

# Daniela Kajihara

Caracterização dos genes *mustang* em gramíneas com ênfase no estudo funcional em cana-de-açúcar

São Paulo 2010

# Daniela Kajihara

# Caracterização dos genes *mustang* em gramíneas com ênfase no estudo funcional em cana-de-açúcar

Dissertação apresentada ao Instituto de Biociências da Universidade de São Paulo, para a obtenção de Título de Mestre em Ciências, na Área de Botânica.

Orientadora: Maria Magdalena Rossi

São Paulo 2010

# Ficha Catalográfica

Kajihara, Daniela

Caracterização dos genes mustang em gramíneas com ênfase no estudo funcional em cana-deaçúcar.

86p.

Dissertação (Mestrado) - Instituto de Biociências da Universidade de São Paulo. Departamento de Botânica.

- Genética molecular de plantas.
   Elementos transponíves.
   Transposases.
   Genes mustang.
   Mutator.
- I. Universidade de São Paulo. Instituto de Biociências. Departamento de Botânica.

# Comissão Julgadora:

Prof(a). Dr(a).		Prof(a). Dr(a).
	Profa. Dra.	Maria Magdalena Rossi

Orientador(a)



Aos meus pais.

# Agradecimentos

À Profa. Dra. Maria Magdalena Rossi pela excelente orientação, incentivo, oportunidade, amizade e por estar presente em todas as etapas desse trabalho.

À Profa. Dra. Marie-Anne Van Sluys pelo incentivo e motivação na pesquisa científica, amizade, ser um exemplo de profissionalismo e por estar sempre de laboratório aberto para a realização deste trabalho.

À Prof. Dra. Hana Masuda pela amizade, ensinamentos, pelas excelentes dicas e ideias que foram muito importantes para este trabalho.

Ao pessoal do GMP lab: Fabiana, Juliana, Junia, Luísa, Sílvia, Thais e ao membro júnior, Santiago.

À Thais Hamaji por participar da temática cana enquanto todo mundo trabalhava com tomate e principalmente pela ajuda com o material vegetal usado neste trabalho.

À Fabiana, Andrea Brandão, Maximiller, Carolina e Renato pelas inumeráveis e indispensáveis dicas sobre clonagens, oligos e real time.

À Juliana Almeida, Robson e Oberdan por me ajudaram na parte de bioinformática e filogenia.

À Sílvia, Douglas, Erika, Marisa e Myna pelos diversos ensinamentos de laboratório e biomol. Ao Edgar, Mayra e Sílvia por estarem sempre dispostos a ajudar.

Aos colegas e amigos do GaTE lab: Andrés, Alessandro, Ana Paula, Breno, Bruno, Claudinha, Daniela M, Dani Q, Douglas, Edgar, Elis, Eliza, Érika, Guilherme, Hana, Juliana Nico, Juliane I, Jonas, Kleber, Leonor, Marisa, Mayra, Myna, Maria Eliza, Marcelo, Nathália, Nilo, Oberdan, Regina, Robson, Úrsula, Vivian, Vanessa e Wanessa pela ótima convivência.

A Profa. Roseli Corazzini, Prof. Paulo Moreno, Ana Paula Santos e Marcos Leite pela amizade, constante incentivo e por me apresentarem o "mundo da ciência".

Ao Waltir do Laboratório de Microscopia Confocal.

A todos do Laboratório de Fisiologia Vegetal principalmente a Ana Maria, Auri e Luciano.

A todos do Laboratório de Biocel. Em especial a Amanda, Fernanda e Júlia.

À Profa. Nanuza e a Paula do Laboratório de Anatomia Vegetal.

Ao pessoal do IQ, Carolina, Max, Paloma e Rodrigo.

Á Ju Ishida, Marcelo, Douglas, Edgar, Mayra e Bruno por deixarem a rotina de trabalho mais leve e alegre. E mais recentemente a Junia e Juliana pelos almoços e conversas. Ao pessoal sempre animado das algas, que também são plantas, Emmanuelle e Khey.

As amigas de estrada Christiane, Eliza, Karen, Keli, Keyde, Giordana, Vanessa e todos os amigos do curso de Biologia.

Aos meus pais e ao meu irmão Guilherme pelo incentivo, dedicação e carinho imprescindíveis. À minha grande família, tias, primas e primos por sempre me apoiarem. Especialmente minhas tias Alice, Lúcia, Maria e Rosa e as primas(os) Erika, Karina, Luciane e Adriano.

A todos os professores e funcionários do Departamento de Botânica do Instituto de Biociências/USP.

Agradeço a todas as pessoas que diretamente ou indiretamente colaboram na realização deste trabalho.

Este trabalho foi desenvolvido pelo apoio da Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

# Índice

Resumo	01
Abstract	03
I. Introdução	05
1. Elementos de transposição	05
2. Impacto e domesticação dos TEs	08
3. Superfamília <i>Mutator</i>	11
4. A cana-de-açúcar	13
5. Cana-de-açúcar e TEs	17
6. Superfamília <i>Mutator</i> em cana-açúcar	19
II. Objetivos	23
III. Materiais e Métodos	24
1. Análise comparativa entre os genes mustang de cana-de-açúcar, Oryza	sativa,
Sorghum bicolor e Zea mays.	24
1.1 Análise filogenética dos genes mustang .	24
1.2. Estudo comparativo entre as seqüências codificantes e regulató	rias dos
genes mustang de cana-de-açúcar, sorgo, milho e arroz.	26
2. Caracterização da atividade transcricional dos genes mustang de	
cana-de-açúcar.	26
2.1. Cepas bacterianas.	26
2.2. Material vegetal.	27
2.3. Vetores.	27
2.4. Meios de cultura para bactérias e células vegetais.	29
2.5. Extrações de DNA.	30
2.5.1. Extração de DNA de BACs	30
2.5.2. Extração de DNA de cana-de-açúcar	31
2.5.3. Extração de DNA de células BY-2.	33
2.5.4. Midi-preparações de DNA plasmidial.	33
2.6. Eletroforese em gel.	34
2.7. Obtenção de bactérias competentes.	34

	2.7.1. Escherichia coli.	34
	2.7.2. Agrobacterium tumefaciens	35
	2.8. Transformação bacteriana	36
	2.8.1. Transformação de Escherichia.coli	36
	2.8.2. Transformação de Agrobacterium tumefaciens por choque térmico.	37
	2.9. Amplificação e clonagem das regiões promotoras.	37
	2.10. Avaliação da presença dos fragmentos de interesse nos	
	sucessivos passos de clonagem.	39
	2.11. Análise das construções por ensaios de bombardeamento.	40
	2.12. Ensaio histoquímico para detecção da atividade da $eta$ -glucuronidase.	41
	2.13. Transformação estável de células de fumo BY-2	42
	2.14. Análise microscópica para visualização da	
	Green Fluorescent Protein (GFP).	43
	2.15. Tratamentos hormonais das linhagens de células transgênicas.	43
	2.16. Extração de RNA e síntese de cDNA.	44
	2.17. PCR em tempo real (Real Time PCR).	45
IV. I	Resultados	47
Сар	ítulo 1. Análise comparativa entre os genes <i>mustang</i> de cana-de-aç	úcar,
Ory	za sativa, Sorghum bicolor e Zea mays.	47
	1.1 Análise filogenética dos genes <i>mustang</i> .	47
	1.2. Estudo comparativo entre as seqüências codificantes e regulatórias	
	dos genes <i>mustang</i> de cana-de-açúcar, sorgo, milho e arroz.	51
Сар	nítulo 2. Caracterização da atividade transcricional de três genes <i>mustal</i>	ng de
can	a-de-açúcar.	55
	2.1 Obtenção das construções de fusão contendo os promotores a	
	montante dos genes repórteres GFP (green fluorecense protein)	
	e GUS ( $\beta$ -glucoronidase).	55
	2.2 Avaliação das construções através de ensaios de expressão	
	transitória por bombardeamento.	58
	2.3 Obtenção de linhagens de Nicotiana tabacum (BY2)	
	transgênicas para as construções de fusão.	59

2.4 Avaliação da indução hormonal dos promotores nas	
linhagens transgênicas.	61
2.5 Avaliação do padrão de expressão espacial dos genes	
endógenos em cana-de-açúcar.	62
V. Discussão	65
1. O genoma das gramíneas e os genes mustang	65
2. O padrão de expressão dos genes <i>mustang</i>	67
VI. Considerações finais	75
VII. Conclusões	77
VIII. Referências bibliográficas	79

#### **RESUMO**

Os elementos transponíveis constituem grande parte do genoma das plantas, particularmente em gramíneas, constituem entre 50 a 80% do conteúdo genômico. Recentemente, foi demonstrado que estes elementos servem como fonte de material genético para a formação de novos genes e novas redes regulatórias. O SUCEST, projeto de seqüenciamento de ESTs de cana-de-açúcar da FAPESP, gerou a següência parcial de 237.954 mRNA de diversos tecidos e condições fisiológicas, fornecendo valiosa informação sobre o transcriptoma deste cultivo. Um levantamento dos elementos transponíveis nesse genoma mostrou que o transposon Mutator é o mais expresso. A superfamília Mutator foi amplamente estudada em cana-de-açúcar, arroz e Arabidopsis thaliana e se constatou que o sistema está composto por dois clados de transposons "verdadeiros" (Classe I e Classe II) e dois clados de transposases domesticadas (Classe III e Classe IV), chamadas mustang. As transposases domesticadas são seqüências derivadas de transposons, que perderam a capacidade de se mobilizar, e adquiriram função celular. Recentemente, foram clonadas e seqüenciadas, pelo nosso grupo, duas cópias genômicas da Classe III e uma da Classe IV. Para somar evidências que permitam desvendar a função das proteínas MUSTANG, este trabalho realizou uma análise comparativa destes genes em gramíneas assim como o estudo da atividade transcricional em cana-de-açúcar. Desta forma, foram identificados os *loci* ortólogos no genoma de sorgo e milho, e foi possível verificar que os genes mustang são altamente conservados. As putativas regiões regulatórias dos genes de cana-de-açúcar apresentaram diversos motivos de união a fatores de transcrição envolvidos na resposta a luz, hormônios e estresse. Fusões com genes repórteres permitiram demonstrar que as regiões estudadas são promotores transcricionais ativos. Adicionalmente, a obtenção de linhagens de células de fumo transgênicas viabilizou experimentos que permitiram revelar que os promotores dos

genes *mustang* são modulados por fitohormônios. O perfil transcricional para ambas as classes revelou que estes genes são expressos de forma ubíqua, sendo o meristema o tecido que apresenta maiores níveis relativos de mRNA. A análise integrada dos resultados obtidos sugere o possível envolvimento das proteínas MUSTANG na manutenção da homeostase da resposta hormonal.

### **Abstract**

Transposable elements constitute a vast quantity of plant genomes, particularly in grasses, they comprise between 50 to 80% of genomic content. Recently, it has been demonstrated that these elements are source of genetic material for new genes creation and new regulatory network establishment. The Brazilian Sugarcane EST Sequencing Project, SUCEST, financed by FAPESP, generated 237.954 mRNA partial sequence derived from several tissues and different physiological conditions, providing a wide range of information of sugarcane transcriptome. A wide spectrum of transposable elements was identified, revealing the *Mutator* transposon as the most abundantly expressed transposable element in sugarcane genome. The Mutator superfamily was deeply explored in Arabidopsis, sugarcane and rice and it was found that the system comprises two clades of bona fide transposons (Class I and Class II), and two clades of domesticated transposases (Class III and Class IV), named mustang. The domesticated transposases are sequences that have lost their movement capacity and, acquired cellular function. Recently, two genomic copies of Class III and one for Class IV have been cloned and sequenced by our group. In order to gain evidences for unraveling the function of MUSTANG proteins, this work performs a comparative sequence analysis of these genes in grass genomes and a transcriptional activity profile study in sugarcane. Thus, the orthologous *loci* from sorghum and maize were identified, and it was verified that *mustang* genes are highly conserved in grass genomes. The putative promoter region of sugarcane genes displayed several transcription factor motifs involved in light, hormone and stress response. Reporter gene fusions showed that the studied regions are indeed transcriptional active promoters. Furthermore, transgenic lines of tobacco BY-2 cells demonstrated that the sugarcane mustang genes are modulated by phytohormones. The expression profile revealed that both classes are ubiquitously

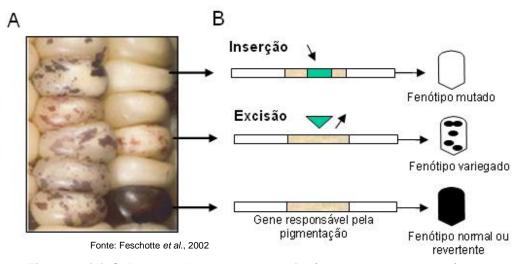
transcribed being the meristem the tissue that shows higher relative expression levels.

The integrated analysis of these results suggests a possible involvement of MUSTANG proteins in the homeostasis maintenance of hormonal response.

## I. Introdução.

## 1. Elementos de transposição (TEs).

Os elementos transponíveis (TEs) são segmentos de DNA (500bp-10.000pb) capazes de mudar de posição e/ou propagar-se no genoma (Okamoto e Hirochika 2001). Os TEs foram primeiramente descritos em milho, pela citogeneticista Barbara McClintock na década de 40, devido ao seu envolvimento na origem de mutações instáveis. Ela os denominou de "elementos controladores do gene" pelo fato de modularem a expressão dos genes aos quais estavam associados, observando a influência dos transposons no fenótipo variegado de grãos de milho (Figura1). Este fenótipo variegado é visível devido à inserção e posterior excisão de um transposon em um gene responsável pela pigmentação dos grãos de milho, causando a interrupção do seu produto gênico e posterior restabelecimento da pigmentação. Desde os primeiros ensaios McClintock já observava a alternância de fases de quiescência e reativação de transposons em milho (McClintock 1956).



**Figura 1.** (A) Grãos de milho apresentando fenótipo variegado devido à atividade de um elemento de transposição. (B) Representação esquemática da inserção de um elemento de transposição (verde) em um gene responsável pela pigmentação. A inserção do TE promove a inativação do gene, produzindo um fenótipo mutado. Dependendo do estágio de desenvolvimento do grão no momento da excisão, o restabelecimento da expressão gênica normal leva à reversão total ou a um fenótipo variegado.

Centenas de genomas eucariotos têm sido seqüenciados, analisando esses dados do ponto de vista filogenético os TEs se mostraram ubíquos, e enquanto o número de genes apresenta pouca diferença, a quantidade e diversidade dos elementos é altamente variável (Pritham 2009). Os TEs constituem a maior parte do DNA repetitivo presente no genoma de plantas e animais, constituindo cerca de 45 % do genoma humano e 50 a 80% em genomas vegetais (Feschotte *et al.* 2002).

Dependendo do intermediário utilizado para sua transposição, os TEs são divididos em duas classes. A Classe I, ou retrotransposons, são elementos que se mobilizam através de um intermediário de RNA que é reverso transcrito em uma cópia de DNA que posteriormente é inserida no genoma. Este mecanismo é chamado de "copia e cola". Os elementos de Classe II, ou transposons, se mobilizam através de um intermediário de DNA, na sua maioria, utilizam um mecanismo do tipo "corta e cola" pelo qual o fragmento de dupla fita de DNA se excisa do genoma e se reinsere em uma nova posição do genoma (Pritham 2009). Em ambas as classes existem elementos não-autônomos e autônomos. Os elementos autônomos codificam todas as enzimas necessárias para a sua mobilização, enquanto que os elementos não-autônomos são mutantes, dependendo inteiramente da maquinaria enzimática dos elementos autônomos para se transpor (Feschotte *et al.* 2002).

Para facilitar a anotação da crescente quantidade de dados que está sendo gerada sobre TEs, os elementos da Classe II ou transposons, em organismos eucarióticos são divididos em duas subclasses. A Subclasse I compreende aqueles transposons clássicos que se mobilizam pelo mecanismo de "corta e cola" gerando quebras na dupla fita de DNA durante o processo de transposição. Os elementos de tipo *Helitrons* e *Mavericks* pertencem à Subclasse II. Enquanto que os *Helitrons* possuem um mecanismo de replicação do tipo círculo rolante, os *Mavericks* utilizam

uma DNA polimerase própria (Sinzelle et al. 2009, Pritham 2009). Os elementos autônomos da Subclasse I são caracterizados por codificar uma transposase e pela presença das inversões terminais repetidas flanqueando o elemento (TIRs). Durante o ciclo de transposição, um domínio de união ao DNA (DBD) no extremo N-terminal da transposase reconhece as TIRs, e o sítio catalítico no extremo C-terminal quebra e mobiliza o elemento. O sítio catalítico está determinado pelo domínio constituído pelos aminoácidos DDE (Haren et al. 1999). Dependendo da seqüência das TIRs, dos motivos protéicos entre elas, e do comprimento das duplicações do sítio de inserção (TSDs) geradas após a transposição, esta subclasse apresenta nove superfamílias descritas (Tabela 1) (Sinzelle et al. 2009). Um estudo recente demonstrou que os genes que codificam transposases são os mais abundantes da natureza (Aziz et al. 2010).

Diversas evidências indicam que o tamanho do genoma está diretamente correlacionado com a presença de retrotransposons, pois de acordo ao seu mecanismo de transposição, após cada ciclo, eles aumentam o número de cópias. No entanto, essa premissa não se aplica para os transposons que se mobilizam, majoritariamente através de mecanismos conservativos (Feschotte e Pritham 2007, Devos 2010).

Tabela 1. Propriedades estruturais e moleculares das nove superfamílias de DNA transposons pertencentes à subclasse I.

Superfamília	Ocorrência	Tamanho (l	(b) TIRs (bp)	TSDs (bp)	Proteínas codificadas	DBD	Núcleo catalítico
Tc1/mariner	Eucariotos	1,2 - 5,0	17-1100	2 (TA)	Tnp	HTH	(DDE) Tnp
hAT	Eucariotos	2,5 - 5,0	5-27	8	Tnp	BED ZnF	(DDE) Tnp
Mutator	Eucariotos	1,3 - 7,4	0 - vários kb	9-11	Tnp	WRKY/GCM1 ZnF	(DDE) Tnp
Merlin	Animais e eubacteria	1,4 - 3,5	21-462	8-9	Tnp	nd	(DDE) Tnp
Transib	Metazoários e fungos	3 - 4	9-60	5	Tnp	nd	(DDE) Tnp
P	Plantas e matazoários	3 - 11	13-150	8	Tnp	THAP ZnF	nd
piggyBac	Eucariotos	2,3 - 6,3	12-19	4 (TTAA)	Tnp	nd	nd
PIF/Harbinger	Eucariotos	2,3 - 5,5	15-270	3 (CWG ou TWA)	Tnp + proteína Myb-like	MyB/SANT	(DDE) Tnp
CACTA	Plantas, metazoários e fungos	4,5 - 15	10-54	2-3	TnpA + TnpD	nd	nd

DBD: Domínio de ligação ao DNA, TSD: sítio alvo duplicado, Tnp: transposase e nd: não determinado. Adaptado de Sinzelle *et al.* 2009.

A sua natureza mutagênica, aliada a sua abundância, levanta o questionamento de como o hospedeiro mantém a sua viabilidade. Na verdade, a maioria dos TEs se encontra em estado quiescente, sendo ativados em condições fisiológicas específicas abióticos. como estresses bióticos ou Complexos mecanismos regulatórios estabelecem um equilíbrio que garantem a perpetuidade desses elementos no genoma hospedeiro (Lisch 2009). A transcrição é o primeiro ponto de controle da atividade dos TEs (Grandbastien et al. 2005). Por outro lado, a transcrição de um TE não se correlaciona necessariamente com a presença de novas inserções já que, para completar o ciclo de transposição, além da transcrição é necessário que ocorra a tradução do mensageiro, transcrição reversa no caso dos retrotransposons, e a integração em um novo local. A regulação em qualquer destes níveis limita a transposição (Slotkin e Martienssen 2007). Os mecanismos moleculares envolvidos na regulação dos TEs envolvem interferência de RNA (RNAi) e alterações epigenéticas como a modificação pós-traducional de histonas levando à remodelação da cromatina (Sinzelle et al. 2009).

