente modificadas e com a proteção do meio ambiente, levantou sérios questionamentos e desafios para os pesquisadores (EAGLESHMAN et al., 2001; NELSON, 2001). Têm se tornado cada vez mais claro que os esforços dedicados para minimizar os potenciais efeitos adversos das plantas geneticamente modificadas em organismos não-alvo devem ser prioridade. Esse

processo deve ser feito por meio do controle do nível apropriado de expressão gênica em certas partes da planta e em determinado período de tempo.

Os esforços realizados na identificação de novos promotores nem sempre garantem que essas sequências regulatórias irão dirigir a expressão adequada do gene de interesse no organismo desejado. Isso ocorre porque os promotores nem sempre são funcionalmente compatíveis entre diferentes organismos. O promotor CaMV 35S do vírus do mosaico da couve-flor tem sido largamente utilizado para dirigir altos níveis de expressão constitutiva em dicotiledôneas. Este, porém, confere baixos níveis de expressão em monocotiledôneas (CHRISTENSEN; QUAIL, 1996).

Em geral, apenas promotores isolados de monocotiledôneas apresentam alta atividade nesse grupo. O promotor da actina de arroz *Act1* (MCELROY et al., 1990, 1991; WANG et al., 1992; ZHANG et al., 1991) e o promotor ubiquitina de milho *Ubi1* (CHRISTENSEN; QUAIL, 1996) alcançaram níveis de expressão muito mais altos do que o promotor CaMV 35S em muitas monocotiledôneas testadas. Entre os promotores testados em cana-de-açúcar, o promotor *Emu* e o da ubiquitina, ambos isolados do milho, mostraram altos níveis de expressão em relação ao CaMV 35S (GALLO-MEAGHER; IRVINE, 1993; MCELROY et al., 1991; RATHUS et al., 1993). No entanto, estes mesmos promotores apresentaram o oposto em relação aos níveis de expressão em dicotiledôneas (CALLIS et al., 1990; MITRA; HIGGINS, 1994). A variação do nível de expressão do transgene entre diferentes espécies e promotores ocorre devido à disponibilidade distinta de fatores de transcrição, variações no reconhecimento de sequências promotoras ou sítios de “splicing” dos introns entre outros fatores (WILMINK et al., 1995).

Além disso, a variação do nível de expressão de um determinado transgene também tem sido uma importante preocupação no melhoramento de plantas com auxílio da manipulação genética. Estudos revelaram que em alguns casos a expressão do transgene foi perdida em uma proporção variável dos descendentes (CHAREONPORNWATTANA et al., 1999). Esse fenômeno chamado de “silenciamento gênico” tem sido extensivamente estudado em plantas dicotiledôneas, no entanto, poucos estudos em monocotiledôneas foram realizados (COOLEY et al., 1995). O silenciamento ocorre devido à homologia de sequências entre o DNA introduzido e as sequências endógenas, a qual desencadeia mecanismos de reconhecimento celular que resultam no silenciamento de genes

repetidos (BHULLAR et al., 2003). Um modo de prevenir o silenciamento do DNA introduzido é o desenvolvimento de cassetes de genes que contêm pouca similaridade a sequências endógenas. Isto pode ser feito por meio da substituição das regiões regulatórias originais por outras que funcionem da maneira desejada, obtendo-se um gene quimérico (HERRERA-ESTRELLA et al., 1984). No caso de sequências promotoras é imprescindível a disponibilidade de uma ampla diversidade de promotores isolados tanto de genomas de plantas, vírus ou a confecção de promotores sintéticos (BHULLAR et al., 2003).

Como a função dos promotores é integrar informações sobre as condições da célula e a taxa de iniciação da transcrição de genes, a recepção de sinais que o promotor é capaz de integrar pode variar bastante. Os promotores de genes expressos durante o desenvolvimento inicial do organismo integram fluxos de informação em um determinado espaço e tempo, que produzem padrões altamente dinâmicos de transcrição em regiões específicas do embrião (DAVIDSON, 2001; WILKINS, 1993). Promotores de genes que codificam proteínas *housekeeping* são constitutivamente ativos, porém podem ser desativados em resposta à condições específicas como mudança brusca de temperatura ou falta de nutrientes (PIRKKALA et al., 2001). Outros promotores são inativos na maior parte dos tecidos e fases do desenvolvimento, mas podem ser ativados em resposta à sinais hormonais, fisiológicos ou ambientais (BENECKE et al., 2001; SHORE; SHARROCKS, 2001). Esses diversos sinais atingem os promotores em forma de fatores de transcrição capazes de alterar as taxas de iniciação transcricional e determinam a regulação da expressão gênica, por isso a importância do estudo destas sequências.

Diante deste cenário, os promotores possuem um papel fundamental para responder questões relacionadas à segurança de organismos transgênicos, evitar o silenciamento gênico nas plantas geneticamente modificadas e permitir um grande avanço no conhecimento sobre a regulação da expressão de genes. Esse conhecimento é precioso tanto para o aperfeiçoamento de ferramentas biotecnológicas de uso comercial, como tem potencial imenso para ser aplicado na pesquisa básica.

2.4 Promotores constitutivos, modificados, regulados e sintéticos

Os promotores são extremamente úteis para a expressão regulada de genes introduzidos em plantas transgênicas e podem ser classificados como constitutivos, regulados, virais e sintéticos (ODELL et al., 1985; POSZKOWSKI et al. 1984). Dependendo das sequências-sinal presentes na região promotora, a expressão do gene associado irá responder a uma ou mais combinações desse sinal, promovendo um determinado padrão de expressão.

2.4.1 Promotores constitutivos

Promotores constitutivos são aqueles capazes de dirigir a expressão gênica em quase todos os tecidos da planta durante todos ou quase todos os estágios de desenvolvimento da planta (BADE et al., 2003). Cada um dos elementos ativadores da transcrição não exibe uma especificidade tecidual e são capazes de mediar a ativação transcricional na maior parte da planta. Vários promotores constitutivos de plantas são provenientes de genes como a actina e ubiquitina. A actina é um componente fundamental do citoesqueleto e é expresso em quase todas células da planta. O promotor do gene da actina 1 de arroz foi desenvolvido e utilizado em diversos cereais (MCELROY et al., 1991; ZHANG et al., 1991) e mostrou ser capaz de dirigir a expressão da β -glucuronidase (GUS) em virtualmente quase todos os tecidos.

A ubiquitina é conhecida como uma proteína altamente conservada e está associada a muitos processos celulares vitais como dobramento protéico, estrutura da cromatina e reparo de DNA. Desta forma, o promotor da ubiquitina 1 de milho (pUbi) é dos mais utilizados para transformação genética de plantas. Estudos iniciais mostram que a expressão do pUbi-1:CAT (cloranfenicol acetiltransferase) foi 10 vezes maior em protoplastos de milho do que o promotor 35S:CAT, porém, foi 10 vezes menor em protoplastos de tabaco em relação ao 35S:CAT, limitando sua utilidade em monocotiledôneas (CHRISTENSEN et al., 1996; GUPTA et al., 2001; WEEKS et al., 1993).

Algumas plantas possuem a habilidade de sintetizar proteínas que permitem o seu crescimento em condições anaeróbicas. Foi demonstrado por Walker et al.

(1987) que esta deficiência de oxigênio possui um efeito dramático no padrão de expressão dos genes em plântulas de milho. Uma das enzimas induzidas neste processo é a álcool desidrogenase 1 de milho (*Adh1*) (SACHS; FEELING, 1980), sua síntese é presumivelmente uma resposta adaptativa que permite a sobrevivência da planta em períodos de alagamento, o qual foi demonstrado experimentalmente em plântulas de milho (SCHWARTZ, 1969). O promotor *Adh-1* isolado do milho tem sido muito utilizado na transformação genética de outros cereais como a cevada, aveia e o próprio milho, porém, em dicotiledôneas como o tabaco, apresenta baixos níveis de expressão (ELLIS et al., 1987).

Promotores derivados de vírus de plantas são largamente empregados para dirigir a superexpressão de transgenes e um exemplo deste tipo é o 35S (CaMV) (ODELL et al., 1985), o qual é considerado um dos promotores mais fortes em dicotiledôneas (GUILLEY et al., 1982). Ele é capaz de intermediar a expressão de genes exógenos em quase qualquer órgão da planta, além de ser comumente utilizado na expressão constitutiva de qualquer gene de interesse em plantas (JEFFERSON et al., 1987; ODELL et al., 1985). Além disso, este promotor em particular é valioso por estar disponível para a utilização em vários vetores de plantas, facilitando a subclonagem do transgene de interesse e a sua ampla aplicação na área da engenharia genética (POTENZA, 2004).

No entanto, o promotor CaMV 35S demonstrou atividade relativamente baixa em monocotiledôneas, como por exemplo, a cana-de-açúcar e o milho (LAKSHMANAN et al., 2005). Estas observações sugerem que devem existir diferenças entre monocotiledôneas e dicotiledôneas com relação aos fatores de transcrição e o reconhecimento de sequências de promotores (LAST et al., 1991). Estas diferenças fazem com que os promotores isolados atuem de maneira diferente quando situados numa outra espécie, gênero e até cultivar de plantas (POTENZA, 2004).

Outro promotor conhecido foi isolado a partir do gene da octopina sintase e é proveniente do plasmídeo indutor de tumor da bactéria patogênica de plantas denominada *Agrobacterium tumefaciens* (DE GREVE et al., 1982); é uma parte do T-DNA a qual é transferida de *A. tumefaciens* para as células das plantas durante a formação do tumor. Esse gene, embora originalmente isolado de um procarioto,

somente é expresso depois de sua introdução em células de plantas e, portanto, deve conter sinais de transcrição eucarióticos (DE GREVE et al., 1982).

Em cana-de-açúcar, a maioria dos promotores isolados e caracterizados são de natureza constitutiva. Wei e colaboradores foram os primeiros a isolar promotores constitutivos de cana-de-açúcar a partir de genes da poliubiquitina (WEI et al., 1999). Outros promotores constitutivos como o gene do fator de alongação α , o gene que codifica uma proteína rica em prolina (YANG et al., 2003) e um vírus baciliforme de cana-de-açúcar (BRAITHWAITE et al., 2004; TZAFRIR et al., 1998) também foram isolados.

No entanto, a expressão constitutiva pode ser problemática já que promotores constitutivos não apresentam padrão de expressão específico e possuem limitações no uso da engenharia genética (GITTINS et al., 2000). Isto ocorre porque a super-expressão de um transgene específico em tecidos no qual normalmente não é expresso, ou alto níveis do produto deste transgene podem causar consequências inesperadas no crescimento e desenvolvimento da planta (POTENZA et al., 2004).

2.4.2 Promotores constitutivos modificados

Aperfeiçoamentos derivados do promotor CaMV 35S ou do promotor do gene da álcool desidrogenase *Adh 1* de milho já foram desenvolvidos. O promotor CaMV 35S teve sua atividade reforçada pela inserção de um intron dentro de sua região transcrita, sendo utilizado com sucesso para dirigir a expressão de genes marcadores de seleção ou genes de interesse agrônomo como a toxina inseticida Bt em arroz transgênico (DATTA et al., 1998; FUJIMOTO et al., 1993)

Já o gene da álcool desidrogenase *Adh 1* foi submetido à uma fusão entre a região promotora de *Adh 1* e o primeiro intron do gene da própria *Adh 1*. Esta construção foi unida à múltiplas cópias de um promotor de elemento de resposta anaeróbico e ao elemento ocs do gene da octopina synthase de *Agrobacterium tumefaciens* para a criação de um promotor modificado denominado pEmu. Este promotor teve sua atividade testada em diferentes sistemas em monocotiledôneas e demonstrou 10 a 100 vezes maior atividade do que o promotor CaMV 35S (CHAIR et al., 1996; LAST et al., 1991; RATHUS et al., 1993; SCHLEDZEWSKI; MENDEL, 1994; WILMINK et al., 1995).

2.4.3 Promotores regulados

Os promotores regulados são capazes de dirigir a expressão gênica através da regulação da expressão temporal e/ou espacial, o que inclui promotores tecido-específicos e promotores induzidos. Desta forma, estes promotores são capazes de dirigir a expressão gênica em diferentes tecidos ou tipos celulares, em diferentes estágios de desenvolvimento ou em resposta a diferentes condições ambientais.

Transgenes que estão sob controle desses tipos de promotores irão apresentar expressão apenas nos tecidos em que o produto do transgene é desejado, desta maneira o restante da planta não sofre alterações pela expressão inadequada do transgene (POTENZA et al., 2004; STITT et al., 1995).

2.4.3.1 Promotores tecido-específicos

O desenvolvimento de promotores tecido-específicos para engenharia genética pode ser bem complexa. O tamanho do fragmento do promotor isolado é capaz de gerar um determinado efeito na expressão do transgene. No entanto, o processo de isolamento de promotores podem causar a eliminação de elementos *enhancer cis* atuantes na região 5', modificando completamente o padrão de expressão do transgene (POTENZA et al., 2004).

A integração ao acaso do transgene no genoma hospedeiro também pode determinar a perda de potenciais interações *trans*, as quais se baseiam na estrutura da cromatina ou devido à própria localização dentro do genoma. Outro fator determinante é a ausência de interação efetiva entre os *cis*-elementos pertencentes ao promotor e os fatores *trans*-atuantes presentes na planta hospedeira transgênica (POTENZA et al., 2004).

Portanto, torna-se essencial os esforços realizados para uma melhor compreensão da funcionalidade de promotores homólogos e heterólogos em plantas hospedeiras. Este empenho seria crucial para aumentar tanto a eficiência da expressão do transgene, como a obtenção de níveis adequados de expressão no tecido desejado.

A expressão tecido-específica em plantas pode ser realizada através da inserção de transgenes que possuem sequências promotoras particulares capazes de serem reconhecidas pelo organismo-alvo e apresentar expressão no local

desejado. Um exemplo com aplicação biotecnológica foi descrito por Muchhal et al. (1996) que foram capazes de isolar cDNA de *Arabidopsis thaliana* em condições de deficiência de fósforo e caracterizaram dois genes transportadores de fosfato: AtPT1 e AtPT2 (*Arabidopsis Phosphate Transporters*). Para comprovar que estes genes possuíam especificidade de expressão na raiz quando submetidos à condições de deficiência de fósforo inorgânico, Karthikeyan et al. (2002) transformaram *Arabidopsis thaliana* com três genes repórteres diferentes, GUS, LUC e GFP sob controle dos promotores dos genes AtPT1 e AtPT2. Desta maneira foi possível comprovar a ativação do promotor do gene AtPT1 como tecido-específico sob estresse do fósforo inorgânico, enquanto que o promotor AtPT2 não mostrou especificidade, sua expressão ocorreu tanto em raízes como em folhas. O esforço em identificar promotores tecido-específicos que respondam ao estresse de fósforo inorgânico (Pi) é extremamente interessante para a geração de transgênicos mais eficientes na utilização do fósforo em solos deficientes, como é o caso dos solos do cerrado (DE LA FUENTE et al., 1997; MUCHAL et al., 1996; RAGHOTHAMA, 1999)

Outro exemplo com aplicação biotecnológica foi descrito por Huang et al. (2001) o qual desenvolveram um método de seleção de plantas transformadas de arroz a partir da identificação de um promotor capaz de regular a expressão especificamente em calos. Deste modo, não haverá expressão do gene de resistência ao antibiótico em outros tecidos e por consequência, os efeitos indesejáveis possivelmente derivados da expressão constitutiva em toda a planta serão ausentes.

Já a análise do promotor isolado do gene 1-aminociclopropano-1-carboxilato (ACC) de tomate, mostrou que os fragmentos que variaram entre -1966 a -1159pb do sítio de início de transcrição possuíam a capacidade de dirigir a expressão específica do amadurecimento do fruto de tomate sem apresentar atividade em outros tecidos (ATKINSON et al., 1998). Outro promotor com atividade tecido-específica foi isolado a partir do gene que codifica a ribulose-1,5-bifosfato carboxilase/oxigenase (Rubisco) ativase (RCA) em *Arabidopsis thaliana*. Este promotor é responsivo à luz e regulado pelo relógio circadiano, além de ser órgão-específico por apresentar atividade nas regiões clorofiladas da planta, sendo ausente em raízes e pétalas (LIU et al., 1996).

Um promotor tecido-específico bem caracterizado e comumente utilizado é o promotor do gene da β -faseolina isolada do feijão (VAN DER GEEST; HALL, 1997). Este gene é responsável por codificar proteínas de reserva e sua região promotora foi analisada funcionalmente através da fusão com gene repórter (BAUMLEIN et al., 1991; SENGUPTA- GOPALAN et al., 1985). Estes estudos mostraram que a região promotora abriga elementos regulatórios capazes de dirigir forte expressão específica em sementes de plantas transgênicas de tabaco (SENGUPTA-GOPALAN et al., 1985). O promotor da faseolina é amplamente utilizado em experimentos de engenharia genética que visam a expressão semente-específica (ALTENBACH et al., 1989).

Neste contexto, a capacidade de direcionar a expressão dos genes de interesse em determinados órgãos da planta através do uso de promotores tecido-específicos constitui uma ferramenta biotecnológica valiosa. Tal inovação viabiliza a obtenção de um padrão de expressão tecido-específico altamente desejável em programas biotecnológicos que visam a produção e liberação comercial de organismos geneticamente modificados.

2.4.3.2 *Promotores induzidos*

Em plantas existem vários tipos de estresses bióticos e abióticos que são capazes de mobilizar um grande número de genes relacionados à defesa e reparo, assim como moléculas sinalizadoras a ataques de patógenos ou tecidos lesionados. A partir destes genes induzidos por estresse e defesa, foi possível o isolamento e caracterização de alguns promotores induzidos valiosos para a engenharia genética (POTENZA et al., 2004). Esses promotores são úteis porque são capazes de dirigir a expressão específica e precisa no ponto da infecção ou lesão da planta quando fusionados a genes de resistência. Um exemplo de um promotor promissor induzido por patógeno, é o promotor do gene *hsr203J* (BELBAHRI et al., 2001; KELLER et al., 1999) o qual foi utilizado para o desenvolvimento de plantas transgênicas que possuíam um gene capaz de codificar um elicitor altamente ativo. Este elicitor era produzido nos locais de infecção e capaz de sinalizar as defesas naturais da planta, impedindo o desenvolvimento do patógeno (DE WIT, 1992).

Seria ideal se os promotores induzidos não apresentassem qualquer nível basal de expressão gênica na ausência do agente indutor e, além disso, a expressão deve ser reversível e dose-dependente. O indutor deve ser barato, fácil de aplicar e não tóxico. Isso é especialmente importante para sistemas a serem utilizados no campo ou comercializados (GATZ; LENK, 1998). A expressão induzida do gene desejado em estágios específicos do desenvolvimento do organismo podem contribuir para o melhor entendimento da correlação entre o fenótipo e a expressão específica do gene. O fluxo de metabólitos nas plantas que apresentam expressão alterada de certas enzimas poderia ser estudada e monitorada com grande precisão (CADDICK et al., 1998).

Sistemas de expressão gênica regulada poderiam também ser valiosos em aplicações biotecnológicas como a expressão condicional de genes “pesticidas” (i.e., *Bacillus thuringiensis* cry genes), expressão de genes-alvo resistentes a herbicidas ou indução de florescimento sincronizado ou amadurecimento (POTENZA et al., 2004). Todos estes exemplos apontam para a necessidade valiosa de identificação de novos sistemas promotores induzidos os quais terão papel fundamental na produção de plantas geneticamente modificadas.

Exemplos de alguns promotores induzidos são o *wun1* e o inibidor da proteinase II (*pin2*) isolados de batata e que são capazes de serem expressos através da injúria e presença de patógenos, porém, não apresentam expressão sem o estímulo específico (KEIL et al., 1990; LOGEMANN et al., 1989; XU et al., 1993). O promotor *pin2* dirige baixa expressão constitutiva, porém é induzido quando submetido a altos níveis de metil jasmonato e injúria quando transformado em arroz (XU et al., 1993).

Muitos dos sistemas de promotores induzidos/repressíveis são formados por dois componentes que apresentam circuitos moleculares do tipo “liga-desliga” que variam em complexidade. Estes componentes incluem fatores de transcrição capazes de se ligarem a promotores-alvo, os quais contêm elementos *cis*- atuantes específicos para os fatores de transcrição em questão. Desta forma, o promotor alvo é fusionado a um transgene de interesse e a ligação do fator de transcrição ao promotor modula a expressão do transgene. Um exemplo deste sistema de dois componentes é o sistema induzido por etanol (FELENBOK et al., 1988), o qual é composto por um promotor do gene da álcool desidrogenase, *alcA* e um gene que

codifica um fator de transcrição ativado por álcool (*alcR*) dirigido pelo promotor CaMV 35S. O fator de transcrição (*ALCR*) na presença de etanol se liga a sítios específicos do promotor quimérico mínimo *alcA*:CaMV 35S (região TATA box), ativando por sua vez a transcrição do gene alvo (CADDICK et al., 1998; ROSLAN et al., 2001; SALTER et al., 1998). O sistema *alc* é extremamente sensível a etanol e induzido rapidamente (ROSLAN et al., 2001; SALTER et al., 1998). O gene-alvo (álcool desidrogenase) possui expressão basal e abriga a vantagem de ser um sistema indutor econômico e não-tóxico em concentrações efetivas. A indução da expressão do gene-alvo foi similar à de genes dirigidos pelo promotor CaMV 35S porém possui a vantagem de ser reversível (ROSLAN et al., 2001). Uma desvantagem é que em condições anaeróbicas, plantas que são conhecidas por expressar o gene da álcool desidrogenase são prováveis de ativar esse sistema, gerando a produção de etanol, um metabólito fermentável.

Apesar deste inconveniente, o sistema parece ser um dos mais promissores para uso no campo e em condições comerciais (POTENZA et al., 2004).

2.4.4 Promotores sintéticos

Um promotor sintético é constituído por uma sequência de DNA que não existe de forma natural e que foi desenhada para controlar a expressão de um determinado gene-alvo. Sequências *cis*-regulatórias derivadas de elementos promotores que ocorrem de maneira natural são utilizadas para construir promotores sintéticos através da técnica de montagem em blocos (Figura 3). O resultado é uma sequência de DNA composta por alguns elementos *cis*-regulatórios distintos em uma orientação completamente nova e que podem atuar como um promotor *enhancer*, capaz de iniciar a transcrição mediada pela RNA polimerase II (Roberts, 2011).

Um típico promotor sintético contém os elementos *cis*-regulatórios e um promotor mínimo fusionados a um gene repórter à jusante, seguido pela presença de um gene de seleção a antibiótico e uma sequência poliA essencial para processamento do sinal pós transcricional (Figura 3). O promotor mínimo sozinho não é capaz de promover a transcrição, apenas a interação entre os elementos *cis*-regulatórios e o promotor mínimo é que tornam possível acionar a transcrição pela RNA polimerase II.

Figura 3 - Estrutura de um típico promotor sintético de eucariotos o qual contém elementos *cis*-regulatórios e um promotor mínimo fusionados a um gene repórter à jusante seguido pela presença de um gene de seleção a antibiótico e uma sequência poliA.

Fonte: Roberts, 2011.

A construção de promotores sintéticos só foi possível devido à natureza modular das regiões regulatórias de genes nativos. Os primeiros estudos utilizando promotores sintéticos foram realizados com o objetivo de identificar as principais características estruturais dos promotores eucarióticos e procarióticos. Desta maneira foi possível caracterizar os elementos essenciais de DNA capazes de dirigir a transcrição (ROBERTS, 2011).

A tecnologia de construção de promotores sintéticos foi rapidamente associada com a possibilidade de aumentar a eficiência da expressão gênica e atingir um outro patamar desejado: o de garantir o controle refinado da expressão do transgene em um determinado tempo e local. O interesse na obtenção de promotores com estas características é ilustrado pelo recente aumento de pedidos de patentes por empresas de agrobiotecnologia interessadas na obtenção de promotores que permitam aumentar a resistência a doenças, ataque de patógenos sem prejudicar o desenvolvimento da planta (GURR; RUSHTON, 2005).

A percepção de que promotores induzidos por patógenos contêm elementos *cis*-atuantes bem conservados em outras espécies (EULGEM et al., 1999), fez com que esforços fossem empregados no refinamento das sequências promotoras através da inclusão seletiva de elementos que contribuíssem para aumentar a força e atividade do promotor (YERI, 2009). Isto foi feito através da construção de promotores sintéticos provenientes de uma variedade de elementos *cis*-atuantes

induzidos por patógenos. Quando retirados do seu contexto de promotor nativo e organizados como componentes de promotores sintéticos, estes elementos *cis*-atuantes eram capazes manter a indução por patógenos (GURR; RUSHTON, 2005).

A atividade do promotor e a força da indução depende de alguns parâmetros como o número, ordem e espaçamento dos elementos *cis*-atuantes. O número de cópias de um elemento individual no promotor parece ser fundamental por influenciar não somente a força e capacidade de indução, mas também por determinar a expressão basal do gene (RUSHTON et al., 2002). O espaçamento entre os elementos *cis*-atuantes e/ou entre estes elementos e o complexo de pré-iniciação também pode ter um profundo efeito, no entanto o espaçamento é de difícil predição (WRAY, 1998) e o ótimo espaçamento assim como o número ótimo de elementos necessários devem ser determinados experimentalmente.

No caso de promotores induzidos por patógenos, estes também devem permanecer inativos na ausência de infecções, para garantir que falsas respostas de defesa não sejam desencadeadas por um escape da expressão do transgene (MCDOWELL; WOFFENDEN, 2003). Além disso, o promotor não deve ser auto-ativável pelo transgene, o que poderia desencadear uma propagação descontrolada da expressão do transgene causando morte celular (YERI, 2009). De forma pouco surpreendente, um número limitado de promotores se encaixam nestes requisitos (HAMMOND-KOSACK; PARKER, 2003; MICHELMORE, 2003).

Uma promessa de promotor induzido por patógeno é proveniente do gene do tabaco *hsc203J* (BELBAHRI et al., 2001; KELLER et al., 1999). Este promotor foi utilizado pela primeira vez por Wit (DE WIT, 1992), o qual notavelmente inseriu em plantas um gene capaz de codificar uma proteína elicitora altamente ativa sob controle de um promotor que é especificamente induzido por um virulento patógeno. A produção do elicitor em sítios infectados é suficiente para desencadear as respostas de defesa naturais da planta e parar a propagação do patógeno (GURR; RUSHTON, 2005).

Deste modo, os promotores sintéticos surgiram como excelentes ferramentas para facilitar a identificação de importantes características estruturais das sequências de promotores que ocorrem de forma nativa, permitindo tanto o aumento como o controle mais refinado da regulação da expressão gênica. Várias instituições têm utilizado a tecnologia de confecção de promotores sintéticos com o intuito de

aplicar na área biofarmacêutica, no campo da agricultura, bioenergia e no desenvolvimento de novas terapias gênicas (ROBERTS, 2011).

O campo da biologia sintética continua a se desenvolver em uma indústria multi-bilionária e é provável que a utilização da tecnologia de confecção de promotores sintéticos permaneça em franca expansão e com promessas estimulantes.

Os diferentes tipos de promotores descritos e seus variados mecanismos de ação refletem a complexidade inerente destas sequências em integrar informações sobre as condições atuais da célula e relacioná-las com a modulação da taxa de iniciação da transcrição gênica de maneira coordenada e específica. Portanto, o desafio no campo biotecnológico reside em adquirir conhecimentos sobre o modo de regulação destas sequências específicas de DNA e integrar esta informação a cassetes de expressão capazes de modular a expressão de genes de interesse num organismo desejado.

2.5 Métodos de identificação e isolamento de regiões promotoras em plantas

A identificação de promotores em plantas é considerada uma tarefa complexa. Isto acontece porque sequências promotoras divergem muito e nem sempre os motivos conhecidos em uma espécie se encontram conservados em outra. Cada gene abriga uma combinação de elementos *cis*-regulatórios única, o que faz com que o controle da transcrição seja modulado de maneira específica num determinado tempo e espaço (QIU, 2003). Apesar das dificuldades de identificação de promotores, Arnone et al. (1997) e Wilkins (1993) sugerem que num estudo de promotores eucarióticos bem caracterizados um número na ordem de 10 a 50 sítios de ligação para 5 a 15 fatores de transcrição diferentes não é incomum.

Uma variedade de metodologias são empregadas para a identificação de regiões promotoras. A forma convencional de isolamento é através da identificação de um clone de cDNA de um gene em particular e a utilização desta região transcrita como sonda para efetuar uma busca em uma biblioteca genômica de DNA por sequências que possuem complementariedade. Estas sequências seriam então clonadas e a região 5' à montante do gene desejado seria identificada através da comparação entre o cDNA e a sequência genômica de DNA (DATLA et al., 1997). A

diferença entre clones genômicos e clones de cDNA é a falta de introns nos clones provenientes do cDNA, o que facilita a sua caracterização e utilização como sonda (AN; KIM, 1993).

Uma das técnicas que utilizam a abordagem acima é denominada “*genome walking*” (DEVIC et al., 1997) a qual emprega o isolamento do DNA genômico seguido pela hidrólise com várias enzimas de restrição de corte abrupto de maneira separada e posterior ligação de adaptadores aos fragmentos de restrição resultantes. Estes fragmentos+adaptadores são utilizados como moldes para amplificações por PCR através de um oligonucleotídeo que abriga uma sequência complementar ao adaptador e uma sequência complementar a um gene cuja expressão tecido-específica e sequência de bases já foram previamente verificadas, seguindo-se da clonagem e sequenciamento dos fragmentos.

Recentemente, avanços no campo da genômica funcional resultaram em uma variedade de técnicas para isolamento de promotores tecido-específicos em plantas. Algumas delas envolve a informação sobre os transcritos para a clonagem do gene e de sua região regulatória (DIGEON et al., 1999; MISHRA et al., 2002) como a técnica “*differential display*” (REUBER; AUSUBEL, 1995) ou a metodologia denominada perfil de expressão baseado em *microarray* (RICHMOND; SOMMERVILLE, 2000; SCHENA et al., 1995).

Outra abordagem interessante para o isolamento de promotores é a inserção aleatória de fragmentos de DNA que contêm à jusante apenas o gene repórter e ausência do promotor. Somente as construções que apresentam a expressão do gene repórter seriam selecionadas. Um exemplo utilizando essa abordagem foi a construção de um vetor chamado pROA97 o qual continha dois marcadores (gene *npt* e gene *GUS*). Ambos marcadores foram posicionados sob o controle mínimo de um promotor 35S que possuía a região TATA box (AN; KIM, 1993). Através dessa metodologia alguns fragmentos promotores foram isolados a partir do DNA genômico de *Arabidopsis thaliana* (OTT; CHUA, 1990) e arroz (CLAES et al., 1991). Tendo à disposição o sequenciamento completo do genoma da planta modelo *Arabidopsis*, é possível a fácil e rápida identificação dos promotores e a análise comparativa com outras espécies.

Após o isolamento de putativas regiões promotoras, a funcionalidade do promotor é geralmente testada pela fusão do respectivo fragmento de DNA a um

gene repórter como o GUS (JEFFERSON et al., 1987), luciferase (DE WET et al., 1987) ou GFP (CHALFIE et al., 1994). Estes genes tem a característica de terem o seu produto protéico facilmente detectável nos tecidos transformados, seja por quimioluminescência ou ensaio histoquímico. Estas construções contendo a putativa região promotora à montante do gene repórter são introduzidas em células de maneira transiente ou permanente. A detecção da atividade da região promotora é então comprovada pela expressão do gene repórter, no entanto, a detecção da expressão do gene GUS possui um limite mínimo. Desta maneira, a ausência da visualização da atividade do gene GUS não significa necessariamente ausência de expressão da região promotora (ROBERTS, 2011).

Além dos métodos de isolamento *in vitro*, existem também numerosos métodos e programas computacionais para análise e reconhecimento de sítios de ligação de fatores de transcrição (CASELLE et al., 2002; PRESTRIDGE et al., 1993; QUANDT et al., 1995, 1997; WOLFERTSTETTER et al., 1996). A predição de sítios e suas combinações são baseadas na procura por “sinais”, através de algoritmos. Estes por sua vez tentam identificar regiões regulatórias baseadas no contexto da sequência do DNA analisado. No entanto, a detecção *in silico* de elementos regulatórios não significa necessariamente que todos os elementos detectados na sequência de DNA sejam funcionalmente relevantes. Uma vez que estas sequências *cis*-atuantes geralmente consistem de segmentos curtos de nucleotídeos, há sempre uma possibilidade aleatória de encontrar estas sequências numa determinada extensão do DNA (BLANCHETTE; SINHA, 2001; HUTCHINSON, 1996; KONDRACHIN et al., 1995; PRESTRIDGE, 1995;). No entanto, a precisão desse reconhecimento ainda é baixa (FICKETT; HATZIGEORGIOU, 1997).

Portanto, atualmente existe um desafio considerável em se detectar com precisão sítios regulatórios responsáveis pela ativação e expressão de grupos de genes funcionalmente ativos (MOHANTY et al., 2005). Torna-se imperativa a identificação de novos promotores derivados de espécies de interesse que apresentem expressão constitutiva, tecido-específica e/ou ativados sob determinadas condições. Desta forma poderão ser produzidas, através de nossa tecnologia, plantas geneticamente modificadas com genes de interesse fusionados à regiões promotoras livres de patentes.

2.6 Identificação de regiões promotoras em cana-de-açúcar

Genomas complexos são caracterizados por possuir grande quantidade de sequências altamente repetitivas, pseudogenes e alta redundância funcional de genes. Exemplos de genomas complexos de gramíneas de interesse comercial abrangem espécies como o trigo, milho, cevada e a cana-de-açúcar. Dentre estas espécies, a única que possui o genoma sequenciado por completo é o milho (WEI et al., 2009), sendo que está em andamento outro grande desafio: o sequenciamento completo do genoma da cana-de-açúcar pelo grupo do BIOEN (Programa FAPESP de Pesquisa em Bioenergia). Este híbrido ocupa posição de destaque no agronegócio do estado de São Paulo, justamente por sua extraordinária capacidade de estocar sacarose nos entrenós. Esta sacarose é matéria-prima para a produção de açúcar e bioetanol, sendo que o último se destaca como principal fonte de energia renovável na matriz energética brasileira.

Apesar dos avanços nos últimos 15 anos relacionados ao aprofundamento do conhecimento da cana-de-açúcar, o desenvolvimento de ferramentas moleculares capazes de auxiliar na compreensão deste complexo genoma ainda estão em desenvolvimento. Estudos de gramíneas filogeneticamente relacionadas à cana como sorgo e arroz, mostram que para melhorar a produtividade da cana é preciso conhecer como a atividade dos genes é controlada. Este controle é exercido por sequências de DNA conhecidos como promotores.

Atualmente existem poucas publicações de promotores isolados em cana-de-açúcar (WEI et al., 1999). Isto se deve ao fato de que este híbrido apresenta uma complexidade genética excepcional e um alto nível de ploidia, o que faz com que um único gene apresente múltiplas cópias (WENDEL, 2000). Esta particularidade dificulta muito o processo de isolamento de promotores ativos, já que existe um alto risco de se isolar promotores não funcionais provenientes de cópias inativas do gene (YANG et al., 2003). Além disso, o genoma da cana-de-açúcar não está completamente sequenciado, o que dificulta ainda mais o processo de identificação de promotores por análise comparativa com outras gramíneas já sequenciadas.

Diante deste cenário, a grande pergunta que surge é a seguinte: Como identificar promotores funcionais em espécies altamente complexas como a cana-de-açúcar?

Os trabalhos realizados para isolar promotores funcionais em cana-de-açúcar demandam uma grande quantidade de trabalho, tempo e gastos elevados. Yang e colaboradores realizaram uma triagem com uma sonda de cDNA marcada radioativamente em uma biblioteca genômica de cana com o objetivo de isolar promotores constitutivos. Um total de onze promotores foram isolados, sendo que dois destes promotores foram capazes de dirigir a expressão do marcador GUS em calos de cana e embriões de trigo em níveis equivalentes à aqueles obtidos com o promotor do milho Ubi1 (YANG et al., 2003). Estes promotores foram isolados dos genes que codificam um fator de alongação e uma proteína rica em prolina (YANG et al., 2003).

Outro recente trabalho realizado para isolar promotores de membros de uma família multigênica de cana-de-açúcar foi desenvolvido por Damaj et al. (2010) o qual utilizou uma biblioteca de BACs para a seleção de um cDNA chamado SHDIR16, que corresponde a um promotor com expressão em específica no colmo da cana-de-açúcar (DAMAJ et al., 2010). Após esta etapa foi adotada uma metodologia de “*genome walking*” que consistiu na construção de bibliotecas “*genome walker*” de SHDIR16 a partir do BAC positivo para a seleção do cDNA de SHDIR16. Este BAC foi digerido e ligado a adaptadores seguido pela amplificação do promotor-alvo com um iniciador específico à região 5’ não traduzida (UTR) do gene SHDIR16 e um iniciador com sequência específica ao adaptador. Esta técnica gerou o isolamento de um promotor de 2.6kb pertencente à subfamília do gene DIRIGENT (DIR), a qual está potencialmente envolvida em respostas de defesa e biosíntese de fibra (CASU et al., 2004; RALPH et al., 2007). Este promotor provou ser funcional em arroz e cana transformados de maneira estável (DAMAJ et al., 2010). No entanto, para a utilização desta metodologia é necessário um pré-requisito básico: a construção de uma biblioteca de BACs para que seja possível o isolamento de promotores do genoma da cana-de-açúcar.

A dificuldade no isolamento de promotores funcionais a partir de um genoma tão complexo como a cana-de-açúcar é evidente, sendo que os poucos promotores descritos neste híbrido já estão protegidos por patentes, o que faz com que seja necessária a autorização dos respectivos autores para o uso destas sequências. Para a comercialização de plantas cujos promotores já são patenteados é

necessário o pagamento de *royalties* aos proprietários das patentes, de modo a obter a permissão no país para a comercialização destas plantas.

Diante deste cenário de escassez de sequências regulatórias de uso público para monocotiledôneas, em especial de cana-de-açúcar, o isolamento e caracterização de novos promotores é de vital importância para o melhoramento genético via transgenia deste híbrido de grande valor econômico, assim como de outras espécies de interesse para o homem. É neste contexto que o presente trabalho se insere e propõe o desenvolvimento de uma nova abordagem para a identificação de promotores, utilizando como modelo o genoma da cana-de-açúcar.

3 OBJETIVOS

3.1 Objetivo geral

Estabelecimento de uma estratégia para clonagem customizada de promotores

3.2 Objetivos específicos

- a. Processamento do genoma para a produção de fragmentos de DNA compatíveis ao sistema de clonagem;
- b. Geração de bibliotecas de vetores de expressão: associação dos fragmentos de DNA preparados no item A a genes repórteres GUS e GFP;
- c. Busca de regiões regulatórias funcionais através da triagem da biblioteca de vetores por biobalística e análise da expressão do gene repórter GUS;
- d. Identificação e análise das regiões regulatórias funcionais;
- e. Validação dos promotores através de transformação permanente em sistema heterólogo e nativo.

4 MATERIAL E MÉTODOS

As atividades referentes ao presente trabalho foram conduzidas no laboratório de Biologia Molecular de Plantas – GaTE (Genomic and Transposable Elements), no Instituto de Biociências, Departamento de Botânica, Universidade de São Paulo.

As propostas descritas nos itens A a E relacionadas aos objetivos específicos serão mantidas na mesma ordem para a apresentação e descrição de Material e Métodos.

O termo “clonagem *customizada*” está relacionado ao delineamento de um processo de clonagem de fragmentos voltada a uma demanda de produtos específica. Os produtos são as regiões promotoras e o processo está relacionado com a clonagem e triagem de fragmentos de DNA para a identificação dos produtos. A demanda se caracteriza pela necessidade crescente de estratégias de identificação de novos módulos regulatórios valiosos como ferramentas biotecnológicas. A metodologia como cada uma destas etapas foi executada será descrita em detalhe nos itens abaixo. A seção 4.6 agrega todos os protocolos de soluções básicas, preparo de meios de cultura, extração de DNA e clonagem.

4.1 Processamento do genoma para a produção de fragmentos de DNA compatíveis ao sistema de clonagem

4.1.1 Extração do DNA genômico e hidrólise com endonucleases de restrição

A extração do DNA genômico de folhas de cana-de-açúcar (cultivares SP80-3280 e R570) foi realizada de acordo com Aljanabi et al. (1999). A hidrólise do DNA genômico das cultivares SP80-3280 e R570 foi padronizada para a combinação das enzimas *Pst*I e *Bam*HI (Biolabs). A reação de hidrólise foi realizada em banho-maria a 37 °C por 16 horas (Tabela 1).

Tabela 1 - Condições das reações de dupla-hidrólise do DNA genômico das cultivares SP80-3280 e R570.

	R570	SP803280
DNA (10 µg)	6,5 µl	11,1
Tampão 3 Biolabs (100x)	2,5 µl	2,5 µl
BSA (100x)	0,25 µl	0,25 µl
<i>Pst</i> I (20U/µl)	4 µl	4 µl
<i>Bam</i> HI (20U/µl)	4 µl	4 µl
Água	12,75µl	8,15µl
Total	30 µl	30 µl

4.1.2 Excisão e purificação de fragmentos de DNA de tamanho desejado e ligação de adaptadores

O DNA digerido foi migrado em gel de agarose 0,8%. Fragmentos de tamanho entre 2 e 4 kb foram excisados do gel com auxílio de um bisturi e a purificação dos fragmentos foi feita com o kit GFXTM PCR DNA and Gel Band Purificaton Kit (GE Healthcare).

A ligação dos adaptadores *adapt_attB1_Pst*I e *adapt_attB2_Bam*HI aos fragmentos de DNA purificado foi realizada utilizando a enzima T4 DNA Ligase (Fermentas) na proporção 3:1, em banho-maria a 25 °C por 16 horas, de acordo com instruções do fabricante (Tabela 2).

Tabela 2 - Condições das reações de ligação dos fragmentos de DNA digeridos das cultivares SP80-3280 e R570.

	R570	SP-803280
DNA (100 ng)	5,9 µl	8 µl
Tampão 10x T4 ligase	2 µl	2 µl
T4 ligase (5U/µl)	0,2 µl	0,2 µl
attB1 <i>Pst</i> I dupla-fita (1ng/µl)	1 µl	1 µl
attB2 <i>Bam</i> HI/ <i>Bgl</i> II dupla-fita (1ng/µl)	1 µl	1 µl
Água	0,9 µl	7,8 µl
Total	20 µl	20 µl

O adaptador adapt_attB1_ *Pst*I (5' TAC AAA AAA GCA GGC TTG CA 3') possui em uma de suas extremidades a sequência complementar ao sítio de restrição gerado pela enzima *Pst*I e na outra extremidade uma sequência parcial da região de recombinação attB1 do sistema Gateway. O mesmo ocorre para o outro adaptador adapt_attB2_ *Bam*HI (5' GAT CAC CCA GCT TTC TTG 3'), o qual possui em uma de suas extremidades a sequência complementar ao sítio de restrição gerado pela enzima *Bam*HI, e a outra extremidade abriga a sequência parcial à região de recombinação attB2.

4.1.3 Primeira amplificação dos fragmentos

Após a ligação de adaptadores nas extremidades digeridas dos fragmentos de DNA, esses fragmentos foram utilizados como molde para amplificação com a enzima *Elongase* (Invitrogen), em uma primeira reação de PCR (Tabela 3). Utilizou-se os iniciadores Adapt_ *Pst*I_attB1 (5' TAC AAA AAA GCA GGC TTG CA 3') e Adapt_ *Bam*/ *Bgl* attB2 low (5' CAA GAA AGC TGG GT 3'), os quais se anelam completamente ao adaptador, gerando milhares de sequências contendo a parte inicial da região de recombinação.

Tabela 3 - Condições da primeira reação de amplificação dos fragmentos de R570 e SP80-3280.

	R570	SP80-3280
Reação de ligação	10 µl	10 µl
Adapt PstI_attB1 (10µM)	2 µl	2 µl
Adapt Bam/Bgl_attB2 (10µM)	2 µl	2 µl
low		
dNTP (10mM)	1 µl	1 µl
Buffer A 5x	5 µl	5 µl
Buffer B 5x	5 µl	5 µl
Mix Elongase	2 µl	2 µl
Água	23 µl	23 µl
Total	50 µl	50 µl

Foi utilizado 10 µl da reação de ligação da seção 4.1.2 como molde para a amplificação. O programa utilizado no termociclador foi o seguinte: 94 °C por 30 S seguido por um ciclo de 10x de 94 °C (30 S), 55 °C (30 S) e 68 °C (6 min) seguido por 4 °C (∞). A reação de PCR foi purificada pelo kit GFX.

4.1.4 Segunda amplificação dos fragmentos

A purificação dos fragmentos de 2 a 4 kb obtidos na primeira reação de PCR foi realizada pelo Kit GFX (HE - Healthcare). Subsequentemente, 40 ng do produto purificado de PCR do cultivar R570 e SP80-3280 foi utilizado como molde para uma segunda reação de PCR (Tabela 4) com os iniciadores AttB1.1 (5' GGG GAC AAG TTT GTA CAA AAA AGC AGG CT 3') e AttB2.2 (5' GGG GAC CAC TTT GTA CAA GAA AGC TGG GT 3'). Esses iniciadores ancoram no início da sequência de recombinação do sistema Gateway e tiveram a função de completar o sítio de recombinação homóloga.

Tabela 4 - Condições da segunda reação de amplificação dos fragmentos de R570 e SP80-3280.

	R570	SP80-3280
Primeira PCR purificada (20ng/μl)	2 μl	2 μl
AttB1.1 (10μM)	2 μl	2 μl
AttB2.2 (10μM)	2 μl	2 μl
dNTP (10mM)	1 μl	1 μl
Buffer A	5 μl	5 μl
Buffer B	5 μl	5 μl
Elongase	2 μl	2 μl
Água	31 μl	31 μl
Total	50 μl	50 μl

As condições do programa utilizado no termociclador foram: 94 °C (30 s); 10 ciclos a 94 °C (30 s), 45 °C (30 s) a e 6 minutos a 68 °C seguido por 25 ciclos a 94 °C 30 (s), 65 °C 30 (s) e 6 minutos a 68 °C. O produto da reação de PCR foi purificado utilizando-se o kit GFX (HE - Healthcare).

4.2 Geração de bibliotecas de vetores de expressão: associação dos fragmentos de DNA preparados na seção 4.1 a genes repórteres GUS e GFP

4.2.1 Reações de Recombinação BP e LR do Sistema Gateway

A partir do momento em que foram gerados os fragmentos amplificados na segunda reação de PCR (os quais contêm em suas extremidades as regiões de recombinação completas), foi possível efetuar a reação de recombinação através do Sistema Gateway. A enzima BP Clonase é responsável por efetuar a primeira reação de recombinação homóloga, a qual resulta na inserção do fragmento contendo as extremidades de recombinação completas num vetor doador chamado pDONR221.

Já a enzima LR Clonase é responsável por realizar a segunda recombinação entre o vetor doador pDONR221 com o vetor de destino chamado pHGWFS7. Esta reação gera diversos clones que contêm uma diversidade de fragmentos de DNA inseridos em vetores de expressão pHGWFS7.

4.2.1.1 Reação de recombinação BP

A primeira reação de recombinação, denominada BP, foi feita por meio da utilização de 75 ng do produto purificado da segunda reação de PCR, tanto para a cultivar R570 como SP80-3280, 150 ng do vetor pDONR221 (Figura 4) e 2 µl de enzima BP ClonaseTM II Enzyme Mix (Invitrogen). Um controle positivo da reação de recombinação também foi adicionado conforme o manual Gateway (Cat no. 11789-020). As reações foram incubadas em banho-maria 25 °C por 16 horas, seguido pela adição de 1 µl de Proteinase K em cada uma das amostras e incubação de 10 min a 37 °C para desativar a enzima BP Clonase.

Figura 4 - Mapa do plasmídeo comercial pDONRTM221. Plasmídeo utilizado para o primeiro passo da clonagem utilizando o sistema Gateway Invitrogen®.

4.2.1.2 Transformação da reação de recombinação BP em *E.coli* seguido por miniprep de placa 96 poços e confirmação da obtenção dos clones recombinantes

Após a desativação da enzima BP Clonase, a reação BP foi utilizada para a transformação de *E.coli* (conforme seção 4.6.7) e os clones gerados tiveram seu DNA plasmidial extraído conforme seção 4.6.8. Para confirmar a inserção dos fragmentos em vetores pDONR221, foi feita uma dupla-hidrólise com as endonucleases de restrição EcoRI e EcoRV as quais permitem verificar o tamanho dos fragmentos inseridos no vetor pDONR221. Para tal, 300 ng de DNA plasmidial de cada clone foram utilizados em um volume final de 10 µl por reação, (tampão 10x

EcoRI NE buffer, 1 U/ μ l de EcoRI e EcoRV, BSA 10x). As reações foram incubadas em banho-maria a 37 °C por 4 horas e visualizadas em gel de agarose numa concentração 0,7% para estimativa do tamanho dos insertos.

4.2.1.3 Reação de recombinação LR

Após a estimativa do tamanho dos fragmentos inseridos no vetor pDONR221, foi feita a segunda reação de recombinação chamada LR, onde ocorre a transferência dos fragmentos inseridos no vetor pDONR221 para o vetor de expressão pHGWFS7, o qual possui os genes repórteres GUS e GFP (Figura 5).

Nesta segunda reação foram utilizados os clones obtidos pela primeira reação de recombinação BP. As reações de recombinação LR e o controle positivo da reação de recombinação foram feitas conforme o manual Gateway (Cat no 11791-020). Após a adição da enzima LR ClonaseTM II Enzyme Mix (Invitrogen) a reação foi incubada em banho-maria a 25°C por 16 horas. Após o tempo de incubação, foi adicionado 1 μ l de Proteinase K em cada amostra e estas foram incubadas a 10 min a 37 °C para desativar a enzima LR.

Figura 5 - Mapa do plasmídeo pHGWFS7. Plasmídeo utilizado para o segundo passo da clonagem utilizando o sistema Gateway Invitrogen®

4.2.1.4 Transformação de *E.coli* com as reações de recombinação LR seguido por extração de DNA plasmidial em placas de 96 poços e sequenciamento

Após a desativação da enzima, cada reação LR foi utilizada para a transformação de *E.coli* e os clones gerados tiveram seu DNA plasmidial extraídos. Esses clones foram sequenciados de acordo com o método descrito por Sanger et al. (1977) e a reação de sequenciamento foi preparada em água milli-Q autoclavada com 1 µl de DNA plasmidial, 0,1 µM de oligo attB2 (reverse), 2 µl de Big Dye Terminator v3.1 (Applied Biosystems) e 2 µl de tampão Big Dye totalizando 8 µl de volume final. A reação foi feita em um termociclador Veriti (Applied Biosystems), no qual as amostras foram submetidas a uma desnaturação inicial a 94 °C por 10 minutos e em seguida, a 40 ciclos de amplificação a 94 °C (45 S), 55 °C por (45 S) e 72 °C por 1 minuto seguido por extensão final a 72 °C por 10 minutos.

Logo em seguida, foi feita a precipitação da reação. Em cada poço foram adicionados 60 µl de isopropanol a 65% (v/v) e as amostras foram deixadas 25 minutos em repouso ao abrigo da luz e então foram centrifugadas a 4.000 rpm por 45 min. O sobrenadante foi descartado e o excesso absorvido por papel toalha, em seguida foi adicionado 170 µl de etanol 60% (v/v) em cada poço e as amostras foram centrifugadas a 10 min a 4.000 rpm e o sobrenadante descartado. As amostras foram levadas à estufa 37 °C por 25 min para que todos os poços ficassem secos.