### 2. Impacto e domesticação dos TEs.

Os TEs são comumente considerados "parasitas genômicos" e foram historicamente chamados de "DNA egoísta". O seu ciclo de vida compreende a dispersão em diversas espécies, aumento do número de cópias, e se perpetuar no genoma hospedeiro até, finalmente, virarem fósseis perdendo a sua capacidade de replicação e mobilização. Durante a sua existência como elementos ativos, ao se transpor, os TEs geram mutações por inserção causando inativação gênica ou mudanças no padrão de expressão. Neste sentido, diversos trabalhos tem descrito

exemplos nos quais TEs aportaram seqüências promotoras ou *enhancers* transcricionais, sítios de *splicing* alternativo ou de poliadenilação, e motivos regulatórios em *cis* (Bowen e Jordan 2002, Marino-Ramírez *et al.* 2005). Por outro lado, a presença de seqüências repetidas ao longo do genoma provê sítios de recombinação ilegítima gerando translocações, inversões e deleções. Um fenômeno interessante e amplamente descrito é a captura de fragmentos duplicados do genoma hospedeiro pelo TE dando origem a quimeras (Saccaro-Jr *et al.* 2007). Devido ao seu mecanismo de transposição replicativa, a atividade dos retrotransposons é um dos maiores fatores de variação no tamanho do genoma em plantas (Devos 2010).

As evidências expostas demonstram claramente o papel crucial que as sequências móveis possuem na evolução dos genomas eucariotos. O mais drástico exemplo deste impacto é o processo descrito como "domesticação molecular" no qual seqüências derivadas de TEs dão origem a genes funcionais do hospedeiro (Benjak 2008). Alguns destes genes surgiram da fusão de um domínio funcional com uma transposase ou, nos casos mais extremos, da domesticação de uma transposase inteira. Neste último caso, o elemento perde as suas TIRs, perdendo assim a capacidade de se transpor, e adquire uma função celular. Estes genes são chamados de "transposases domesticadas".

Como mencionado anteriormente, as transposases possuem dois domínios funcionais: o DBD e o catalítico. Embora muitas proteínas domesticadas provêem de uma transposase inteira e consequentemente mantém ambos os domínios, o DBD é o preferencialmente incorporado pelo hospedeiro e consequentemente a maioria das transposases domesticadas desempenham função de reguladores transcricionais (Feschotte 2008). Neste contexto, têm sido caracterizadas proteínas que atuam como fatores de transcrição (Yamashita *et al.* 2007, Butter *et al.* 2010), remodeladores da

cromatina (Macfarlan *et al.* 2005), e até reguladores da tradução (Liu *et al.* 2003). Regularmente, o processo evolutivo de domesticação vem acompanhado da modificação do domínio catalítico DDE e as proteínas derivadas de transposons têm demonstrado ser defectivas para complementar em *trans* a transposição de elementos não-autônomos. A única exceção até o momento descrita é o caso da proteína RAG1 envolvida na recombinação somática responsável pela diversidade de imunoglobulinas das células T em humanos e outros vertebrados (Agrawal *et al.*1998, Kapitonov e Jurka 2005). Levando em consideração os potenciais efeitos deletérios da atividade transposicional, tem sido proposto que as transposases domesticadas possam participar da regulação dos TEs correlatos, seja desencadeando o mecanismo de RNAi ou como antagonistas da atividade das transposases selvagens através da heterodimerização e conseqüente complementação negativa dominante (Rossi *et al.* 2004, Sinzelle *et al.* 2009).

Quando se discute sobre a origem de genes do hospedeiro derivados de elementos móveis é importante mencionar que a evolução frequentemente atua de maneira convergente e, conseqüentemente, é teoricamente possível que as transposases tenham surgido a partir de genes do hospedeiro. O argumento mais contundente para estabelecer esta relação evolutiva provê da árvore filogenética. Por exemplo, enquanto que as transposases domesticadas apresentam origem relativamente recente, os elementos relacionados possuem uma distribuição filogenética muito mais ampla podendo assim se inferir que os genes do hospedeiro surgiram a partir de uma transposase e não *vice versa* (Sinzelle *et al.* 2009).

Embora a maioria das transposases domesticadas não tenha sido completamente caracterizada ainda, em particular nas plantas alguns exemplos interessantes foram reportados. Como o gene *Daysleeper* descrito em *A. thaliana*,

derivado do transposon Activator da superfamília hAT. Esta transposase domesticada reconhece o motivo Kubox1 presente em vários promotores vegetais, incluindo o gene da maquinaria de reparo de DNA Ku70. Plantas silenciadas ou que super expressam Daysleeper apresentam um padrão de desenvolvimento anormal e o perfil de expressão gênica completamente alterado. Assim, foi demonstrado que este gene é um fator de transcrição mestre do qual depende o padrão morfogenético normal da planta (Bundock e Hooykaas 2005). O fitocromo A (phyA) media diversas respostas das plantas a estímulos luminosos. Em A. thaliana foram caracterizados os genes fhy3 e far1 que codificam proteínas relacionadas à transposase da superfamília Mutator. Estes dois genes atuam conjuntamente modulando a transdução de sinal do phyA ativando a transcrição dos genes fhy1 e fhl, os quais são necessários para a acumulação do phyA ativado no núcleo da célula vegetal. Por sua vez, phyA regula negativamente a expressão de fhy3 e far1. Este sistema apresenta um exemplo no qual uma proteína com atividade de união a DNA, altamente eficiente foi capturada pelo genoma vegetal para manter a homeostase da resposta á luz mediada pelo phyA (Lin et al. 2007). Finalmente, o produto do gene rosina (rsi) de Antirrhinum majus foi identificado como uma proteína de união a DNA, que regula genes cujos promotores possuem o motivo CArG-box, como é o caso do gene deficiens que controla o desenvolvimento floral nessa espécie. RSI é uma proteína do transposon Tam pertencente à superfamília CACTA (Roccaro et al. 2007). Plantas transgênicas silenciadas para rsi apresentaram alterações morfológicas nas sementes, relacionando mais uma vez o desenvolvimento vegetal aos elementos genéticos móveis.

## 3. Superfamília Mutator.

A superfamília *Mutator* (*Mu*) foi primeiramente relatada por Robertson (1978) em milho, e tem sido descrito como o mais ativo e mutagênico transposon em plantas descoberto até o momento. Devido a esta característica, *Mu* tem sido usado como uma ferramenta para a clonagem de genes por "transposon tagging" (Walbot 2000). Nos últimos dez anos, o sistema *Mutator* tem sido caracterizado em diversas espécies vegetais como *Arabidopsis thaliana*, arroz e cana-de-açúcar (Yu *et al.* 2000, Saccaro-Jr *et al.* 2007).

O elemento autônomo da superfamília *Mutator, MuDR*, responsável pela mobilização do sistema, possui dois genes que são transcritos em sentido inverso: *mudrA* e *mudrB* (Figura 2). A região codificante está flanqueada por TIRs de 220 pb onde se localizam os promotores. O gene *mudrA* codifica para a transposase MURA, enzima responsável pela transposição, enquanto o produto de *mudrB* é requerido para a integração. No entanto, a presença de *mudrB* não é uma característica geral dentro da superfamília *Mutator* e foi encontrado somente nos elementos do gênero *Zea*.

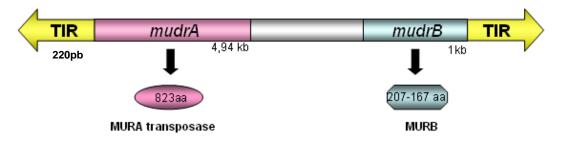


Figura 2. Estrutura do elemento autônomo MuDR de milho

Além do elemento autônomo, o sistema *Mutator* compreende um conjunto de elementos não autônomos altamente variáveis em seqüência e estrutura denominados *MuLEs* (*Mutator-like elements*). Estes representam à maioria dos elementos do sistema e dependem do autônomo para a sua transposição (Lisch 2002) (Figura 3). Enquanto que poucas linhagens de milho possuem elementos ativos, todos os genótipos portam

elementos homólogos a *MuDR*, chamados de *hMuDRs* cujas regiões codificantes são entre 80 e 90 % idênticas ao elemento autônomo. Estes elementos são expressos e participam da regulação negativa do sistema através de mecanismos de silenciamento gênico transcricional (TGS) e pós-transcricional (PTGS) (Rudenko & Walbot 2001, Slotkin *et al.* 2003). Quando além de *mudrA*, os elementos possuem um fragmento de DNA hospedeiro são chamados de *Transduplicated MuLEs* (Juretic *et al.* 2005). Em alguns casos, os elementos apresentam, apenas DNA do hospedeiro entre TIRs, neste caso, são definidos como *Pack-MuLEs* (Jiang *et al.* 2004). Finalmente, os genes *mustang*, transposases domesticadas da superfamília *Mutator*, completam a diversidade do sistema. Estes genes, ao contrário de *far1* e *fhy3*, foram descritos tanto em espécies de mono quanto eudicotiledôneas (Cowan *et al.* 2005, Saccaro-Jr *et al.* 2007).

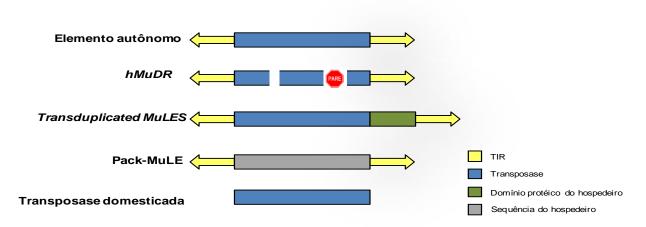


Figura 3. Os diferentes tipos de elementos que compõem o sistema Mutator

## 4. A cana-de-açúcar.

A cana-de-açúcar (Saccharum spp.) é uma gramínea cultivada em regiões tropicais e subtropicais destinadas principalmente a produção de açúcar e biocombustíveis, principalmente bioetanol. É uma das culturas mais importantes

economicamente, e está juntamente com o milho, arroz e trigo, no topo mundial das commodities agrícolas com respeito a produtividade (Devos 2010).

Assim como o sorgo (*Sorghum bicolor*) e o milho (*Zea mays*), a cana-de-açúcar pertence à família Poaceae, subfamília Panicoideae, tribo Andropogoneae. Já outras gramíneas de importância agronômica como o arroz (*Oryza sativa*) e o trigo (*Triticum aestivum*) pertencem às subfamílias Ehrhartoideae e Pooideae, respectivamente. Seis espécies de *Saccharum* (*S. spontaneum, S. robustum, S. officinarum, S. barberi, S. sinense e S. edule*) juntamente com espécies de gêneros próximos, formam o "complexo *Saccharum*", a partir do qual surgiu a cana de açúcar cultivada (Mukerjee 1957, Daniels *et al.* 1975). De um ponto de vista prático, os recursos genéticos da cana-de-açúcar podem ser divididos em três grupos: (1) cultivares tradicionais, (2) espécies selvagens aparentadas, e (3) cultivares modernos (Grivet *et al.* 2004).

- (1) Os cultivares tradicionais, descendentes dos primeiros domesticados, já não são cultivados para produção, mas utilizados como genitores dos cultivares modernos e constituem importantes fontes de germoplasma para a introdução de caracteres de interesse agronômico. Estes incluem:
- O cultivar nobre S. officinarum (2n=80), rico em açúcar e ainda utilizado na agricultura tradicional especialmente na Melanésia.
- As espécies *S. barberi* e *S. sinense* que eram comumente encontradas na Índia e China, respectivamente, onde provavelmente tenha-se originado a indústria açucareira. Estas possuem mais de 80 cromossomos e menores conteúdos de açúcar que *S. officinarum*.
- S. edule utilizado para consumo humano na Melanésia.
- (2) As espécies selvagens aparentadas, genitores dos cultivares tradicionais incluem:

- *S. spontaneum* (2n=40-128) possui baixo teor de açúcar e ampla distribuição geográfica, desde algumas ilhas do Pacífico, Melanésia e Ásia Tropical, passando pelo Médio Oriente até a África.
- S. robustum (2n=60 ou 2n=80) praticamente não acumula açúcar e é encontrada na Melanésia.
- Espécies dos gêneros *Erianthus, Miscanthus, Sclerostachya* e *Narenga*. Estes gêneros possuem ampla distribuição desde o Himalaia até a Melanésia. Uma particular concentração de espécies destes gêneros é encontrada no nordeste da Índia.
- (3) Os cultivares modernos, que substituíram os tradicionais ao longo do século 20, foram produzidos através de cruzamentos artificiais entre *S. officinarum*, com alto teor de açúcar, e *S. spontaneum*, selvagem e vigorosa, seguidos de várias gerações de seleção clonal. Assim, foi possível aliar características como resistência a doenças e alta produção de açúcar. Desta maneira, os cultivares modernos possuem genomas complexos com entre 100 a 130 cromossomos, dos quais 15-25% provém de *S. spontaneum* e o restante de *S. officinarum*, são altamente poliplóides e apresentam aneuplóidias (Figura 4) (D'Hont 2005).

A poliploidia tem sido o maior fator na evolução das linhagens de gramíneas devido ao aumento do número de genes, ativação de elementos transponíveis, alteração do panorama epigenético acarretando novos padrões de expressão gênica e potencial criação ou perda gênica (Kashkush *et al.* 2005). Apesar do alto nível de ploidia, o tamanho dos genomas monoplóides de *S. officinarum* e *S. spontaneum* está em torno de 930 e 750 Mpb, respectivamente; enquanto que o dos cultivares híbridos é de 1000 Mpb (D'Hont e Glaszmann 2001). Assim, para cada *locus* de cana-de-açúcar há em torno de 12 cópias ou haplótipos, a maioria proveniente do genitor *S. officinarum*, alguns de *S. spontaneum* e outros recombinantes.

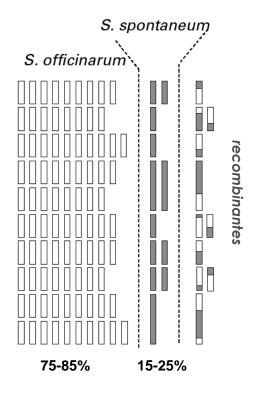


Figura 4. Representação esquemática do genoma dos cultivares modernos de cana-de-açúcar. Barras brancas e cinzas correspondem aos cromossomos provenientes de *S. officinarum* e *S. Spontaneum*, respectivamente (Adaptado deD'Hont, 2005).

Em 1999, uma iniciativa da Fundação de Amparo à Pesquisa no Estado de São Paulo (FAPESP) deu o primeiro grande passo para o conhecimento do genoma da cana-de-açúcar abrindo um amplo panorama para o desenvolvimento científico e tecnológico em torno desse cultivo de imensa importância econômica mundial. O Projeto SUCEST (Sugarcane Expressed Sequence Tags) foi o responsável pelo següenciamento do transcriptoma da cana-de-açúcar. Foram parcialmente següenciados mais de 260.000 clones de cDNA, provenientes de 26 bibliotecas de cDNA, geradas a partir de RNA extraído de diferentes tecidos e condições fisiológicas. Com isso foram produzidas 237.954 ESTs (Expressed Sequence Tags) de alta qualidade. Estas ESTs foram montadas (sobrepostas) resultando em 43.141 transcritos consenso. A anotação dessas següências associou guase 50% destes transcritos com metabolismo protéico, comunicação celular e transdução de sinal, bioenergética e resposta a estresse. Os elementos de transposição representaram 2,3% das seqüências anotadas, enquanto que 16,8% das seqüências não apresentaram homologia com nenhuma seqüência de DNA previamente caracterizada. Uma análise comparativa entre os 43.141 transcritos revelou uma redundância de 22%, indicando que foram identificados 33.620 genes em cana de açúcar. Uma vez que em arroz ou tomate existem aproximadamente 35.000 genes, número também estimado para cana, o número obtido equivale a mais de 90% dos genes da planta (Vettore *et al.* 2003).

# 5. Cana-de-açúcar e TEs.

A partir dos dados produzidos no projeto SUCEST, foi realizado um levantamento a procura de elementos transponíveis revelando uma grande quantidade e diversidade de TEs no transcriptoma de cana-de-açúcar (Rossi et al. 2001). Através de uma busca adstringente foram identificados 276 cDNAs representando 21 elementos diferentes. 54% das següências resultaram ser transcritos de transposons enquanto que 46% de retrotransposons. As superfamílias mais representadas em cana-de-açúcar foram Mutator e Hopscotch para transposons e retrotransposons, respectivamente. Em 2005, Araújo e Rossi et al. realizaram um estudo transcricional com 68 clones de cDNA de cana-de-açúcar, previamente descritos por Rossi et al. (2001) demonstrando que diversas superfamílias de TEs são expressos em calos, meristema apical, folhas e flores em desenvolvimento. Dentre essas superfamílias, Mutator representou 38% dos TEs expressos, seguido por 13% do retrotransposon Hopscotch. Analisando a distribuição da expressão nos diferentes tecidos testados, foi possível demonstrar que a cultura de tecidos induz drasticamente a transcrição tanto de transposons quanto de retrotransposons. Experimentos de transformação transitória e estável demonstraram, para três dos elementos Hopscotch caracterizados, que os

promotores são transcricionalmente ativos, confirmando o padrão de expressão observado através de macroarranjos.

Análises genômicas revelaram que durante sua história evolutiva, 70% das angiospermas sofreram poliploidização. As espécies com origem poliplóide geralmente provêem de genitores diplóides, assim apresentam níveis mais altos de diversidade genética e genômica que seus ancestrais. Após a hibridação, o genoma percorre um processo de reorganização que inclui deleções, translocações, inversões e duplicações que resulta em um novo padrão de expressão gênica e até no surgimento de novas funções (Soltis e Soltis 2000, He et al. 2003). Também foi demonstrado que a hibridação interespecífica ativa a retrotransposição e consequentemente os rearranjos cromossômicos vinculando assim a atividade dos TEs com a reestruturação do genoma após o processo de alopoliploidização (Kashkush et al. 2003, Bennetzen 2007). Por exemplo, os trigos poliplóides possuem mais cópias de TEs que a soma dos seus genitores. No entanto, quando são analisados híbridos sintéticos recentes não são detectadas diferenças significativas (Li et al. 2004). Da mesma maneira, não foram reportadas evidências de transposição em alopoliplóides sintéticos resultantes do cruzamento de Arabidopsis thaliana e Arabidopsis lyrata até a terceira geração (Beaulieu et al. 2009). Enquanto que não foi demonstrado um aumento da transposição em curto prazo após a alopoliploidização, as mudanças epigenéticas drásticas, como a alteração no padrão de metilação, levam a uma consequente ativação da transcrição dos TEs (Parisod et al. 2010). Neste sentido, a cana-de-açúcar é um perfeito exemplo do "estresse genômico" produzido pela hibridação recente reforçado pela sua propagação vegetativa. Os altos níveis de TEs presentes no transcriptoma descritos reforçam a hipótese de que o progressivo aumento da expressão após a alopoliploidização poderia resultar num aumento da transposição e conseqüente aumento no número de cópias no genoma. Apesar disso, a "explosão" no número de cópias não parece ser uma regra e pode estar restrita exclusivamente a algumas famílias de elementos (Parisod *et al.* 2010).

## 6. Superfamília Mutator em cana-açúcar.

Nos genomas de arroz, cana, milho e sorgo a superfamília de transposons *Mutator* é uma das mais representadas (Rossi *et al.* 2001; Paterson *et al.* 2009). Em 2004, Rossi *et al.* realizaram um estudo aprofundado sobre a diversidade das seqüências de cana-de-açúcar homólogas ao gene que codifica para a transposase do elemento autônomo do sistema *Mutator*, *mudrA*. A análise filogenética comparativa com *Arabidopsis thaliana* e *Oryza sativa* mostrou a existência de quatro clados que foram denominados Classe I, Classe II, Classe III e Classe IV. A identificação de seqüências ortólogas de arabidopsis e cana-de-açúcar em cada um dos ramos permitiu concluir que a diversificação do sistema *Mutator* ocorreu precocemente nas Angiospermas previamente a divergência de mono e eudicotíledoneas. Tanto o domínio de união a DNA quanto o sítio catalítico da transposase MURA foi identificado nas següências pertencentes às quatro classes (Rossi *et al.* 2004).