As amostras foram ressuspendidas em 10 µl de formamida deionizada (Hi-Di™, Applied Biosystems) e imediatamente antes do sequenciamento, as amostras foram desnaturadas em termociclador a 90 °C por 4 min e imediatamente colocadas sobre gelo por 5 min. As amostras foram sequenciadas no 3730 DNA Analyzer (Applied Biosystems) e as sequências obtidas foram analisadas pelo 3730/3730xl Data Collection Software v 3.0 (Applied Biosystems).

4.3 Busca por regiões regulatórias funcionais através da triagem da biblioteca de vetores por biobalística e análise da expressão do gene repórter GUS

4.3.1 Ensaios de expressão transiente em epitélio de cebola

Para realizar o bombardeamento, partículas de tungstênio (48 mg, 1,1 µm de diâmetro, M17, Bio-Rad) foram lavadas duas vezes em 1 ml de etanol absoluto e ressuspendidas em 500 µl de glicerol 50%. As partículas foram sonicadas em aparelho de ultra-som Ultrasonic processor (Cole Parmer) por 1 minuto com os seguintes parâmetros: amplitude 70%; pulso ligado e temperatura 10 °C.

Para cada preparo a ser bombardeado, foi feito um tubo Eppendorf de 1,5 ml onde foram adicionados 22,5 µl de partículas em suspensão mantidas sob agitação constante em vortex (Daigger Vortex Gene 2). Foram acrescentados a este tubo em agitação constante, 22,5 µl de uma solução de DNA contendo no total 2 µg dos vetores de interesse (500 ng de cada um dos 4 clones). Em seguida, adicionou-se 37,5 µl de CaCl₂ gelado e 15 µl de espermidina (“free base” – Sigma) e os tubos permaneceram sob agitação constante por 5 minutos e após este tempo, foram

centrifugados 5 minutos à temperatura ambiente para sedimentação destas partículas. O sobrenadante foi retirado e as partículas foram lavadas em 200 µl de etanol 70% e ressuspendidas em 25 µl de etanol absoluto. Alíquotas de 6 µl de solução de etanol, DNA e partículas de tungstênio foi depositada em discos carreadores (disco de Kapton™ de 24 mm de diâmetro e 50 µm de espessura). Utilizando-se um acelerador de partículas produzido pelo CENARGEN (Dr. Elíbio Rech), os catáfilos de *Allium cepa* foram bombardeados com essas partículas sob vácuo de 27 polegadas de Hg (1 polegada= 2,5 cm) e 1100 psi de pressão, rompendo os discos de ruptura (disco de Kapton™ de 13,2 mm) e atingindo o material vegetal.

Cada preparo foi bombardeado em triplicata. Um controle positivo denominado pHGWFS7_sat6 (promotor ativo isolado de batata fusionado ao gene repórter GUS) e um controle negativo pHGWFS7 (vetor vazio) para cada experimento de biobalística foi incluído. O material vegetal foi deixado no escuro e em temperatura ambiente por cerca de 24 horas e, em seguida, colocado em imersão no ensaio histoquímico contendo x-gluc.

4.3.2 Ensaio histoquímico

O material vegetal bombardeado foi submerso em um tampão histoquímico ($\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ 100mM, $\text{Na}_2\text{EDTA} \cdot 2\text{H}_2\text{O}$ 10mM, $\text{K}_4\text{Fe}(\text{CN})_6 \cdot 3\text{H}_2\text{O}$ 0,5mM e triton 0,1% pH ajustado para 7,0) acrescido de x-gluc (dissolvido em DMSO ou dimetil formamida) na concentração final de 1mM e mantido na estufa a 37 °C por pelo menos 18 horas. Em seguida, retirou-se o tampão e etanol 70% foi adicionado para interromper a reação.

4.4 Identificação e análise das regiões regulatórias funcionais

4.4.1 Hidrólise enzimática dos clones identificados na análise de expressão transiente

Os clones que apresentaram expressão transiente positiva através do ensaio histoquímico foram submetidos à hidrólise enzimática com a endonuclease de

restrição *SacII*, a qual libera o fragmento recombinado por clivar o vetor em duas regiões flanqueantes ao local de recombinação e inserção do fragmento. O objetivo foi identificar o tamanho da região inserida no vetor de expressão pHGWFS7. Foram utilizados 1.5 µg de DNA plasmidial de cada clone, 1 µl de tampão 10x NE buffer 4, 0,5 µl de *SacII* e 4,5 µl de água milli-Q, totalizando 10 µl de volume final. As reações foram incubadas em banho-maria a 37 °C por 4 horas e visualizadas em gel de agarose numa concentração 0,7%.

4.4.2 Sequenciamento completo dos clones com expressão do gene GUS

Os clones que apresentaram expressão do gene GUS no ensaio de expressão transiente possuem fragmentos inseridos pela reação de recombinação do Sistema Gateway que variam entre 2,5kb a 3,3kb. Para o sequenciamento completo destes clones foi necessário desenhar iniciadores na região interna dos clones e realizar a técnica de *primer walking*.

A reação de sequenciamento e precipitação das amostras foram realizadas conforme a seção 4.2.1.4 A reação foi feita em um termociclador Veriti 96 well (Applied Biosystems), no qual as amostras foram submetidas a uma desnaturação inicial a 96 °C por 1 minuto e em seguida a 40 ciclos de amplificação a 96 °C por 1 min, 57 °C (20 s) e 60 °C por 3 minutos.

4.4.3 Análise *in silico* das putativas regiões promotoras identificadas

Após o sequenciamento completo dos clones que apresentaram expressão transiente positiva em epitélio de cebola, foi feita a busca por elementos *cis*-regulatórios em bancos de dados distintos de elementos regulatórios de plantas através do PLACE (HIGO et al., 1999) e Plantcare (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>), assim como análises de alinhamento nucleotídico através do Phytozome (<http://www.phytozome.net/>) e ncbi (<http://www.ncbi.nlm.nih.gov/>).

4.4.4 Extração de DNA plasmidial em pequena escala (miniprep) dos clones com expressão transiente positiva

Em fluxo laminar, uma colônia bacteriana contendo o plasmídeo foi inoculada em 2 ml de meio LB líquido com o antibiótico seletivo, as bactérias foram incubadas a 37 °C com agitação de 288 RPM por 18 horas. A cultura celular foi transferida para tubos Eppendorfs de 1,5 ml e sedimentadas por centrifugação a 12.000 g por 1 minuto a 4 °C e o sobrenadante descartado. As bactérias foram ressuspensas em 100 µl de solução P1 (glicose 50 mM, Tris 25 mM, EDTA 10 mM), logo em seguida em 200 µl da solução P2 (NaOH 10 N e SDS 20%) foi adicionada e misturada por inversão suave. A solução P3 (KoAc 5M e ácido acético glacial) gelada foi adicionada e misturada por 10 segundos, logo em seguida os tubos foram incubados no gelo por 5 minutos. Após a incubação os tubos foram centrifugados a 12.000 g por 5 minutos a 4 °C, sendo o sobrenadante transferido para novo eppendorf. Uma solução de fenol/clorofórmio na proporção 1:1 foi adicionada em novo eppendorff e os tubos foram agitados no vórtex seguidos por centrifugação a 12.000g por 2 minutos a 4 °C. O sobrenadante foi transferido cuidadosamente para novo eppendorf e o DNA precipitado com 2 volumes de etanol 100% gelado, seguidos por incubação a 2 minutos no gelo.

Os tubos foram novamente centrifugados a 12.000 g por 5 minutos a 4 °C, o sobrenadante foi descartado e 1ml de etanol 70% foi adicionado. Novamente os tubos foram centrifugados a 12.000 g por 5 minutos a 4 °C, o sobrenadante removido com pipeta e o pellet seco na speed-vac por cerca de 15 minutos. O DNA foi ressuspensado em 50 µl de TE/RNase (20 µg/ml) e sua integridade averiguada em gel de agarose 0,7%, assim como a quantificação do material extraído no Nanodrop.

4.4.5 Ensaios de expressão transiente em arroz

4.4.5.1 Obtenção de protoplastos de arroz (*Oryza sativa*)

Os protoplastos foram isolados da cultivar Kitaake (subespécie *Japonica*), cujas sementes foram colocadas para germinar em meio de cultura MS (ver seção 4.6.5.2). Plantas de 15 dias tiveram sua parte aérea cortada com bisturi em pedaços de 0,5 mm e imersa em 10 ml de solução enzimática (ver seção 4.6.4).

Foi aplicado vácuo de 28 mmHg por 1 hora na solução enzimática contendo o tecido cortado de arroz, seguido por 4 horas de incubação a 28 °C sob agitação de 56 rpm. Após este tempo, com o auxílio de uma pipeta (com ponteira cortada) foi removido toda solução enzimática e adicionado com cuidado 25 ml de meio W5 (ver seção 4.6.4).

O preparo foi incubado a 28 °C sob agitação de 98 rpm por 1 hora para liberação dos protoplastos e em seguida o preparo foi transferido para um novo becker contendo um filtro no qual foi possível reter os pedaços de tecido de arroz e permitir a passagem da solução contendo protoplastos. Esta solução foi filtrada através de um filtro de nylon BD Falcon de 100 µm de diâmetro (BD Biosciences) e a solução resultante foi transferida para 2 tubos Falcon de 50 ml (BD Falcon) e centrifugada a 1.000 rpm por 5 minutos numa centrífuga Sorvall Legend RT+ (Thermo Scientific). O sobrenadante foi descartado e 300 µl de meio de suspensão de protoplastos (ver seção 4.6.4) foram adicionados em cada tubo Falcon. Com o auxílio de uma pipeta (com ponteira cortada) foi retirado 10 µl desta solução de protoplastos e estes foram visualizados em uma câmara de Neubauer (PRECISS France) para o cálculo do número total de protoplastos viáveis. O número de células viáveis foi ajustado para 1.5×10^6 células/ml através da diluição com meio de suspensão para protoplastos (Figura 6).

Figura 6 - Esquematização do processo de obtenção de protoplastos viáveis de arroz utilizados na transformação transiente com PEG.

4.4.5.2 Transformação transiente de protoplastos de arroz por PEG

Após o cálculo de protoplastos viáveis, foi preparado em tubos eppendorf de 1,5 ml uma solução de 10 µg de DNA plasmidial na concentração de 1 µg/µl nos quais foram adicionados gentilmente 200 µl (1.5×10^5 células/ml) da solução de protoplastos que estavam em meio de suspensão. Em seguida, foram adicionados gentilmente 250 µl de solução PEG 40% (ver seção 4.6.4) em cada tubo e estes foram misturados por completo através de agitação suave e incubados à temperatura ambiente por 10 minutos.

Após este período foram adicionados 1 ml de meio W5 em cada tubo para diluir a solução de PEG e cada tubo teve seu conteúdo transferido para uma placa de cultura de tecidos (TPP), a qual foi envolvida em papel alumínio e incubada a 28 °C no escuro por 16 horas. Como controle positivo da transformação transiente por PEG foi utilizado o vetor pART7 HA GFP (Figura 7), o qual possui o gene repórter GFP fusionado a um promotor 35S e como controle negativo, o vetor pHGWS7 vazio (Figura 5). A estimativa da eficiência da transformação dos protoplastos foi feita de acordo com Bart et al. (2006).

Figura 7 - Mapa do vetor pART7 HA GFP (s65t) utilizado como controle positivo na transformação transiente de protoplastos de arroz.

4.4.5.3 Análise da expressão transiente através da visualização da expressão de GFP

A placa de cultura de tecidos que contém as amostras transformadas foi transferida para novos tubos eppendorf de 1,5 ml os quais foram centrifugados a 1.000 rpm por 3 minutos numa centrífuga de tubos MIKRO 22R (Uniscience). Com auxílio de uma pipeta com ponteira cortada, 10 µl de cada tubo eppendorf foram coletados e dispostos numa lâmina para análise em microscopia de fluorescência refletida (Olympus CKX41) e as fotos foram feitas pela câmera Olympus Q-Color 3™.

4.5 Validação da funcionalidade dos promotores através de transformação permanente em sistema nativo

4.5.1 Ensaio de expressão permanente em cana-de-açúcar

4.5.1.1 Transformação genética de calos e discos de cana-de-açúcar via biobalística

Esta etapa foi realizada em colaboração com o laboratório de Genômica e Biologia Molecular do Centro de Biotecnologia Agrícola Cebotec na ESALQ (USP), o qual tem como responsável a prof^a Dra. Helaine Carrer. Esta técnica modificada de transformação permanente em cana-de-açúcar foi desenvolvida no projeto de mestrado do aluno André Luiz Barbosa e está em processo de submissão pelo grupo da prof^a Helaine Carrer.

Segundo Gambley et al. (1993) a escolha do alvo a ser transformado é um dos fatores fundamentais neste processo. Dentre os testes realizados em cana, foi observado que os calos embriogênicos apresentaram bons resultados devido à facilidade em regenerar plantas de cana-de-açúcar, facilidade de manuseio no processo de seleção em meio de cultura e grande quantidade de material produzido para bombardeamento em curto período de tempo (BIRCH et al., 1995) e, portanto, foram escolhidos para a realização da transformação permanente. Além dos calos, discos de cana também foram selecionados para bombardeio (Figura 8).

Figura 8 - Esquema representativo do método de transformação genética por biobalística utilizando como tecidos-alvo calos e discos de cana-de-açúcar.

O acelerador de partículas utilizado para o bombardeamento foi o modelo PDS-1000/He *Biolistic*^R *Particle Delivery System* (BIO-RAD). A assepsia das partículas foi feita da seguinte forma: foram pesadas 25 mg de partículas de tungstênio em um tubo eppendorf e lavadas em 500 µl de etanol absoluto gelado. Este tubo contendo as partículas foi colocado em agitador por 6 minutos e incubado em banho-maria a 95 °C durante 1 hora. Centrifugou-se a 6797 x g por 2 segundos e a fase aquosa foi removida cuidadosamente. Adicionou-se 500 µl de etanol absoluto gelado e as partículas foram sonicadas em aparelho de ultra-som Ultrasonic processor (Cole Parmer) por 15 minutos com os seguintes parâmetros: amplitude 75%; 3 segundos ligado e 1 segundo desligado; temperatura 20 °C. Centrifugou-se a 6797 x g por 2 segundos e a solução aquosa foi removida. O precipitado foi lavado três vezes com 500 µl de ddH₂O (resfriada) e centrifugou-se a 6797 x g por 2 segundos com a remoção da solução aquosa em cada lavagem. O precipitado foi diluído em 500 µl de ddH₂O (resfriada).

Para a precipitação do DNA plasmidial nas partículas de tungstênio, foram adicionados 50 µl de suspensão de tungstênio, 10 µg de DNA plasmidial (na co-transformação utilizou-se 5 µg de cada plasmídeo), 50 µl de CaCl₂ 2,5 M gelado (-20 °C), 20 µl de espermidina 0,1 M. A suspensão foi incubada em agitador durante 30

minutos a 4 °C. Foi adicionado 200 µl de etanol absoluto gelado e as amostras foram centrifugadas a 10621 x g durante 1 segundo e a fase aquosa foi retirada. O precipitado foi ressuspendido em 200 µl de etanol absoluto gelado e centrifugado a 10621 x g durante 1 segundo e a fase aquosa foi retirada. Esta etapa de lavagem do precipitado foi repetida por três vezes. O precipitado foi diluído em 30 µl de etanol absoluto gelado. Alíquotas de 5 µl de partículas com DNA foram utilizadas nos disparos.

Todo o material utilizado: membranas carreadoras, telas de retenção, membranas de ruptura e suportes para membranas carreadoras (BIO-RAD) foram desinfetados em etanol absoluto por 3 minutos e secos em condições assépticas (fluxo laminar). Para os disparos, 5 µl da suspensão de partículas com o DNA foram colocados no centro da membrana carreadora.

As condições de bombardeamento para todos os experimentos de transformação foram as seguintes:

- distância do alvo: a placa de Petri contendo o tecido alvo foi colocada em prateleira com distância de 11 cm da membrana carreadora;
- pressão de disparo: membrana de ruptura de 1100 psi;

4.5.1.2 Co-transformação com o vetor *pHA9*

No preparo de precipitação de partículas foi feita a co-transformação de dois plasmídeos: 5 µg do plasmídeo *pHA9*, (figura 9) o qual contém o gene *neo* (neomicina fosfotransferase) fusionado sob controle do promotor *Ubi* de milho e 5 µg do vetor isolado neste trabalho: *pHGWFS7_F6*.

Figura 9 - Esquema representativo do plasmídeo pH A9 o qual contém o cassete de expressão do gene *neo* que codifica a neomicina fosfotransferase (NPTII) sob controle do promotor da ubiquitina do milho (*Ubi-1*) e o terminador da nopalina sintetase T-Nos.

As setas indicam a posição dos primers *nptII* (Forward) e *nptII* (Reverse). Sítios de restrição: B, *Bam*HI; BgII; E, *Eco*RI; H, *Hind*III; N, *Nae*I; P, *Pst*I; Sa, *Sal*I; Sp, *Sph*I; T, *Tth*1; X, *Xba*I.
 Fonte: Melotto-Passarini, 2009.

Como controle positivo do bombardeamento foi utilizado o vetor pFF19G_GUS, (figura 10) o qual contém o cassete de expressão do gene repórter GUS sob controle do promotor CaMV 35S.

Figura 10 - Esquema representativo do plasmídeo pFF19G o qual contém o cassete de expressão do gene repórter *uid-A* (GUS) sob o controle do promotor *CaMV35S* e o terminador 35S *polyA*.

Sítios de restrição: B, *Bam*HI; E, *Eco*RI; H, *Hind*III; K, *Kpn*I; N, *Nco*I; P, *Pst*I; S, *Sac*I; Sa, *Sal*I; Sm, *Sma*I; X, *Xba*I.
 Fonte: Melotto-Passarini, 2009.

4.5.1.3 Seleção, regeneração e obtenção de plantas transgênicas

Os calos e discos bombardeados foram transferidos para frascos contendo meio MRP acrescido de 30 mg.L⁻¹ de antibiótico geneticina. Este material foi incubado a 27 °C e fotoperíodo de 16 horas e a cada 25 dias foi feita a transferência dos calos embriogênicos e discos para novo meio de cultura contendo 30 mg.L⁻¹ de antibiótico geneticina.

No total foram feitas 3 transferências dos calos e discos de cana-de-açúcar bombardeados para novo meio MRP sólido contendo geneticina, com intervalo de um mês entre as transferências. As plântulas que enraizaram foram individualmente transferidas para novo meio MRP sólido contendo geneticina e posteriormente analisadas por métodos moleculares.

4.5.1.4 Extração de DNA total e análises de PCR

A extração de DNA total de folhas de plântulas de cana-de-açúcar resistentes a geneticina foi realizado de acordo com Aljanabi et al. (1999). Para a confirmação da presença do transgene *neo* (presente no vetor pHA9) cujo tamanho é de 410 pb foram utilizados os iniciadores *nptII* Forward (5' GAG GCT ATT CGG CTA TGA CTG 3') e *nptII* Reverse (5' TCG ACA AGA CCG GCT TCC ATC 3'). Para cada reação de PCR foi utilizado cerca de 130 ng de DNA total, 1 µl de cada iniciador a 10 µM, 0,5 µl de dNTP a 10 mM, 2,5 µl de tampão *Dream taq* (Fermentas) 10x, 0,12 µl de *Dream taq* e água para completar 25 µl do volume total de cada reação. O protocolo para amplificação do gene *neo* foi: desnaturação inicial por 5 minutos a 95 °C; 35 ciclos: 1 minuto a 95 °C, 1 minuto a 60 °C, 1 minuto a 72 °C; extensão final de 5 minutos a 72 °C. Os pares de iniciadores utilizados para a confirmação da presença do gene *neo* estão descritos abaixo:

Para detecção da presença do outro vetor co-transformado, pHGWFS7_F6, foram feitas reações de PCR das plantas supostamente transgênicas de cana com iniciadores *eGFP* Forward (5' CCT ACG GCA AGC TGA CC 3') *eGFP* Reverse (5' CGC TTC TCG TTG GGG 3') nas mesmas condições das reações realizadas para amplificar o gene *neo*.

O gene GFP está presente à jusante da sequência F6 e o tamanho amplificado é de 500pb. O protocolo para amplificação deste gene foi: desnaturação inicial por 5 minutos a 95 °C; 35 ciclos: 1 minuto a 95 °C, 1 minuto a 55 °C, 50 (S) a 72 °C; extensão final de 4 minutos a 72 °C. Em todas as reações de PCR foram utilizados como controles: DNA plasmidial como controle positivo (pHA9 ou pHGWFS7 vazio), DNA da planta não transformada como controle negativo e ddH₂O (controle da reação).

4.6 Material vegetal, soluções estoque, meios de cultura e procedimentos de biologia molecular

4.6.1 Material vegetal

Os cultivares de cana-de-açúcar utilizados neste trabalho foram: SP80-3280 e R570. O primeiro foi desenvolvido pelo Centro de Tecnologia Canavieira (Piracicaba-SP), amplamente comercializado como sendo uma cultivar produtiva, robusta e tolerante à condições adversas de solo e clima. O segundo é um cultivar moderno resistente à ferrugem, derivado do cruzamento entre dois cultivares modernas: H32-8560 e R445. Foi desenvolvido pelo CERF (Centred'Essai de recherche et de formation, Réunion) e também foi utilizado na preparação de uma biblioteca comercial de clones genômicos de BAC de cana-de-açúcar (Tomkins et al., 1999).

4.6.2 Soluções estoque

As soluções foram preparadas de acordo com os procedimentos descritos em Sambrook et al. (1989).

- água destilada esterilizada por autoclavagem (120 °C 30 minutos): utilizada para a preparação de todas as soluções.
- Dodecil sulfato de sódio (SDS) 20%: O SDS deve ser dissolvido em água destilada previamente esterilizada em autoclave (120 °C, 20 minutos).
- RNaseA (10 mg/mL): A ribonuclease A foi dissolvida em solução de Tris pH 7,5 (10 mM) e Cloreto de Sódio (15mM) utilizando as soluções citadas. Em seguida, é

incubada a 100 °C por 15 minutos e mantida à temperatura ambiente até esfriar e ser estocada a -20 °C.

- 5-bromo-4-cloro-3-indolyl (β -D-glucuronide, cyclohexylammonium salt) conhecido como x-gluc: 50mg dissolvidos em 1 ml de dimetilformamida (50 mg/mL).
- Hidróxido de sódio – NaOH 10N: Pastilhas de hidróxido de sódio (400 g/L) são dissolvidos em água destilada previamente esterilizada.
- Ácido etilenodiamino tetra-acético – EDTA 0,5 M (pH 8,0): O EDTA (186,1 g/L) é dissolvido em água ultra-filtrada e pastilhas de hidróxido de sódio são adicionadas para acertar o pH. Esterilizado em autoclave (120 °C, 15 minutos).
- Cloreto de sódio – NaCl 5 M: O cloreto de sódio (292,2 g/L) é dissolvido em água ultra-filtrada e esterilizado em autoclave (120 °C, 15 minutos).
- Cloreto de magnésio – $MgCl_2$ 1 M: O cloreto de magnésio 6 H_2O (202,3 g/L) é dissolvido em água ultra filtrada e esterilizado em autoclave (120 °C, 15 minutos).
- Tampão TE 1x: Tris pH 8,0 (10 mM) e EDTA (1 mM). Preparado a partir das soluções estoque de Tris e EDTA, utilizando água filtrada previamente esterilizada em autoclave (120 °C, 20 minutos).
- Brometo de etídeo (10 mg/mL): Preparado em água destilada previamente esterilizada.

4.6.3 Antibióticos

- Kanamicina 50 mg/mL – Dissolvido em água MilliQ
- Higromicina 50 mg/ml – Dissolvido em água MilliQ
- Gentamicina 50 mg/mL – Dissolvido em água MilliQ
- Geneticina 50 mg/mL – Dissolvido em água MilliQ
- Streptomicina 20 mg/mL – Dissolvido em água MilliQ
- Spectinomicina 50 mg/mL – Dissolvido em água MilliQ
- Tetraciclina 20 mg/mL – Dissolvido em DMSO

Todas as soluções foram preparadas e esterilizadas em filtro Millipore 0,22 μ m (Corning).

4.6.4 Soluções para isolamento de protoplastos:

- Solução enzimática para preparo de protoplastos: 1,5 g de celulase, 0,3 g de macerozima e 12 g de sorbitol são dissolvidos num volume final de 100 ml de água destilada e alíquotas de 10 ml são distribuídos em tubos Falcon de 15 ml (BD Falcon) e estocados a -20 °C.

- Solução PEG 40% (Polietileno glicol 4'000 SIGMA): 8 g de PEG, 1,45 g de Manitol e 0,29 g de cloreto de cálcio dihidratado são dissolvidos em um volume final de 20 ml contendo água ultra-filtrada e armazenados num tubo Falcon de 50 ml (BD Falcon) e estocado a -20 °C.

- Meio W5: 4,5 g de cloreto de sódio; 9,2 g de cloreto de cálcio; 0,1875 g de cloreto de potássio 5mM e 0,1952 g de MES foram dissolvidos em 500 ml de água ultra-filtrada e o pH ajustado para 5.7 com KOH 1M, autoclavado e estocado a 4 °C.

- Meio de suspensão de protoplastos: 18,22 g de manitol, 0,735 g de cloreto de cálcio, 0,244 g de MES foram dissolvidos em 250 ml de água ultra-filtrada e o pH foi ajustado para 5.7 com KOH 1M, a solução foi autoclavada e estocada a 4 °C.

4.6.5 Meios de cultura

4.6.5.1 Meio para crescimento de *Escherichia coli*

A cepa DH10B (Invitrogen) foi mantida e repicada em meio Luria-Bertani (LB) sólido (Sambrook et al., 1989) e quando transformadas com algum plasmídeo utilizou-se o antibiótico seletivo adequado.