Com o intuito de conhecer o número de cópias e a estrutura dos elementos de cada classe no genoma de arroz e cana-de-açúcar, Saccaro-Jr et al. (2007) adotou três estratégias: i) a triagem de uma biblioteca genômica de cana-de-açúcar utilizando sondas específicas para cada classe, ii) o mapeamento in silico de todas as seqüências homólogas à transposase mudrA no genoma de arroz e, iii) a caracterização da estrutura de cada elemento de arroz identificado. Os resultados revelaram dados interessantes sobre a história evolutiva do sistema Mutator em gramíneas. Ponderando os genomas monoploides, e levando em consideração que a hibridação molecular é

mais adstringente que o mapeamento in silico, número de cópias semelhantes foram observadas para cada classe em ambas as espécies (Tabela 2). No entanto, entre as classes foram observadas diferenças de uma ordem de magnitude, onde Classe II é a mais abundante seguida de Classe I revelando que a amplificação diferencial das classes foi anterior à divergência entre as espécies. As análises estruturais permitiram revelar que enquanto as Classes I e II possuem TIRs e constituem elementos de transposição verdadeiros, os elementos de Classe III e IV são transposases domesticadas do tipo mustang (Cowan et al. 2005) não apresentando mais estrutura de transposon. A análise filogenética dos genes mustang descritos em arroz, arabidopsis, e cana-de-açúcar revelou um único evento de domesticação seguido de uma duplicação que deu origem às classes III e IV (Figura 5). Os genes de Classe III apresentam dois clados enquanto que os de Classe IV estão divididos em três ramos (Tabela 2). A topologia da árvore está perfeitamente de acordo com o número de cópias de cada classe indicadas na Tabela 2. A presença de mais de uma següência de cana dentro de cada clado é devida à poliploidia, lembrando que o genoma de canade-açúcar apresenta aproximadamente 12 haplotipos para cada loci.

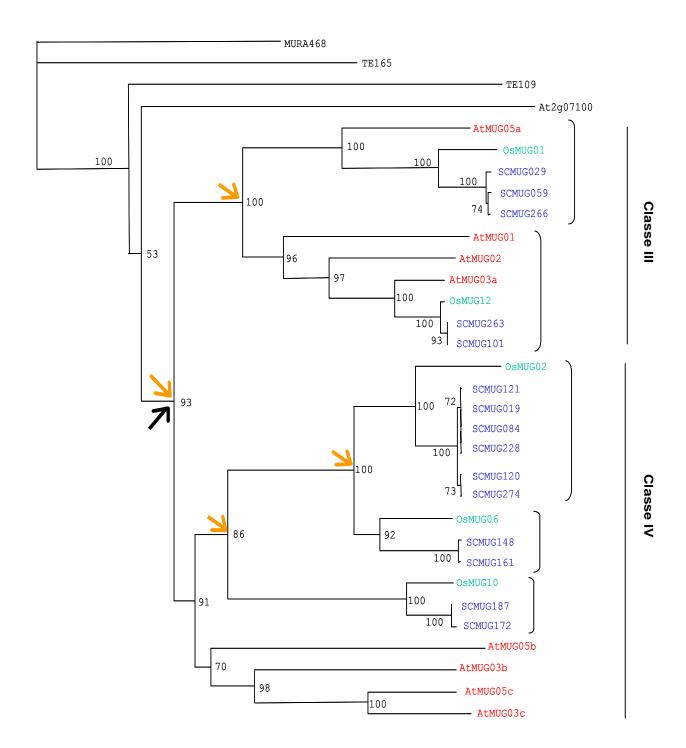
Tabela 2. Número de cópias das sequências semelhantes a MudrA em cana-de-açúcar e arroz.

	Número de cópias no genoma monoplóide (1000 Mb) de cana-de-açúcar	Número de cópias no genoma monoplóide (430 Mb) de arroz
Classe I	28	50
Classe II	172	386
Classe III	~1	2
Classe IV	3	3

Finalmente, Saccaro-Jr (2007) sequenciou duas cópias genômicas dos genes mustang de cana-de-açúcar de Classe III provindas de cada um dos genitores: *S. officinarum* e *S. spontaneum*. As análises comparativas com arroz demonstraram uma perfeita colinearidade entre as regiões genômicas que contém estes genes. Desta

maneira, as relações filogenéticas e de sintenia suportam sua ortologia e descartam a possibilidade de que a ampla distribuição destes genes seja devida à transferência horizontal. A ampla diversidade e distribuição dos genes *mustang* gera perguntas sobre a possível função desta família multigênica. Há especificidade funcional entre as classes? Existe especialização funcional entre os diferentes clados de cada classe? Há mais de uma cópia ativa dentro de cada classe ou clado?

Com o intuito de acrescentar informações sobre a funcionalidade das transposases domesticadas MUSTANG em cana-de-açúcar este trabalho propôs ampliar a análise comparativa para outras espécies de gramíneas cujos genomas foram recentemente seqüenciados e, estudar a atividade transcricional destes genes.



**Figura 5**. Árvore filogenética dos genes *MUSTANG* em gramíneas. Árvore publicada por Saccaro-Jr *et al.* (2007). Setas pretas e laranjas indicam eventos de domesticação e duplicação, respectivamente.

## II. Objetivo

O objetivo geral do projeto foi aprofundar no estudo dos genes *mustang* em gramíneas. Por um lado, foram realizadas análises comparativas com os genomas recentemente seqüenciados de sorgo e milho. Além disso, foi proposta a caracterização funcional com o intuito de somar dados sobre o possível papel destes genes e avaliar se as diferentes cópias de um mesmo gene são ativas no genoma poliplóide da cana de açúcar.

Desta maneira foram propostos os seguintes objetivos específicos:

- Realizar uma análise comparativa entre os genes mustang de cana-de-açúcar,
   Oryza sativa, Sorghum bicolor e Zea mays.
- 1.1. Acrescentar à filogenia dos genes *mustang* existente as sequências genômicas de cana-de-açúcar, sorgo e milho.
- 1.2. Realizar um estudo comparativo entre as seqüências codificantes e regulatórias dos genes *mustang* de cana-de-açúcar, sorgo, milho e arroz.
- 2. Caracterizar a atividade transcricional de três genes *mustang* de cana-de-açúcar, dois Classe III e um de Classe IV.
  - 2.1 Obter construções de fusão contendo os promotores dos genes *mustang* de cana-de-açúcar a montante dos genes repórteres GFP (*green fluorescent protein*) e GUS (β-glucoronidase).
  - 2.2 Avaliar as construções através de ensaios de expressão transitória por bombardeamento.
  - 2.3 Obter linhagens de células de *Nicotiana tabacum* (BY-2) transgênicas para as construções de fusão.
  - 2.4 Avaliar a indução hormonal dos promotores nas linhagens transgênicas.
  - 2.5 Avaliar o padrão de expressão espacial e temporal dos genes endógenos em cana-de-açúcar.

### III. Materiais e Métodos

1. Análise comparativa entre os genes *mustang* de cana-de-açúcar, *Oryza sativa*, Sorghum bicolor e Zea mays.

# 1.1. Análise filogenética dos genes mustang.

Os genes *mustang* objeto de estudo deste trabalho, foram identificados por Saccaro-Jr *et al.* (2007) a partir de uma biblioteca genômica da variedade de cana-de-açúcar R570, construída no vetor pBeloBAC11 pelo CUGI (http://www.genome.clemson.edu/protocols/), utilizando como sonda os clones de cDNA de *mudrA* TE266(Classe III) e TE148 (Classe IV) (Rossi *et al.* 2004). Posteriormente, as seqüências completas dos genes foram obtidas por Saccaro-Jr (2007).

Os três loci *mustang* de cana foram chamados de: SCMUG266 BAC148 (GU080318), haplótipo proveniente do genitor *Saccharum spontaneum*; SCMUG266 BAC095 (GU080321), haplótipo proveniente do genitor *Saccharum.officinarum*; e o SCMUG148 BAC249 (GU080320) haplótipo proveniente do genitor *Saccharum.officinarum*.

A análise filogenética dos genes *mustang* foi realizada com as sequências codificantes completas dos genes de Classe III e Classe IV de cana-de-açúcar, previamente descritas, as de arroz e arabidopsis já analisadas por Saccaro-Jr *et al.* (2007), e as de sorgo e milho identificadas neste trabalho. Os genes *mustang* de sorgo e milho foram identificados nos respectivos bancos de dados através do algoritmo BLAST (Altschul *et al.* 1990, Zhang *et al.* 1997) utilizando as seqüências de cana-de-açúcar e arroz.

As sequências de arroz foram obtidas do banco de dados TIGR Rice (http://rice.plantbiology.msu.edu), os loci são: LOC\_Os01g41210 (OsMUG01), LOC\_Os02g09900 (OsMUG02), LOC\_Os06g42640 (OsMUG06), LOC\_Os10g01550 (OsMUG10) e LOC\_Os12g40530 (OsMUG12).

As sequências de Arabidopsis foram retiradas do banco de dados TAIR (http://www.arabidopsis.org) com o seguinte loci, AT5G16505.1 (AtMUG5a), AT5G34853.1 (AtMUG5b), AT2G30640.1 (AtMUG2), AT1G06740.1 (AtMUG01), AT3G04605.1 (AtMUG3a) AT3G05850.1 (AtMUG3b), AT3G06940.1 (AtMUG3c) e AT2G07100.

A partir do banco de dados do Phytozome (http://www.phytozome.net) foram identificadas as sequências de sorgo dos seguintes loci: Sb01g027410.1 (SbMUG01), Sb03g002010.1 (SbMUG03), Sb04g006220.1 (SbMUG04), Sb05g027680 (SbMUG05), Sb08g020220.1 (SbMUG08) e Sb10g024700.1 (SbMUG10).

No banco de dados MaizeGDB de milho (http://www.maizegdb.org) foram identificadas as sequências de milho dos seguintes loci: AC177909 (246.725-245.351 parcial, ZmMUG01a) AC190982 (202.769-204551, ZmMUG03a), AC201818 (110.632-112.416, ZmMUG08), AC204517 (72.081-74.315, ZmMUG01b), AC194150 (47.510-49.765, ZmMUG10), AC221000 (28.198-30.453, ZmMUG03b), AC202094 (94.687-100.808, ZmMUG01c), AC215910 (131.798-132.870 parcial, ZmMUG09a) e AC209850 (21.913-24.150, ZmMUG09b).

As seqüências foram alinhadas com o programa ClustalW (Higgins *et al.* 1994). A matriz de distância foi construída com o algoritmo de *neighbor-joining* (Saito e Nei 1987) utilizando o programa MEGA 3.1 (Kumar *et al.* 2004). A análise de bootstrap foi realizada com 1000 replicatas.

1.2. Estudo comparativo entre as seqüências codificantes e regulatórias dos genes mustang de cana-de-açúcar, sorgo, milho e arroz.

O grau de identidade entre as sequências dos genes ortólogos de cana-deaçúcar, sorgo, milho e arroz foi calculado por alinhamento global com o EMBOSS Pairwise Alignment Algorithm (http://www.ebi.ac.uk/Tools/emboss/align/).

A busca por motivos de união para fatores de transcrição sobre as regiões promotoras dos genes foi realizada com o programa PLACE (Higo *et al.* 1999).

2. Caracterização da atividade transcricional dos genes *mustang* de cana-de-açúcar.

#### 2.1. Cepas bacterianas.

Para a obtenção das construções de fusão entre os promotores dos genes de cana-de-açúcar e os genes repórteres foi utilizada a cepa de *Escherichia.coli* DH10B. Para o cultivo das bactérias foi utilizado o meio líquido Luria Bertani (LB) (Sambrook *et al.* 1989).

Para a transformação de células de fumo BY-2 foi utilizada a cepa de Agrobacterium tumefaciens GV3101. A mesma foi cultivada em meio LB com rifampicina (20 μg/ml), gentamicina (50 μg/ml).

As cepas bacterianas, assim como todos os clones obtidos foram estocados a -80°C com 25% de glicerol.

#### 2.2. Material vegetal.

Para o estudo da expressão em sistema heterólogo, foram transformadas células de *Nicotiana tabacum L. cv. bright yellow-2 (BY-2)* (Nagata *et al.* 1992). As células foram mantidas em meio BY-2 em estufa com agitação (130rpm) a 28°C no escuro, e repicadas semanalmente.

Para o estudo do perfil transcricional dos genes *mustang* foram cultivadas plantas de cana-de-açúcar da variedade R-570. Fragmentos de colmo contendo um nó foram cultivados em casa de vegetação em vasos plásticos contendo terra vegetal e vermiculita (1:1). Amostras de tecidos de 15 e 30 dias, após emergência da gema, foram coletadas. Foram coletados os seguintes tecidos: folha mais jovem totalmente expandida (2º metáfilo e 4º metáfilo para 15 e 30 dias, respectivamente), meristema caulinar contendo parte da bainha da folha e do caule, e as raízes.

As amostras de tecidos adultos foram coletadas de plantas adultas de aproximadamente 3 anos e 1,70 m de altura, mantidas no Jardim do Instituto de Biociências da Universidade de São Paulo. Foram coletados os seguintes tecidos: a folha mais jovem totalmente expandida, meristema caulinar e raízes.

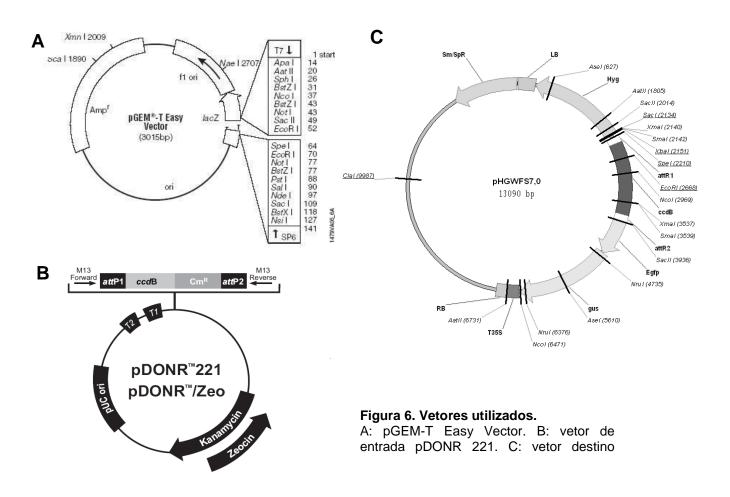
Todas as amostras foram limpas com etanol 70% e/ou lavadas com água para a retirada da terra vegetal. Rapidamente foram congelados em nitrogênio líquido no local da coleta e armazenados em freezer -80°C até o momento da extração de RNA.

#### 2.3. Vetores.

Durante o processo de obtenção das construções de fusão entre os promotores dos genes *mustang* e os genes repórteres foram utilizados três vetores plasmidiais (Figura 6). O pGEM T-Easy (Promega Cat. No. A1360) com resistência a ampicilina

(100 μg/ml). O vetor de entrada do sistema GATEWAY de recombinação pDONR<sup>TM</sup>221 (Invitrogen, Cat. No. 12536-017) que possui resistência a canamicina (50 μg/ml). Finalmente, o vetor de destino pHGWFS7 (Karimi *et al.* 2002) com resistência a estreptomicina (20 μg/ml) e espectinomicina (50 μg/ml) que contém o sitio de recombinação a montante de uma proteína de fusão entre os genes repórteres da *Green Fluorescent Protein* (*GFP*) e da β-glucuronidase (*GUS*). Este último vetor é binário possibilitando a transformação genética de células vegetais *via Agrobacterium tumefaciens*.

Como controle positivo dos experimentos de bombardeamento foi utilizado o vetor pCambia1391Z.35S que possui resistência a canamicina (50 µg/ml) e contém o gene repórter *GUS* sob controle do promotor 35S do vírus do mosaico da couve-flor.



Materiais e Métodos

2.4. Meios de cultura para bactérias e células vegetais.

**LB** (Luria Bertani): 10 g/l de bactotriptona, 5 g/l de extrato de levedura e 10 g/l de NaCl. Ajustar o pH 7,0 com NaOH 5N. Esterilizar na autoclave por 20 minutos a 120°C.

SOB: 20 g/l de bactotriptona, 5 g/l de extrato de levedura, 0,5 g/l NaCl, 2,5 mM KCl e 1

mM MgCl<sub>2</sub>. Esterilizar na autoclave por 20 minutos a 120°C.

SOC: 20 g/l de bactotriptona, 5 g/l de extrato de levedura, 0,5 g/l NaCl, 2,5 mM KCl, 1

mM MgCl<sub>2</sub> e 20 mM de glicose. Esterilizar na autoclave por 20 minutos a 120°C.

YEB: 5 g/l de extrato de carne, 1 g/l de extrato de levedura, 5 g/l de peptona, 5 g/l de

sacarose e 240 mg/l de MgSO<sub>4</sub>. Ajustar o pH 6,8 com NaOH 0,1N. Esterilizar na

autoclave por 20 minutos a 120°C.

Meio BY-2: 500 mg/l de MES, 4,3 g/l de sais básicos do meio MS (Murashige e Skoog

1962), 30 g/l de sacarose, 10 ml/l de B1- Inositol, 3 ml/l de vitamina Millers, 1 µM de

2,4-D, 6 g/l de phytoágar (para meio sólido). Ajustar o pH 5,7 com KOH. Esterilizar na

autoclave por 20 minutos a 120°C.

Soluções estoque:

B1-Inositol: 0,1 g/l Tiamina; 10 g/l de mio-inositol. Dividir a solução em tubos de 10 ml.

Congelar.

Vitamina Millers: 60 g/l de KH4PO4.Colocar em tubos de 1,5 ml e congelar.

29

## 2.5. Extrações de DNA.

## 2.5.1. Extração de DNA de BACs (adaptado de Sambrook et al. 1989).

- 1. Colocar um pré-inóculo de bactéria para crescer em 2 ml de meio LB com 12,5 mg/l de cloranfenicol durante 4-6 horas.
- 2. Inocular 100 µl do pré-inóculo em 100 ml de meio LB com antiobiótico e deixar crescer durante 14-16 horas.
- 3. Após período, centrifugar por 10 minutos a 6500 RPM (Centrífuga MPW-350R, rotor nº 11469) a temperatura de 4° C.
- 4. Descartar sobrenadante.
- 5. Ressuspender o precipitado através de agitação leve, sem pipeta, em 4 ml de GTE [Glicose 50 mM, Tris (trishidroximetilaminometano)-HCl 25 mM e EDTA (ácido etilenodiamino tetra-acético) 10 mM]. Transferir o tubo para o gelo por 10 minutos.
- 6. Adicionar 8 ml de uma solução NaOH 0,5N e SDS (dodecil sulfato de sódio) 1%. Misturar e deixar no gelo por 10 minutos.
- 7. Adicionar 6 ml de solução KOAc (Para 60 ml de acetato de potássio 5M adicionar 11,5 ml de ácido acético glacial e 28,5 ml de água. A concentração final será 3 M para potássio e 5 M para acetato.) e deixar novamente no gelo por 10 minutos.
- 8. Centrifugar a 4 °C por 50 minutos a 11000 RPM (Centrífuga MPW-350R, rotor nº 11454).
- 9. Transferir para outro tubo e adicionar um volume de isopropanol correspondente a 80% do volume contido no tubo. Incubar de 30 a 60 minutos à temperatura ambiente.
- 10. Centrifugar por 20 minutos a 7000 RPM (Centrífuga MPW-350R, rotor nº 11454).à temperatura ambiente. Descartar o sobrenadante.
- 11. Ressuspender o precipitado, gentilmente e sem micropipeta, em 500 µl de tampão TE (Tris-HCl 10 mM e EDTA 1 mM).

- 12. Transferir para um microtubo e adicionar 500 μl de uma solução de cloreto de lítio 5M a 4°C.
- 13. Centrifugar a 4 °C por 10 minutos a 14000 RPM (Centrífuga eppendorf, rotor F45-24-11).
- 14. Transferir o sobrenadante para tubos novos e adicionar 500 µl de isopropanol. Incubar por 10 minutos à temperatura ambiente.
- 15. Centrifugar por 10 minutos a 14000 RPM (Centrífuga eppendorf 5415D, rotor F45-24-11) à temperatura ambiente. Descartar o sobrenadante.
- 16. Ressuspender o DNA em 200 µl de TE.
- 17. Adicionar 4 µl de RNAse A 10 mg/ml e incubar por 15 minutos.
- 18. Adicionar 400 µl de NaCl 1,6M/PEG8000 13% gelado. Agitar gentilmente.
- 19. Incubar no gelo por 30 minutos.
- 20. Centrifugar por 5 minutos a 10000 RPM (Centrífuga eppendorf, rotor F45-24-11) à temperatura ambiente.
- 21. Descartar o sobrenadante, sem inversão, com ajuda de uma micropipeta.
- 22. Ressuspender o DNA em 200 µl de TE.
- 22. Adicionar 150 μl de solução NH<sub>4</sub>OAc 7,5 M, misturar por agitação.
- 23. Adicionar 1100 µl de etanol absoluto à temperatura ambiente.
- 22. Deixar durante 14-16 horas a -20 °C.
- 23. Centrifugar e diluir em 100 µl de tampão TE.
- 24. A integridade do DNA foi conferida por eletroforese em gel de agarose.