Meio LB (pH5,7): Triptona (1%), Extrato de levedura (0,5%), Cloreto de Sódio (1%) e Ágar (Himedia). O pH foi corrigido com ácido clorídrico (HCl) ou hidróxido de sódio (NaOH).

4.6.5.2 Meio de cultura para germinação de sementes de *Oryza sativa* (arroz)

MS (pH5,7): Sais MS (*) Sigma M-5524 (4,3 g/L), Sacarose (30 g/L), Vitamina Murashigue & Skoog 1000x (1 mL/L) e Phytigel (2,5 g/L).

(*) São os sais do meio MS descrito por Murashige e Skoog (1962), com redução na concentração dos macronutrientes (metade da concentração original).

4.6.5.3 *Meio osmótico para calos de cana-de-açúcar*

Sais MS Sigma M-5524 (4,3 g/L), ácido cítrico (50 mg/L), 2,4D (Ácido 2,4-dicloro-fenóxi-acético) (3 mg/L), cinetina (0,1 mg/L), manitol (73 g/L) e sorbitol (73 g/L).

4.6.5.4 *Meio para seleção de calos e discos transformantes de cana-de-açúcar*

MRP G30 pH: 5,7: Sais MS Sigma M-5524 (4,3 g/L), Sacarose (30 g/L), NAA (ácido naftaleno acético) (3,72 mg/L), BAP (6-benzilaminopurina) (0,45 mg/L) e ácido cítrico (50 mg/L).

MS30 G30: pH: 5,7: Sais MS Sigma M-5524 (4,3 g/L), Sacarose (30 g/L),), Vitamina Murashigue e Skoog 1000x (1 mL/L) e Phytigel (2,3 g/L). Após a esterilização dos meios acima, adicionou-se geneticina para a seleção dos calos e discos transformantes de cana-de-açúcar.

Todos os meios foram preparados em água destilada e esterilizados em autoclave a 120 °C por 15 minutos.

4.6.6 **Bactérias eletrocompetentes**

Escherichia coli cepa: DH10B – genótipo F- mcrA Δ (mrr-hsdRMS-mcrBC) Φ 80d lacZ Δ MIS Δ lacX74 enA1 recA1 deoR Δ (ara,leu) 7697 araD139 galU galK nupG rpsL foi utilizada para o preparo de bactérias eletrocompetentes. As bactérias foram mantidas em placas de meio sólido LB e repicadas mensalmente. Para a conservação da cepa por longos períodos, foram feitos estoques de glicerol, os quais são obtidos adicionando-se 500 μ l da cultura líquida bacteriana crescida por no máximo 12 horas em LB, a 500 μ l de glicerol 50%, e armazenados a -70 °C.

Para a obtenção de bactérias eletrocompetentes, um inóculo de colônia isolada de DH10B foi colocado para crescer em 4ml de meio LB e a cultura foi mantida sob agitação durante 18 horas a 37 °C. Após este período, o inóculo foi transferido para 400ml de meio SOB seguido por nova incubação a 37 °C durante 2 h até que a medida de absorbância em 600nm atingisse uma densidade óptica (DO) em torno de 0,60. Em seguida as bactérias foram incubadas no gelo por 20 min, com subsequente centrifugação a 7.000g por 5 min, e o precipitado foi ressuspensionado em água milli-Q. O processo foi repetido duas vezes e o precipitado foi ressuspensionado em glicerol 10%. Alíquotas de 50 µl foram distribuídas em tubos de 1,5 mL e armazenadas em freezer -70 °C. Toda a manipulação foi realizada em câmara de fluxo de ar laminar, com uso de material estéril.

4.6.7 Transformação de bactérias eletrocompetentes

Uma alíquota de 50 µl de bactéria eletrocompetente foi adicionada em 1 µl de reação de recombinação diluído em 4 µl de água e acondicionada em uma cuveta de eletroporação (0,2 cm de largura) previamente resfriada no gelo. O eletroporador Gene Pulser® (Biorad) foi ajustado às seguintes condições: capacitância: 25µF; resistência: 200 Ω e 800 Ω; voltagem: 1,80KV.

Imediatamente após a eletroporação, foi adicionado 200 µl de meio SOC na cuveta e todo o preparado foi transferido cuidadosamente para um falcon de 15 ml (BD Falcon) o qual já possuía 750 µl de meio SOC e levado ao shaker 37 °C, rotação de 200 rpm por 30 minutos. Em seguida, 200 µl do preparado foi plaqueado em meio LB sólido contendo o antibiótico de seleção apropriado. As colônias capazes de crescer no meio seletivo foram selecionadas e tiveram seu material plasmidial extraído em uma miniprep de placa 96 poços.

4.6.8 Extração de DNA plasmidial em placa de 96 poços

Em fluxo laminar, 96 colônias de bactérias isoladas foram coletadas e dispostas individualmente em poços de uma placa de crescimento de 96 poços, contendo 1 ml de meio LB líquido e antibiótico de seleção em cada poço. A placa de crescimento foi levada ao shaker 37 °C numa rotação de 200 rpm por 16 horas.

Após a incubação, a placa foi centrifugada a 4.000 rpm por 10 minutos e o sobrenadante descartado, seguido pela adição de 240 µl de solução I (glicose 50 mM, Tris 25 mM, EDTA 10 mM). A placa foi selada com selos (Thermo Scientific) e a ressuspensão foi feita num vortex (Fischer Scientific) seguido por centrifugação a 4.000 rpm por 10 min. O sobrenadante foi descartado e 80 µl de solução I foi adicionada à placa seguido pela ressuspensão das bactérias num vortex novamente. Em sequência foi adicionado 2,5 µl de RNase 10 mg/ml em uma placa de fundo em U (Costar®) e 50 µl da cultura de bactérias da placa foi adicionado nesta placa em fundo em U contendo RNase. Logo em seguida, foi adicionado 50 µl de solução II (NaOH 10N e SDS 20%) e a placa foi selada e invertida 10 vezes seguida por incubação a 10 min em temperatura ambiente. Passado o tempo de incubação, foi adicionado 50 µl de solução III (KoAc 5M e ácido acético glacial) e a placa foi selada e invertida 10 vezes seguida por incubação em temperatura ambiente por 10 min. Após o tempo de incubação, foi retirado o selo da placa e esta foi incubada na estufa a 90 °C por 30 minutos, seguido por 10 minutos de incubação no gelo e centrifugação a 4.000 rpm por 8 minutos. Foi preparado o ajuste de uma placa de filtro (Millipore) sob uma placa nova de fundo "V" (Costar®) e todo sobrenadante advindo da centrifugação foi colocado nesta nova placa de filtro e centrifugado a 4.000 rpm por 6 minutos. A placa de filtro foi descartada e 110 µl de isopropanol 100% foi adicionado à placa de fundo em V, esta foi selada, invertida 20 vezes e centrifugada a 4.000 rpm por 45 min. O sobrenadante foi descartado e 200 µl de etanol 70% gelado foi adicionado seguido por centrifugação a 4.000 rpm por 5 minutos e o sobrenadante foi descartado novamente. A placa foi incubada numa estufa 37 °C por 20 minutos e o DNA plasmidial foi ressuspendido em 40 µl de água.

5 RESULTADOS E DISCUSSÃO

5.1 Processamento do genoma para a produção de fragmentos de DNA compatíveis ao sistema de clonagem

Para uma estratégia de clonagem customizada de sequências regulatórias do genoma da cana-de-açúcar, o DNA genômico das cultivares SP80-3280 e R570 foi extraído. Após verificar a integridade do DNA, realizou-se a hidrólise do DNA genômico das duas cultivares com uma combinação de duas endonucleases de restrição de corte raro: *Pst*I/*Bam*HI (Figura 11). A escolha da combinação de enzimas foi baseada no conhecimento de que *Pst*I é sensível à metilação, portanto, não cliva as regiões metiladas do DNA. Já a enzima *Bam*HI não é sensível à metilação e é uma enzima que digere regiões ricas em AT, clivando regiões metiladas e não metiladas do genoma.

Espera-se que a combinação deste par de endonucleases gere um padrão de fragmentos diversificados que possam abrigar regiões potencialmente promotoras de expressão gênica. A metilação da molécula de DNA geralmente está associada à inativação gênica, (KLOSE et al. 2006) (MATZKE et al. 1995) porém ainda há fenômenos associados à metilação do DNA que não são integralmente compreendidos (SNUSTAD; SIMMONS, 2008).

Figura 11 - A: Hidrólise do DNA genômico de cana-de-açúcar das cultivares R570 e SP80-3280 através da combinação de duas endonucleases de restrição: *Pst*I/*Bam*HI.

Passo 2

A: 1: Hidrólise de R570 com *Pst*I/*Bam*HI; 2: Hidrólise de SP80-3280 com *Pst*I/*Bam*HI e L: Ladder 1 kb (Fermentas) 500 ng. 500 ng aplicado nos poços 1 e 2. Gel 0,8%, TAE 0,5x, 30mA; B: Análise dos fragmentos purificados a partir da hidrólise-dupla do DNA genômico. 1: R570; 2: SP80-3280; L: Ladder 1 kb (Fermentas). Gel 0,8%, TAE 0,5x, 30 mA. 60ng aplicado nos poços 1 e 2; C: Marcador Ladder de 1 kb GeneRuler Fermentas.

Após a hidrólise do DNA genômico, foi feita a excisão e purificação de fragmentos em torno de 2 a 4 kb do gel de agarose (Figura 11). De acordo com An; e Kim (1993), o tamanho da região regulatória é difícil de ser estimada e portanto, optamos por iniciar com fragmentos em torno de 2 kb para a caracterização de promotores. Outros trabalhos isolaram promotores entre 2 a 2,5 kb (DE PATER et al., 1992; HE et al., 2009), sendo que no presente trabalho optamos por isolar fragmentos de 2 a 4 kb.

Os adaptadores *adapt_attB1_Pst*I e *adapt_attB2_Bam*HI foram ligados nas extremidades digeridas do DNA purificado. Nesta etapa, é importante destacar que a ligação dos adaptadores fornece uma clonagem direcional por definir as extremidades 5' e 3' que serão amplificadas. Após a ligação, foi feita a primeira reação de PCR para a geração de fragmentos contendo a parte inicial da região de recombinação (Figura 12), essencial para a reação de recombinação no sistema Gateway.

Figura 12 - A: Análise dos fragmentos gerados na primeira reação de PCR os quais contêm a parte inicial da região de recombinação, essencial para a reação de recombinação Gateway.

Passo 4

A: 1: R570; 2: SP80-3280; 3: Controle negativo da reação de PCR; L: Ladder 1kb (Fermentas). 4 µl aplicado nos poços 1 a 3; B: Análise dos fragmentos purificados após a primeira reação de PCR contendo a parte inicial da região de recombinação. 1: R570; 2: SP80-3280; L: Ladder 1kb (Fermentas). 70 ng aplicado no poço 1 e 50 ng no poço 2; C: Marcador Ladder de 1 kb GeneRuler Fermentas. Gel 0,8%, TAE 0,5x, 30mA.

Uma etapa importante neste processo de amplificação dos fragmentos aleatórios foi gerar uma quantidade concentrada de fragmentos na região definida entre 2 a 4 kb, tanto na primeira reação de PCR como na segunda. Este passo é fundamental para definir a diversidade e o tamanho de fragmentos que serão inseridos em vetores de expressão pelo Sistema Gateway, através das reações de recombinação BP e LR. Estas reações devem, portanto, gerar o maior número de clones contendo fragmentos distintos com tamanhos entre 2 a 4 kb.

Na Figura 12 é possível observar uma concentração de fragmentos amplificados que contêm a parte inicial da região de recombinação attB no tamanho desejado entre 2 a 4 kb, mostrando que a estratégia utilizada para isolar essa região foi bem sucedida. Estes fragmentos foram purificados para serem utilizados como molde para a segunda reação de PCR com os iniciadores AttB1.1 e AttB2.2 para completar o sítio de recombinação homóloga.

Figura 13 - A: Análise dos fragmentos gerados na segunda reação de PCR contendo a região de recombinação completa.

Passo 5

A: 1: R570; 2: SP80-3280; L: Ladder 1 kb (Fermentas). Gel 0,8%, TAE 0,5x, 30 mA. 4 µl aplicado nos poços 1 e 2. B: 1: R570; 2: SP80-3280; L: Ladder 1 kb (Fermentas). Gel 0,8%, TAE 0,5x, 30 mA. 80 ng aplicado nos poços 1 e 2.

A Figura 13 ilustra a segunda amplificação por PCR e purificação dos amplificados. Nessas reações, foi obtida uma grande quantidade de fragmentos purificados entre 2 a 4 kb das duas cultivares (R570 e SP80-3280) de cana-de-açúcar. Portanto, a etapa de obtenção de fragmentos de DNA que possuem regiões de recombinação em suas extremidades foi realizada com sucesso para as duas cultivares de cana-de-açúcar.

5.2 Geração de bibliotecas de vetores de expressão: associação dos fragmentos de DNA preparados no item 5.1 a genes repórteres GUS e GFP

5.2.1 Clonagem dos fragmentos de DNA em vetores intermediários pDONR221 através da reação de recombinação BP

A partir do momento em que foram gerados os fragmentos de DNA que possuem a região de recombinação em suas extremidades, esses foram submetidos à primeira reação de recombinação do Sistema Gateway. Uma única reação de recombinação BP foi realizada para a geração de clones que contêm fragmentos de 2 a 4kb da cultivar R570 inseridos num vetor intermediário denominado pDONR221. O mesmo foi feito para a cultivar SP80-3280.

O passo seguinte foi a transformação de células de *E.coli* com as duas reações de recombinação BP seguido pela seleção dos clones recombinantes para a extração do DNA plasmidial. Isso permitiu a geração de duas placas de 96 poços: Placa 1 (clones de SP80-3280) e Placa 2 (clones de R570). Estas duas placas abrigam fragmentos diversificados da cultivar SP80-3280 e R570 inseridos em vetores pDONR221 respectivamente.

Para uma análise da integridade desses clones, no total foram visualizados em um gel de agarose 8 clones da miniprep da Placa 1 e 8 clones da Placa 2 (Figura 14).

Figura 14 - Análise da integridade dos vetores pDONR221 recombinados contendo fragmentos aleatórios de R570 (Placa 1) e SP80-3280 (Placa 2).

Poços: 1 a 8: clones recombinados de R570; 9 a 16: clones recombinados de SP80-3280; 17: vetor pDONR221 vazio; L: Ladder 1 kb (Fermentas). Gel 0,7%, TAE 0,5x, 70 mA. 300 ng aplicado nos poços 1 a 17.

Em cada clone observado na Figura 14, é possível identificar uma diversidade de bandas nos poços 1 a 16, as quais estão relacionadas com as diferentes e complexas conformações do DNA plasmidial. No entanto, esta variedade de formas não compromete a confirmação da integridade dos vetores extraídos por miniprep.

Os 16 clones recombinantes foram analisados com relação ao tamanho dos fragmentos inseridos no vetor pDONR221 através da reação de recombinação BP. A determinação do tamanho dos fragmentos inseridos no vetor pDONR221 foi feita por dupla-hidrólise com as enzimas de restrição *EcoRI* e *EcoRV* e visualização em gel de agarose. A análise dos resultados permitiu constatar uma variação no tamanho dos fragmentos inseridos no vetor pDONR221 entre 300 a 3300 pb, comprovando desta forma, que a padronização das reações de PCR e seleção de fragmentos em torno de 2 a 4 kb foi feita de maneira adequada. Esta importante padronização permitiu resgatar nas duas reações de recombinação realizadas os clones na faixa do tamanho desejado. No entanto, também foram detectados fragmentos recombinados abaixo do tamanho esperado, possivelmente devido ao fato de que a purificação de fragmentos em torno de 2 a 4kb através do gel não elimina completamente a possibilidade de alguns fragmentos de tamanho menor serem recuperados também.

5.2.2 Clonagem dos fragmentos de DNA genômico em vetores de expressão os quais abrigam os genes repórteres GUS e GFP

De posse de clones que abrigam regiões genômicas de cana-de-açúcar no vetor intermediário pDONR221, o passo seguinte foi transferir estas regiões para o vetor binário pHGWFS7 através de uma reação de recombinação denominada LR. No vetor pHGWFS7, os genes repórteres GUS e GFP estão posicionados à jusante da região em que o fragmento aleatório irá se inserir, de modo que a sequência inserida irá regular a expressão destes genes repórteres.

Foram selecionados dois conjuntos de clones da placa 1, os quais foram submetidos à reação de recombinação LR em duas reações independentes. O primeiro conjunto continha 41 clones e o segundo 42 clones provenientes de SP80-3280 (Figura 15). Para R570 foram selecionados dois conjuntos de 48 clones da placa, os quais foram submetidos à reação de recombinação LR em reações independentes (Figura 16).

Figura 15 - Placa 1 abriga clones de SP80-3280 gerados na primeira reação de recombinação BP.

PLACA 1: CLONES SP80-3280

PLACA SP80-3280	coluna 1	coluna 2	coluna 3	coluna 4	coluna 5	coluna 6	coluna 7	coluna 8	coluna 9	coluna 10	coluna 11	coluna 12
Linha A		Clone 7	Clone 14	Clone 21	Clone 28	Clone 35	Clone 1	Clone 8		Clone 21	Clone 27	Clone 35
Linha B	Clone 1		Clone 15	Clone 22	Clone 29	Clone 36	Clone 2	Clone 9	Clone 15		Clone 28	Clone 36
Linha C	Clone 2	Clone 8		Clone 23	Clone 30	Clone 37	Clone 3	Clone 10	Clone 16	Clone 22	Clone 29	Clone 37
Linha D	Clone 3	Clone 9	Clone 16		Clone 31	Clone 38	Clone 4	Clone 11		Clone 23	Clone 30	Clone 38
Linha E		Clone 10	Clone 17	Clone 24		Clone 39	Clone 5	Clone 12	Clone 17	Clone 24	Clone 31	Clone 39
Linha F	Clone 4	Clone 11	Clone 18	Clone 25	Clone 32		Clone 6	Clone 13	Clone 18	Clone 25	Clone 32	Clone 40
Linha G	Clone 5	Clone 12	Clone 19	Clone 26	Clone 33	Clone 40		Clone 14	Clone 19	Clone 26	Clone 33	Clone 41
Linha H	Clone 6	Clone 13	Clone 20	Clone 27	Clone 34	Clone 41	Clone 7		Clone 20		Clone 34	Clone 42

41 clones

Recombinação LR	SP80-3280
pDONR + 2 a 4kb	6,6 µl (150ng)
pHGWFS7	1,15µl (150ng)
Tampão TE	0,26µl
Enzima LR Clonase	2µl
Total	10µl

42 clones

Recombinação LR	SP80-3280
pDONR + 2 a 4kb	6,41 µl (150ng)
pHGWFS7	1,15µl (150ng)
Tampão TE	0,44µl
Enzima LR Clonase	2µl
Total	10µl

Clones de SP80-3280 foram utilizados para a segunda reação de recombinação LR. As cores amarelo e azul identificam os dois conjuntos de clones utilizados em duas reações LR independentes. As caixas em branco são amostras não utilizadas na reação LR.

Figura 16 - Placa 2 abriga clones de R570 gerados na primeira reação de recombinação BP.

PLACA 2: CLONES DE R570

PLACA R570	coluna 1	coluna 2	coluna 3	coluna 4	coluna 5	coluna 6	coluna 7	coluna 8	coluna 9	coluna 10	coluna 11	coluna 12
Linha A	Clone 1	Clone 9	Clone 17	Clone 25	Clone 33	Clone 41	Clone 1	Clone 9	Clone 17	Clone 25	Clone 33	Clone 41
Linha B	Clone 2	Clone 10	Clone 18	Clone 26	Clone 34	Clone 42	Clone 2	Clone 10	Clone 18	Clone 26	Clone 34	Clone 42
Linha C	Clone 3	Clone 11	Clone 19	Clone 27	Clone 35	Clone 43	Clone 3	Clone 11	Clone 19	Clone 27	Clone 35	Clone 43
Linha D	Clone 4	Clone 12	Clone 20	Clone 28	Clone 36	Clone 44	Clone 4	Clone 12	Clone 20	Clone 28	Clone 36	Clone 44
Linha E	Clone 5	Clone 13	Clone 21	Clone 29	Clone 37	Clone 45	Clone 5	Clone 13	Clone 21	Clone 29	Clone 37	Clone 45
Linha F	Clone 6	Clone 14	Clone 22	Clone 30	Clone 38	Clone 46	Clone 6	Clone 14	Clone 22	Clone 30	Clone 38	Clone 46
Linha G	Clone 7	Clone 15	Clone 23	Clone 31	Clone 39	Clone 47	Clone 7	Clone 15	Clone 23	Clone 31	Clone 39	Clone 47
Linha H	Clone 8	Clone 16	Clone 24	Clone 32	Clone 40	Clone 48	Clone 8	Clone 16	Clone 24	Clone 32	Clone 40	Clone 48

48 clones

Reação LR	R570
pDONR + 2 a 4kb	5,2µl (150ng)
pHGWFS7	0,65µl (150ng)
Tampão TE	2,15µl
Enzima LR Clonase	2µl
Total	10µl

48 clones

Reação LR	R570
pDONR + 2 a 4kb	5,5µl (150ng)
pHGWFS7	0,65µl (150ng)
Tampão TE	1,85µl
Enzima LR Clonase	2µl
Total	10µl

Clones de R570 foram utilizados para a segunda reação de recombinação LR. As cores verde e rosa identificam os dois conjuntos de clones utilizados em duas reações LR independentes.

Quatro reações de recombinação LR independentes foram realizadas (Figura 15 e Figura 16). Cada uma das quatro reações de recombinação foi realizada da seguinte maneira:

Foi feito um mix total dos 41 clones de SP80-3280 obtidos da placa 1 em um único tubo eppendorf contendo 400 ng de cada clone. Esse mix foi quantificado e diluído 10x para ser utilizado na reação LR. O mesmo foi feito para os outros 3 mix.

O objetivo deste mix total foi tentar padronizar a distribuição da quantidade de cada clone no mix, para que quando adicionados à reação LR, não ocorresse a priorização da reação de recombinação de certos clones devido à sua maior quantidade em massa (nanogramas) na reação LR. Se não fosse feita esta padronização, as chances de recuperar uma alta variabilidade de clones em vetores

de expressão pHGWFS7 seriam muito menor, conforme testes realizados anteriormente (dados não mostrados).

As reações de recombinação LR efetuadas para os clones provenientes de R570 e SP80-3280, seguido pela transformação em *E.coli*, possibilitou a geração de duas bibliotecas que contêm fragmentos de DNA de cana-de-açúcar inseridos em vetores binários pHGWFS7. A primeira biblioteca contém vetores pertencentes à cultivar SP80-3280 em placas de miniprep denominadas 3 e 4 e, a segunda biblioteca abriga vetores da cultivar R570 em placas de miniprep 5 e 6 (Tabela 5).

Tabela 5 - Relação de placas de miniprep geradas pelo Sistema Gateway de Clonagem

Placas de miniprep	Descrição do conteúdo
Placa 1	Clones de SP80-3280 inseridos no vetor pDONR221
Placa 2	Clones de R570 inseridos no vetor pDONR221
Placa 3	Clones de SP80-3280 inseridos no vetor pHGWFS7 (a partir de 41 clones da placa 1)
Placa 4	Clones de SP80-3280 inseridos no vetor pHGWFS7 (a partir de 42 clones da placa 1)
Placa 5	Clones de R570 inseridos no vetor pHGWFS7 (a partir de 48 clones da placa 2)
Placa 6	Clones de R570 inseridos no vetor pHGWFS7 (a partir de 48 clones da placa 2)

Foram selecionados aleatoriamente 4 clones da miniprep da Placa 3 e da Placa 4 e, 5 clones da miniprep das Placa 5 e 6, totalizando 18 clones para uma análise da integridade destes clones em um gel de agarose 0,7% (dados não mostrados).

5.2.3 Identificação das sequências obtidas das cultivares R570 e SP80-3280.

Após confirmar a integridade do DNA das amostras das Placas 3 a 6, os insertos de todos os clones das quatro placas foram sequenciados em suas extremidades pelo método Sanger e as sequências geradas forneceram a eficiência em porcentagem de cada reação de recombinação LR (tabela 6):

Tabela 6 - Listagem das placas de miniprep 3 a 6 as quais contêm os clones de pHGWFS7 recombinados com fragmentos provenientes das placas de miniprep de clones de pDONR221 recombinados. A eficiência da reação de recombinação LR das placas 3 a 6 é dada em porcentagem.

Placas de miniprep	Quantidade de clones utilizados na reação LR	Quantidade de clones recuperados após a reação LR
Placa 3	41 clones provenientes da placa 1	33 clones (80% de eficiência)
Placa 4	42 clones provenientes da placa 1	33 clones (78% de eficiência)
Placa 5	48 clones provenientes da placa 2	39 clones (81% de eficiência)
Placa 6	48 clones provenientes da placa 2	37 clones (77% de eficiência)

O sequenciamento das placas 3 a 6 permitiu quantificar a eficiência do sistema de recombinação LR, o qual partiu de uma reação contendo um número conhecido de clones e permitiu identificar quantos destes clones foram recuperados após a reação LR. É possível notar pela Tabela 6 que as eficiências (em termos de porcentagem) das reações de recombinação independentes são similares, em torno de 80%, o que mostra um alto grau de recuperação de clones distintos em cada reação de recombinação, tornando possível a padronização desta etapa. Os clones redundantes em uma mesma placa foram agrupados, otimizando o processo de bombardeamento destes e minimizando o tempo de gasto de consumíveis.

De acordo com os resultados apresentados neste item podemos concluir que a formação de conjuntos de clones (quatro placas) utilizados para efetuar a reação LR seguido pelo sequenciamento da extremidade de todos os vetores foi uma abordagem extremamente interessante e valiosa. Com ela, foi possível determinar a eficiência da reação de recombinação LR, assim como otimizar o processo de triagem dos clones que serão analisados através da expressão transiente.