#### 2.5.2. Extração de DNA de cana-de-açúcar

Para a extração de DNA de cana-de-açúcar foi usado o método de CTAB (Bernatzky e Tanksley 1986):

- 1. Macerar o material vegetal em cadinho com nitrogênio líquido e PVPP (polivinilpolipirrolidona).
- 2. Com o auxílio de uma espátula transferir o macerado para um tubo falcon contendo o tampão CTAB (vide receita abaixo). Para cada 3 g de material vegetal usar de 18 ml de tampão CTAB + 36 μl de β-mercaptoetanol.
- 3. Colocar em banho-maria a 65°C por 30 minutos. Invertendo o tubo algumas vezes.
- 4. Retirar o tubo do banho-maria e colocar na estante.
- 5. Não realizar movimentos bruscos com o tubo.
- 6. Colocar 18 ml de clorofórmio: álcool isoamílico (24:1).
- 7. Fechar o tubo e misturar levemente. Levar para a centrífuga a 5000 RPM (Centrífuga MPW-350R, rotor nº 11469) por 20 minutos em 22°C.
- 8. Transferir a fase aquosa para um tubo novo.
- 9. Acrescentar o dobro do volume de fenol equilibrado: clorofórmio: álcool isoamílico (25:24:1).
- 10. Colocar novamente na centrífuga para mais 20 minutos 5000 RPM (Centrífuga MPW-350R, rotor nº 11469) a 22°C.
- 11. Retirar a primeira camada e colocar em um tubo de ensaio (10 ml).
- 12. Colocar o dobro de volume da solução de fenol equilibrado: clorofórmio: álcool isoamílico (25:24:1).
- 13. Colocar novamente na centrífuga para mais 20 minutos 5000 RPM (Centrífuga MPW-350R, rotor nº 11454) a 22°C.
- 14. Retirar a primeira camada e colocar em um tubo de ensaio (10 ml).
- 15. Colocar o dobro de volume da solução de clorofórmio: álcool isoamílico. (24:1)
- 16. Centrifugar novamente por mais 20 minutos 5000 RPM (Centrífuga MPW-350R, rotor nº 11454) a 22°C
- 17. Retirar a primeira camada e colocar em um tubo de ensaio.

- 18. Colocar 2/3 de isopropanol e inverter o tubo.
- 19. Ocorrendo a visualização do enovelado de DNA, retirar com uma pipeta Pasteur com a ponta torta colocando em um microtubo com Etanol 70%.
- 20. Colocar a amostra para centrifugar por 5 a 10 minutos a 3000 RPM (Centrífuga MPW-350R, rotor nº 11454) a 22°C.
- 21. Retirar da centrífuga. Inverter o tubo para a retirada do isopropanol.
- 22. Colocar a amostra na centrífuga a vácuo, para evaporação do etanol.
- 23. Depois dissolver a amostra com 100 µl de TE (10 mM de Tris-HCl (pH 8,0), 1mM de EDTA (pH 8,0).
- 24. A integridade do DNA purificado foi conferida por eletroforese em gel de agarose.

Tampão CTAB: 4 g de CTAB (brometo de cetil-trimetilamônio), 56 ml de cloreto de sódio 5 M, 8 ml de EDTA 0,5 M, 20 ml de Tris pH 8 e 2 g de PVP (polivinilpirrolidona). Completar o volume para 200 ml. Esterilizar na autoclave e depois adicionar β-mercaptoetanol 400 μl com a concentração final de 10 mM.

#### 2.5.3. Extração de DNA de células BY-2.

Para conferir a presença do transgene por PCR nas linhagens de células de fumo BY-2, o DNA foi purificado usando o kit DNeasy Plant Minikit QIAGEN (Cat. No. 69104).

## 2.5.4. Midi-preparações de DNA plasmidial.

Para a purificação de DNA plasmidial, tanto no processo de obtenção das construções de fusão assim como para bombardeamento, 2 ml de LB foram inoculados

com uma colônia bacteriana e mantidos a 37°C com agitação por 16 a 18 horas. Depois desse período, o cultivo bacteriano foi transferido para um microtubo de 2 ml, centrifugado por 1 minuto a 7000 RPM (Centrífuga eppendorf, rotor F45-24-11) e o meio de cultura foi retirado. O precipitado foi usado para a extração de DNA plasmidial com o kit QIAprep Spin MiniprepKit - QIAGEN (Cat. No. 27104). A integridade do DNA purificado foi conferida por eletroforese em gel de agarose.

## 2.6. Eletroforese em gel.

Para a análise de DNA genômico, plasmidial ou produtos de PCR foi realizado eletroforese em gel de agarose. Dependendo do tamanho dos fragmentos, a concentração de agarose foi entre 0,8% a 1,0%, dissolvida em tampão TBE 1/2X (TBE 10x: 108 g/l de Tris-base, 55 g/l de ácido bórico, 40 ml de EDTA 0,5M pH 8,0).

Para carregar as amostras no gel foi utilizada uma solução contendo: 50 mg azul de bromofenol, 50 mg xileno cianol, 6 ml de glicerol e 14 ml de água deionizada.

Para a visualização, os géis foram corados em uma solução de brometo de etídeo (Fluka, Cat. No. 46067) de 5 μg/ml, lavados em água destilada e visualizados em um transiluminador da marca Vilber Lourmart.

#### 2.7. Obtenção de bactérias competentes.

#### 2.7.1. Escherichia coli.

A obtenção de bactérias competentes foi realizada segundo descrito por Hanahan et al. (1983):

- 1. A partir de uma colônia de E.coli DH10B cultivar um inóculo de 2 ml em meio SOB durante 18 horas a 37°C sob agitação constante.
- 2. Inocular 50 ml de meio SOB com 2 ml de cultivo e incubar a 37°C sob agitação constante até atingir 0,6 de densidade óptica a 600 nm.
- 3. Adicionar 0,5 ml de MgCl<sub>2</sub> e incubar no gelo por 15 minutos.
- 4. Centrifugar a 3000 RPM (Centrífuga MPW-350R, rotor nº 11469) por 20 minutos a 4°C.
- 5. Descartar o sobrenadante, ressuspender as bactérias em 15 ml de solução RFI e manter por 10 minutos no gelo.
- 6. Centrifugar a 3000 RPM (Centrífuga MPW-350R, rotor nº 11469) por 20 minutos a 4°C.
- 7. Descartar o sobrenadante e ressuspender as bactérias em 2 mL da solução RFII.
- 8. Aliquotar as bactérias em tubos com de 50 µl.
- 9. Congelar com nitrogênio líquido e manter em freezer a -80°C.

**Solução RFI:** Acetato de potássio pH 6,9 (30 mM), cloreto de manganês (50 mM), cloreto de potássio (100 mM), cloreto de cálcio (10 mM) e glicerol (15%). Esterilizar em autoclave por 15 minutos a 120°C.

**Solução RFII:** MOPS pH 7,0 (10 mM), cloreto de cálcio (75 mM), cloreto de potássio (10 mM) e glicerol (15%). Esterilizar em autoclave por 15 minutos a 120°C.

#### 2.7.2. Agrobacterium tumefaciens.

A obtenção de bactérias competentes foi realizada segundo descrito por Brasileiro et al. 1998.

- 1. A partir de uma colônia de Agrobacterium cultivar um inóculo de 3 ml em meio YEB durante 16-18 horas a 28°C sob agitação constante.
- 2. Inocular 50 ml de meio YEB com 2 ml de cultivo e incubar a 28°C sob agitação constante até atingir 0,5-1,0 de densidade óptica a 600 nm.
- 3. Centrifugada a 5000 RPM (Centrífuga MPW-350R, rotor nº 11469) a 4ºC por 5 minutos.
- Descartar o sobrenadante e ressuspender as bactérias em 1 ml de solução de CaCl<sub>2</sub>
   mM e glicerol 10%.
- 5. Aliquotar as bactérias em tubos com de 100 μl.
- 6. Congelar com nitrogênio líquido e manter em freezer a -80°C.

#### 2.8. Transformação bacteriana

## 2.8.1. Transformação de Escherichia.coli

A transformação de *E. coli* foi realizada por choque térmico (adaptado de Hanahan *et al.*1983):

- 1. Descongelar uma alíquota de 50µl de bactérias competentes (DH10B) no gelo.
- 2. Adicionar 2 µl de reação de ligação ou recombinação.
- 3. Incubar em gelo por 30 minutos.
- 4. Colocar no banho a 42°C por 45 segundos.
- 5. Colocar imediatamente no gelo por 2 minutos.
- 6. Adicionar 500 µl de meio SOC e incubar com agitação a 37°C por 40 minutos.
- 7. Centrifugar rapidamente e descartar aproximadamente 300 µl de meio de cultura.

- 8. Distribuir o resto em uma placa de Petri contendo meio LB com o antibiótico apropriado, X-GAL (5-bromo-4-cloro-3-indolil-β-D-galactopiranosídeo, 40 μg/ml) e IPTG (Isopropil-β-D-tiogalactopiranosídeo, 200 μg/ml) segundo corresponda.
- 9. Incubar as placas a 37°C por 18 horas.

#### 2.8.2. Transformação de *Agrobacterium tumefaciens* por choque térmico.

A transformação de *A. tumefaciens* foi realizada por choque térmico (adaptado de Brasileiro *et al.* 1998):

- 1. Descongelar uma alíquota de 100 µl de bactérias competentes no gelo.
- 2. Adicionar 1 µg de DNA plasmidial e misturar delicadamente.
- 3. Incubar no gelo por 30 minutos.
- 4. Incubar no nitrogênio líquido por 2 minutos.
- 5. Incubar a 37°C por 5 minutos.
- 6. Adicionar 1 ml de meio YEB e incubar por 2 horas a 28°C sob agitação de 140 RPM.
- 7. Plaquear 100 μl da cultura em meio LB sólido contendo rifampicina (20 μg/ml), gentamicina (50 μg/ml), estreptomicina (300 μg/ml) e espectinomicina (100 μg/ml).
- 8. Incubar a 28°C durante 2-3 dias.

#### 2.9. Amplificação e clonagem das regiões promotoras.

Para a amplificação da possível região promotora dos genes *mustang* foram desenhados iniciadores, baseados nas sequências obtidas por Saccaro-Jr (2007). Os iniciadores foram ancorados de forma que amplificassem um fragmento de aproximadamente 2000 pb a montante do início da transcrição, determinado pelo começo do cDNA, ou até a presença do gene vizinho.

Para ambos haplótipos de Classe III, SCMUG266 BAC148 e SCMUG266 BAC095, a seqüência dos iniciadores foi: senso 5' CCTTGTGAGAAACCTTCAAG 3', e reverso 5' CGACTCCGCTCCACTCTC 3'. Para o gene de Classe IV, SCMUG148 BAC249, os iniciadores foram: senso 5' CACCTGTACATGATCCAAGGAGT 3', e reverso 5' GGTCTGGCTCGGGGATAAG 3'.

Utilizando como molde DNA dos respectivos clones de BAC, BAC148 e BAC095, foram amplificadas por PCR (*Polymerase Chain Reaction*) as putativas regiões promotoras. As condições da PCR foram: 200 ng de DNA, 400 nM de cada iniciador, 200 nM de dNTPs, 2,5 U de Taq DNA Polymerase (Invitrogen, Cat. No. 11615-010), 2,5 mM de MgCl<sub>2</sub>, 1X buffer. O programa de PCR utilizado foi: 94°C por 10 minutos; 30 ciclos de 94°C por 30 segundos, 50-65°C (dependendo dos iniciadores) por 30 segundos. 72°C por 2 minutos (1 minuto por Kb); 72°C por 10 minutos.

Os fragmentos amplificados foram clonados no vetor pGEM-T Easy Vector. A partir dos clones obtidos, uma segunda PCR dos fragmentos de interesse foi realizada. Desta vez, além de conter as sequências específicas, os iniciadores continham parte da seqüência do vetor doador pDNOR 221 que permitiria recombinar os fragmentos incorporando-os ao novo vetor. Para Classe III os iniciadores foram: senso 5' TACAAAAAGCAGGCTCCTTGTGAGAAACCTTCAAG 3', 5 е reverso CAAGAAAGCTGGGTCTCCGCTCCCTCTCCG 3'. Para Classe IV os iniciadores foram: TACAAAAAGCAGGCTATGATCCAAGGAGTTTTAC 3', senso 5'CAAGAAAGCTGGGTGTCTGGCTCGGGATAAG 3'. Os fragmentos amplificados foram purificados com o kit GFX (GE Healthcare, Cat. No. 28-9034-70). Após uma nova amplificação com os iniciadores AttB1 (5' GGGGACAAGTTTGTACAAAAAAGCAGGCT 3') e AttB2 (5' GGGGACCACTTTGTACAAGAAAGCTGGGT 3') os fragmentos de interesse continham em ambos os lados as regiões de recombinação completas para serem introduzidos no pDNOR 221. A reação de recombinação foi realizada com a enzima ClonaseBP® (Invitrogen, Cat. No. 11789-020). O produto dessa reação foi utilizado para a transformação de *E.coli* por choque térmico. DNA plasmidial dos clones obtidos em pDONOR 221 foi utilizado para, finalmente, incorporar os putativos promotores ao vetor de destino pHGWFS7. Esta segunda recombinação foi realizada pela enzima ClonaseLR® (Invitrogen, Cat. No. 11791-019). Assim, finalmente, foram obtidos os plasmídios: pHGWFS7prom095, pHGWFS7prom148 e pHGWFS7prom249.

2.10. Avaliação da presença dos fragmentos de interesse nos sucessivos passos de clonagem.

Após cada passo da clonagem, a presença dos fragmentos de interesse contendo os putativos promotores nos respectivos vetores foi conferida por digestão com endonucleases de restrição, por PCR e finalmente por sequenciamento.

Para o vetor pGEM-T Easy foi utilizada a enzima Notl que libera o inserto. Para o vetor de entrada pDONR 221, o DNA plasmidial foi digerido com a endonuclease Ndel. Finalmente, para o vetor pHGWFS7 foi realizada uma dupla digestão com Ndel e Sacl. Todas as digestões foram realizadas com aproximadamente 300 ng de DNA e quantidade de enzima e tampão de acordo às especificações do fabricante (New England Biolabs).

Para a verificação das clonagens por PCR foram usados os iniciadores específicos utilizados na amplificação dos putativos promotores assim como as condições de reação descritos no item 2.9.

As reações de seqüenciamento foram realizadas com 200-350 ng de DNA, 5 pmols de iniciadores (T7 para pGEM-T Easy e iniciador 5' específico para cada promotor para pDONOR 221 e pHGWFS7), 2 µl de Big Dye Terminator V3.1 Cycle Sequencing Ready kit (Applied Biosystems, Cat. No.4336919), e água deionizada para

completar o volume final de 10µl. Os tubos foram colocados no termociclador Bioclycler MJ96G (Biosystems) com a seguinte programação 96°C por 3 minutos e 25 ciclos de 96°C por 10 segundos, 48°C por 20 segundos e 60°C por 4 minutos. Posteriormente, os produtos de reação foram precipitados de acordo com o protocolo abaixo:

- Adicionar 80 μl de isopropanol 65% e agitar por 10 segundos.
- Incubar por 15 minutos no escuro a temperatura ambiente.
- Centrifugar por 25 minutos a 14000 RPM (Centrifuga Eppendorf 5415D, rotor F45-24-11).
- Descartar o sobrenadante.
- Adicionar 150 µl de etanol 70 % gelado.
- Centrifugar 15 minutos a 14000 RPM (Centrifuga Eppendorf 5415D, rotor F45-24-11).
- Descartar o sobrenadante e deixar as amostras no escuro por 1 hora a temperatura ambiente para a evaporação do álcool.
- Ressuspender as amostras em 10 μl de formamida ("Hi Di formamide Perkin Elmer Biosystems).
- Desnaturar o DNA por 4 minutos a 96°C e deixar 10 minutos no gelo.

As amostras foram lidas em um aparelho ABI PRISM 3700 DNA Analyzer (Applied Biosystems/Hitashi). O BioEdit 7.0 foi o programa usado para a análise das sequências obtidas.

#### 2.11. Análise das construções por ensaios de bombardeamento.

Para a análise através de ensaios de expressão transitória das construções contendo os putativos promotores dos genes *mustang* foi utilizada a técnica de bombardeamento gênico. Catáfilos de cebola (*Allium cepa*) foram bombardeadas com

partículas de tungstênio cobertas com DNA dos plasmídios a serem testados: pHGWFS7prom095, pHGWFS7prom148 e pHGWFS7prom249. Como controle positivo foi utilizado o vetor pCambia1391Z.35S que possui o gene *GUS* sob controle do promotor 35S e o vetor pHGWFS7 vazio como controle negativo.

Para a preparação das partículas, 48 mg de partículas de tungstênio (BIORAD M-10) foram lavadas duas vezes com 1 ml de etanol 100% e resuspendidas em 500 μL de glicerol 50%. Para cada vetor a ser utilizado no bombardeamento 22,5 μl de partículas foram misturadas sob agitação constante por 5 minutos a 2 μg de plasmídeo (22,5 μl), 37,5 μl de CaCl<sub>2</sub> 2,5 M (8°C) e 15 μl de espermidina (Sigma) 0,1 M. Após centrifugação a 12000 RPM (Centrífuga eppendorf 5415D, rotor F45-24-11) por 5 minutos, o sobrenadante foi retirado, o precipitado lavado com 500 μl de etanol 70%, e as partículas ressuspendidas em etanol 100%. Para cada bombardeamento foram usadas alíquotas de 6 μl da preparação.

As partículas foram depositadas sobre a membrana carregadora (24 mm de diâmetro e 50μm de espessura) e disparadas sob vácuo de 27 Hg, com 1200 Ψi de pressão de hélio gerada com 4 discos de ruptura (13,2 mm, 300 Ψi/membrana). O bombardeador utilizado foi fabricado e adquirido no CENARGEN (Dr. Elíbio Rech).

Foram realizados três bombardeamentos com cada vetor. O material depois de bombardeado foi deixado no escuro à temperatura ambiente por de 24h. Posteriormente, a epiderme dos catáfilos foi removida para a realização do ensaio histoquímico.

## 2.12. Ensaio histoquímico para detecção da atividade da $\beta$ -glucuronidase.

O ensaio histoquímico da  $\beta$ -glucuronidase (GUS) descrito por Jefferson *et al.* (1987) é um método qualitativo, baseado na clivagem do substrato 5-bromo-4-cloro-3-

indol- $\beta$ -D-glucoronídeo (X-gluc) pela  $\beta$ -glucuronidase. O produto dessa reação na presença de oxigênio, forma dímeros, resultando em um precipitado insolúvel de cor azul.

O material vegetal, epiderme de cebola ou células BY-2, foi colocado em solução tampão contendo X-gluc (0,5 mg/ml) durante 24h a 37°C no escuro. Após este período, retirou-se o tampão, lavou-se o material com etanol 70% e observou-se na lupa.

Tampão de reação para ensaio histoquímico (McCabe et al. 1988):

 $NaH_2PO_4H_2O$  100 mM

 $Na_2EDTA 2H_2O$  10 mM

 $K_4Fe(CN)_6 3H2O$  0,5 mM

Triton X-100 (Sigma) 0,1 %

Água ultrafiltrada completar 100 ml

Ajustar o pH em 7,0, com NaOH, esterilizar a solução por filtração e armazenar a -20°C em alíquotas de 10 ml.

## 2.13. Transformação estável de células de fumo BY-2.

A transformação de células BY-2 foi realizada segundo descrito por An (1985).

- 1. Quatro mililitros de células BY-2 de 3 ou 4 dias foram dispostos em placas de Petri as quais foram acrescentados 100 µl de cultura de *Agrobacterium* contendo cada um dos plasmídios de interesse (pHGWFS7prom095, pHGWFS7prom148 e pHGWFS7prom249 e pHGWFS7 vazio).
- 2. As placas foram vedadas com papel filme e incubadas em local escuro durante dois dias.