No total, duas bibliotecas de vetores de expressão pHGWFS7 pertencentes às variedades R570 (não-comercial) e SP80-3280 (comercial) foram geradas, totalizando 142 clones que abrigam uma diversidade de sequências regulatórias provenientes do genoma da cana-de-açúcar com grande potencial biotecnológico a ser explorado. Estas duas bibliotecas estão prontas para serem utilizadas em diversos sistemas de triagem funcional, os quais serão capazes de identificar novas regiões regulatórias expressas em diferentes condições de tratamento. O processo descrito mostrou ser um método confiável, com alto grau de reprodutibilidade e passível de ser realizado não só para a espécie de cana-de-açúcar mas para

qualquer outro genoma de interesse. Lembrando que no presente trabalho a padronização do processo foi validada para duas cultivares distintas, SP80-3280 e R570, reforçando a reprodutibilidade do método.

5.3 Busca por regiões regulatórias funcionais por meio da triagem da biblioteca de vetores por biobalística e análise da expressão do gene repórter GUS

Após a geração das bibliotecas de vetores de expressão pertencentes as duas cultivares R570 e SP80-3280, é necessário avaliar quais desses vetores abrigam sequências capazes de ativar a expressão gênica. A triagem por meio de análise da expressão transiente permite avaliar de maneira rápida quais são os vetores que abrigam sequências regulatórias funcionais, capazes de serem reconhecidas pela RNA polimerase, e dirigir a transcrição e tradução de genes repórteres como o GUS e GFP. Para realizar esta etapa, as duas bibliotecas de vetores de expressão os quais contêm os clones pertencentes às placas 3 a 6 foram utilizados nos ensaios de biobalística.

Com o objetivo de agilizar o processo de triagem funcional foi adotada a estratégia da formação de mix, ou seja, o agrupamento de clones distintos em um mix seguido pelo ensaio de biobalística. A Figura 17 representa a distribuição dos clones pertencentes à Placa 3 e a formação do mix. A Placa 3 abriga 33 clones únicos identificados a partir do sequenciamento, esses clones foram divididos em 8 mix, sendo que cada mix contém 4 clones (com exceção de um mix que continha 5 clones) (Figura 17). Os mixes com expressão transiente positiva tiveram seus clones bombardeados individualmente. Cada mix foi bombardeado em triplicata em epitélio de cebola e a Tabela 7 mostra quantos mix foram bombardeados de cada placa e quantos clones foram identificados com expressão transiente positiva.

Figura 17 - Representação esquemática da distribuição dos 33 clones da Placa 3 em oito mix.

No primeiro bombardeamento, cada um dos oito mixes foram bombardeados separadamente. No segundo bombardeamento, os mixes que apresentaram com expressão transiente positiva tiveram cada um de seus clones bombardeados individualmente.

Tabela 7 - Relação das placas de miniprep submetidas a bombardeio em epitélio de cebola e quantidade de mix bombardeados por placa seguido pelo número de clones identificados com expressão transiente positiva.

Placas	Quantidade de mix bombardeados	Quantidade de mix com expressão transiente	Clones ativos identificados
Placa 3	8	2	clones posição F7, B4
Placa 4	8	2	mix 9 e F6
Placa 5	10	1	mix col 2
Placa 6	9	1	clone posição C5
Total	35	6	6

No total foram realizados cerca de 450 tiros em epitélio de cebola para gerar a identificação total de 6 clones com expressão transiente positiva, sendo que foi impossível obter um clone com expressão transiente individual para o mix 9 e o mix col 2. Uma dificuldade encontrada no processo de bombardeio de epitélio de cebola foi a padronização e reprodutibilidade dos bombardeios, visto que muitos fatores

influenciam para que ocorra a expressão transiente do vetor de expressão, entre eles, precipitação do DNA sobre as micropartículas de tungstênio, homogeneização das micropartículas, estágio de desenvolvimento celular do material a ser bombardeado, pressão de gás hélio, níveis de vácuo, distância entre a onda de choque e a membrana carreadora, distância entre a membrana carreadora e a tela de retenção, distância entre a tela de retenção e o material a ser bombardeado (BRASILEIRO; CARNEIRO, 1998). Todos estes parâmetros estão inter-relacionados com os níveis de penetração das micropartículas nas células e sua sobrevivência e consequentes níveis de expressão do vetor (BRASILEIRO; CARNEIRO, 1998). Foram testadas várias modificações nos parâmetros descritos acima, porém não foi possível a padronização e reprodutibilidade ao longo dos experimentos.

De acordo com Birch e Franks, (1991); Klein et al. (1988) e Sanford (1988), existe uma grande variação nos resultados entre bombardeamentos consecutivos usando tratamentos aparentemente idênticos, o que parece ser uma característica comum entre aparatos de transferência de genes via microprojéteis. Portanto, por meio do bombardeio dos clones pertencentes às Placas 3 a 6 seguido por ensaios histoquímicos foram isolados quatro clones com expressão transiente positiva em epitélio de cebola. Esses clones são F7, B4 e 6F do cultivar SP80-3280 e C5 do cultivar R570 (Figura 18).

Figura 18 - Análise histoquímica dos clones que apresentaram expressão transiente positiva 24 horas após o bombardeamento em epiderme de cebola.

Controle positivo: pHGWFS7_U3A2MR; Controle negativo: pHGWFS7 vazio. Visualização em lupa Nikon 5MZ 800 aumento de 10x.

Todos os clones que apresentaram expressão transiente positiva foram quantificados quanto ao número de pontos azuis presentes no epitélio, sendo observados valores similares para todos os clones. Entre um a três pontos azuis foram contabilizados por epitélio para a maioria dos eventos de transformação independentes dos clones gerados pela clonagem customizada, sendo que em geral, o controle positivo sempre apresentou entre dois a 50 pontos azuis em cada experimento de bombardeio (Figura 18).

É importante lembrar que a estratégia empregada para a identificação dos clones com atividade funcional foi testada apenas em epitélio de cebola por meio da detecção da expressão do gene repórter GUS. Este sistema limitou a detecção da expressão de outras possíveis regiões regulatórias que poderiam ser ativadas apenas de maneira tecido-específica ou em determinados momentos do desenvolvimento da planta e até mesmo, em condições particulares desencadeadas por algum fator biótico ou abiótico.

O desenvolvimento da clonagem customizada de fragmentos de DNA e elaboração de um sistema de triagem voltado à rápida identificação de sequências regulatórias funcionais foi estabelecido com sucesso. Essa estratégia permite

investigar a presença de promotores por meio de uma abordagem direta de análise do produto, neste caso, a detecção da expressão do gene GUS. De posse do produto é possível examinar os segmentos genômicos correspondentes que estão envolvidos em sua síntese. Essa forma direta de avaliar a funcionalidade de certas regiões genômicas é oposta a abordagens computacionais, as quais se baseiam apenas em predições que sugerem a presença de putativos segmentos funcionais, dispersos entre regiões não codificantes.

Desta forma, a estratégia de clonagem customizada possibilita a geração de bibliotecas de vetores de expressão os quais abrigam uma diversidade de fragmentos fusionados a genes repórteres GUS e GFP. Estes fragmentos podem ser triados quanto ao seu potencial de indução da expressão de ambos os genes repórteres em diversos sistemas de expressão transiente e/ou permanente.

5.4 Identificação e análise das regiões regulatórias funcionais

5.4.1 Identificação de regiões regulatórias funcionais por meio de hidrólise enzimática e sequenciamento

5.4.1.1 Hidrólise enzimática dos clones com expressão transiente positiva

Os quatro clones que apresentaram expressão transiente positiva em epitélio de cebola através da expressão do gene repórter GUS, foram submetidos à hidrólise enzimática com a endonuclease de restrição SacII, a qual cliva em duas regiões do vetor pHGWFS7: a primeira na posição em 2.014 pb e a segunda na posição 3.936 pb. Estas duas regiões flanqueiam os sítios de recombinação do vetor, liberando, portanto, o fragmento recombinado cujo tamanho está indicado na tabela abaixo para os quatro clones analisados (Tabela 8).

Tabela 8 - Estimativa do tamanho dos fragmentos liberados pela hidrólise enzimática com *SacII* dos clones identificados por apresentar expressão transiente positiva em epitélio de cebola.

Clones	Estimativa do tamanho do inserto	Cultivar
F7	3kb	SP80-3280
C5	2,3kb	R570
F6	750pb	SP80-3280
B4	3,3kb	SP80-3280

Por meio da análise do padrão de fragmentos gerados pela hidrólise com a enzima *SacII*, o clone F7 sofreu a recombinação e inserção de um fragmento de cerca de 3 kb; C5 a inserção de cerca de 2,3 kb; F6 a inserção de 750 pb e B4 a inserção em torno de 3,3 kb. A metodologia de clonagem aleatória e triagem de clones com expressão transiente positiva recuperou mais fragmentos provenientes da cultivar SP80-3280 do que de R570 (Tabela 8).

5.4.1.2 Sequenciamento completo dos clones com expressão positiva em epitélio de cebola: B4, C5 e F7

Para sequenciar pelo método de Sanger os clones identificados com expressão transiente positiva em epitélio de cebola, foi necessária a utilização do método de *primer walking*. Este método é capaz de determinar a sequência completa do fragmento de DNA clonado no vetor pHGWFS7 através de ciclos repetidos que envolvem a síntese de iniciadores e sequenciamento do fragmento inserido no vetor. Desta maneira é possível obter o sequenciamento completo do clone desejado em uma das fitas do DNA. O mesmo processo é repetido para a outra direção do fragmento até que a sequência completa do inserto de DNA em ambas as fitas seja finalizado (Figura 19).

Figura 19 - Representação esquemática do processo de *primer walking* utilizado para o sequenciamento completo dos clones F7.

A sequência inicial dos clones F7, B4 e C5 foi obtida com o uso do iniciador attB (ver seção 4.1.4) para a primeira reação de sequenciamento. Baseado nos dados obtidos a partir dessa primeira reação foram desenhados iniciadores para uma segunda reação de sequenciamento. Os dados obtidos desta segunda reação se sobrepõem com os dados iniciais e desta forma é possível expandir a análise da sequência à jusante do iniciador attB.

Para os clones F7 e B4 foi necessário o desenho de 2 pares de iniciadores para realizar o sequenciamento completo destes dois clones (Tabela 9), sendo que a Figura 19 representa o sequenciamento do clone F7. Para o clone C5 apenas o desenho de um par de iniciadores foi suficiente para obtenção da sequência completa (Tabela 9).

Tabela 9 - Relação de oligonucleotídeos utilizados para sequenciar por completo os clones F7, C5 e B4.

Nome do oligonucleotídeo	Sequência
F7_Aleat_Forward	CCT CAT CCT CCA CTC TTC CAT CCG T
F7_Aleat_Reverse	GCG TGG GCA AGA AGT GCA AGT
F7_Aleat_Forward_Interno	GGT GAG TTG TTT ACC CAA CTC G
F7_Aleat_Reverse_Interno	GGA ACT CTA GAT ATG CAT CAT TGT C
B4_Aleat_Forward	CCA AAC AAG TGA GAC ATT GCC AAA TGG
B4_Aleat_Reverse	AGG CAA GCA GTG GCA GTT GTC AG
B4_Aleat_Forward_Interno	CAT CCA TGT GCT ATA GAG GTA G
B4_Aleat_Reverse_Interno	CTT ATA GAC CCT CAA TTC TGA TTT C
C5_Aleat_Forward	CCA TGC CCA GAT CCA GGA GGC A
C5_Aleat_Reverse	GTG CAC TTG AAA ATG ACG ATC TAA ACC C

5.4.2 Análise *in silico* dos clones F6, B4, F7 e C5

5.4.2.1 Análise da sequência completa dos clones F6, C5, B4 E F7 no banco de dados PLACE e PlantCare

Sabe-se que o ponto de início de transcrição determina o posicionamento relativo dos elementos *cis*-atuantes. Apesar de não se conhecer esse início de transcrição nos clones isolados neste trabalho, foi realizada uma busca por sítios de ligação a fatores de transcrição em toda a extensão dos quatro clones isolados por meio do banco de dados PLACE e PlantCare. Por via de regra, sabe-se que os programas de predição de sítios regulatórios não são acurados, o que gera um alto índice de falsos-positivos (BRANDALISE, 2007). No entanto, essa busca teve como objetivo criar um panorama geral dos putativos motivos *cis*-regulatórios que poderiam estar presentes nas sequências regulatórias extraídas de cana-de-açúcar. A identificação destes motivos e da sua provável função poderia fornecer pistas sobre o papel destas sequências regulatórias em promover a transcrição e expressão gênica em cana-de-açúcar.

De acordo com as análises *in silico* realizadas junto ao banco PLACE e PlantCARE ao longo da putativa região promotora do clone F6 identificado neste trabalho (385 pb), foram encontrados no total 17 putativos elementos *cis*-regulatórios de monocotiledôneas e dicotiledôneas. Deste total, sete elementos são encontrados em gramíneas: o primeiro deles é o motivo G (G-box) CACGTT e está descrito na literatura como um elemento envolvido na responsividade à luz (CHANG et al., 2012) e foi isolado em milho. Segundo Guiltinan et al. (1990), este elemento está presente

em uma variedade de genes responsivos a outros estímulos fisiológicos como luz (GIULIANO et al., 1988; SCHULZE-LEFER et al., 1989), auxina (LIU et al., 1994), ácido jasmônico (MASON et al., 1993) e ácido salicílico (QIN et al., 1994). O segundo elemento cis-regulatório identificado é o Skn-1 (GTCAT), o qual exerce importante papel em conferir especificidade na expressão do gene da glutelina *GluB-1* em endosperma de arroz (WASHIDA et al., 1999). O terceiro elemento chamado LTR (*low temperature response*), é responsivo à injúria causada por baixas temperaturas e à desidratação proveniente de ambientes secos e de alta concentração salina (HUGHES; DUNN, 1996) e foi identificado no promotor de cevada por Dunn et al. (1998). O quarto e quinto elementos identificados, motivo CGTCA e TGACG respectivamente, estão envolvidos na responsividade a moléculas sinalizadoras como o metil jasmonato (MeJA) (ROUSTER et al., 1997) em cevada. O MeJA é conhecido por estar associado ao acúmulo de alguns metabólitos secundários, os quais fazem parte de respostas de defesa desencadeadas pelo organismo em questão. Os sexto e sétimo elementos foram isolados em milho, porém não se sabe quais são as suas funções.

As análises realizadas no clone C5 (2.520 pb), identificaram 49 putativos elementos *cis*-regulatórios provenientes de monocotiledôneas e dicotiledôneas, sendo que 22 foram identificados em gramíneas. A maioria dos elementos descritos em gramíneas e identificados no clone C5 estão relacionados à responsividade à luz, como o elemento CATT (DICKEY; PETRACEK, 1998), motivo G (G-box) (CHANG et al., 2012), motivo I (I-box) (ARGUELLO-ASTORGA; HERRERA-ESTRELLA, 1996), Sp1 (YANAGISAWA; SHEEN, 1998) e MNF1 (MORISHIMA, 1998), todos identificados em milho. Isto sugere um compartilhamento interessante de sinalização para este estresse específico ao longo da sequência do clone C5. Outro elemento responsivo à luz identificado é o elemento chalcone sintase *chs-CMA2a*, isolado em cevada (ARGUELLO-ASTORGA; HERRERA-ESTRELLA, 1996). Além destes elementos também foram identificados o motivo CCAAT o qual é um sítio de ligação ao MYBHv1 (ROHDE et al., 1991) que está relacionado ao controle do desenvolvimento celular (STRACKE et al., 2001) e responsividade a reguladores de crescimento (MARTIN; PAZ-ARES, 1997); o motivo AAGAA com função desconhecida identificado em arroz e trigo (BAEK et al., 2006) e os motivos TGACG e CGTCA já descritos anteriormente.

Para o clone B4 (3.100 pb) foi possível identificar 62 elementos *cis*-regulatórios pertencentes à monocotiledôneas e dicotiledôneas, sendo que 22 são pertencentes à gramíneas. Além dos motivos já identificados em outros clones como o ARE, Skn1, SP1, motivo G, MNF1, CGTCA, foram identificados dois elementos únicos neste clone: o elemento da amilase (TATCCAT) presente em promotores de cereais regulados pela presença de metabólitos compostos por carbono, nitrogênio e fósforo. A presença destes metabólitos desencadeiam a mobilização da fonte de carbono durante a germinação e estágios subsequentes de desenvolvimento da plântula (HWANG et al., 1998). O segundo elemento denominado lib foi recentemente isolado em arroz e é responsivo a ácido absísico (JIANG et al., 2012). Dois elementos pertencentes à milho também foram identificados, porém, sem função conhecida.

As análises *in silico* realizadas ao longo do clone F7 (3.136 pb) identificaram 71 elementos *cis*-regulatórios putativos pertencentes à monocotiledôneas e dicotiledôneas, sendo que 30 destes elementos foram identificados em gramíneas.

Em comparação com a análise dos outros três clones, o clone F7 apresentou a maior diversidade de elementos *cis*-regulatórios e apontou a presença de seis elementos únicos que não foram identificados nos outros três: o primeiro, é o elemento ABRE, responsivo ao ácido absísico presente em promotores de cevada (CASARETTO; HO, 2003). Este elemento apresenta papel essencial em processos fisiológicos como maturação de sementes, fechamento de estômatos, regulação do desenvolvimento da planta e por desencadear respostas adaptativas a mudanças de condições ambientais (NAMBARA; MARION-POLL, 2005). O segundo elemento denominado AP2 atua principalmente na regulação do desenvolvimento floral em *Oryza sativa* (ZHAO et al., 2006); o terceiro elemento denominado Chalcone Sintase Chs-Unit1 m1 (CHS) é induzido pelo fornecimento de luz UV-B/UV-A em milho (CASATI; WALBOT, 2004). O quarto elemento denominado O2 (*opaque 2*) isolado em milho (MAURI et al., 1993) está envolvido com a regulação do metabolismo da proteína zeína. Esta é a principal proteína de armazenamento contida no endosperma do grão de milho. O quinto elemento denominado CE3 é responsivo à ABA e foi originalmente isolado em cevada (HOBBO et al., 1999). O sexto elemento, P-box é responsivo à giberelina e foi isolado em arroz (MENG et al., 2006).

O elemento GC foi identificado nos clones F7 e B4 e está envolvido na responsividade à anoxia e foi primeiramente isolado em milho (MANJUNATH; SACHS, 1997).

A análise de sítios putativos de ligação a fatores de transcrição nos clones F6, B4, C5 e F7 mostrou que todos estes clones compartilham quatro elementos regulatórios em comum: CGTCA e TGACG (responsivos à MeJA), G-box (responsivo à luz) e Skn-1 (especificidade na expressão do gene da glutelina *GluB-1*).

Já os clones B4, C5 e F7 compartilham dois elementos regulatórios em comum: o elemento ARE que é essencial na indução anaeróbica e frequentemente encontrado nas regiões promotoras de genes induzidos anaerobicamente. Este elemento foi reportado em promotores de milho e arabidopsis (DOLFERUS et al., 1994; HOEREN et al., 1998; OLIVE et al., 1990). O segundo elemento denominado CATT é responsivo à luz e foi reportado em milho (DICKY; PETRACEK, 1998).

É importante deixar claro que a análise de motivos *cis*- regulatórios não fornece a informação transcricional de um promotor. A posição relativa, orientação e sequências de nucleotídeos destes sítios, além dos perfis de expressão dos fatores de transcrição e co-fatores, todos estes elementos contribuem para determinar qual o perfil de transcrição de um gene (WRAY et al., 2003). Estas interações são complexas, não lineares e frequentemente contexto-dependentes. Muitos elementos *cis*-regulatórios putativos identificados não se ligam à proteínas *in vivo* e não possuem influência na transcrição (BIGGIN; MCGINNIS, 1997). A identificação de potenciais sítios de ligação que realmente se ligam a proteínas requer ensaios bioquímicos e experimentais. Um exemplo de método de identificação de sítios de ligação com papel na regulação da transcrição é modificar o sítio de ligação e realizar um teste de transcrição *in vivo*, por transformação transiente ou estável com um gene repórter.

5.4.2.2 *Análise de elementos regulatórios flanqueantes aos clones isolados: F6, B4, C5 e F7*

Tendo em vista uma caracterização mais completa dos motivos relacionados a estas sequências F6, B4, C5 e F7, foram realizadas análises de elementos regulatórios flanqueantes aos clones referidos. Esta análise complementou a busca por elementos *cis*-regulatórios localizados dentro das sequências dos clones. Foi feita uma busca por similaridade nucleotídica entre o genoma completo de *Sorghum bicolor*, utilizando o banco de dados do *Phytozome* v 8.0 (<http://www.phytozome.net/>). A região genômica de sorgo que apresentou a melhor identidade com o clone F6 foi utilizada como ponto de referência para a extração da sequência nucleotídica adjacente a esta região. Foram extraídos 385 pb (tamanho completo do clone F6) de sequência nucleotídica do genoma de sorgo adjacente ao clone F6, tanto a jusante como a montante do referido clone. Estes dois blocos de sequências de 385 pb foram utilizadas como referência para uma busca por elementos *cis*-regulatórios através do banco PLACE e PlantCare.

Esta análise foi realizada com o objetivo de verificar a riqueza de elementos *cis*-regulatórios circundante à região em que o clone F6 se encontra, já que este clone possui um tamanho limitado de 385 pb e possivelmente abriga o núcleo do promotor e sequências regulatórias proximais. Foram identificados à montante do clone F6 um total de cinco motivos *cis*-regulatórios, enquanto que à jusante deste mesmo clone foram identificados sete motivos *cis*-regulatórios. Uma análise anterior realizada com a sequência nucleotídica do clone F6 identificou 17 motivos *cis*-regulatórios dentro do clone. Esta quantidade de motivos é bem superior quando comparado às regiões genômicas flanqueantes ao clone F6 e corrobora a hipótese de que através da clonagem customizada efetuada neste trabalho, foi possível o isolamento de uma região rica em elementos controladores da transcrição, os quais são capazes de serem reconhecidos por uma RNA polimerase II e promover a expressão gênica.

Segundo Potenza et al. (2004), a região que compõe o núcleo do promotor e seus elementos regulatórios proximais é estimada em 200 pb. Portanto, clones que apresentam tamanhos muito superiores a 400 pb dificultam a análise de elementos regulatórios flanqueantes, como é o caso dos clones B4, C5 e F7 os quais possuem

tamanhos que variam entre 2 a 3.1 kb. Isto ocorre porque a extensão de nucleotídeos a ser analisada é muito maior do que a região que compõe o núcleo do promotor e seus elementos regulatórios proximais. O clone F6 é o único que apresenta um tamanho reduzido de 385 pb o qual foi apropriado para realizar a análise de elementos regulatórios flanqueantes.

Desta forma, a metodologia descrita neste trabalho possibilitou o isolamento e recuperação de blocos enriquecidos de elementos regulatórios a partir do genoma da cana-de-açúcar, os quais poderão ser estudados e ter sua função desvendada. Estes elementos serão valiosos como ferramentas biotecnológicas no aprimoramento da regulação refinada da expressão de genes de interesse em cana-de-açúcar e outros sistemas heterólogos.

5.4.2.3 Predição do TATA e CAAT para os clones F6, C5, F7 e B4

A região promotora de um gene eucarioto em geral possui uma sequência conservada denominada TATA a aproximadamente 30 pb à montante do ponto de início da transcrição e elementos proximais localizados a 80 pb (CAAT Box) à montante do ponto de início da transcrição (SNUSTAD; SIMMONS, 2008). Estes dois elementos TATA e CAAT apresentam uma distância média em torno de 40 pb (SNUSTAD; SIMMONS, 2008) e normalmente determinam o ponto correto de início da transcrição, bem como o local e momento em que este processo biológico irá ocorrer. Devido à importância destes dois elementos no processo de iniciação da transcrição, os quatro clones isolados F6, B4, C5 e F7 foram analisados quanto aos putativos motivos TATA e CAAT presentes em suas sequências.

O clone F6 inteiro (385 pb) foi analisado no banco de dados PlantCare por motivos TATA e CAAT e foi possível identificar um putativo TATA predito na posição 245 pb de sequência TTTTA, distante 140 pb do ponto de início da tradução (Figura 20). Também foi possível identificar um putativo CAAT predito na posição 224 pb de sequência CAATT, distante 161 pb do ponto de início da tradução (Figura 20). Estes dois elementos possuem uma distância de 30 pb entre si e estão de acordo com a descrição de Snustad e Simmons (2008), além de estarem alocados na fita + da sequência do clone F6.

Figura 20 - Predição dos elementos TATA e CAAT na sequência completa do clone F6.

Esta é uma predição da posição dos dois elementos TATA e CAAT, visto que não se sabe o ponto de início da transcrição na sequência do clone F6. Sabe-se que o ponto de início da tradução ocorre no códon de início ATG situado dentro da sequência do gene marcador GFP (Figura 20), portanto estima-se que o ponto de início da transcrição está situado entre a posição 245pb até o limite do início da sequência do gene GFP.

De acordo com Wray et al. (2003), o espaçamento entre o ponto de início da transcrição e tradução difere consideravelmente entre genes, variando entre 10^1 a 10^4 pb. A região 5' não traduzida UTR também pode conter introns que alteram seu comprimento após a transcrição. As consequências funcionais das diferenças no comprimento da região 5' UTR ainda não são bem compreendidas (WRAY et al., 2003).

Para o clone C5 de 2.520pb foi analisada uma região de 400 pb à montante dos genes repórteres GUS e GFP, em que foi possível identificar um putativo TATA localizado a 62 pb e um único putativo CAAT localizado a 70 pb ambos a montante do início da tradução. Os dois elementos estão alocados na fita + da sequência do clone C5, no entanto a distância entre o TATA e CAAT é de 8 pb e está em desacordo com o descrito por Snustad e Simmons (2008).

O elemento F7 de 3.136 pb também foi analisado em uma região de 400 pb à montante dos genes repórteres GUS e GFP. Foi possível a identificação de um putativo TATA localizado a 86pb e um CAAT a 132 pb do início da tradução. Ambos

elementos estão alocados na fita + da sequência do clone F7, sendo que a distância entre o TATA e CAAT é de 46 pb e está de acordo com o descrito por Snustad e Simmons (2008).

Já para o clone B4 cujo tamanho é de 3.100 pb, foi analisado uma região de 400 pb à montante dos genes repórteres GUS e GFP. Não foi possível a identificação de putativos elementos TATA. No entanto, foi identificado um motivo GC de sequência CCCCCG localizado a 4 pb do início da tradução. Este motivo está associado à ativação da transcrição por recrutamento de um fator de transcrição ubíquo denominado Sp1 (BRANDEIS et al., 1994; MACLEOD et al., 1994). Como não foram encontrados potenciais motivos TATA na região de 400pb à montante dos genes GUS e GFP, a busca foi expandida para 750 pb. Dentro deste segmento foi identificado um putativo TATA de sequência TAATA localizado a 650 pb do início da tradução e um putativo CAAT localizado a 684 pb do início da tradução. Ambos elementos estão alocados na fita + da sequência do clone B4. A distância entre o TATA e CAAT é de 34 pb e está de acordo com o descrito por Snustad e Simmons (2008).

5.4.2.4 Análise do clone F6

A sequência F6 foi utilizada em uma busca por alinhamento local através da ferramenta BLAST (Basic Local Alignment Search Tool), selecionando como base de dados sequências genômicas de referência do NCBI (*refseq_genomics*). A sequência com o melhor alinhamento foi o acesso NW 002994245.1, localizada no cromossomo 1 de *Sorghum bicolor*, que apresentou 100% de cobertura e 83% de identidade nucleotídica com a sequência de 385 pb do clone F6. Esta alta identidade com o genoma de sorgo era esperada, uma vez que esta espécie é uma gramínea filogeneticamente relacionada à cana-de-açúcar e já possui o genoma completamente sequenciado.

Foi realizada também uma busca por similaridade de nucleotídeos contra o genoma completo do *Sorghum bicolor* utilizando o banco de dados do *Phytozome* v 8.0. Foi utilizada a versão não mascarada do genoma, ou seja, a versão que inclui todos os elementos de transposição presentes (Figura 21). Com esta análise

procuramos identificar no genoma de sorgo as regiões flanqueantes àquela que apresentou identidade com o clone F6.

Figura 21 - Alinhamento entre o clone F6 (em rosa) e o cromossomo 1 de *Sorghum bicolor* através do *Phytozome*.

A região à qual se alinhou com o clone F6 corresponde ao gene que codifica a proteína hipotética Sb01g048200.1. O clone F6 apresentou 83% de identidade nucleotídica com o primeiro íntron da proteína. As sequências estão representadas em escala.

A sequência do clone F6 alinhou com o cromossomo 1 de *Sorghum bicolor*, numa região em que está localizado um potencial gene da proteína Sb01g048200.1, com uma cobertura de 100% e identidade de 83% (figura 21). O clone F6 está posicionado no primeiro íntron do gene da hipotética proteína, porém, na orientação inversa à da proteína. Para determinar se a proteína hipotética Sb01g048200.1 é conservada em outras gramíneas, sua sequência de aminoácidos foi utilizada como referência em uma busca por peptídeos homólogos no *Phytozome*. Foi encontrado um homólogo com similaridade significativa para cada um dos genomas sequenciados de quatro gramíneas, *Oryza sativa*, *Zea mays*, *Setaria italica* e *Brachypodium distachyon*. As quatro proteínas homólogas à proteína Sb01g048200.1 de *Sorghum bicolor* estão indicadas na tabela 10.

Tabela 10 - Proteínas homólogas à proteína Sb01g048200.1 de *Sorghum bicolor*.

Organismo	Identificação do gene	Similaridade (%)
<i>Sb</i>	Sb01g048200.1	-
<i>Osa</i>	LOC_0s03g04000.1	91
<i>Zma</i>	GRMZM2G096020_T01	93,7
<i>Sit</i>	Si035024m	89,7
<i>Bdi</i>	Bradi1g76280.1	89,9

Com o objetivo de verificar a conservação entre o primeiro exon da proteína Sb01g048200.1 de *Sorghum bicolor* e o primeiro exon da mesma nas outras quatro gramíneas identificadas, foi feito um alinhamento múltiplo da sequência de nucleotídeos deste exon. Este alinhamento múltiplo mostrou uma alta conservação do primeiro exon entre as 5 gramíneas analisadas.

A conservação do intron e a presença da sequência do clone F6 também foram verificadas para as outras quatro gramíneas, através de uma busca nos seus respectivos genomas no Phytozome. Foi identificada a presença do clone F6 no primeiro intron de *Zea mays* (75% de identidade) e de *Setaria italica* (74% de identidade), porém, nenhuma identidade nucleotídica significativa foi encontrada em *Oryza sativa* e *Brachypodium distachyon*. Desta maneira, foi identificada a presença do clone F6 no primeiro intron de *Zea mays* e *Setaria italica*, além de também estar presente em *Sorghum bicolor* (Figura 22 B). A comparação das regiões intrônicas destas três gramíneas revelou uma conservação de sequências, porém não de uma forma extensiva.

Portanto, os resultados obtidos na análise *in silico* do clone F6 demonstraram que ele está presente no primeiro intron da proteína hipotética Sb01g048200.1 do cromossomo 1 de *Sorghum bicolor*, em direção oposta à da putativa ORF. O clone F6 também foi encontrado no locus correspondente à proteína hipotética Sb01g048200.1 presente em *Zea mays* e *Setaria italica*. O mesmo gene da predita proteína está presente também em outras duas gramíneas, *Oryza sativa* e *Brachypodium distachyon*, porém numa versão que não possui o clone F6 em seu intron. A filogenia das gramíneas analisadas indica que *Sorghum bicolor*, *Zea mays* e *Setaria italica* são mais relacionadas entre si, compondo a subfamília Panicoideae (Figura 22 A). Já *Oryza sativa* pertence à subfamília Ehrhartoideae, e *Brachypodium distachyon* pertence à subfamília Pooideae (Kellogg, 2001) (Figura 22 A).

Figura 22 - Filogenia das gramíneas analisadas quanto à presença do clone F6.

A: Filogenia das gramíneas retirada de Kellogg, 2001. B: Árvore gerada através do alinhamento da proteína Sb01g048200.1 de *Sorghum bicolor* e de quatro proteínas homólogas de *Zea mays*, *Setaria italica*, *Oryza sativa* e *Brachypodium distachyon*, respectivamente. As caixas em cinza representam os exons da proteína, e as linhas pretas representam os introns. As setas em vermelho mostram a localização do clone F6 no primeiro intron de *Sorghum bicolor*, *Zea mays* e *Setaria italica*. A estrela em vermelho representa o momento em que a sequência do clone F6 apareceu no ancestral comum da subfamília Panicoideae. A árvore das espécies foi baseada em Kellogg (2001).

O alinhamento da proteína hipotética Sb01g048200.1 de sorgo com seus homólogos nas outras quatro gramíneas resultou em uma árvore concordante com a filogenia descrita por Kellogg (2001) (Figura 22 A). A sequência de nucleotídeos do clone F6 está presente nas espécies da subfamília *Panicoideae*, o que sugere que esta região regulatória surgiu em um ancestral comum desta subfamília, após a separação das outras duas subfamílias, *Ehrhartoideae* e *Pooideae* (Figura 22).

O clone F6 pode atuar como um promotor alternativo, ativado em situações específicas gerando um produto protéico alternativo, ou mesmo um promotor de RNAs regulatórios não codificantes, uma vez que está situado na direção oposta à da ORF. No entanto, futuros experimentos são necessários para testar estas hipóteses. Além disso, o clone F6 identificado neste trabalho abriu a perspectiva de estudos futuros com o objetivo de explorar o seu potencial biotecnológico. O fato de ter tamanho reduzido (385 pb) e já possuir a sua sequência fusionada a genes

repórteres (GUS e GFP), permitirá a realização de experimentos de deleções e mutações pontuais. Estes ensaios irão possibilitar a delimitação da região promotora mínima contida na sequência do clone F6, bem como os motivos essenciais para manter esta função. Esta sequência promotora mínima poderia ser fusionada a genes heterólogos de interesse para promover a sua expressão ou mesmo ser utilizada para dirigir a expressão de genes na própria cana-de-açúcar. Plantas transformadas de arroz contendo o clone F6 como região promotora foram obtidas em colaboração com o grupo da prof.^a Márcia Margis da UFRGS (Universidade Federal do Rio Grande do Sul). Estas plantas poderiam ser submetidas a diversos tratamentos como estresses abióticos e bióticos com o objetivo de determinar em quais situações estas regiões regulatórias poderiam desencadear uma expressão mais intensa. Desta forma, este promotor poderia ser utilizado para modular a expressão de genes de interesse em situações específicas.

5.4.2.5 Análise do clone C5

Assim como foi feito para o clone F6, a sequência completa do clone C5 foi utilizada em uma busca por alinhamento local através do BLAST, selecionando como base de dados sequências genômicas de referência do NCBI (*refseq_genomics*). A sequência com o melhor alinhamento foi o acesso NW 002994245.1, localizado no cromossomo 1 de *Sorghum bicolor*, o qual apresentou 99% de cobertura e 95% de identidade com a sequência de 2.520 pb do clone C5 e está localizada no cromossomo 1 de *Sorghum bicolor*. Foi feito também uma busca por similaridade de nucleotídeos contra o genoma completo de *Sorghum bicolor* (não mascarado) através do *Phytozome* (Figura 23).

Figura 23 - Alinhamento entre o clone C5 (em rosa) e o cromossomo 1 de *Sorghum bicolor* através do *Phytozome*.

Os primeiros 4 exons (em cinza) da proteína Sb01g038990.1 localizada no cromossomo 1 de sorgo mostram um alinhamento de 95% de identidade nucleotídica com o clone C5. Foi identificado à jusante do clone C5 um EST (em verde), o qual poderia ter sua regulação associada ao putativo promotor do clone C5. Uma hipotética proteína Sb01g039000.1 (rosa claro) também foi identificada à jusante do clone C5 porém no sentido contrário a ele.

A sequência do clone C5 alinhou com o cromossomo 1 de *Sorghum bicolor* numa região em que está localizado um potencial gene da proteína Sb01g038990.1 com uma identidade nucleotídica acima de 90%. O clone C5 está posicionado nos primeiros 4 exons e 2 íntrons desta proteína e está situado na mesma orientação desta hipotética proteína (Figura 23).

O clone C5 também está posicionado à montante do EST (*Expressed sequence tag*) *assembl_5215.p5sorghum0*, o que sugere que motivos *cis*-regulatórios contidos no clone C5 poderiam influenciar na regulação da expressão do EST em questão. Foi observado também, a presença de uma hipotética proteína Sb01g039000.1 numa orientação contrária ao clone C5 e a uma distância de 2.4kb do referido clone.

5.4.2.6 Análise do clone F7

A sequência completa do clone F7 foi utilizada em uma busca por alinhamento local através do BLAST selecionando como base de dados, sequências genômicas de referência do NCBI (*refseq_genomics*). A sequência com o melhor alinhamento foi o acesso NW 002994229.1, localizado no cromossomo 9 de *Sorghum bicolor* o qual apresentou 80% de cobertura e 73% de identidade nucleotídica com a sequência de 3.136 pb do clone F7.

Foi feita uma busca por similaridade a nível de sequência nucleotídica utilizando o banco de dados do *Phytozome* v8.0 contra o genoma não mascarado de *Sorghum bicolor* (Figura 24).

Figura 24 - Alinhamento entre o clone F7 (em rosa) e o cromossomo 9 de *Sorghum bicolor* através do *Phytozome*.

O clone F7 apresenta 73% de identidade nucleotídica com o cromossomo 9 de sorgo e está à montante da proteína Sb09g023600.2 (em azul) porém no sentido contrário à ela. O clone F7 também está à montante de outra proteína Sb09g023590.1. (roxo) e no mesmo sentido que a sequência codificante desta proteína.

A Figura 24 mostra em escala o alinhamento entre a sequência do clone F7 e o cromossomo 9 do genoma de *Sorghum bicolor*, este clone está situado à montante do gene da hipotética proteína Sb09g023600.2, porém, no sentido oposto à ela. O clone F7 está posicionado a aproximadamente 5 kb à jusante de outra proteína hipotética denominada Sb09g023590.1 (Figura 24). Como o clone F7 está posicionado entre duas proteínas hipotéticas, isso nos leva à suposição de que motivos regulatórios contidos no clone F7 poderiam influenciar na regulação da expressão de ambas as proteínas, caracterizando este promotor como bidirecional.

No entanto são necessários ensaios funcionais para comprovar a atividade bidirecional do promotor F7 na regulação da expressão de ambas as proteínas.

5.4.2.7 Análise do clone B4

A sequência completa do clone B4 (3.100 pb) foi utilizada em uma busca por similaridade de nucleotídeos contra o genoma completo de *Sorghum bicolor* (não mascarado) através do *Phytozome* (Figura 25). Três segmentos do clone B4 apresentaram identidade nucleotídica acima de 78% com o cromossomo 1 de *Sorghum bicolor*. O primeiro segmento abrange a posição 4 pb até 1.105 pb, o segundo a posição 1165 pb a 2023 pb e o terceiro segmento a posição 2.153 pb a 3.084 pb (Figura 25). O terceiro segmento do clone B4 abriga uma identidade nucleotídica acima de 70% com um potencial gene da proteína Sb01g002730.1, porém o clone B4 está no sentido oposto à ela. O clone B4 também está à jusante de duas proteínas hipotéticas Sb01g002740.1 e Sb01g002750.1 e também está à jusante de outra proteína hipotética Sb01g002723.1, porém no sentido oposto à ela (Figura 25).

Figura 25 - Alinhamento entre o clone B4 (em rosa) e o cromossomo 1 de *Sorghum bicolor* através do *Phytozome*.

Parte da sequência do clone B4 possui identidade nucleotídica de 87% com a proteína Sb01g002730.1. O clone B4 está à jusante da proteína Sb01g002740.1 (em marrom) e da proteína Sb01g002750.1 (em verde) porém no sentido contrário à sequência codificante das mesmas. O clone B4 está à jusante de uma outra proteína Sb01g002723.1 (roxo) e no mesmo sentido que a sequência codificante.

Os quatro clones isolados através da metodologia descrita neste trabalho abrigam regiões promotoras que foram capazes de serem mapeadas em regiões codificantes no genoma de *Sorghum bicolor*. Isto comprova a eficácia empregada na metodologia em isolar módulos regulatórios funcionais no genoma de cana-de-açúcar.

A estratégia também permitiu a descoberta de promotores denominados não-canônicos, os quais podem estar posicionados dentro de introns, dentro de proteínas e situados entre proteínas hipotéticas. O termo “promotor canônico” se refere ao tradicional posicionamento dos elementos regulatórios à montante em relação ao gene a ser regulado. No entanto, conceitos que pareciam ser tradicionais até pouco tempo atrás, estão sofrendo uma mudança de perspectiva pela constatação de que grande parte dos promotores é capaz de atuar de forma não-canônica (GERSTEIN et al, 2007), corroborando os resultados obtidos neste trabalho. Esta constatação permite afirmar o que outros estudos também têm encontrado: uma menor distinção deve ser feita entre regiões gênicas e intergênicas no genoma. Os genes parecem se estender no que uma vez foi chamado de espaço intergênico e a descoberta crescente de novos transcritos que possuem origem a partir de sítios regulatórios adicionais apontam para esta nova realidade (GERSTEIN et al., 2007).

É importante também ressaltar o papel dos elementos regulatórios situados dentro de introns, os quais podem influenciar a expressão dos genes hospedeiros de forma direta ou indireta (GERSTEIN et al., 2007). Esse pode ser o caso dos promotores não-canônicos identificados neste trabalho e que foram mapeados em regiões intrônicas no genoma de sorgo. Outros trabalhos identificaram elementos regulatórios que residiam em regiões transcritas como o operador lac (JACOB; MONOD, 1961), um *enhancer* do gene que regula a beta-globina (TUAN et al., 1989) e um sítio de ligação ao DNA do fator YY1 (SHI et al., 1991), o que permite afirmar que muitos elementos regulatórios na realidade residem dentro do primeiro exon, introns ou no corpo inteiro de um gene (CAWLEY et al., 2004; EUSKIRCHEN et al., 2004; GERSTEIN et al., 2007; KIM et al., 2005; ZHANG et al., 2007), o que confirma os achados deste trabalho.

Portanto, a descoberta de novos sítios regulatórios funcionais os quais estão posicionados em regiões gênicas e intergênicas no genoma de sorgo apresentam mais uma evidência de que há muito material ativo em regiões consideradas por

possuir pouca atividade (intergênicas) e que esta atividade pode representar um papel fundamental na regulação da expressão gênica.

5.4.3 Ensaios de expressão transiente em arroz

5.4.3.1 Análise da expressão transiente de protoplastos através da visualização da expressão de GFP

Os quatro clones identificados por possuírem expressão transiente positiva em epitélio de cebola: B4, C5, F6 e F7 foram transformados em protoplastos de arroz por PEG.

A figura 26 revela a expressão da proteína GFP em todas as construções analisadas: pHGWFS7_F6, pHGWFS7_C5, pHGWFS7_B4, pHGWFS7_F7, além do controle positivo pART7_HA_GFP. A eficiência da transformação dos protoplastos com as quatro construções foi estimada em torno de 10% e o controle positivo em torno de 20%. Esta diferença de eficiência pode ter ocorrido devido à variação de tamanho entre o controle positivo (6 kb) e os vetores testados (11 kb), já que de acordo com Bart et al. (2006), plasmídeos de tamanho menor possuem uma eficiência de transformação maior.

Desta forma, além da expressão transiente em epitélio de cebola, a análise da expressão transiente em protoplastos de arroz também foi confirmada para as quatro construções testadas. Isto demonstra que as sequências inseridas no vetor pHGWFS7 através da clonagem customizada são capazes de serem reconhecidas pela RNA polimerase II e desencadear a transcrição do gene repórter GFP. O resultado é a produção de proteínas fluorescentes capazes de serem detectadas pela emissão de fluorescência em comprimento de onda conhecido.

Figura 26 - Análise da expressão transiente em protoplastos de arroz através de visualização de GFP por microscopia de fluorescência.

A: vetor pHGWFS7_6F, luz transmitida; B: vetor pHGWFS7_6F, GFP; C: vetor pHGWFS7_C5, luz transmitida; D: vetor pHGWFS7_C5, GFP; E: vetor pHGWFS7_B4, luz transmitida; F: vetor pHGWFS7_B4, GFP; G: vetor pHGWFS7_F7 luz transmitida; H: vetor pHGWFS7_F7, GFP; I: controle positivo vetor pSAT, luz transmitida; J: controle positivo vetor pSAT, GFP; K: Controle negativo vetor pHGWFS7 vazio, luz transmitida; L: Controle negativo vetor pHGWFS7 vazio, GFP. Visualização em microscópio de fluorescência refletida Olympus CKX41. Aumento de 20x.

A confirmação da expressão transiente das quatro construções em uma espécie monocotiledônea como o arroz, que é uma gramínea filogeneticamente

relacionada à cana-de-açúcar, nos leva a acreditar que essas construções poderiam também ter sua atividade detectada na própria cana-de-açúcar. Para avaliar se estas sequências regulatórias são funcionais no próprio organismo modelo, ou seja, são capazes de dirigir a expressão de genes repórteres, foi realizada a transformação permanente de uma das construções na própria cana-de-açúcar.

5.5 Validação dos promotores através da transformação permanente em sistema heterólogo e nativo

5.5.1 Ensaios de expressão permanente em cana-de-açúcar

5.5.1.1 Transformação genética de calos e discos de cana-de-açúcar via biobalística, seleção e regeneração de plantas transformadas

A metodologia de clonagem customizada desenvolvida neste projeto gerou a identificação de potenciais sequências regulatórias utilizando como modelo, o genoma da cana-de-açúcar. Após a confirmação da expressão transiente dos quatro clones os quais abrigam putativas regiões promotoras de cana-de-açúcar em sistemas heterólogos (cebola e protoplastos de arroz), partimos para a transformação permanente de um destes clones em sua própria espécie nativa: a cana-de-açúcar. É importante lembrar que a expressão do gene GUS e GFP nos ensaios de expressão transiente é considerada apenas uma evidência da transformação, não sendo suficiente como prova definitiva da integração da sequência de interesse no genoma da planta (POTRYKUS, 1990).

Para validação dos promotores isolados é necessário avaliar a manutenção do potencial de indução de expressão destes promotores. Para atingir este objetivo, ensaios funcionais tanto em sistemas heterólogos como não-heterólogos devem ser conduzidos por meio da transformação permanente. Desta forma, é possível avaliar em quais tecidos da planta existe expressão dos promotores, assim como as condições em que estas sequências regulatórias são capazes de serem expressas.

Nesta etapa do trabalho foi realizada a transformação permanente de cana-de-açúcar como forma de validar um dos promotores identificados neste trabalho, o clone F6. Este clone está inserido no vetor pHGWFS7 e foi co-transformado com o vetor pH9 (Figura 9), o qual possui o promotor *ubi* fusionado ao gene *neo*

(neomicina fosfotransferase). Este gene atua na inativação de antibióticos como canamicina e geneticina e portanto, as plantas que forem transformadas com este vetor serão resistentes a estes antibióticos. O *ubi* é um promotor forte capaz de expressar constitutivamente genes em monocotiledôneas, de forma que as proteínas recombinantes possam ser produzidas em altos níveis (CORNEJO et al., 1993). A escolha do agente de seleção das plantas transformadas é extremamente importante e optamos por utilizar a geneticina. Segundo Nehra et al. (1994), este antibiótico não interfere no processo de regeneração e pode, portanto, ser utilizado para a seleção de tecido transformado em estágios iniciais. A canamicina tem se mostrado pouco eficaz como agente seletivo quando comparada a sua atuação em espécies dicotiledôneas. Seu uso em monocotiledôneas tem sido desencorajado devido ao alto nível de resistência endógena e pela interferência na regeneração de células transformadas (HAUPTMANN et al., 1988; NEHRA et al., 1994).

A co-transformação é amplamente utilizada em ensaios de expressão permanente, já que em muitas ocasiões não é possível selecionar diretamente os transformantes com a sequência de interesse. Uma possibilidade é ter em uma mesma construção plasmidial o gene de interesse e o gene marcador (FALCO, 1998). Outra opção (a qual foi escolhida para este trabalho), é misturar em uma mesma suspensão de partículas os dois vetores, um contendo a sequência de interesse e o outro contendo o gene marcador, neste caso, um gene de resistência a aminoglicosídeos. Neste último processo, uma proporção das plantas transformadas e selecionadas pela resistência ao antibiótico terão incorporado também a sequência de interesse, em um fenômeno conhecido como co-transformação (BIRCH, 1995).

O clone F6 foi co-transformado em seis bombardeios, sendo que três tiros foram disparados em três placas contendo calos embriogênicos de cana da cultivar SP80-3280 e os outros três tiros em três placas contendo discos da mesma cultivar. Os discos foliares são derivados de secções transversais de folhas imaturas e segundo Lakshmanan et al. (2006), a regeneração direta de plantas através destas secções se mostrou favorável. Três bombardeios com o controle positivo de bombardeio pFF19G (figura 10) foram feitos em folhas de tabaco.

As outras construções contendo putativas regiões promotoras não foram utilizadas neste experimento devido à geração de grande volume de material

supostamente transformado a ser manuseado, além do curto período de tempo disponível para a execução destes experimentos.

Foram utilizadas como controle, plantas que não sofreram bombardeio e, portanto, não são transformadas. Essas plantas passaram pelos mesmos processos de regeneração que as plantas transformadas e possuem a mesma idade.

Após o período de recuperação dos calos e discos na placa, cada placa foi transferida para vidros contendo meio MS e geneticina e foi realizado a cada 30 dias a passagem do material transformante para novos vidros, sendo mantida a seleção por geneticina em todas as passagens. Após quatro meses de seleção, foram regeneradas 26 plântulas resistentes ao antibiótico geneticina e possivelmente transformadas com o clone F6.

Para se ter controle dos eventos de transformação, os calos e discos regenerados foram nomeados de acordo com a placa de origem para a identificação dos eventos de bombardeamento e mantidos separadamente. Assim, torna-se possível saber exatamente qual o evento em que a planta transformada pertence.

O fato da existência de plantas que regeneraram e cresceram em meio de seleção contendo o antibiótico geneticina não é suficiente para a confirmação de que estas plantas estão transformadas com a construção pHGWFS7_F6. Após o bombardeio, o DNA de interesse introduzido via micropartículas poderá estar integrado ao genoma das plantas ou não. A permanência do agente seletivo após inúmeras divisões celulares se caracteriza apenas como um elemento de transformação positiva, sendo necessária a utilização de técnicas de biologia molecular para a confirmação da integração do DNA desejado no genoma das plantas transformadas (FALCO, 1998).

5.5.1.2 Análises Moleculares

Análises moleculares do material transformado foram realizadas em um estágio inicial para determinar se o material é ou não transgênico. Entende-se por célula transformada aquela que teve a introdução de uma sequência de DNA. Portanto, o termo transformada não deve ser confundido com transgênica, que se refere à célula em que o DNA de interesse foi integrado no genoma (FALCO, 1998).

5.5.1.3 Extração de DNA genômico das plântulas transformadas com o clone F6 em cana-de-açúcar

Foi feita a extração de DNA genômico a partir das folhas de 26 plântulas de cana-de-açúcar resistentes ao antibiótico geneticina e a confirmação da integridade das amostras foi realizada através da aplicação de alíquotas do DNA extraído em gel de agarose 0,7%. No total foram extraídos DNA genômico de qualidade de 25 plântulas de cana-de-açúcar os quais foram utilizados para as análises moleculares posteriores.