- 3. As células foram transferidas para uma placa contendo meio BY-2 sólido com higromicina (30 μg/ml) e carbenicilina (100 μg/ml) e mantidas em estufa a 28°C durante duas semanas.
- As colônias de células desenvolvidas foram transferidas para meio seletivo fresco.
   Cada colônia foi considerada uma linhagem independente.
- 5. As células foram repicadas a cada duas semanas.

A presença do transgene foi confirmada por PCR utilizando os iniciadores específicos dos putativos promotores assim como as condições de reação descritos no item 2.9. Também a expressão do gene repórter *GUS* foi verificada através de ensaios histoquímicos como descrito no item 2.12.

## 2.14. Análise microscópica para visualização da Green Fluorescent Protein (GFP).

Para a detecção de *GFP* foi utilizado o microscópio confocal LSM (Laser Scan Microscopy) 410 Zeiss, que possui um filtro de excitação de 488nm e filtro de emissão LP515. Para a visualização das imagens foi usado o programa Axionvision (Zeiss). Foram preparadas lâminas (76 x 26 mm, Knittel Glaser) com 100 µl de células contendo cada uma das linhagens de células transformadas assim como um controle sem transformar.

#### 2.15. Tratamentos hormonais das linhagens de células transgênicas.

As linhagens transformadas de células de fumo foram tratadas com ácido indolbutírico (AIB), isopentenil adenina (IP) e acido abscisíco (ABA). Para este ensaio células de 6 a 7 dias foram tratadas com doses 10 µM de AIB,10 µM de IP e 20 µM

ABA durante 24h. O esquema de coletas está apresentado na figura 7. As células amostradas (2 ml) foram precipitadas, congeladas em nitrogênio líquido e armazenada em freezer a -80°C até o momento da extração de RNA. O experimento foi realizado em triplicata.



Figura 7. Sequência de coletas das amostras de células BY-2 tratadas com AIB,IP e ABA.

#### 2.16. Extração de RNA e síntese de cDNA.

Para a extração de RNA de células BY-2 e tecido de cana-de-açúcar foi utilizado o método de Trizol (Invitrogen, Cat. No.15596-026). As amostras de RNA foram ressuspendidas em 30 µL de água DEPC 0,1%. A concentração e a integridade do foram avaliadas por espectrofotômetro Nanodrop RNA ND-100 (Nanodrop Technologies) e por eletroforese em gel de agarose. Para a remoção de DNA das amostras foi usada a DNAse I (Invitrogen, Cat. No.18068-015). Para analisar a efetividade da DNAse nas amostras de RNA de células de fumo foi realizada uma PCR com iniciadores que se anelam em éxons diferentes do gene de actina de batata (senso 5' TGGCATCATACCTTTTACAA 3' e reverso 5' TCCGGGCATCTGAACCCTCT 3'). O fragmento amplificado a partir de DNA genômico é de 812 pb e a partir de cDNA de 521 pb. Já para as amostras de RNA de cana-de-açúcar foi realizada uma PCR com os iniciadores que amplificam os fragmentos dos putativos promotores. A síntese de cDNA foi feita com o kit SuperScript® III Reverse Transcriptase (Invitrogen, Cat. No. 18080-044).

## 2.17. PCR em tempo real (Real Time PCR).

Para os experimentos com células BY-2, foram desenhados iniciadores sobre a sequência do gene repórter *GUS* (senso 5' GGACAAGGCACTAGCGGGACTT 3', reverso 5' CCGACGCGAAGCGGGTAGATAT 3', amplicon de 131 pb,) e dois genes endógenos: proteína ribossomal L25 (senso 5' CCCCTCACCACAGAGTCTGC 3', reverso 5' AAGGGTGTTGTTCCTCAATCTT3', amplicon de 51 pb) e o fator de elongação 1α (senso 5' TGAGATGCACCACGAAGCTC 3', reverso 5' CCAACATTGTCACCAGGAAGTG 3', amplicon de 51 pb,) como controle constitutivo.

Para avaliar o padrão de expressão espacial e temporal dos genes endógenos em cana-de-açúcar foram desenhados um par de iniciadores para Classe III (senso 5' CCGGACAGGACCCACTGG 3', reverso 5' CGACCACACTGAACGATCCG 3', amplicon de 170 pb) e outro para Classe IV (senso 5' TTAGCTGTTCTTCAAATAGGTGGAG 3', reverso 5' GAATGCGAACTGGTCAGGC 3', amplicon de 171 pb). Como gene de expressão constitutiva foi utilizado o que codifica para ubiquitina (senso 5' CGTCCGCAGTCCCCAAT 3', reverso 5' TGAGAGGATCGCGAGGATTC 3', amplicon de 100 pb).

Os cinco pares de iniciadores foram avaliados utilizando DNA genômico para verificar a presença de um único amplicon e estabelecer a temperatura de anelamento.

Para otimizar a reação de PCR em Tempo Real, foram testadas diferentes concentrações de iniciadores (100, 200 e 300 nM) assim como diferentes diluições de cDNA (1/2, 1/3 e 1/4) através de RT-PCR. Estabelecendo-se como concentração ótima de iniciadores para ubiquitina, Classe III e IV 200 Nm, 500 nM para *GUS*, 78nM para proteína ribossomal L25 e 313nM para o fator de elongação 1α. A diluição de cDNA usada para todos os experimentos foi 1/3.

Para cada gene foi construída uma curva de calibração para calcular a quantidade de cDNA em cada amostra. A mesma foi realizada utilizando uma mistura de todas as amostras de cDNA de cada experimento e as seguintes diluições seriadas: 1/3, 1/9, 1/27 e 1/81. As amostras foram amplificadas nas mesmas condições ideais acima citadas. A partir dos valores de Ct (*cycle threshold*) obtidos e as respectivas diluições de cDNA foi calculada a equação de regressão e o R<sup>2</sup> utilizando o programa Excel (Microsoft).

As condições da PCR em Tempo Real foram: 1 μL de uma diluição 1/3 de cDNA (~15ng) , 200 a 300 nM de iniciadores (dependendo do par de iniciadores) e 12,5 μL de SYBR Green® Master Mix (Cat. No. 4309155, Applied Biosystems), completando com água deionizada para um volume final de 25 μL. O programa utilizado foi: 10 minutos a 95°C, 40 ciclos de 20 segundos de 95°C, 20 segundos a temperatura de anelamento (específica para cada par de iniciadores) e 60 segundos a 72°C. O equipamento utilizado foi o 7500 PCR Real Time da Applied Biosystems.

Os dados obtidos em forma de valores de Ct (*cycle threshold*) foram convertidos a quantidade de cDNA através da curva de calibração e analisados estatisticamente com o programa InfoStat (www.infostat.com Grupo InfoStat, FCA, Universidad Nacional de Córdoba, Argentina) utilizando o teste não paramétrico de Kruskal Wallis.

.

#### IV. Resultados

Capítulo 1. Análise comparativa entre os genes *mustang* de cana-de-açúcar, Oryza sativa, Sorghum bicolor e Zea mays.

### 1.1 Análise filogenética dos genes mustang.

Saccaro-Jr et al. (2007) realizaram uma análise filogenética da família de genes mustang utilizando as seqüências de arabidopsis, arroz e cDNAs de cana-de-açúcar previamente identificados por Rossi et al. (2004). Neste trabalho, com o objetivo de aprofundar o conhecimento sobre os genes mustang em gramíneas, foi realizada uma nova análise onde foram incluídos os genes de milho e sorgo cujos genomas foram recentemente disponibilizados. Ainda, nesta análise foram utilizadas as seqüências codificantes completas dos genes, e não apenas um domínio conservado como nas análises publicadas anteriormente pelo grupo (Saccaro-Jr et al. 2007, Rossi et al. 2004), a fim de avaliar se os clados previamente descritos eram mantidos. Assim, de cana-de-açúcar foram utilizados apenas os três genes mustang totalmente seqüenciados caracterizados por Saccaro-Jr (2007): SCMUG266 BAC148 (haplótipo de Classe III proveniente do genitor Saccharum spontaneum), SCMUG266 BAC095 (haplótipo de Classe III proveniente do genitor Saccharum officinarum) e o SCMUG148 BAC249 (haplótipo de Classe IV proveniente do genitor Saccharum officinarum).

Utilizando as seqüências de cana de açúcar e de arroz foi realizada uma busca in silico nos bancos dos genomas de milho (Maize GDB Database) e sorgo (Phytozome). A Tabela 1 apresenta o número de cópias por classe identificadas nos genomas monoplóides de milho e sorgo, comparando com os números previamente descritos para cana e arroz descritos por Saccaro-Jr et al. (2007). De forma geral, os números revelam a presença de 1 a 2 cópias de Classe III, e 3 a 4 cópias para Classe

IV nas quatro espécies comparadas. Estes números coincidem com os clados identificados na análise de Saccaro-Jr *et al.* (2007). O único número que se destaca é o referente às 6 cópias de Classe III identificadas em milho.

Tabela 3. Número de cópias de genes mustang em arroz, milho, sorgo e cana-de-açúcar.

	Número de cópias po		
Espécies / Tamanho do genoma monoploide	mustang Classe III	mustang Classe IV	Referência
Arroz (Oryza sativa) (430Mb)	2	3	In silico <sup>1</sup>
Milho <i>(Zea mays)</i> (3283Mb)	6	3	In silico <sup>2</sup>
Sorgo (Sorghum bicolor) (760Mb)	2	4	In silico <sup>2</sup>
Cana-de-açúcar (Saccharum sp.) (~1000Mb)	~1	3	Hibridização molecular 1

<sup>1</sup> Saccaro-Jr et al. (2007).

A topologia da árvore obtida é semelhante à descrita anteriormente por Saccaro-Jr et al. (2007) (Figura 8). O filograma apresenta as Classe III e Classe IV como grupos irmãos, sugerindo apenas um evento de domesticação antes da divergência entre Mono e Eudicotiledôneas. O gene ancestral domesticado sofreu um evento de duplicação dando origem a ambas as classes e a história evolutiva dentro de cada uma é diferente. A Classe III apresenta dois clados que divergiram antes da separação de Mono e Eudicotiledôneas. Diferentemente, a Classe IV de genes *mustang* em gramíneas apresenta três clados que divergiram mais tardiamente já que não contém os ortólogos correspondentes de arabidopsis. O estudo permitiu acrescentar aos grupos de genes ortólogos os representantes de sorgo e milho. É interessante ressaltar que milho apresenta 3 genes dentro de cada clado de Classe III refletindo, provavelmente, diferentes haplótipos presentes no ancestral poliplóide desta espécie. Por outro lado, um dos clados de Classe IV não apresenta o correspondente ortólogo de milho. Foram identificados os genes ortólogos de sorgo para todos os clados. O número de cópias presentes em sorgo coincide com arroz, com exceção do clado1 de Classe IV para o qual em sorgo foram identificados dois genes. Os genes de cana-deaçúcar apresentaram o posicionamento esperados: as cópias genômicas SCMUG266

<sup>2</sup> Presente trabalho.

BAC148 e SCMUG266 BAC095 situaram-se no clado do cDNA SCMUG266 utilizado como sonda na identificação dos clones genômicos. Da mesma forma, o gene SCMUG148 BAC249 se localizou no mesmo clado que o cDNA SCMUG148 com o qual foi identificado (Saccaro-Jr *et al.* 2007).

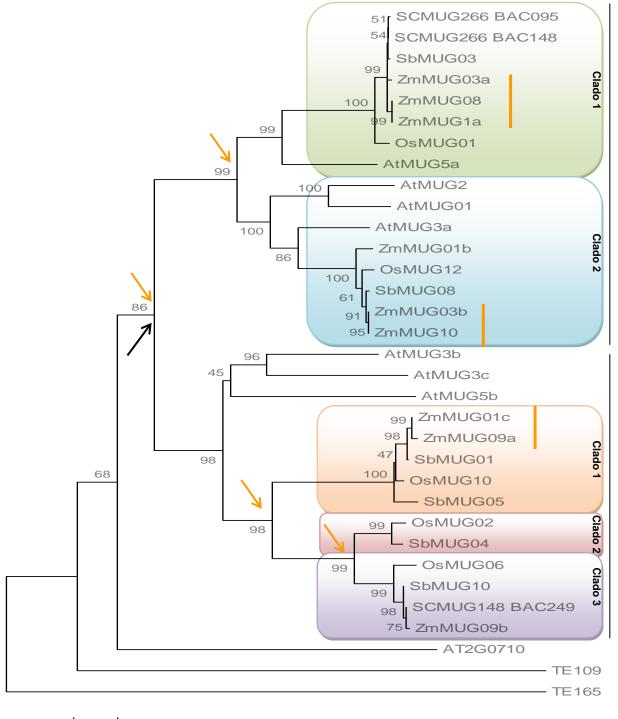


Figura 8. Árvore filogenética dos genes mustang de gramíneas.

(A) AtMUGXX, OsMUGXX, SCMUGXXX, SbMUGXX e ZmMUGXX são genes mustang de arabidopsis, arroz, cana-de-açúcar, sorgo e milho, respectivamente. Todas as sequências de arroz e arabidopsis usadas foram descridas por Saccaro-Jr et al. (2007). As sequências de sorgo e milho disponíveis no site: http://www.phytozome.net/sorghum e http://www.maizegdb.org/ respectivamente. Retângulos indicam grupos ortólogos da Classe III e Classe IV. A árvore foi enraizada nas següências mudrA não MUSTANG TE165, TE109 e At2G7100 também utilizadas por Saccaro-Jr et al. (2007). As retas laranjas indicam os genes de mustang de milho provavelmente originados a partir da poliploidização.

# 1.2. Estudo comparativo entre as seqüências codificantes e regulatórias dos genes *mustang* de cana-de-açúcar, sorgo, milho e arroz.

Identificados os grupos ortólogos de genes *mustang* em gramíneas foi realizada uma análise comparativa das regiões regulatórias e codificantes dentro do clado 1 de Classe III e clado 3 de Classe IV (Figura 1), para os quais foram identificadas as cópias genômicas das quatro espécies: cana-de-açúcar, sorgo, milho e arroz. Como região regulatória, promotora da transcrição, foi considerado o fragmento de 2000 pb a montante do início de transcrição de acordo com o cDNA de cada gene, ou até o começo da região codificante do gene adjacente.

As sequências codificantes de cana, sorgo e milho apresentam uma alta porcentagem de identidade acima de 90%. Já a identidade das seqüências destas espécies com arroz diminui para ~70%. Como esperado, as regiões regulatórias são menos conservadas e a identidade diminui para 40-50 % entre as diferentes espécies. Excepcionalmente, as regiões regulatórias do *mustang* da Classe IV de milho e sorgo apresentaram identidade de 59% (Tabela 4).

Tabela 4. Porcentagem de identidade entre os genes ortólogos de *mustang* em cana-de-açúcar, (SCMUGXX), sorgo (SbMUG), arroz (OsMUGXX) e milho (ZmMUGXX).

Classe III	SCMUG266. BAC095 promotor(1685pb)/ sequência codificante (1785bp)	SCMUG266. BAC148 promotor(1921pb)/ sequência codificante (1785bp)	SbMUG03 promotor (2000pb)/ sequência codificante (1785bp)	ZmMUG08 promotor (2000pb)/ sequência codificante (1725pb)	OsMUG01 promotor (2000pb)/ sequência codificante (2070bp)
SCMUG266 BAC095	100/100	81/98	44/97	42/91	45/72
SCMUG266 BAC148	-	100/100	45/97	44/91	46/73
SbMUG03	-	-	100/100	43/90	42/72
ZmMUG08	-	-	-	100/100	42/69
OsMUG01	-	-	-	-	100/100

Classe IV	SCMUG148. BAC249 promotor(2000pb)/ sequência codificante (2237bp)	SbMUG10 promotor(2000pb)/ sequência codificante (2163bp)	ZmMUG09b promotor(2000pb)/ sequência codificante (2233pb)	OsMUG06 promotor(1444pb)/ sequência codificante (2169bp)
SCMUG148. BAC249	100/100	45/95	46/93	40/81
SbMUG10	-	100/100	59/96	43/84
ZmMUG09b	-	-	100/100	42/84
OsMUG06	-	-	-	100/100

Um dos objetivos deste trabalho é a caracterização da atividade transcricional dos mustang de cana-de-açúcar assim, foi realizada uma busca para a identificação dos elementos em cis de união a fatores de transcrição que pudessem aportar evidências do perfil de expressão dos três genes sob estudo. Para definir os motivos em cis foi usado o programa PLACE que é um banco de dados específico de plantas. Foram considerados apenas aqueles previamente caracterizados funcionalmente. Os mesmos foram agrupados nas seguintes categorias: início de transcrição, resposta à luz, resposta a hormônios, resposta a frio e desidratação, tecido-específico e outros. A figura 9 apresenta um gráfico com a diversidade de elementos encontrados para cada categoria. A análise permitiu identificar a presença dos elementos essenciais presentes nos promotores eucariotas como o CAATbox, iniciador (Inr) e TATAbox, este último elemento não foi encontrado na região promotora do gene *mustang* de SCMUG266.BAC095 da Classe III. A maioria dos motivos é comum aos três promotores analisados. Em relação aos elementos diferenciais, é interessante notar que existem motivos específicos de classe ou de cópia. Os motivos CATATG (responsivo a auxina), CPBCSPOR (responsivo a citocinina), DRE (responsivo a desidratação e frio) e IBOX (responsivo a luz) mostraramse presentes em pelo menos um dos promotores de Classe III e ausentes na Classe IV. Elementos presentes somente na Classe IV não foram identificados, todos os elementos são compartilhados com uma das cópias da Classe III. Dentro da Classe III alguns elementos responsivos a ácido giberélico, ciclo celular, frio e elementos tecido-específicos (raiz) foram encontrados unicamente em uma das cópias (Tabela 5).

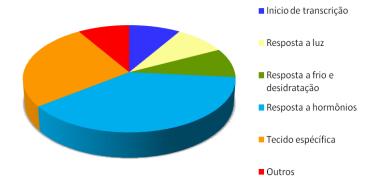


Figura 9 Categorização funcional dos elementos em *cis* das regiões promotoras dos genes *mustang* de cana-de-açúcar.

Todos os motivos de ligação a fatores de transcrição identificados in silico nos promotores dos genes mustang de cana-de-açúcar (Tabela 3) foram agrupados de acordo a sua função.

Tabela 5. Elementos em cis presentes nas regiões promotoras dos genes mustang de cana-deaçúcar.

Elemento em <i>cis</i>	Potencial função	Sítio	095	148	249
CAAT box	Sítio comum de início de transcrição	CAAT	Χ	Χ	Χ
TATA box	Sítio comum de início de transcrição	TTATTT		Χ	X
Inr	Sítio comum de início de transcrição	YTCANTYY	Χ	Χ	Χ
CURE	Resposta a cobre e O2	GTAC	Χ	Χ	Χ
BIHD1	Resposta a defesa a doenças	TGTCA		Χ	X
E2F	Encontrada em promotores de genes envolvidos na progressão do ciclo celular e metabolismo de DNA	AAAG	Χ		Х
-10PEHVPSBD	Resposta a luz	TATTCT	Χ	Χ	Χ
CGCG box	Resposta a luz	VCGCGB	X	Χ	X
GATA	Resposta a luz	GATA	X	Χ	X
GT-1 CONSENSUS	Indução por sal e patógenos, resposta a luz	GRWAAW	Χ	Χ	Χ
IBOX	Resposta a luz	GATAA	Χ	Χ	
DRE	Responsivo a desidratação e frio	RCCGAC		Χ	
LTRE	Responsivo a frio	CCGAC	Χ		Χ
MYC	Responsivo a desidratação	CANNTG	Χ	Χ	Χ
ABRE	Responsivo a desidratação, escuro e ABA	ACGTG	Χ	Χ	Χ
ARF	Responsivo a auxina	TGTCTC	Χ	Χ	Χ
ARR1	Responsivo a citocinina	NGATT	X	Χ	X
ASF1	Responsivo a auxina e ácido salicílico	VCGCGB	X	Χ	X
CATATG	Responsivo a auxina	CATATG	X	Χ	
CGCG box	Etileno e ABA	VCGCGB	Χ	Χ	Χ
CPBCSPOR	Responsivo a citocinina	TATTAG	X	Χ	
DOFCORE	Responsivo a hôrmonios (GA) e sinais de estresse.	AAAG	X	Χ	X
DPBF	Responsivo a ABA	ACACNNG	X	Χ	X
GARE	Responsivo a GA	TAACAAR		Χ	X
MYBCORE	Estresse hídrico, ABA e sinal de açúcar	CNGTTR	X	Χ	X
SORLIAP	Responsivo a ABA.	GCCAC	X	Χ	X
W-box	Ácido salicílico, ferimento e sinal de açúcar	TGACY	Χ	Χ	Χ
WRKY	Responsivo para ABA e reprimido por GA	TGAC	Χ	Χ	Χ
CACTF	Tecido especificidade (mesófilo)	YACT	Χ	Χ	Χ
DOFCORE	Germinação, tecido especificidade (endosperma, células-guarda e folhas)	AAAG	Χ	Χ	Χ
EBOX	Tecido especificidade (pólen)	CANNTG	Χ	Χ	Χ
GTGANT	Tecido especificidade (pólen)	GTGA	Χ	Χ	Χ
POLLEN1LELAT52	Tecido especificidade (pólen)	AGAAA	Χ	Χ	Χ
RAVIAT	Tecido especificidade (folhas e raízes)	CAACA	Χ	Χ	Χ
RHEs	Tecido especificidade (diferenciação de raízes)	KCACGW		Χ	Χ
ROOT motif	Encontrado em promotores de raiz	ATATT	Χ	Χ	Χ
SORLIAP	Tecido especificidade (embrião)	GCCAC	Χ	Χ	Χ

As letras N: A, C, G ou T; V: A, C ou G; W: A ou T; Y: C ou T e R: A ou G

095: região promotora do gene *mustang* SCMUG266 BAC095, 148: região promotora do gene *mustang* SCMUG266 BAC148
e, 249: região promotora do gene *mustang* SCMUG148 BAC249.
As cores correspondem as categorias funcionais de acordo à figura 2.