5.5.1.4 Análise por PCR com iniciadores *nptII*

O vetor pHA9 o qual abriga o gene *neo* utilizado na co-transformação de cana-de-açúcar, foi apenas um instrumento para seleção do material transformado “in vitro”. No entanto, optou-se por confirmar a presença do mesmo no genoma das plântulas de cana-de-açúcar através de ensaios de PCR com os iniciadores *nptII* descritos na seção 4.5.1.4. Estes ensaios de amplificação do gene *neo* foram realizados nas 25 plântulas de cana.

Das 25 plântulas de cana-de-açúcar analisadas com relação à presença do gene *neo*, foi possível confirmar a inserção deste gene em 7 plântulas (amostras 3, 4, 5, 6, 7, 11 e 15) portanto, 28% de eficiência na detecção deste gene. Na figura 27 está a representação do padrão de amplificação esperado do gene *nptII* (410 pb), utilizando como molde o DNA genômico extraído da plântula 15. Este mesmo padrão de amplificação foi detectado nas outras 6 plântulas com variações na intensidade da visualização da banda.

Figura 27 - Análise da presença do gene *neo* no possível transformante de cana-de-açúcar.

1: planta 15 transformada por biobalística; 2: Controle negativo da reação de PCR; 3: controle positivo (pHA9); 4: planta não transformada; L: Ladder 1 kb (Fermentas) 500 ng. 5 µl aplicado nos poços 1 a 3. Gel 1%, TAE 0,5x, 40 mA.

5.5.1.5 Análise por PCR com iniciadores *eGFP*.

Após a confirmação da presença do gene marcador *npII* em 7 plântulas de cana-de-açúcar, procedemos para a confirmação da presença do outro gene marcador, o GFP presente no vetor pHGWFS7. É importante lembrar que este é o vetor que abriga o clone F6 fusionado ao gene GFP. No entanto, não é possível uma confirmação direta por PCR da inserção do clone F6, visto que sua sequência teve origem no DNA genômico de cana e foi isolada através da clonagem customizada realizada neste trabalho. Desta forma, existe a possibilidade de que esta região do clone F6 esteja presente em outro locus do genoma da cana, o que representaria um falso positivo na confirmação por PCR da planta transformada.

Por este motivo, foi escolhido confirmar a presença do gene GFP ao invés da presença da sequência do clone F6 e para isso foram utilizados os iniciadores *eGFP* descritos na seção 4.5.1.4. A confirmação deste gene no genoma da cana-de-açúcar é um indício de que a sequência promotora do clone F6 também foi inserida com sucesso.

Foram realizadas várias tentativas para amplificar o gene GFP nas 25 plântulas de cana-de-açúcar, no entanto, não foi possível um resultado claro da inserção deste gene no genoma desta cultivar. A etapa de cultura de tecidos e seleção dos transformantes de cana foi baseada na presença do vetor co-transformado pHA9 ou

seja, na presença do gene *neo*, o qual confere resistência ao antibiótico geneticina e não na presença do vetor de interesse, pHGWFS7. Desta forma, foi realizada uma transferência das 26 plântulas de cana-de-açúcar para um novo meio de crescimento. Este meio ao invés de possuir geneticina como antibiótico de seleção continha higromicina a 20 µg/ml, concentração utilizada para seleção de tecido transformado de cana-de-açúcar (HOSSAIN et al., 2007) e óleo de palma (PARVEEZ et al., 2007). Desta forma, apenas as plantas que possuíam a inserção do vetor pHGWFS7 apresentariam a resistência ao antibiótico higromicina.

O resultado foi que após 20 dias todas as 26 plântulas de cana-de-açúcar morreram no meio de seleção com higromicina, o que comprova que as plantas não possuíam a presença do vetor desejado, pHGWFS7_F6. A transformação permanente de cana-de-açúcar foi realizada com sucesso para a seleção do vetor co-transformado pH9, porém não foi eficiente para a seleção do vetor de interesse pHGWFS7. Este relato é contrário ao que já foi descrito por Birch et al. (1995) em que relataram taxas acima de 60% de co-expressão de genes co-precipitados em transformação de cana-de-açúcar.

Já os experimentos de co-transformação realizados por Melotto (Melotto-Passarim, 2009), foi observado que 83% das plantas de cana-de-açúcar que apresentaram PCR positivo para o gene *neo* também apresentaram PCR positivo para o gene em estudo. Não se sabe o porquê da ineficiência de inserção do vetor pHGWFS7 no presente trabalho, visto que todos os controles positivos e negativos realizados nas etapas de bombardeio e cultura de tecidos funcionaram conforme o esperado, porém, alguns apontamentos podem ser feitos. Pode ser que devido ao baixo número de cópias integradas no genoma da cana-de-açúcar transformada não foi possível a obtenção da expressão de um nível apropriado do gene de resistência à higromicina.

Portanto, quando todo o material foi transferido para meio com higromicina, algumas plantas que poderiam abrigar o vetor desejado não conseguiram sobreviver devido aos baixos níveis de expressão dos transgenes inseridos por bombardeio. Em alguns casos é necessário a seleção de centenas de linhagens transformantes para a obtenção de uma capaz de apresentar altos níveis de expressão do transgene (HOSHINO, 2007). Outro alerta diz respeito ao número de experimentos necessários quando se visa a obtenção de plantas transformadas, já que muitas

células são atingidas (transformadas) mas poucas sobrevivem e se dividem regerando plantas, por isso a necessidade de se trabalhar com grande quantidade de explantes (FALCO, 1998). O presente trabalho não manipulou esta grande quantidade de material, já que realizou seis tiros, sendo três em placas de calos e três em placas que continham discos foliares. Talvez seja esse um dos motivos do insucesso na obtenção de transformantes de cana-de-açúcar com o vetor desejado, porém, o tempo hábil disponível para a execução dos experimentos era o maior limitante naquele estágio do projeto e não permitia a manipulação de grande quantidade de explantes.

5.5.2 Ensaios de expressão permanente em arroz realizado em colaboração com o Núcleo de Genômica Funcional de Plantas (UFRGS)

Ensaios de expressão transiente confirmaram a presença de promotores dentro dos fragmentos F6, B4, C5 e F7 isolados do genoma da cana-de-açúcar. Em seguida foi realizado um ensaio de transformação permanente no arroz com o objetivo de consolidar a funcionalidade destes promotores nesta monocotiledônea. Este procedimento foi realizado através de uma colaboração com o Núcleo de Genômica Funcional de Plantas da UFRGS (Universidade Federal do Rio Grande do Sul), coordenado pela prof^a Marcia Margis e em colaboração com a aluna de doutorado Fernanda Lazzarotto. Foram obtidas 12 plantas de arroz transgênicas sete delas com a construção F6, duas com a construção B4 e três com a construção F7. Nenhum regenerante contendo a construção C5 foi obtido.

A presença do transgene foi confirmada nessas plantas através de PCR sobre DNA genômico com o par de iniciadores eGFP desenhado sobre o gene GFP (seção 4.5.1.4). Adicionalmente foram realizados ensaios histoquímicos para determinar a expressão do gene marcador GUS em diferentes tecidos (Figura 28). A expressão do gene GUS foi positiva em todos os calos de arroz analisados, transformados com os clones F6, B4 e F7. Nas plantas regenerantes foi possível observar a expressão de GUS apenas em folhas do clone F6. O clone B4 não apresentou expressão do gene GUS em folhas ou raízes. De forma contrastante, o clone F7 apresentou forte expressão do gene GUS em folhas e raízes. Estas diferenças na intensidade e especificidade do padrão de expressão nos diferentes tecidos de arroz podem ser

atribuídas em parte, pela influência dos diferentes sítios de integração no genoma da planta e possíveis interações entre múltiplas cópias integradas do DNA introduzido (BIRCH et al., 2000). Outro fator importante, que pode causar mudança na intensidade do padrão de expressão em diferentes tecidos de arroz, são as diferenças intrínsecas entre as sequências regulatórias contidas nos diferentes clones e a interação destas sequências com o contexto genômico no qual foram inseridas (BIRCH et al., 1997).

Figura 28 - Expressão do gene GUS em tecidos de arroz transformados com as construções F6, B4 e F7.

A detecção foi realizada através de ensaio histoquímico em tampão X-gluc. Foram analisados os seguintes tecidos: A: calo; B: folhas, C: raízes.

Estas 12 plantas foram transplantadas do meio *in vitro* para vasos contendo terra e vermiculita na proporção 1:1, com o objetivo de coletar sementes e realizar a análise da geração T1. Esta análise será valiosa por permitir um monitoramento da

expressão dos genes repórteres GUS e GFP desde o estágio inicial da plântula até o completo desenvolvimento do arroz. Isto possibilitará a avaliação dos tecidos e condições ideais de atividade dos promotores F6, B4 e F7.

O desenvolvimento do presente trabalho permitiu estabelecer e validar uma estratégia de recuperação de sequências regulatórias funcionais a partir de um genoma complexo como o da cana-de-açúcar. Esta técnica pode constituir uma poderosa ferramenta de prospecção de promotores, também em outros genomas. Além disso, foram isoladas quatro regiões genômicas distintas que abrigam promotores do genoma da cana-de-açúcar.

O perfil de expressão de genes repórteres dirigidos por estes promotores foi validado em dois sistemas heterólogos, sendo que em um deles foi confirmada a expressão tanto transiente quanto permanente. O clone F7, por exemplo, pode ser empregado em construções que visam uma expressão forte de genes de interesse em folhas e raízes devido à alta atividade desta região regulatória nestes tecidos. Por outro lado, ensaios de mutação pontual dentro dos motivos regulatórios podem elucidar nucleotídeos chave para a funcionalidade e especificidade dos mesmos.

Estas regiões regulatórias apresentam, portanto, ampla possibilidade para serem exploradas na área biotecnológica e no campo da ciência básica. A metodologia desenvolvida neste trabalho permitiu a disponibilização de módulos biológicos, que podem ser empregados para o melhoramento desta cultura de grande importância econômica que é a cana-de-açúcar.

6 CONCLUSÕES

- A estratégia de clonagem customizada e geração de bibliotecas de vetores contendo fragmentos de DNA associados a genes repórteres GUS e GFP foi estabelecida com sucesso.

- O processamento do genoma através da utilização de endonucleases de restrição específicas foi adequado para o êxito na seleção de fragmentos de DNA com tamanho adequado para a identificação de potenciais regiões promotoras.

- A estratégia de triagem funcional associou os processos de identificação e ensaio funcional das regiões promotoras, acelerando a obtenção de regiões promotoras da expressão gênica.

- Quatro regiões regulatórias funcionais (F6, B4, F7 e C5) foram isoladas do genoma da cana-de-açúcar mediante expressão transitória em epitélio de cebola e confirmada em protoplastos de arroz.

- Foi possível observar diferenças na intensidade de indução da expressão do gene GUS dirigida pelos promotores F6, B4 e F7 em plantas transgênicas de arroz.

- A disponibilização de módulos regulatórios já com informação funcional abre amplas perspectivas a serem exploradas no campo da biotecnologia e da ciência básica.

- A utilização do genoma da cana-de-açúcar como modelo para a estratégia de clonagem customizada foi aplicada com êxito e pode ser estendida a qualquer genoma de interesse.

7 PERSPECTIVAS

Foi disponibilizada uma diversidade de regiões regulatórias isoladas através da estratégia inovadora descrita neste trabalho, as quais servirão como base para:

- Caracterização funcional das regiões promotoras por análises de deleções, substituições de bases seguido por expressão em genes repórteres
- Testar a responsividade dos 4 clones isolados a tratamentos bióticos e abióticos: identificação de regiões regulatórias expressas de maneira específica
- Fusão dos clones F6, B4, C5 e F7 à genes de interesse para dirigir a expressão na própria cana-de-açúcar e em sistemas heterólogos

REFERÊNCIAS*

- ALJANABI, S. M.; FORGET, L.; DOOKUN, A. An improved and Rapid Protocol for the Isolation of Polysaccharide- and Polyphenol-Free Sugarcane DNA. **Plant Molecular Biology Reporter**, v. 17, 1999.
- ALTENBACH, S. B. et al. Enhancement of the methionine content of seed proteins by the expression of a chimeric gene encoding a methionine-rich protein in transgenic plants. **Plant Molecular Biology**, v. 13, p. 513-522, 1989.
- AN, G.; KIM, Y. S.; Techniques for isolating and characterizing plant transcription promoters, enhancers and terminators. In: GLICK, B.; THOMPSON, J. **Methods in plant molecular biology and biotechnology**. London: CRC Press, 1993. p. 155-166.
- ARBEITMAN, M. N.; FURLONG, E. E.; IMAM, F. et al. Gene expression during the life cycle of *Drosophila melanogaster*. **Science**, p. 2270-2275, 2002
- ARGUELLO-ASTORGA, G. R.; HERRERA-ESTRELLA, L. R. Ancestral multipartite units in light-responsive plant promoters have structural features correlating with specific phototransduction pathways. **Plant Physiol.**, v. 3, p. 1151-1166, 1996.
- ARMELIN CORREA, L. M. **Estudo da região promotora do gene do colágeno XVIII humano**. 2007. 204 f. Tese (Doutorado em Ciências) – Instituto de Biociências, Universidade de São Paulo, São Paulo, 2007.
- ARNONE, M. I.; DAVIDSON, E. H. The hardwiring of development: organization and function of genomic regulatory systems. **Development**, v. 124, p. 1851-1864, 1997.
- ATKINSON, R. G. et al. ACC-oxidase and polygalacturonase: ripening-specific gene expression and promoter analysis in transgenic tomato. **Plant Mol. Biol.**, v. 38, p. 449-460, 1998.
- BADE, J. et al. T-DNA tagging in *Brassica napus* as an efficient tool for the isolation of new promoters for selectable marker genes. **Plant Molecular Biology**, v. 52, p. 53-68, 2003.
- BAEK, K. et al. Molecular structure and organization of the wheat genomic manganese superoxide dismutase gene. **Genome**, v. 49, p. 209-218, 2006.
- BART, R. et al. A novel system for gene silencing using siRNAs in rice leaf and stem-derived protoplasts. **Plant Methods**, v. 2, p.13, 2006.

*De acordo com:

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 6023**: informação e documentação: referências: elaboração. Rio de Janeiro, 2002.

BAUMIEIN, H. et al. A novel seed protein gene from *Vicia faba* is developmentally regulated in transgenic tobacco and *Arabidopsis* plants. **Molecular and General Genetics**, v. 225, p. 459-467, 1991.

BELBAHRI, L. et al. A local accumulation of the *Ralstonia solanacearum* PopA protein in transgenic tobacco renders a compatible plant-pathogen interaction incompatible. **Plant J.**, v. 28, p. 419-430, 2001.

BENECKE, A.; GAUDON, C.; GRONEMEYER, H. Transcriptional integration of hormone and metabolic signals by nuclear receptors. **Academic Press**, p. 167-214, 2001.

BHULLAR, S. et al. Strategies for development of functionally equivalent promoters with minimum sequences homology for transgene expression in plants: cis-Elements in a novel DNA context versus domain swapping. **Plant Physiology**. v. 132, p. 988-998, 2003.

BIGGIN, M. D.; MCGINNIS, W. Regulation of segmentation and segmental identity by *Drosophila* homeoproteins: the role of DNA binding in functional activity and specificity. **Development**, v. 124, p. 4425-4433, 1997.

BIRCH, R. G.; FRANKS, T. Development and optimization of microprojectile systems for plant genetic transformation, **Australian Journal of Plant Physiology**, v. 18, p. 453-469, 1991.

BIRCH, R.G. et al. Expression of foreign genes in sugarcane. **Proc. Int. Soc. Sugarcane Technol.**, v. 22, p. 368-372, 1995.

BIRCH, R. G. Plant transformation: problems and strategies for practical application. **Annu Rev. Plant Physiol. Plant Mol. Biol.**, v. 48, p. 297-326, 1997.

BLANCHETTE, M.; SINHA, S. Separating real motifs from their artifacts. **Bioinformatics** v. 17, p. 30-38. 2001.

BUCHANAN, B. B.; GRUISSEM, W.; JONES, R. L. **Biochemistry & molecular biology of plants**. Rockville: American Society of Plant Physiologists, 2001.

BRAITHWAITE, K. S.; GEIJSKES, R. J.; SMITH, G. R.; A variable region of the sugarcane bacilliform virus (SCBV) genome can be used to generate promoters for transgene expression in sugarcane. **Plant Cell Rep.**, v. 23, p. 319-326. 2004.

BRANDALISE, M. **Isolamento e caracterização de promotores tecido-específicos de raiz e folha de *Coffea arabica***. 2007. 141 f. Tese (Doutorado em Genética) – Universidade Estadual Paulista, Botucatu, São Paulo, 2007.

BRANDEIS, M. et al. Sp1 elements protect a CpG island from de novo methylation. **Nature**, v. 371, p. 435-438, 1994.

BRASILEIRO, A. C. M.; CARNEIRO, V. T. C. **Manual de transformação genética de plantas**. Brasília: Embrapa, 1998.

BUTLER, J. E.; KADONAGA, J. T. The RNA polymerase II core promoter: a key component in the regulation of gene expression. **Genes Dev.**, v. 16, p. 2583-2592, 2002.

CADDICK, M. X. et al. An ethanol inducible gene switch for plants used to manipulate carbon metabolism. **Nature Biotechnol.**, v. 16, p. 177-180, 1998.

CALLIS, J.; RAASCH, J. A.; VIERSTRA, R. D. Ubiquitin extension proteins of *Arabidopsis thaliana*. Structure, localization and expression of their promoters in transgenic tobacco. **J. Biol. Chem.**, v. 265, p. 12486-12493, 1990.

CHANG, M. L. et al. Effect of cadmium on peroxidase isozyme activity in roots of two *Oryza sativa* cultivars, **Botanical Studies**, v. 53, p. 31-44, 2012.

CASU, R. E. et al. Identification of differentially expressed transcripts from maturing stem of sugarcane by in silico analysis of stem expressed sequence tags and gene expression profiling. **Plant Mol. Biol.**, v. 54, p. 503-517, 2004.

CLAES, B. et al. Organ-dependent regulation of a plant promoter isolated from rice by "promoter-trapping" in tobacco. **The Plant J.**, p. 15-26. 1991.

CORNEJO, M. J. et al. Activity of a maize ubiquitin promoter in transgenic rice. **Plant Mol. Biol.**, v. 23, p. 567-587, 1993.

CASARETTO, J. A.; HO, T. H. D. The transcription factors HvABI5 and HvVP1 are required for the ABA induction of gene expression in barley aleurone cells. **Plant Cell**, v.15, p. 271-284, 2003.

CASATI, P.; WALBOT, V.; Rapid transcriptome responses of maize (*Zea mays*) to UV-B in irradiated and shielded tissues. **Genome Biol.**, v. 5, 2004.

CASELLE, M. et al. Correlating overrepresented upstream motifs to gene expression: a computational approach to regulatory element discovery in eukaryotes. **BMC Bioinformatics**, v. 3, 2002.

CASU, R. E. et al. Identification of differentially expressed transcripts from maturing stem of sugarcane by in silico analysis of stem expressed sequence tags and gene expression profiling. **Plant Mol. Biol.**, v. 54, p. 503-517, 2004.

CAWLEY, S. et al. Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of noncoding RNA. **Cell**. v. 116, p. 49-509, 2004.

CHAIR, H.; LEGAVRE, T.; GUIDERDONI, E. Transformation of haploid, microspore derived cell suspension protoplasts of rice (*Oryza sativa* L.), **Plant Cell Rep.**, v. 15, p. 766-770, 1996.

CHANG, M. L. et al. Effect of cadmium on peroxidase isozyme activity in roots of two *Oryza sativa* cultivars. **Botanical Studies**, v. 53, p. 31-44, 2012.

CHALFIE, M. et al. Green fluorescent protein as a marker for gene expression. **Science**, v. 263, p. 802-805, 1994.

CHAREONPORNWATTANA, S, et al. Inheritance, expression and silencing of a chitinase transgene in rice. **Theor. Applied Genet.**, v. 98, p. 371-378, 1999.

CHEN, P. W. et al. Rice α -amylase transcriptional enhancers direct multiple mode regulation of promoters in transgenic rice. **J. Biol. Chem.** v. 16, 2002.

CHUNG, J. H.; WHITELEY, M.; FELSENFELDK, G. A 5' element of the chicken β -globin domain serves as an insulator in human erythroid cells and protects against position effect in *Drosophila*. **Cell**, v. 74, p. 505-504, 1993.

CHRISTENSEN, A.H.; QUAIL, P.H.; Ubiquitin promoters based vectors for high level expression of selectable and/or screenable marker genes in monocotyledonous plants. **Transgenic Research**. v. 5, p. 213-218, 1996.

CLAES, B. J. et al. Organ-dependent regulation of a plant promoter isolated from rice by promoter-trapping in tobacco. **Plant J.**, v. 1, p. 15-26, 1991.

COOLEY, J.; FORD, T.; CHRISTOU, P.; Molecular and genetic characterization of elite transgenic rice plants produced by electric-discharge particle acceleration. **Theoretical and Applied Genetics**, v. 90, p. 97-104, 1995.

DAMAJ, M. B. et al. Sugarcane dirigent and o-methyltransferase promoters confer stem-regulated gene expression in diverse monocots. **Planta**, v. 231, p. 1439-1458, 2010.

DATTA, K. et al. Constitutive and tissue-specific differential expression of CryIA (b) gene in transgenic rice plants conferring enhanced resistance to insect pests. **Theor. Appl. Genet.**, v. 97, p. 20-30, 1998.

DATLA, R.; ANDERSON, J. W.; SELVARA, J. G. Plant promoters for transgene expression. **Biotechnology Annual Review**, v. 3, p. 269-296, 1997.

DAVIDSON, E. H. **Genomic regulatory systems**: development and evolution. San Diego, Calif.: Academic Press, 2001. 261 p.

DE LA FUENTE, J. M. et al. Aluminum tolerance in transgenic plants by alteration of citrate synthesis, **Science**, v. 276, p.1566-1568, 1997.

DE GREVE, H. J. et al. Regeneration of normal and fertile plants that express octopine synthase from tobacco crown galls after deletion of tumor controlling functions. **Nature**, v. 300, p. 752-755, 1982.

DE PATER, B. S. et al. The promoter of the rice gene GOS2 is active in various different monocot tissues and binds rice nuclear factor. ASF-1. **Plant J.**, v. 2, p. 837-844, 1992.

DEVIC, M. et al. Efficient PCR walking on plant genomic DNA. **Plant Physiol. Biochem.**, v. 35, p. 1-9, 1997.

DE WET, J. R. et al. Firefly luciferase gene: structure and expression in mammalian cells. **Mol. Cell Biol.**, v. 7, p. 725-737, 1987.

DE WIT, P. J. G. M. Molecular characterization of gene-for-gene systems in plant-fungus interactions and the application of avirulence genes in control of plant pathogens. **Annu. Rev. Phytopathol.**, v. 30, p. 391-418, 1992.

DICKEY, L. F. et al. Light regulation of Fed-1 mRNA requires an element in the 5' untranslated region and correlates with differential polyribosome association. **The Plant Cell**, v. 10, p. 465-484, 1998.

DIGEON, J. F. et al. Cloning of a wheat puroindoline gene promoter by IPCR and analysis of promoter regions required for tissue-specific expression in transgenic rice seeds. **Plant Molecular Biology**, v. 39, p. 1101-1112, 1999.

DOLFERUS R. ET AL. Differential interactions of promoter elements in stress responses of the Arabidopsis Adh gene. **Plant Physiology**, v. 105, p. 1075-1087, 1994.

DUNN, M. A. et al. Identification of promoter elements in a low-temperature-responsive gene (blt4.9) from barley (*Hordeum vulgare* L.). **Plant Mol. Biol.**, v. 38, p. 551-564, 1998.

EAGLESHAM, A.; PUEPPKE, S. G.; HARDY, R. W. F. (Ed.). **Genetically modified food and the consumer**. New York: National Agricultural Biotechnology Council, 2001.

ELLIS, J. G. et al. Maize *Adh-1* promoter sequences control anaerobic regulation: addition of upstream promoter elements from constitutive genes is necessary for expression in tobacco. **EMBO J.**, v. 6, p. 11-16, 1987.

EULGEM, T. et al. Early nuclear events in plant defence signaling: rapid gene activation by WRKY transcription factors. **EMBO J.**, v. 18, p. 4689-4699, 1999.

EUSKIRCHEN, G. et al. CREB binds to multiple loci on human chromosome 22. **Mol. Cell. Biol.**, v. 24, p. 3804-3814, 2004.

FALCO, M. C. **Transformação de cana-de-açúcar para resistência do herbicida glufosinato de amônio, via biobalística**. 129 f. Tese (Doutorado em Genética e Melhoramento de Plantas) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 1998.

FELENBOK, B. et al. The ethanol regulon in *A. nidulans*: characterization and sequence of the positive regulatory gene *alcR*. **Gene**, v. 73, p. 385-396, 1988.

FICKETT, J. W.; HATZIGEORGIOU, A. G. Eukaryotic promoter recognition. **Genome res.** v. 7, p. 861-878, 1997.

FUJIMOTO, H. et al. Insect resistant rice generated by introduction of a modified α -endotoxin gene of *Bacillus thuringiensis*. **Bio/Technology**, v. 11, p. 1151-1155, 1993.

GALLO-MEAGHER, M.; IRVINE, J. E. Effects of tissue type and promoter strength on transient GUS expression in sugarcane following particle bombardment. **Plant Cell.Rep.**, v. 12, p. 666-670, 1993.

GAMBLEY, R. L. et al. Cytokinin enhanced regeneration of plants from microprojectile bombardment sugarcane meristematic tissue. **Australian Journal of Plant Physiology**, v. 21, p. 603-612, 1994.

GATZ, C.; LENK, I. Promoters that respond to chemical inducers. **Trends Plant Sci.**, v. 3, p. 352-358, 1998.

GERSTEIN, M. B. et al. What is a gene, post-ENCODE? History and updated definition. **Genome Res.**, p. 669-681, 2007.

GITTINS, J. R. et al. Transgene expression driven by heterologous ribulose-1,5-bisphosphate carboxylase/oxygenase small-subunit gene promoters in the vegetative tissues on apple (*Malus pumila* Mill.). **Planta**, v. 210, p. 232-240, 2000.

GIULIANO, G. et al. An evolutionarily conserved protein binding sequence upstream of a plant light-regulated gene. **Proc. Natl Acad. Sci.**, v. 85, p. 7089-7093, 1988.

GUILLEY, H. et al. Transcription of cauliflower mosaic virus DNA; detection of promoter sequences and characterization of transcripts. **Cell**, v. 30, p. 763-773.

GUILTINAN, M. J.; MARCOTTE, W. R.; QUATRANO, R. S. A plant leucine zipper protein that recognizes an abscisic acid response element. **Science**, v. 250, p. 267-271, 1990.

GUPTA, C. P. et al. Antibiosis mediated necrotrophic effect of *Pseudomonas* GRC2 against two fungal plant pathogens. **Curr. Sci.**, v. 81, p. 91-94, 2001.

GURR, S. J.; RUSHTON, P. J. Engineering plants with increased disease resistance: how are we going to express it? **Trends Biotechnol.**, v. 23, p. 283-290, 2005.

HAMMOND-KOSACK, K. E.; PARKER, J. E. Deciphering plant-pathogen communication: fresh perspectives for molecular resistance breeding. **Curr. Opin. Biotechnol.**, v. 14, p. 177-193, 2003.

HAUPTMANN, R.M. et al. Evaluation of selectable markers for obtaining stable transformants in the Gramineae. **Plant Physiology**, v. 86, p. 602-606, 1988.

HE, C.; LIN, Z.; MCELROY, D.; WU, R. Identification of a rice *actin2* gene regulatory region for high-level expression of transgenes in monocots. **Plant Biotechnol. J.**, v. 7, p. 227-239, 2009.

HERRERA-ESTRELLA, L. A light inducible and chloroplast associated expression of a chimeric gene introduced into *Nicotiana tabacum* using a Ti plasmid vector. **Nature**, v. 310, p. 115-120, 1984.

HIGO, K. et al. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. **Nucleic Acids Research**, v. 27, p. 297-300, 1999.

HOBBO, T. et al. ACGT-containing abscisic acid response element (ABRE) and coupling element 3 (CE3) are functionally equivalent. **The Plant J.**, v. 19, p. 679-689, 1999.

HOEREN F. U. et al. Evidence for a role of AtMYB2 in the induction of the Arabidopsis alcohol dehydrogenase (ADH1) gene by low oxygen. **Genetics**, v. 149, p. 479-490, 1998.

HOSSAIN, M. A. et al. Transformation of Foreign gene in sugarcane variety Isd 28 using *Agrobacterium*-mediated method. **Bangladesh J. Sugarcane**, v. 29, p. 1-6, 2007.

HOSHINO A. A. **Isolamento e caracterização de promotores tecido-específicos a partir das informações do SUCEST ("Sugarcane expressed sequence tags")**, 2007. 144 f. Tese Doutorado (Doutor em Ciências) - Universidade Estadual Paulista Júlio de Mesquita Filho, Botucatu, São Paulo, 2007.

HUANG, N. et al. The tissue-specific activity of a rice beta-glucanase promoter (*Gns9*) is used to select rice transformants. **Plant Sci.**, v. 161, p. 589-595, 2001.

HUGHES, M. A.; DUNN, M. A. The molecular biology of plant acclimation to low-temperature. **J. Exp. Bot.**, v. 47, p. 291-305, 1996.

HUTCHINSON, G. B. The prediction of vertebrate promoter regions using differential hexamer frequency analysis. **Comput. Appl. Biosci.**, v. 12, p. 391-398, 1996.

HWANG, Y. S. et al. Three *cis*-elements required for rice α -amylase *Amy3D* expression during sugar starvation. **Plant Molecular Biology**, v. 36, p. 331-341, 1998.

ILUMINA. 2012. Disponível em: <<http://www.illumina.com.cn/applications/epigenetics.asp>>. Acesso em: 9 Nov. 2012.

ISHII, K. et al. Chromatin boundaries in budding yeast: the nuclear pore connection. **Cell**, v. 109, p. 551-562, 2002.

IYER, V. R. et al. Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. **Nature**, v. 409, p. 533-538, 2001.

JACOB, F.; MONOD, J. Genetic regulatory mechanisms in the synthesis of proteins. **J. Mol. Biol.**, v. 3, p. 318-356, 1961.

JEFFERSON, R. A.; KAVANAGH, T. A.; BEVAN, M. W. GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. **EMBO J.**, v. 6, p. 3901-3907, 1987.

JIANG, Y. et al. Analysis of GDSL lipase (GLIP) family genes in rice (*Oryza sativa*). **Plant Omics Journal**, v. 5, p. 351-358, 2012.

JOAQUIM, L. M.; EL-HANI, C. N. A genética em transformação: crise e revisão do conceito de gene. **Scientle Studia**, v. 8, p. 93-128, 2010.

KARTHIKEYAN, et al. Regulated expression of Arabidopsis phosphate transporters. **Plant Physiol.**, v. 130, p. 221-233, 2002.

KAYO, T. et al. Influences of aging and caloric restriction on the transcriptional profile of skeletal muscle from rhesus monkeys. **Proc. Natl. Acad. Sci.**, v. 98, p. 5093-5098, 2001.

KEIL, M. et al. Localization of elements important for the wound-inducible expression of a chimeric potato proteinase inhibitor II-CAT gene in transgenic tobacco plants. **Plant Cell**, v. 2, p. 61-70, 1990.

KELLER, H. et al. Pathogen-induced elicitor production in transgenic tobacco generates a hypersensitive response and non-specific disease resistance. **Plant Cell**, v. 11, p. 223-235, 1999.

KELLOGG A. E. Evolutionary History of the Grasses. **Plant Physiology**, v. 125, p. 1198-1205, 2001.

KIM, T. H. et al. A high-resolution map of active promoters in the human genome. **Nature**, v. 436, p. 876-880, 2005.

KLEIN, T. M. et al. Factors influencing gene delivery into *Zea mays* cells by high-velocity microprojectiles. **Bio/Technology**, v. 6, p. 559-563, 1988.

KLOSE, V. et al. Development of a competitive exclusion product for poultry meeting the regulatory requirements for registration in the European Union. **Nutr. Food Res.** v. 50, p. 563-571, 2006.

KONDRAHIN, Y. V. et al. Eukaryotic promoter recognition by binding sites for transcription factors, **Comput. Appl. Biosci.**, v. 11, p. 477-488, 1995.

LAKSHMANAN, P. et al. Sugarcane Biotechnology: The Challenges and Opportunities. **In Vitro Cell Dev. Biol-Plant**, v. 41, p. 345-363, 2005.

LAKSHMANAN, P. et al. Developmental and hormonal regulation of direct shoot organogenesis and somatic embryogenesis in sugarcane (*Saccharum* spp. Interspecific hybrids) leaf culture. **Plant Cell Reports**, v. 25, p. 1007-1015, 2006.

LAN, N.; JANSEN, R.; GERSTEIN, M. Towards a systematic definition of protein function that scales to the genome level: Defining function in terms of interactions. **Proc. IEEE**, v. 90, p. 1848-1858, 2002.

LAN, N.; MONTELIONE, G. T.; GERSTEIN, M. Ontologies for proteomics: Towards a systematic definition of structure and function that scales to the genome level. **Curr. Opin. Chem. Biol.** v. 7, p. 44-54, 2003.

LAST, D. I. et al. pEmu: an improved promoter for gene expression in cereal cells. **Theor. Appl. Genet.**, v. 81, p. 581-588, 1991.

LEVINE, M.; TJIAN, R. Transcriptional regulation and animal diversity. **Nature**, v. 424, p. 147-151, 2003.

LIU, Z. B. et al. Soybean *GH3* promoter contains multiple auxin-inducible elements. **Plant Cell**, v. 6, p. 645-657, 1994.

LIU, Z.; TAUB, C. C.; MCCLUNG, C. R. Identification of an *Arabidopsis thaliana* Ribulose-1-5-Bisphosphate Carboxylase/Oxygenase Activase (RCA) minimal promoter regulated by light and the circadian clock. **Plant Physiol.**, v. 112, p. 43-51, 1996.

LOGEMAN, J. et al. 5' upstream sequences from the *wun1* gene are responsible for gene activation by wounding in transgenic plants. **Plant Cell**, v. 1, p.151-158, 1989.

MACLEOD, D. et al. Sp1 sites in the mouse *aprt* gene promoter are required to prevent methylation of the CpG island. **Genes & Dev.**, v. 8, p. 2282-2292, 1994.

MANJUNATH, S.; SACHS, M. M. Molecular characterization and promoter analysis of the maize cytosolic glyceraldehydes 3-phosphate dehydrogenase gene family and its expression during anoxia. **Plant Molecular Biology**, v. 33, p. 97-112, 1997.

MARTIN, C.; PAZ-ARES, J. MYB transcription factors in plants. **Trends in Genetics**, v. 13, p. 67-73, 1997.

MASON, H. S.; DEWALD, D. B. MULLET, J. E. Identification of a methyl jasmonate-responsive domain in the soybean *vspB* promoter. **Plant Cell**, v. 5, p. 241-251, 1993.

MATZKE, M.; MATZKE, A. J. M. Homology-dependent gene silencing in transgenic plants what does it really tell us? **Trends in Genetics**, v. 11, p. 1-3, 1995.

MAURI, I. et al. Functional expression of the transcriptional activator Opaque-2 of *Zea mays* in transformed yeast. **Molecular and General Genetics**, v. 241, p. 319-326, 1993.

MCDOWELL, J. M.; WOFFENDEN, B. J. Plant disease resistance genes: recent insights and potential applications. **Trends Biotechnol.**, v. 21, p. 178-183, 2003.

MCELROY, D. et al. Construction of expression vectors based on the rice actin 1(Act1) 5' region for use in monocot transformation. **Mol. Gen. Genet.**, v. 231, p. 150-160, 1991.

MCELROY, D. et al. Isolation of an efficient actin promoter for use in rice transformation. **Plant Cell**, v. 2, p. 163-171, 1990.

MELOTTO-PASSARIN, D. M. **Transfomação genética de cana-de-açúcar por biobalística e *Agrobacterium tumefaciens* visando estudar o mecanismo de morte celular programada**. 2009. 147 f. Tese (Doutorado em Ciências) - Escola Superior de Agricultura Luiz de Queiroz, Piracicaba, São Paulo, 2009.

MENG, C. X. et al. Characterization of carotenoid hydroxylase gene promoter in *Haematococcus pluvialis*. **Indian J. Biochem. Biophys.**, v. 43, p. 284-294, 2006.

MICHELMORE, R. W. The impact zone: genomics and breeding for durable disease resistance. **Curr. Opin. Plant Biol.**, v. 6, p.397-404, 2003.

MISHRA, S. K.; WATKINS, S. C.; TRAUB, L. M. The autosomal recessive hypercholesterolemia (ARH) protein interfaces directly with the clathrin-coat machinery. **Proc. Natl. Acad. Sci.**, v. 99, p. 16099-16104, 2002.

MITRA, A.; HIGGINS, D. W. The Chlorella virus adenine methyltransferase gene promoter is a strong promoter in plants. **Plant Mol. Biol.**, v. 26, p. 85-93, 1994.

MODY, M. et al. Genome-wide gene expression profiles of the developing mouse hippocampus. **Proc. Natl. Acad. Sci.**, USA, v. 98, p. 8862-8867, 2001.

MOHANTY, B. et al. Detection and preliminary analysis of motifs in promoters of anaerobically induced genes of different plant species. **Annals of Botany**, v.96, p. 669-681, 2005.

MORISHIMA, A. Identification of preferred binding sites of a light-inducible DNA-binding factor (MNF1) within 5'- upstream sequence of C4-type phosphoenolpyruvate carboxylase gene in maize. **Plant Mol. Biol.**, v. 38, p. 633-646, 1998.

MUCHHAL, U. S.; PARDO, J. M.; RAGHOTHAMA, K. G. Transcriptional regulation of plant phosphate transporters. **Proc. Natl. Acad. Sci.**, v. 93, p. 10519-10523, 1996.

MURASHIGE, T.; SKOOG, F. A. A revised medium for a rapid growth and bioassays with tobacco tissues cultures. **Plant Physiol.**, v. 15, p. 473-479, 1962.

NAMBARA, E.; MARION-POLL, A.; Absciscic acid biosynthesis and catabolism. **Annul Review of Plant Biology**, v. 56, p. 165-185, 2005.

NEHRA, N. D. et al. Self-fertile transgenic wheat plants regenerated from isolated scutellar tissues following microprojectile bombardment with two distinct gene constructs. **Plant J.**, v. 5, p. 285-297, 1994.

NELSON, G. C. (Ed.) **Genetically modified organisms in agriculture: economics and politics**. San Diego, CA: Academic Press, 2001.

ODELL, J. T.; NAGY, F.; CHUA, N. H. Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. **Nature**, v. 6, p. 810-812, 1985.

OLIVE M. R. et al. Functional properties of the anaerobic response element of the maize Adh1 gene. **Plant Molecular Biology**, v. 15, p. 593-604, 1990.

OTT, R. W.; CHUA, N. H. Enhancer sequences from *Arabidopsis thaliana* obtained by library transformation of *Nicotiana tabacum*. **Mol. Gen. Gent.**, v. 223, p. 169-179, 1990.

PARVEEZ, G. K. A. et al. Determination of minimal inhibitory concentration of selection agents for selecting transformed immature embryos of oil palm. **Asia Pacific Journal of Molecular Biology and Biotechnology**, v. 15, p. 133-146, 2007.

PEARSON, H. Genetics: what is a gene? **Nature**, v. 441, p. 398-401, 2006.

PHILLIPS, T.; HOOPES, L. Transcription factors and transcriptional control in eukaryotic cells. **Nature Education**, v. 1, n. 1, 2008.

PIRKKALA, L.; NYKANEN, P.; SISTONEN, L. Roles of the heat shock transcription factors in regulation of the heat shock response and beyond. **FASEB J.**, v. 15, p. 1118-1131, 2001.

POTRYKUS, I. Gene transfer to plants: assessment and perspectives. **Physiologia Plantarum**, Copenhagen, v. 79, p. 125-134, 1990.

POSZKOWSKI, J. et al. Direct gene transfer to plants. **EMBO J.**, v. 3, p. 2717-2722, 1984

POTENZA, C.; ALEMAN, L.; SENGUPTA-GOPALAN, C.; Invited review: Targeting transgene expression in research, agricultural and environmental application: Promoters used in plant transformation. **In Vitro Cellular and Developmental Biology-Plant**, v. 40, 2004.

PRESTRIDGE, D. S.; STORMO, G. Signal scan 3.0: new database and program features. **Comput. Appl. Biosci.** v. 9, p. 113-115, 1993.

PRESTRIDGE, D. S. Predicting Pol II promoter sequences using transcription factor binding sites. **J. Mol. Biol.**, v. 249, p. 923-932, 1995.

QIN, X. F. et al. Immediate early transcription activation by salicylic acid via the cauliflower mosaic virus as-1 element. **Plant Cell**, v. 6, p. 863-874, 1994.

QIU, P. Recent advances in computational promoter analysis in understanding the transcriptional regulatory network. **Biochemical and Biophysical Research Communications**, v. 309, p. 495-501, 2003.

QUANDT, K. et al. MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. **Nucleic Acids Res.**, v. 23, p. 4878-4884, 1995.

QUANDT, K.; FRECH, K. WERNER, T. Analysis of transcription regulatory regions based on recognition of transcription factor binding sites and relative arrangement. **Molec. Biol. (Mosk)**, v. 31, p. 749-758.

RAGHOTAMA, K. G. Phosphate acquisition. **Annu Rev. Plant Physiol**, v. 50, p. 665-693, 1999.

RALPH, S. G.; JANCSEK, S.; BOHLMANN, J. Dirigent proteins in conifer defense II: extended gene discovery, phylogeny, and constitutive and stress-induced gene expression in spruce (*Picea* spp.). **Phytochemistry**, v. 68, p. 1975-1991, 2007.

RATHUS, C.; BOWER, R.; BIRCH, R. G. Effects of promoter, intron and enhancer elements on transient gene expression in sugarcane and carrot protoplasts. **Plant Mol. Biol.**, v. 23, p. 613-618, 1993.

REUBER, T. L.; AUSUBEL, F. M. Differential mRNA display. In: GALBRAITH, D. W.; BOHNERT, H.; BORQUE, D. P. **Methods in cell biology**. New York: Academic Press, 1995. p. 431-440.

RICHMOND, T. A.; SOMERVILLE, C. R. The cellulose synthase superfamily, **Plant Physiol**, v. 124, p. 495-498, 2000.

ROBERTS, M. L. The use of functional genomics in synthetic promoter design. **Computational Biology and Applied Bioinformatics**, 2011.

ROHDE, W. et al. Structure of a chalcone synthase gene from *Hordeum vulgare*. **Plant Molecular Biology**, v. 16, p. 1103-1106, 1991.

ROSLAN, H. A. et al. Characterization of the ethanol-inducible *alc* gene-expression system in *Arabidopsis thaliana*. **Plant J.**, v. 28, p. 225-235, 2001.

ROUSTER, J. et al. Identification of a methyl jasmonate responsive region in the promoter of a lipoxygenase 1 gene expressed in barley grain. **Plant J.**, v. 11, p. 513-523, 1997.

RUSHTON, P. J. et al. Synthetic plant promoters containing defined regulatory elements provide novel insights into pathogen and wound-induced signaling. **The Plant Cell**, v. 14, p. 749-762, 2002.

SACHS, M. M.; FEELING, M.; OKIMOTO, R. The anaerobic proteins of maize. **Cell**, v. 20, p. 761-767, 1980.

SALTER, M. G. et al. Characterisation of the ethanol-inducible *alc* gene expression system for transgenic plants. **Plant J.**, v. 16, p. 127-132, 1998.

SAMBROOK, J.; FRITSCHI, E. F.; MANIATIS, T. **Molecular cloning**: a laboratory manual. New York: Cold Spring Harbor Laboratory Press, 1989.

SANFORD, J. C. The biolistic process. A new concept in gene transfer and biological delivery, **TIBTech**, v. 6, p. 299-302, 1988.

SANGER, F.; NICKLEN, S.; COULSON, A. R. DNA sequencing with chain-terminating inhibitors. **Proceedings of National Academy of Sciences of the United States of America**, v. 74, p. 5463-5467, 1977.

SCHENA, M. et al. Brown Quantitative monitoring of gene expression patterns with a complementary DNA microarrays, **Science**, v. 270, p. 467-470, 1995.

SCHLEDZEWSKI, K.; MENDEL, R. R. Quantitative transient gene expression: comparison of the promoters for maize polyubiquitin 1, rice actin1, maize-derived and CaMV 35S in cells of barley, maize and tobacco. **Transgen. Res.**, v. 3, p. 249-255, 1994.

SCHULZE-LEFERT, P. et al. Functional architecture of the light-responsive chalcone synthase promoter from parsley. **Plant Cell**, v. 1, p. 707-714, 1989.

SCHREM, H.; KLEMPNAUER, J.; BORLAK, J. Liver-enriched transcription factors in liver function and development. Part I: the hepatocyte nuclear factor network and liver-specific gene expression. **Pharmacol. Rev.**, v. 54, p. 129-158, 2002.

SCHWARTZ, D. Alcohol dehydrogenase in maize: genetic basis for multiple isozymes. **Science**, v. 164, p. 585-586, 1969.

SENGUPTA-GOPALAN, C. et al. Developmentally regulated expression of the bean β -phaseolin gene in tobacco seed. **Proc. Natl Acad. Sci.**, v. 82, p. 3320-3324, 1985.

SHI, Y. et al. Transcriptional repression by YY1, a human GLI-Kruppel-related-protein, and relief of repression by adenovirus E1A protein. **Cell**, v. 67, p. 377-388, 1991.

SHORE, P.; SHARROCKS, A. D. Regulation of transcription by extracellular signals. In: LOCKER, J. **Transcription factors**. San Diego: Academic Press, 2001. p. 113-135.

SMALE, S. T.; KADONAGA, J. T. The RNA polymerase II core promoter. **Annu. Rev. Biochem.**, v. 72, p. 449-479, 2003.

SNUSTAD, D. P.; SIMMONS, M. J. **Fundamentos de genética**. 4. ed. Rio de Janeiro: Guanabara Koogan, 2008. 903 p.

STITT, M. The use of transgenic plant to study the resulation of plant carbohydrate metabolism. **Australian Journal of Plant Physiology**, v. 22, p. 635-646, 1995.

STRACKE, R.; WERBER, M.; WEISSHAAR, B. The R2R3-MYB gene family in *Arabidopsis thaliana*. **Current Opinion in Plant Biology**, v. 4, p. 447-456, 2001.

THE ENCODE PROJECT CONSORTIUM. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. **Nature**, v. 447, p. 799-816, 2007.

TJIAN, R. Molecular machines that control genes. **Scientific American**, v. 272, p. 38-46, 1995.

TOMKINS, J. P. et al. Physical map and gene survey of the *Ochrobactrum anthropi* genome using bacterial artificial chromosome contigs. **Microb. Comp. Genomics**, v. 4, p. 203-217, 1999.

TUAN, D. Y. et al. An erythroid-specific, developmental-stage-independent enhancer far upstream of the human "β-like globin" genes. **Proc. Natl. Acad. Sci.**, v. 86, p. 2554-2558, 1989.

TZAFRIR, I. et al. The sugarcane bacilliform virus is active in both monocots and dicots. **Plant Mol. Biol.**, v. 38, p. 347-356, 1998.

VAN DER GEEST, A. H. M.; HALL, T. C. The β-phaseolin 5' matrix attachment region acts as an enhancer facilitator. **Plant Mol.**, v. 33, p. 553-557, 1997.

WALKER, J. C. DNA sequences required for anaerobic expression of the maize alcohol dehydrogenase 1 gene. **Proc. Natl. Acad. Sci. USA**, v. 19, p. 6624-6628, 1987.

WANG, Y. et al. Characterization of cis-acting elements regulating transcription from the promoter of a constitutively active rice actin gene. **Mol. Cell. Biol.**, v.12, p.3399-3406, 1992.

WASHIDA H. et al. Identification of cis-regulatory elements required for endosperm expression of the rice storage protein glutelin gene GluB-1. **Plant Mol. Biol.**, v. 40, p. 1-12, 1999.

WEEKS, J. R. et al. Locus-specific variation in phosphorylation state of RNA polymerase II in vivo: correlations with gene activity and transcript processing. **Genes Dev.**, v.7, p. 2329-2344, 1993.

WEI, F. et al. The Physical and Genetic Framework of the Maize B73 Genome. **Plos Genet.**, v. 5, 2009.

WEI, H.; ALBERT, H. H.; MOORE, P.H.; Differential expression of sugarcane polyubiquitin genes and isolation of promoters from two highly-expressed members of the gene family. **J. Plant Physiol.**, v. 155, p. 513-519, 1999.

WENDEL, J. Genome evolution in polyploids. **Plant Molecular Biology**, v. 42, p.225-249, 2000.

WHITE, K. P. et al. Microarray analysis of *Drosophila* development during metamorphosis. **Science**, v. 286, p. 2179-2184, 1999.

WILKINS, A. S. Genetic analysis of animal development. **Journal of Evolutionary Biology**, v. 6, p. 617-618, 1993.

WILMINK, A.; VAN DE VEN, B. C.; DONS, J. J. Activity of constitutive promoters in various species from the Liliaceae. **Plant Mol. Biol.**, v. 28, p. 949-955, 1995.

WOLFERTSTETTER, F. et al. Identification of functional elements in unaligned nucleic acid sequences by a novel tuple search algorithm. **Comput. Appl. Biosci.**, v. 12, p. 71-80, 1996.

WRAY, G. A. Promoter logic. **Science**, v. 279, p. 1896-1902, 1998.

WRAY, G. A. The evolution of transcriptional regulation in Eukaryotes. **Mol. Biol. Evol.**, v. 20, n. 9, p. 1377-1419, 2003.

XU, D. et al. Systemic induction of a potato *pin2* promoter by wounding, methyl jasmonate, and abscisic acid in transgenic rice plants. **Plant Mol. Biol.**, v. 22, p. 573-588, 1993.

YANAGISAWA, S.; SHEEN, J. Involvement of Maize dof zinc finger proteins in tissue-specific and light-regulated gene expression. **The Plant Cell**, v. 10, p. 75-89, 1998.

YANG, M. et al. A rapid and direct approach to identify promoters that confer high levels of gene expression in monocots. **Crop Science**, v. 43, p. 1805-1913, 2003.

YERI, S. **Functional Analysis of synthetic promoters and construction of expression cassettes with *xynA***. 2009. 51 p. Masters thesis (Science) - Department of Biotechnology College of Agriculture, University of Agricultural Sciences, Dharwad, 2009.

YILMAZ, A.; GROTWOLD, E. Componentes and Mechanisms of Regulation of Gene Expression. **Methods in Molecular Biology**, v. 674, p. 23-32, 2010.

ZHANG, W.; MCELROY, D.; WU, R. Analysis of rice *Act1* 5' region activity in transgenic rice plants. **Plant Cell**, v. 3, p. 1155-1165, 1991.

ZHANG, Z. D. et al. Statistical analysis of the genomic distribution and correlation of regulatory elements in the ENCODE regions. **Genome Res.**, v. 17, p. 787-797, 2007.

ZHAO, L. et al. OsAP2-1, an AP2-like gene from *Oryza sativa*, is required for flower development and male fertility. **Sex Plant Reprod.**, v. 19, p. 197-206, 2006.