Capítulo 2. Caracterização da atividade transcricional de três genes *mustang* de cana-de-açúcar.

2.1 Obtenção das construções de fusão contendo os promotores a montante dos genes repórteres GFP (*green fluorecense protein*) e GUS (*β-glucoronidase*).

A partir do DNA dos BACs que contém as cópias genômicas de cada um dos três genes *mustang* de cana-de-açúcar, e utilizando pares de iniciadores específicos para cada classe foram amplificadas as possíveis regiões promotoras. O tamanho esperado de cada fragmento é apresentado na Tabela 6.

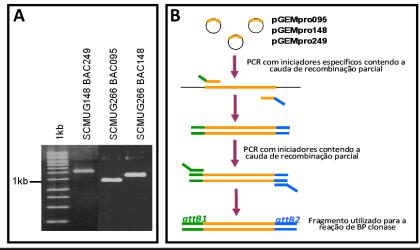
Tabela 6. Tamanho da região promotora amplificada de BACs

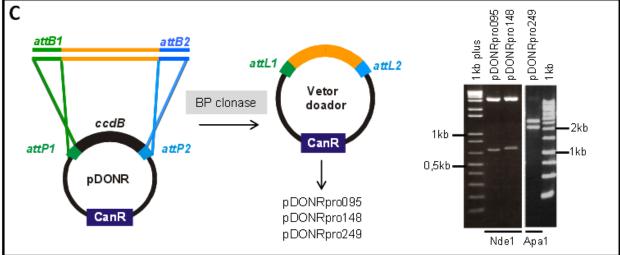
BAC	Tamanho	Classe
SCMUG266 BAC095	1612pb	III
SCMUG266 BAC148	1849pb	Ш
SCMUG148 BAC249	1983pb	IV

Os produtos de amplificação foram analisados por eletroforese em gel de agarose confirmando-se, em cada caso, a presença do amplicon esperado (Figura 10A). Os fragmentos purificados foram clonados no vetor pGEM T-Easy obtendo-se os plasmídios recombinantes pGEMpro095, pGEMpro148 e pGEMpro0249.

A partir deste ponto, e até a obtenção do vetor binário final para a transformação genética, foi utilizado o sistema GATEWAY de clonagem pelo qual, os fragmentos são introduzidos nos vetores através de recombinação genética (Karimi *et al.* 2002). Para isso, aos fragmentos específicos foram acrescentadas as seqüências de recombinação AttB1 e AttB2 através de duas PCRs sucessivas (Figura 10B). Esses fragmentos foram então inseridos no vetor doador pDONR <sup>TM</sup>221 obtendo-se os vetores pDONRpro095, pDONRpro148 e pDONR249 (Figura 10C). Finalmente, os promotores foram incorporados ao vetor binário pHGWFS7 gerando-se os plasmídios pHGWFS7pro095,

pHGWFS7pro148 e pHGWFS7pro249. Os mesmos foram, então, utilizados para a transformação genética vegetal.





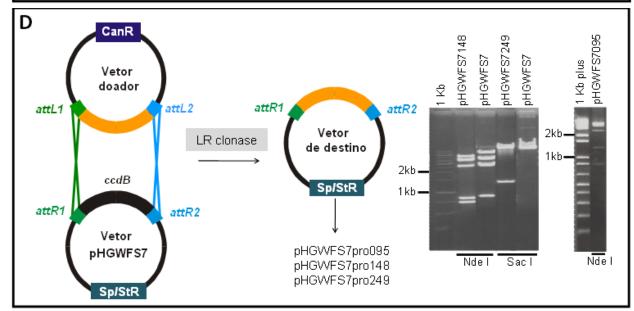


Figura 10. Clonagem dos promotores no vetor binário pHGWFS7.

(A) Gel de agarose mostrando os três promotores amplificados a partir dos clones de BACs respectivos. (B) Preparação dos fragmentos para a recombinação com o vetor pDONOR TM221 através de duas PCRs utilizando iniciadores contendo sequências da região de interesse (laranja) e específicas da região de recombinação (attB1:verde e attB:azul). (C) Recombinação dos fragmentos com o vetor pDONOR. (D) Incorporação dos fragmentos no vetor binário pHGWFS7. CanR: resistência a canamicina. Sp/StR: resistência a espectinomicina e estreptomicina. ccdB:gene letal. Tamanhos dos fragmentos esperados para digestão do pDONRpro095: 692pb e 3730pb; pDONRpro148: 711pb e 4014pb; pDONRpro249: 2161pb e 2639pb; pHGWFS7pro095: 809pb, 1275pb, 2486pb, 3525pb5 4906pb; pHGWFS7pro148: 797pb, 711pb, 809pb, 2486pb, 3525pb e 4910pb e pHGWFS7pro249: 1323pb e 12575pb.

## 2.2 Avaliação das construções através de ensaios de expressão transitória por bombardeamento.

Obtidas as construções contendo as possíveis regiões promotoras a montante dos genes repórteres GFP (*Green Fluorescent Protein*) e *GUS* (β-glucuronidase), foi realizado um ensaio de expressão transitória por bombardeamento, para verificar a atividade promotora dos fragmentos clonados.

Catáfilos de cebola foram bombardeados com partículas de tungstênio cobertas DNA plasmidial dos vetores pHGWFS7prom095, pHGWFS7prom148, de pHGWFS7prom249, pCambia1391Z.35S (controle positivo) e pHGWFS7 vazio (controle negativo) utilizando canhão gênico. um Foram realizados bombardeamentos independentes com cada vetor.

A figura 11 apresenta imagens representativas dos tecidos bombardeados após a revelação da atividade de *GUS*. O experimento demonstrou que os três fragmentos clonados possuem atividade promotora. É interessante ressaltar que ambas as cópias do gene *mustang* de Classe III (pHGWFS7prom095 proveniente do genitor *Saccharum officinarum* e pHGWFS7prom148 proveniente do genitor *Saccharum spontaneum*) apresentaram promotores transcricionalmente ativos.

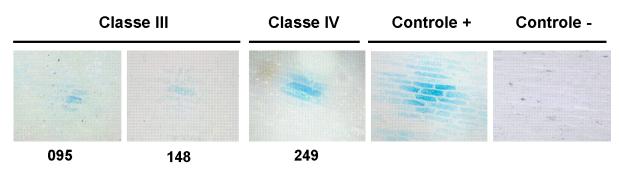


Figura 11. Atividade do gene repórter GUS em ensaio de expressão transitória por bombardeamento.

Catáfilos de cebola foram bombardeados com partículas cobertas por DNA plasmidial das construções contendo as possíveis regiões promotoras dos genes *mustang:* pHGWFS7prom095 (095), pHGWFS7prom148 (148) e pHGWFS7prom249 (249). Os vetores pCambia1391Z.35S (controle +) e pHGWFS7 vazio (controle -) foram utilizados como controle positivo e negativo, respectivamente.

## 2.3 Obtenção de linhagens de *Nicotiana tabacum* (BY-2) transgênicas para as construções de fusão

Com o intuito de avaliar a atividade dos promotores em ensaio de expressão estável, foram obtidas linhagens de células de fumo BY-2 transgênicas. Primeiramente, os vetores binários pHGWFS7prom095, pHGWFS7prom148, pHGWFS7prom249 e pHGWFS7 vazio foram introduzidos na cepa de *Agrobacterium tumefaciens* GV3101 por choque térmico. Os clones foram selecionados na presença dos antibióticos específicos e a confirmação da incorporação dos respectivos vetores verificada por PCR.

Células BY-2 foram inoculadas e co-cultivadas com as diferentes linhagens de *Agrobacterium* durante dois dias, após este período as células foram transferidas para um meio seletivo permanecendo até o crescimento de calos. Depois de dois meses de seleção, foram obtidos 5 clones contendo a região promotora do gene *mustang* SCMUG266 BAC095, 3 clones contendo a região promotora de SCMUG266 BAC148, 2 clones contendo a região promotora do gene *mustang* SCMUG148 BAC249 e 1 clone contendo o vetor pHGWFS7 vazio. A presença do transgene nas linhagens selecionadas foi confirmada por PCR utilizando iniciadores específicos da região promotora e do gene repórter *GUS* e correlacionou com a atividade das proteínas codificadas por ambos os genes repórteres (Tabela 7). Para isto, as células foram transferidas para meio líquido até a obtenção de uma cultura de células em suspensão homogênea. A figura 12 apresenta as imagens de três linhagens visualizadas no microscópio confocal onde é possível identificar a fluorescência da proteína GFP. A figura 13 apresenta o resultado da revelação da atividade da proteína GUS para outras três linhagens. A maior intensidade de GFP, assim como da atividade de GUS, para a

construção contendo o promotor do gene de Classe IV, SCMUG148 BAC249, sugere que, entre os três promotores estudados, este apresenta maior atividade transcricional.

Tabela 7. Clones transgênicos de células BY-2.

Linhagens	Clones	PCR 1	GUS <sup>2</sup>	GFP <sup>3</sup>
	BY- 2.III.095.1	+	+	+
	BY- 2.III.095.2	+	+	+
Classe III SCMUG266 BAC095	BY- 2.III.095.3	+	+	+
	BY- 2.III.095.4	+	+	+
	BY- 2.III.095.5	+	+	+
	BY- 2.III.148.1	+	+	+
Classe III SCMUG266 BAC148	BY- 2.III.148.2	+	+	+
	BY- 2.III.148.3	+	+	+
Classe IV	BY- 2.IV.249.1	+	+	+
SCMUG148 BAC249	BY- 2.IV.249.2	+	+	+
pHGWFS7 vazio	vetor Ø	+	-	-

<sup>1</sup> identificação da presença do transgene nas limnhagens de células BY-2 transformadas.

<sup>3</sup> detecção da atividade da proteína GFP através da visualização no microscópio confocal

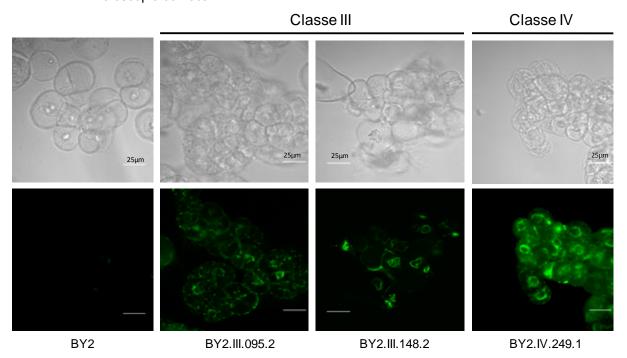
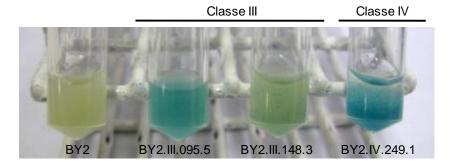


Figura 12. Microscopia confocal das linhagens de células BY-2.

BY-2: células não transformadas. BY-2. III.095.2, BY-2.III.148.2 e BY-2.IV.249.1: indicam as diferentes linhagens transgênicas. Painéis superiores: imagens de luz transmitida. Painéis inferiores:

diferentes linhagens transgênicas. Painéis superiores: imagens de luz transmitida. Painéis inferiores: imagens de fluorescência. Barras brancas: 25µm.

<sup>2</sup> detecção da atividade da β-glucuronidase pelo ensaio histoquímico.



**Figura** 13. **Ensaio** histoquímico para detecção da atividade de **β-glucoronidase** células BY-2. BY-2: células não transformadas. BY-2.III.095.5, 2.III.148.3 e BY-2.IV.249.1: indicam diferentes as linhagens transgênicas.

#### 2.4 Avaliação da indução hormonal dos promotores nas linhagens transgênicas.

Tendo demonstrado que as putativas regiões regulatórias dos genes *mustang* são de fato promotores transcricionais, somado à identificação de motivos de união a fatores de transcrição responsivos a diferentes fitohormônios, foi investigado o efeito da auxina AIB, da citocinina IP e do ácido abscísico (ABA) sobre a atividade dos promotores. Células de 6-7 dias foram tratadas por 24 h com cada regulador de crescimento e a quantidade de mRNA do gene repórter *GUS* foi avaliado por PCR em tempo real. Foram realizados três experimentos independentes com cada hormônio. Para a análise de expressão foram utilizados iniciadores específicos para o gene *GUS* e para os genes de referência L25 e *elongation factor* 1a (EF1a). A figura 7 os níveis de expressão relativa de *GUS* após tratamento hormonal normalizados com os valores da cultura sem tratar.

Interessantemente, os promotores dos genes SCMUG266 BAC148 e SCMUG148 BAC249, de Classe III e IV respectivamente, apresentaram um comportamento similar com uma redução do 30-40 % no nível de expressão em resposta ao tratamento com auxina. Já o promotor do gene SCMUG266 BAC095 mostrou uma discreta modulação frente ao tratamento com citocinina e ácido abscísico reduzindo a atividade em aproximadamente 10 %.

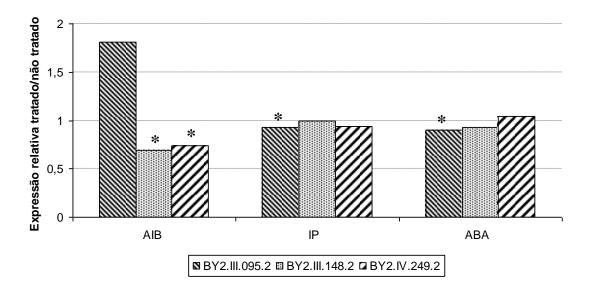


Figura 14. Expressão relativa do gene *GUS* nas linhagens transgênicas de células BY-2 após tratamento com auxina, citocinina e ácido abscísico. Linhagens transgênicas de células BY-2 contendo o gene repórter *GUS* sob controle dos três promotores dos genes *mustang* foram tratadas por 24 h com ácido indolbutírico (AIB), isopentenil adenina (IP) e acido abscisíco (ABA). O experimento foi repetido três vezes. Os asteriscos indicam diferenças estatisticamente significativas ao controle sem tratar segundo o teste não paramétrico de Krustal Wallis (p<0,05).

# 2.5 Avaliação do padrão de expressão espacial dos genes endógenos em canade-açúcar.

Com a intenção de investigar a expressão dos genes *mustang* das Classes III e IV nos seus correspondentes contextos genômicos, foram desenhados iniciadores específicos sobre as regiões codificantes para cada uma das classes, e foi avaliado o perfil transcricional por PCR em tempo real na variedade R-570 de cana-de-açúcar. Os níveis de mRNA foram quantificados em folhas, meristemas caulinares e raízes coletados ao longo do desenvolvimento da planta. Como controle constitutivo de expressão foi quantificado o mRNA do gene da ubiquitina para a normalização dos dados. As figuras 9 e 10 apresentam os níveis relativos de expressão para as Classes III e IV, respectivamente. É importante ressaltar que os iniciadores não discriminaram entre as diferentes cópias gênicas de cada classe. De maneira geral as diferenças nos

níveis de expressão apresentaram pouca significância estatística, o qual foi mais acentuado para Classe IV.

Os genes de Classe III apresentaram o maior nível de expressão no meristema em todos os estágios de desenvolvimento estudados. Em folhas e meristemas foi observado um pico de expressão nas plantas de 30 dias. Já na raiz o maior nível de expressão foi observado em plantas adultas (Figura 15).

Para os genes de Classe IV, o perfil tecidual de expressão foi semelhante, sendo o meristema o tecido com maiores níveis de expressão relativa. Ao longo do desenvolvimento, as folhas e meristemas mais jovens apresentaram as maiores quantidades de mRNA as quais foram caindo ao longo do tempo (Figura 16).

A Classe IV apresentou um valor médio de expressão maior que Classe III. Esta diferença é claramente evidenciada em plantas de 15 dias, onde todos os tecidos apresentaram o dobro da expressão relativa de genes de Classe IV comparada com os genes Classe III.

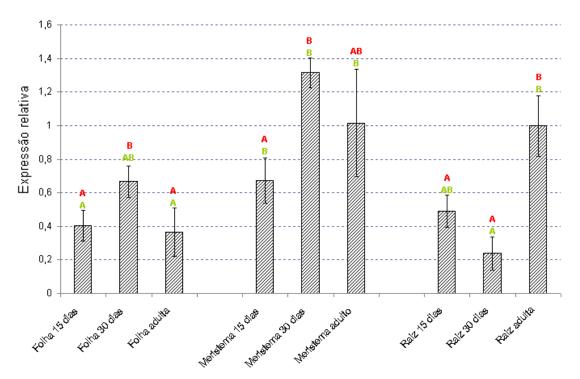


Figura 15. Expressão relativa dos genes *mustang* da Classe III ao longo do desenvolvimento de plantas de cana-de-açúcar.

O experimento foi repetido três vezes. Letras diferentes indicam diferenças estatisticamente significativas segundo o teste não paramétrico de Krustal Wallis (p<0,05). Letras vermelhas indicam comparação entre o mesmo tecido de diferentes idades. Letras verdes indicam comparação entre diferentes tecidos da mesma idade.

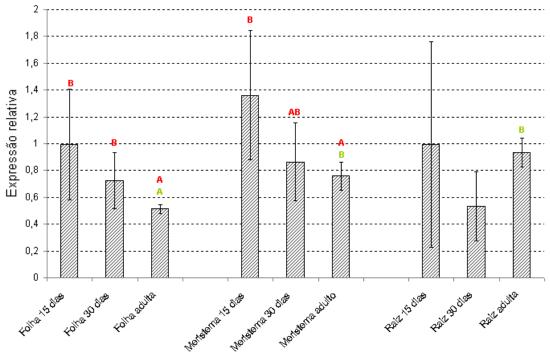


Figura 16. Expressão relativa dos genes *mustang* da Classe IV ao longo do desenvolvimento de plantas de cana-de-açúcar.

O experimento foi repetido três vezes. Letras diferentes indicam diferenças estatisticamente significativas segundo o teste não paramétrico de Krustal Wallis (p<0,05). Letras vermelhas indicam comparação entre o mesmo tecido de diferentes idades. Letras verdes indicam comparação entre diferentes tecidos da mesma idade. A ausência de letras representam resultados sem diferença estatística significativa.

#### V. Discussão

## 1. O genoma das gramíneas e os genes mustang

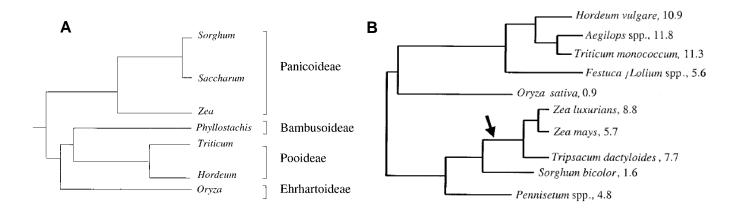
Em 2004, Rossi *et al.*, analisando seqüências de arabidopsis, arroz e cana-de-açúcar, descreveram a existência de quatro classes de seqüências homólogas à transposase *mudrA* em plantas que foram nomeadas como Classe I, II, III e IV. Posteriormente, Saccaro-Jr *et al.* (2007), demonstraram que as classes III e IV correspondem a transposases domesticadas do tipo *mustang* (Cowan *et al.* 2005).

As gramíneas apresentam genomas com um diverso espectro de tamanho e complexidade (Gaut et al. 2000) (Figura 17). Por um lado se encontra o arroz (Oryza sativa) diplóide de 12 cromossomos (2n=24) e com genoma monoploide de 430 Mb. No outro extremo, está o gênero Saccharum altamente poliplóide, com mais de 100 cromossomos e 1.000 Mb. No meio se encontra o milho, Zea mays, diplóide (2n=20) e com 3.283 Mb. Esta ampla diversidade, somada à recente disponibilidade das sequências genômicas de cana-de-açúcar, milho e sorgo incentivaram a realização de um estudo mais representativo sobre a distribuição dos genes mustang no genoma das gramíneas. De tal forma, foram identificados e caracterizados nove e seis genes ortólogos em Zea mays e Sorghum bicolor, respectivamente. A topologia da árvore apresentada neste trabalho corrobora a existência de dois clados para a Classe III e três para a Classe IV (Figura 8). Dentro de cada clado a posição dos genes ortólogos, assim como a porcentagem de identidade entre suas seqüências (Tabela 4), está de acordo com as relações filogenéticas entre as espécies analisadas (Roulin et al. 2009). Como esperado, de forma geral, a distribuição dos genes *mustang* de sorgo acompanham os órtologos de arroz, ambas as espécies diplóides e sem eventos de duplicação recente (Adams e Wendel 2005). Contrariamente, assim como em cana-deaçúcar são observados diferentes haplótipos dentro de cada clado devido à ploliploidia

(Saccaro-Jr et al. 2007), em milho foram identificados três genes para cada clado da Classe III (Figura 8). Este resultado é claramente explicado devido à origem do genoma de milho. Os genomas das plantas modernas apresentam evidências de eventos de poliploidização seguidos de silenciamento massivo e eliminação de genes duplicados (Adams e Wendel 2005). No caso particular do milho, as evidências indicam que os seus genitores diplóides divergiram há ~20,5 milhões de anos atrás. Posteriormente, entre 16,5 e 11,4 milhões de anos, e depois da divergência de Sorghum, o milho teria surgido a partir de um evento de alotetrapoliploidização segmental seguida da rediploidização (Figura 17). Desta forma, o 72 % dos genes de cópia única em arroz ou sorgo estão duplicados no genoma diplóide de milho sendo que foram mantidos os genes duplicados que diversificaram a sua função. É interessante ressaltar que genes cujos produtos desempenham funções de fator de transcrição ou transdução de sinal são predominantemente retidos após a duplicação do genoma, aumentando a plasticidade das respostas e, consequentemente, a capacidade de adaptação do organismo (Gaut et al. 2000, Adams e Wendel 2005). Isto explicaria o maior número de genes mustang – que possuem domínios de união a DNA - de Classe III existentes no genoma de milho. Por outro lado, a cana-de-açúcar, híbrido poliplóide recente, apresenta o mesmo número de cópias por classe que as espécies diplóides, no entanto uma célula somática possui 5 e 26 cópias de Classe III e IV, respectivamente (Saccaro-Jr et al. 2007). Deste modo, os genes de Classe III teriam sido preferencialmente retidos durante o surgimento do milho.

Os argumentos expostos indicam que a poliploidização envolve muito mais do que a simples mistura de dois genomas, mas um amplo repertório de ajustes moleculares e fisiológicos. Estes abrangem rearranjos genômicos, perda gênica e mudança no padrão de expressão. Ainda mais, para alguns sistemas tem sido demonstrado que a poliploidização protege o genoma frente às mutações aumentando

a taxa de transposição (Gaut *et al.* 2000). Contudo, surgem organismos com novos fenótipos expandindo a diversificação evolutiva. Fica evidente, então, que a poliploidização é uma força importante na evolução vegetal e que o conteúdo gênico diferencial observado entre as espécies reflete a história evolutiva do genoma.



**Figura 17. Filogenia das gramíneas.**A: Filogenia segundo Roulin *et al.* (2009). B: Filogenia segundo Gaut *et al.* (2000). A seta assinala o evento de alopoliploidização segmental. Os números indicam o valor do conteúdo do genoma 2C.

#### 2. O padrão de expressão dos genes mustang

A transcrição é um processo celular muito importante no controle de regulação da expressão gênica e consequentemente o estudo dos promotores pode proporcionar valiosa informação sobre a funcionalidade dos genes. Os elementos regulatórios em *cis* presentes nos promotores são reconhecidos por fatores de transcrição de forma específica. De tal modo, agem como interruptores sincronizando redes gênicas, e controlando assim os mais diversos processos biológicos, em resposta a estímulos internos (ex. reguladores de crescimento) e do meio ambiente (ex. estresses bióticos e abióticos) (Yamaguchi-Shinozaki e Shinozaki 2005). Esses elementos se encontram organizados em módulos regulatórios onde, a combinação deles e a interação dos fatores de transcrição que os reconhecem determinam o padrão de expressão

temporal-espacial do gene (Yamaguchi-Shinozaki e Shinozaki 2005). Assim, neste trabalho foi realizado um estudo in silico das regiões regulatórias dos três genes mustang de cana-de-açúcar: SCMUG266 BAC095 e SCMUG266 BAC148 de Classe III, e SCMUG148 BAC249 de Classe IV. Foram consideradas como região regulatória as següências não transcritas a montante do início da transcrição até o fim do gene vizinho, resultando em fragmentos de aproximadamente 2.000 pb. Embora muitos estudos sobre expressão gênica em plantas têm mostrado que a maioria dos elementos regulatórios se encontra imediatamente a montante do sítio de início de transcrição, usualmente nas primeiras 1.000 pb (Xiao et al. 2010), é importante salientar que existem motivos de união a fatores de transcrição importantes para a atividade do promotor que podem estar localizados em posições distantes do genoma (Stam et al. 2002). O recrutamento dos fatores de transcrição e a interação destes com o complexo da RNA polimerase determina o posicionamento da maquinaria transcricional sobre o TATA box. No entanto, a presença de elementos TATA não é imprescindível para a transcrição dos genes. Em arabidopsis menos de 39% dos genes apresentam este motivo, sendo que outros sítios de início de transcrição, como CAAT box e Inr, podem substituir a atividade do TATA box (Bernard et al. 2010). Este parece ser o caso do gene SCMUG266 BAC095.

Uma ampla gama de elementos em *cis* responsivos a hormônios, luz, e tecido específicos foi identificada nos três genes sob estudo. Os estudos funcionais sobre transposases domesticadas têm focado no papel regulatório da transcrição destas proteínas, mas nenhum trabalho tem explorado profundamente o perfil de expressão das mesmas, nem estudado os seus promotores. Desta maneira, não há trabalhos que associam os motivos de união a fatores de transcrição presentes nos promotores com dados funcionais. No entanto, a partir da análise integrada dos motivos presentes nos promotores dos genes estudados neste trabalho, junto aos dados funcionais dos genes

onde esses motivos foram identificados e de transposases domesticadas já caracterizadas funcionalmente, é possível revelar interessantes evidências sobre a função das proteínas MUSTANG. A seguir são discutidos alguns pontos que apresentaram maior interesse no contexto deste trabalho.

A região 5' não transcrita do SCMUG266 BAC095 e do SCMUG148 BAC249 apresentou o elemento E2F, presente em promotores de genes envolvidos na progressão do ciclo celular e metabolismo de DNA. Estudos funcionais de duas transposases domesticadas da superfamília *P* têm demonstrado o seu envolvimento na regulação do ciclo celular. A proteína LIN-36, de *C. elegans*, atua como regulador negativo na progressão para a fase G1 do ciclo celular (Boxem e van den Heuvel 2002). Por sua vez THAP1, de vertebrados, possui função de união a DNA participando nos processos de proliferação celular e remodelação de cromatina (Clouaire *et al.* 2005).

Outro elemento identificado nos três promotores estudados foi o ARF, presente em genes responsivos a auxina como os pertencentes à família gênica GH3 (Hagen e Guilfoyle 2002). Foi demonstrado que produtos protéicos desta família gênica estão relacionados com a na transdução do sinal do phyA durante a fotomorfogênese, da mesma maneira que as transposases domesticadas FAR1 e FHY3. O papel destas proteínas estaria associado à comunicação entre a sinalização do phyA e a resposta a auxina mantendo a homeostase hormonal através da conjugação do AIA com aminoácidos (Hsieh *et al.* 2000, Liscum e Reed 2002, Terol *et al.* 2006, Lin *et al.* 2007).

Dois motivos presentes em genes responsivos a citocininas foram identificados nos genes *mustang*: o ARR1 e o CPBCSPOR. O primeiro foi caracterizado em promotores de genes envolvidos no transporte de oxigênio (Ross *et al.* 2004), e o segundo está presente no promotor do gene da enzima protoclorofilídeo oxidoredutase envolvida na rota biossintética de clorofila a (Fusada *et al.* 2005). Estas evidências

indicariam a participação dos genes *mustang* na regulação do estabelecimento da autonomia metabólica da célula.

Um dos maiores desafios da genômica atual é decifrar o papel biológico de cada gene dentro de um organismo, e para isso é fundamental conhecer o perfil de expressão. Desta maneira, a identificação e estudo dos motivos de união a fatores de transcrição colaboram para o entendimento das redes regulatórias que determinam a transcrição dos genes (Vandepoele *et al.* 2006). Experimentos de mutagênese dirigida poderão ser desenvolvidos futuramente para demonstrar definitivamente e delimitar os elementos em *cis* funcionalmente ativos nos genes *mustang* de cana-de-açúcar.

Quando as regiões propostas como regulatórias dos três genes estudados foram fusionadas a genes repórteres, experimentos de expressão transitória demonstraram que os fragmentos analisados são promotores ativos. Este resultado está de acordo com estudos anteriores que mostraram que a maioria dos elementos regulatórios se encontra usualmente nas primeiras 1.000 pb a montante da região codificante, como mencionado anteriormente (Xiao *et al.* 2010). Ainda mais, isto demonstra que além de ambas as classes serem transcricionalmente ativas, mais de um gene por classe é funcional já que ambos os promotores de Classe III promoveram a transcrição. Neste estudo foi avaliada a expressão de dois genes repórteres: *GUS* (β-glucuronidase) e *GFP* (*Green Flurescent Protein*). Tanto os ensaios de expressão transitória quanto de expressão em células BY-2 transgênicas revelaram maior atividade de ambas as proteínas repórteres quando a transcrição estava dirigida pelo promotor do gene SCMUG148 BAC249 de Classe IV (Figuras 11, 12 e 13).

O estabelecimento de linhagens de células transgênicas permitiu testar se os elementos em *cis* identificados nas regiões promotoras são ativos modulando a funcionalidade dos promotores estudados. Devido à existência de dados sobre o papel das transposases domesticadas no desenvolvimento vegetal (Bundock e Hooykaas

2005), como também a identificação de diversos motivos responsivos a hormônios vegetais e relacionados à progressão do ciclo celular, a atividade dos promotores foi avaliada em resposta ao tratamento com AIB, IP e ABA, que são os hormônios mais abundantes e de ocorrência natural em plantas (Davies 2010). Interessantemente, os três promotores estudados demonstraram ser reprimidos quando as células de fumo BY-2 foram tratadas por 24 h com qualquer dos três hormônios testados. Não foi observada resposta diferencial para cada classe. Pelo contrário, enquanto que o promotor do gene SCMUG266 BAC095 da Classe III apresentou uma queda da atividade transcricional após o tratamento com citocinina e ácido abscísico, o promotor do outro gene da Classe III, SCMUG266 BAC148, apresentou o mesmo padrão que o do gene de Classe IV, SCMUG148 BAC249, sendo reprimidos pela auxina. Estes resultados levantam uma importante discussão a respeito da diferenciação funcional dos genes mustang. A análise filogenética apresentada neste trabalho e análises evolutivas dos genes *mustang* realizadas pelo nosso grupo (dados não publicados) demonstraram que estes genes são altamente conservados e possuem forte seleção purificadora. Assim como foi discutido anteriormente a integração das evidências sugere que a diversificação funcional das diferentes classes e haplótipos desses genes contribuíram para a sua conservação.

O papel inibitório de ABA na transição da fase G1 para a fase de síntese (S) do ciclo celular, através da regulação negativa de genes requeridos para o início de replicação de DNA, tem sido bastante descrito, inclusive em culturas de células BY-2 de fumo (Swiatek et al. 2002). Pelo contrário, as auxinas e citocininas estimulam o acúmulo de ciclina G1 para o começo de um novo ciclo e induzem a transcrição da fase G2 para a mitose (M) (Wolters e Jurgens 2009). O fato do promotor SCMUG266 BAC095 ter sido reprimido pelo ABA e pela citocinina pode, a priori, parecer contraditório, entretanto, ambos fitohormônios controlam pontos chave de regulação do

ciclo celular (transição G1-S e G2-M), chamados *checkpoints*, que poderiam ter elementos regulatórios comuns. A presença do motivo E2F e de vários responsivos a ABA, auxinas e citocininas somada à repressão da atividade transcricional dos promotores estudados, associa a atividade dos genes *mustang* com o controle hormonal da divisão celular. Neste sentido, futuros experimentos poderiam ser desenhados para investigar detalhadamente a expressão dos *mustang* durante o ciclo celular e células sincronizadas e submetidas a diferentes tratamentos hormonais.

Os baixos níveis da repressão observada podem ser devidos a características intrínsecas do sistema utilizado que mascarem o efeito dos tratamentos hormonais testados. Estudos com células BY-2 demonstraram que estas possuem mecanismos para minimizar as flutuações exógenas de citocinina, sendo este controle alcançado por um aumento da atividade de citocinina oxidase, enzima que degrada a citocinina (Motyka et al 2003; Hartig & Beck, 2005). No caso da resposta a auxina, o crescimento normal da cultura de células BY-2, requer da suplementação desse hormônio para garantir a progressão do ciclo podendo assim reduzir o efeito do tratamento (Del Pozo et al. 2005). O estudo funcional de genes em sistemas heterólogos é uma prática amplamente empregada (Rossi et al. 1998, Yesilirmak e Sayers 2009), em particular nos casos onde a transformação genética da espécie original não é possível ou, como no caso da cana-de-acúcar, onde devido à complexidade do genoma, o transgene é alvo de intrincados sistemas de silenciamento gênico (Kajihara, Rossi e Van Sluys dados não publicados). Desta maneira foi escolhido o sistema das células de fumo BY-2 pela sua simples propagação e rápida transformação compatível com o presente projeto. No entanto, é importante ter conhecimento que a expressão em sistemas heterólogos pode não refletir fielmente o padrão de expressão endógeno devido a problemas no reconhecimento dos promotores pela maguinaria transcricional heteróloga (Diener e Hirschi 2000).

O perfil de expressão dos genes *mustang* em diferentes tecidos de diferentes idades foi analisado utilizando a técnica de PCR em tempo real de maneira de mensurar a quantidade de mRNA de Classe III e Classe IV separadamente mas sem discriminar ambos haplótipos de Classe III. Pode-se observar que o tecido que apresenta maiores níveis de expressão de ambas as classes é o meristema. Os genes de Classe III apresentaram um pico de expressão em folhas e meristemas de plantas de 30 dias, enquanto que os genes de Classe IV apresentaram os níveis mais altos de expressão nos mesmos tecidos de plantas de 15 dias caindo ao longo do desenvolvimento. De maneira geral as diferenças nos níveis de expressão apresentaram pouca significância estatística, o qual foi mais acentuado para Classe IV. A principal justificativa para esta observação é o fato dos iniciadores não discriminarem entre as diferentes cópias de cada classe, sendo que no genoma da cana-de-açúcar existem 6 e 26 cópias de Classe III e Classe IV, respectivamente. No entanto, um estudo de expressão mais refinado seria interessante uma vez que os dados apresentados indicam a diversificação funcional dos haplótipos de Classe III sendo que um deles possui perfil de modulação da expressão semelhante ao MUSTANG de Classe IV analisado. Por outro lado, é conhecido que os genes que participam da diferenciação e desenvolvimento são expressos em poucas quantidades e com localizações muito determinadas (Wolters e Jurgens 2009), desta maneira, a dissecção pouco precisa dos tecidos poderia ter levado à diluição das quantidades de RNA mensageiros mascarando as diferenças. Os dados de expressão dos genes endógenos estão de acordo com os níveis de expressão de GUS e GFP, obtidos por testes de expressão transitória e expressão estável em células BY-2, que indicaram o promotor de Classe IV como o mais ativo.

O silenciamento e/ou a super expressão dos genes *mustang* seria uma estratégia muito informativa para acrescentar dados funcionais sobre o papel destes

genes na fisiologia vegetal. No entanto, seria conveniente a utilização de arroz como planta modelo já que, como mencionado anteriormente, o complexo genoma da canade-açúcar constitui um empecilho para a análise da expressão de transgenes. Por sua vez, experimentos de hibridização *in situ* assim como de localização subcelular também aportariam dados sobre a distribuição da expressão ao longo dos tecidos e a organela onde as proteínas desenvolvem a sua função. Finalmente, ensaios de precipitação de cromatina ou de interação DNA-proteína esclareceriam o mecanismo molecular de ação destas proteínas.

Desde a sua descoberta os TEs têm suscitado inúmeras questões polêmicas. No começo foram chamados de "DNA egoísta" ou "DNA lixo" (Volff 2010). No entanto, recentemente foi demonstrado que os genes que codificam para transposases são os mais abundantes da natureza e possuem um importante papel no nascimento de novos genes (Sorek 2007, Aziz et al. 2010). Sua ampla distribuição na árvore da vida indica que há um longo tempo residem em organismos eucarióticos e, as evidências da forte interação com o genoma hospedeiro demonstraram que a sua atividade é determinante na estrutura, evolução e dinâmica funcional dos genomas. Desta maneira, hoje são "respeitosamente" chamados de "simbiontes genômicos" (Pennisi 2007; Feschotte 2008).

## Considerações finais

Embora ainda muitas perguntas permaneçam sem resposta sobre a função das transposases domesticadas, desde o seu descobrimento, diversas evidências vêm sendo coletadas. Primeiramente todos os genes derivados de transposases conservam seus domínios de união a DNA intactos (Sinzelle *et al.* 2009), em particular as proteínas MUSTANG de cana-de-açúcar conservam tanto o domínio de união a DNA quanto o sítio catalítico (Rossi *et al.* 2004). Estudos de transcriptômica demonstraram que os produtos dos genes *Daysleeper*, *far1* e *fhy3* são **fatores de transcrição** que regulam a expressão de diversos outros genes. Em particular, *Daysleeper* está diretamente envolvido na **morfogênese**, sendo que o seu silenciamento provoca graves alterações no desenvolvimento da planta perdendo a sua organização órgãotecidual (Bundock e Hooykaas 2005). Já as proteínas FAR1 e FHY3 modulam a resposta à luz dependente de phyA. Ambos os genes são regulados negativamente pelo phyA ativado mantendo a **homeostase do sistema sinalizador** (Lin *et al.* 2007).

Para os *mustang*, originalmente descritos em arroz (Cowan *et al.* 2005), as análises filogenéticas apresentadas neste trabalho demonstram **alto grau de conservação em diferentes espécies vegetais** e foram caracterizados os genes ortólogos de cana-de-açúcar, sorgo e milho. Estudos evolutivos realizados pelo nosso grupo revelaram que os genes *mustang* em gramíneas apresentam seleção purificadora sinalizando a sua necessidade para o normal funcionamento do hospedeiro (Manetti *et al.* dados não publicados).

Por outro lado, os resultados deste trabalho demonstraram que estes genes são de **expressão ubíqua**, em concordância com o observado para os ortólogos de *A. thaliana* (Schmid *et al.* 2005); e **regulados negativamente pelos fitohormônios** testados (AIB, IP e ABA). É importante lembrar aqui, que os promotores dos genes

*mustang* de cana-de-açúcar apresentam motivos regulatórios para estes três hormônios e para vários outros estímulos morfogenéticos.

Assim, integrando todos os resultados discutidos propomos um modelo preliminar hipotético no qual as proteínas MUSTANG funcionariam como fatores de transcrição, mantendo a homeostase da sinalização hormonal garantindo a capacidade de resposta e desta maneira o padrão normal de desenvolvimento (Figura 18). As diversas abordagens experimentais mencionadas ao longo da discussão permitirão testar o modelo e refinar os mecanismos envolvidos.

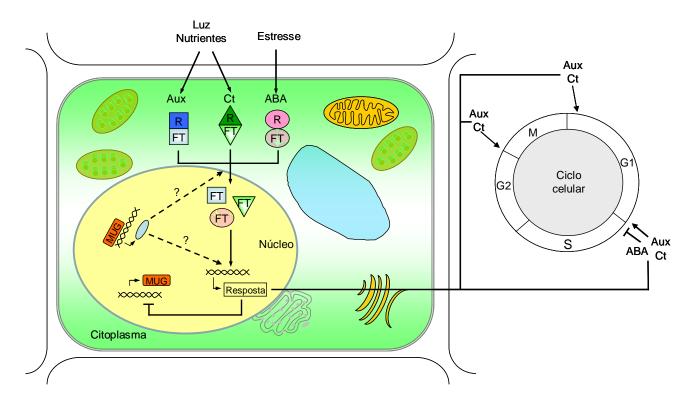


Figura 18. Esquema hipotético de ação das proteínas MUSTANG na célula vegetal.

Estímulos do meio interno ou externo são sinalizados através dos hormônios (Aux, Ct, ABA). Após o reconhecimento do sinal pelo receptor específico (R), a resposta é desencadeada pela mudança no perfil de expressão gênica através dos fatores de transcrição (FT). Neste processo é necessária a participação do produto protéico de um gene cuja transcrição é ativada pela proteína MUSTANG (MUG). Entre as possíveis respostas aos estímulos hormonais está o controle do ciclo celular regulando o padrão de desenvolvimento. Como conseqüência da resposta, a transcrição dos genes *mustang* é inibida. Assim os níveis de MUSTANG caem, baixando a intensidade da resposta e restabelecendo as quantidades de MUSTANG aos níveis iniciais. Desta maneira é mantida a homeostase do sistema e a célula está pronta para perceber um novo estímulo. MUG; proteína MUSTANG. Aux: auxina, Ct: citocinina e ABA: ácido abscísico. FT: fator de transcrição específico a cada hormônio. M: mitose. S: síntese. G1 e G2: intervalos do ciclo celular. Setas indicam estímulo e barras repressão.

#### VI. Conclusões

- Foram identificados os genes *mustang* de Classe III e IV de milho e sorgo.
- -A análise filogenética das seqüências codificantes completas dos genes *mustang* em gramíneas confirmaram a presença de dois clados de Classe III e três clados de Classe IV.
- A identificação de motivos de união a fatores de transcrição revelou a presença de sítios exclusivos de Classe III responsivos a auxina, citocinina, desidratação, frio e luz; e diferenciais entre ambas as cópias responsivos a ácido giberélico, ciclo celular, frio e elementos tecido-específico.
- As três següências clonadas demonstraram ser promotores transcricionais.
- A presença da construção de fusão entre os promotores e os genes repórteres correlacionou com a atividade de ambas as proteínas codificadas em linhagens de células de fumo transgênicas.
- Os promotores dos genes SCMUG266 BAC148 e SCMUG148 BAC249, de Classe III e IV respectivamente, mostraram ser modulados negativamente pela auxina; enquanto que o promotor do gene SCMUG266 BAC095 mostrou uma discreta modulação frente ao tratamento com citocinina e ácido abscísico.
- Os genes *mustang* de Classe III e Classe IV apresentaram perfis de expressão tecidual semelhantes, sendo o meristema o tecido com maior expressão.
- Os genes de Classe III apresentaram um pico de expressão em plantas de 30 dias, enquanto que os genes de Classe IV apresentaram uma queda nos níveis de expressão ao longo do desenvolvimento.
- Os experimentos de expressão transitória, assim como estável em células BY-2
   mostraram que o promotor de Classe IV estudado possui maior atividade transcricional.
   De acordo com os dados de expressão endógena, em plantas de 15 dias os genes da

Classe IV apresentaram o dobro de expressão relativa comparada com os genes da Classe III.

## VII. Referências bibliográficas

- Adams KL, Wendel JF. (2005) Novel patterns of gene expression in polyploid plants. Trends Genet. 21:539–43.
- Agrawal A, Eastman QM, Schatz DG. (1998) Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. Nature 394:744-751.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. (1990) Basic local alignment search tool. J Mol Biol. 215:403–410.
- An G. (1985) High Efficiency Transformation of Cultured Tobacco Cells. Plant Physiology 79:568-570
- Araujo PG, Rossi M, de Jesus EM, Saccaro-Jr NL, Kajihara D, Massa R, de Felix JM, Drummond RD, Falco MC, Chabregas SM, Ulian, EC, Menossi, M, Van Sluys, M.A (2005) Transcriptionally active transposable elements in recent hybrid sugarcane. Plant Journal. 44:707-717.
- Aziz RK, Breitbart M, Edwards RA. (2010) Transposases are the most abundant, most ubiquitous genes in nature. Nucleic Acids Reasearch 38:4207-4217.
- Beaulieu J, Jean M, Belzile F. (2009) The allotetraploid *Arabidopsisthaliana-Arabidopsis lyrata* subsp. *petraea* as an alternative model system for the study of polyploidy in plants. Mol. Genet. Genomics 281: 421–435.
- Bennetzen J. (2000) Transposable element contributions to plant gene and genome evolution. Plant Mol Biol.42:251-269.
- Benjak A, Forneck A, Casacuberta JM. (2008) Genome-wide analysis of the "cut and paste" transposons of grapevine. PLoS ONE. 3:3107.
- Bernard V, Brunaud V, Lecharny A. (2010) TC-motifs at the TATA-box expected position in plant genes: a novel class of motifs involved in the transcription regulation. BMC Genomics. 11:166.
- Bernatzky R, Tanksley SD (1986) Genetics of actin-related sequences in tomato. Theoretical Applied Genetics 72:314–315.
- Bowen J & Jordan K. (2002) Transposable elements and the evolution of eukaryotic complexity. Curr Issued Mol. Biol. 4: 65-76.
- Boxem M, van den Heuvel S. (2002) *C. elegans* class B synthetic multivulva genes act in G(1) regulation. Curr Biol. 12:906–911.

- Brasileiro ACM, Carneiro VTC. (1998) Manual de transformação de genética de plantas. Brasília, DF: Embrapa. 99-107.
- Bundock P, Hooykaas P. (2005) An arabidopsis hAT-like transposase is essential for plant development. Nature. 436: 282-284.
- Butter F, Kappei D, Buchhloz F, Vermeulen M, Mann M. (2010) A domesticated transposon mediates the effects of a single-nucleotide polymorphism responsible for enhanced muscle growth. EMBO Rep 11: 305–311.
- Clouaire T, Roussigne M, Ecochard V, Mathe C, Amalric F, Girard JP. (2005) The THAP domain of THAP1 is a large C2CH module with zinc-dependent sequence specific DNA-binding activity. Proc Natl Acad Sci. USA 102:6907–12
- Cowan RK, Hoen DR, Schoen DJ, Bureau TE. (2005) MUSTANG is a novel family of domesticated transposase genes found in diverse Angiosperms. Molecular Biol Evol 22:2084-2089.
- Daniels J, Smith P, Paton N, Williams CA. (1975) The origin of the genus Saccharum. Sugarcane Breeding. Newsletter 36:24-39.
- Davies P. (2010) The Plant Hormones: Their Nature, Occurrence, and Functions. Plant hormone Biosynthesis, Signal transduction and action. 3º Ed. 1-15.
- del Pozo JC, Lopez-Matas MA, Ramirez-Parra E, Gutierrez C. (2005) Hormonal control of the plant cell cycle. Physiol Plant. 123, 173–183.
- Devos K. (2010) Grass genome organization and evolution. Curr Opin Plant Biol. 13:139-145.
- D'Hont A, Glaszmann JC. (2001) Sugarcane genome analysis with molecular markers, a first decade of research. Proc. Int. Soc. Sugarcane. Technol. 24: 556-559.
- D'Hont A. (2005) Unraveling the genome structure of polyploids using FISH and GISH; examples of sugarcane and banana. Cytogenet Genome Res. 109: 27-33.
- Diener A, Hirschi KI. (2000) Heterologous expression for dominant-like gene activity. Trends Plant Sci. 5:11, 10-11.
- Feschotte C, Pritham EJ. (2007) DNA transposons and the evolution of eukaryotic genomes. Annu Rev Genet. 41: 331–368.
- Feschotte C, Jiang N, Wessler SR. (2002) Plant transposable elements: where genetics meets genomics. Nature Genet. 3:329-341.
- Feschotte C. (2008) Transposable elements and the evolution of regulatory networks.

  Nature Rev Genet 9:397-405

- Fusada N, Masuda T, Kuroda H, Shimada H, Ohta H, Takamiya K. (2005) Identification of a novel cis-element exhibiting cytokinin-dependent protein binding in vitro in the 5'-region of NADPH-protochlorophyllide oxidoreductase gene in cucumber. Plant molecular biology. 59(4):631-45.
- Gaut BS, Le Thierry d'Ennequin M, Peek AS, Sawkins MC. (2000) Maize as a model for the evolution of plant nuclear genomes. *Proc. Natl Acad. Sci. USA*. 97:7008–7015.
- Grandbastien MA, Audeon C, Bonnivard E, Casacuberta JM, Chalhoub B, Costa APP, Le QH, Melayah D, Petit M, Poncet C, Tam SM, Van Sluys MA, Mhiri C. (2005) Stress activation and genomic impact of Tnt1 retrotransposons in Solanaceae. Cytogenet Genome Res. 110:229–41.
- Grivet L, Daniels C, Glaszmann JC, D'Hont A. (2004) A review of recent molecular geneticas evidence for sugarcane evolution and domestication. Ethnobot Res Appl. 2:9-17.
- Hanahan D. (1983) Studies on transformation of E.coli with plasmids. J. Mol. Biol.156:557-580.
- Hagen G, Guilfoyle T. (2002) Auxin-responsive gene expression: genes, promoters and regulatory factors. Plant Mol. Biol. 49: 373–85.
- Hartig K, Beck E. (2004) Endogenous cytokinin oscillations control cell progression of tobacco BY-2 cells. Plant Biol. 7: 33-40
- Haren L, Ton-Hoang B, Chandler M. (1999) Integrating DNA: transposases and retroviral integrases. Annu Rev Microbiol. 53:245-281.
- He P, Friebe BR, Gill BS, Zhou J-M. (2003) Allopolyploidy alters gene expression in the highly stable hexaploid wheat. Plant Mol Biol. 52:401-414
- Higgins D, Thompson J, Gibson T, Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673-4680.
- Higo K, Ugawa Y, Iwamoto M e Korenaga T. (1999) Plant cis-acting regulatory DNA elements (PLACE) database. Nucleic Acids Res. 27: 297-300.
- Hsieh HL., Okamoto H, Wang M, Ang LH, Matsui M, Goodman H e Deng XW. (2000) FIN219, an auxin-regulated gene, defines a link between phytochrome A and the downstream regulator COP1 in light control of Arabidopsis development. Genes Dev. 14:1958–1970.

- Hudson ME, Lisch DR, Quail PH. (2003) The FHY3 and FAR1 genes encode transposase-related proteins involved in regulation of gene expression by the phytochrome A-signaling pathway. Plant Journal 34:453–471.
- Jefferson RA, Kavanagh TA, Bevan MW. (1987) GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J 6:3901-3907.
- Jiang N, Bao Z, Zhang X, Eddy SR, Wessler SR. (2004) Pack-MULE transposable elements mediate gene evolution in plants. Nature 43:569-573.
- Juretic N, Hoen DR, Huynh ML, Harrison PM, Bureau TE. (2005) The evolutionary fate of MULE-mediated duplications of host gene fragments in rice. Genome Res. 15:1292-1297.
- Kapitonov VV, Jurka J. (2005) RAG1 core and V(D)J recombination signal sequences were derived from Transib transposons. PLos Biol. 3:998-1011.
- Kashkush K, Feldman M, Levy AA. (2002) Gene loss, silencing and activation in a newly synthesized wheat allotetraploid. Genetics. 160:1651-1659.
- Karimi M, Inze D, Depicker A. (2002) Gateway vectors for Agrobacterium-mediated plant transformation. Trends Plant Sci. 7:193-195.
- Kumar S, Tamura K, Nei M. (2004) MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Briefings in Bioinformatics 5:150-163.
- Li WL, Zhang P, Fellers JP, B Friebe, Gill BS. (2004) Sequence composition, organization, and evolution of the core Triticeae genome. Plant Journal 40:500–511.
- Li H, Cheng Y, Murphy A, Hagen G, Guilfoyle TJ. (2009) Constitutive repression and activation of auxin signaling in Arabidopsis. Plant Physiol. 149:1277-88.
- Liscum, E e Reed, JW. (2002) Genetics of Aux/IAA and ARF action in plant growth and development. Plant Mol. Biol. 49:387–400.
- Lin R, Ding L, Casola C, Ripoll DR, Feschotte C, Wang H. (2007) Transposase-derived transcription factors regulate light signaling in Arabidopsis. Science 318:1302-1305.
- Lisch D. (2002) Mutator transposons. Trends Plant Sci.7:498-504.
- Lisch D. (2009) Epigenetic regulation of transposable elements in plants. Annu Rev Plant Biol. 60:43-66.

- Liscum E, Reed JW. (2002) Genetics of Aux/IAA and ARF action in plant growth and development. Plant Mol Biol. 49:387-400.
- Liu W, Seto J, Sibille E, Toth M. (2003) The RNA binding domain of Jerky consists of tandemly arranged helix-turn-helix/homeodomain-like motifs and binds specific sets of mRNAs. Mol Cell Biol. 23:4083-4093.
- Macfarlan T, Kutney S, Altman B, Montross R, Yu J, Chakravarti D. (2005) Human THAP7 is a chromatin-associated, histone tail-binding protein that repress transcription via recruitment of HDAC3 and nuclear hormone receptor corepressor. J Biol Chem. 280:7346-7358.
- Marino-Ramirez L, Lewis KC, Landsman D, Jordan IK. (2005) Transposable elements donates lineage specific regulatory sequences to host genomes. Cytogenet. Genome Res. 110:333-341.
- McCabe DE. Swan WF, Matineli BJ. (1988) Stable transformation of soybean (Glycine max) by particle bombarment. BioTechnology 6:923–926.
- McClintock B. (1956) Controlling Elements and the gene. Cold Spring Harbor Symp. Quant Biol. 21:197-216.
- Motyka V, Vankova R, Capkova V, Petrasek J, Kaminek M, Schmulling T. (2003) Cytokinin-induced upregulation of cytokinin oxidase activity in tobacco includes chances in enzyme glycosylation and secretion. Phys Plantarum 117:11-21
- Mukherjee, S.K. (1957) Origin and distribution of Saccharum. Botanical Gazette 19:55-61.
- Murashige T, Skoog F. (1962) A revised medium for rapid growth and bioassays with tobacco tissue culture. Phys Plantarum. 15:473-497.
- Nagata T, Nemoto Y, Hasezawa S. (1992) Tobacco BY-2 cell line as the 'HeLa' cells in the cell biology of higher plants. Int. Rev. Cytol. 132:1–30.
- Okamoto H, Hirochika H. (2001) Silencing of transposable elements in plants. Trends Plant Sci. 6:527-534.
- Parisod C, Alix K, Just J, Petit M, Sarilar V, Mhiri C, Ainouche M, Chalhoub B, Grandbastien MA. (2010) Impact of transposable elements in organization and functioning of allopolyploid genomes. New Phytologist 186:37-45.
- Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer, G, Hellsten U, Mitros T, Poliakov A, Schmutz J, Spannagl M, Tang H, Wang X, Wicker T, Bharti AK, Chapman J, Feltus FA, Gowik U, Grigoriev IV, Lyons E, Maher CA, Martis M, Narechania A, Otillar RP, Penning BW, Salamov AA, Wang

- Y, Zhang L, Carpita NC, Freeling M, Gingle AR, Hash CT, Keller B, Klein P, Kresovich S, McCann MC, Ming R, Peterson DG, Mehboobur R, Ware D, Westhoff P, Mayer KF, Messing J, Rokhsar DS. (2009) The Sorghum bicolor genome and the diversification of grasses. Nature 457:551-556.
- Pennesi E. (2007) Jump genes hop into the evolutionary limelight. Science 317:894-895.
- Pritham EJ. (2009) Transposable elements and factors influencing their success in eukaryotes. J. Hered.100:648-55.
- Robertson DS. (1978) Characterization of a Mutator system in maize. Mutat. Res. 51:21-28.
- Roccaro M, Li Y, Sommer H, Saedler H. (2007) ROSINA (RSI) is part of a CACTA transposable element, *TamRSI*, and links flower development to transposon activity. Mol Genet Genomics 278:243-254.
- Ross EJ, Stone JM, Rockel CG, Arredondo-Peter R, Klucas RV, Sarath G. (2004) Activation of the *Oryza sativa* non-symbiotic haemoglobin-2 promoter by the cytoknin-regulated transcription factor, ARR1. J. Exp. Bot. 55:1721-1731.
- Rossi M, Carrari F, Ponce JLC, Rovere CV, Estrella LH, Gudesblat G, Iusem N. (1998)

  Analysis of an abscisic acid (ABA)-responsive gene promotor belonging to the

  Asr gene family from tomato in homologous and heterologous systems. Mol Gen.

  Genetics 258:1-8.
- Rossi M, Araújo PG, Van Sluys MA. (2001) Survey of transposable elements in sugarcane expressed sequence tags (ESTs). Genet. Mol. Biol. 24:147-154.
- Rossi M, Araújo PG, Jesus EM, Varani AM, Van Sluys MA. (2004) Comparative analysis of sugarcane Mutator-like transposases. Mol. Genet. Genomics 272:194-203.
- Roulin A, Piegu B, Fortune PM, Sabot F, D'Hont A, Manicacci D, Panaud O. (2009) Whole genome surveys of rice, maize and sorghum reveal multiple horizontal transfers of the LTR-retrotransposon Route66 in Poaceae. BMC Evol. Biol. 9:58.
- Rudenko GN, Walbot V. (2001) Expression and post-transcriptional regulation of maize transposable element *MuDR* and its derivatives. Plant Cell 13:553-570.
- Saccaro-Jr NL, Van Sluys MA, Varani AM, Rossi M. (2007) MudrA-like sequences from rice and sugarcane cluster as two bona fide transposon clades and two domesticated transposases. Gene 392:117-125.

- Saccaro-Jr NL. (2007) O sistema Mutator em cana-de-açúcar: uma análise comparativa com arroz. Tese de Mestrado. Instituto de Biociências, Universidade de São Paulo, 121p.
- Saitou N Nei M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4:406–425.
- Sambrook J, Fritsch EF, Maniatis T. (1989) Molecular Cloning, a laboratory manual. 2°Ed., Cold Spring Harbor Laboratory Press, NY.
- Sinzelle L, Izsvak Z, Ivics Z. (2009) Molecular domestication of transposable elements: From detrimental parasites to useful host genes. Cell Mol. Life Sci. 66:1073-1093.
- Schmid M *et al.* (2005) A gene expression map of *Arabidopsis thaliana* development. Nature Genet. 37:501–506.
- Slotkin RK, Freeling M, Lisch D. (2003) Mu killer causes the heritable inactivation of the Mutator family of transposable elements in Zea mays. Genetics 165:781–797.
- Slotkin RK, Martienssen R. (2007) Transposable elements and the epigenetic regulation of the genome. Nat. Rev. Genet. 8:272–285.
- Soltis PS e Soltis DE. (2000) The role of genetic and genomic attributes in the success of poliploids. Proc. Natl. Acad. Sci. USA 97:7051-7057.
- Sorek R. (2007) The birth of new exons: mechanisms and evolutionary consequences. RNA.10:1603-1608.
- Stam M, Belele C, Dorweiler JE, Chandler VL. (2002) Differential chromatin structure within a tandem array 100 kb upstream of the maize *b1* locus is associated with paramutation. Genes Dev. 16:1906-18.
- Swiatek A, Lenjou M, Van Bockstaele D, Inzé D, Van Onckelen H. (2002) Differential effect of jasmonic acid and abscisic acid on cell cycle progression in tobacco BY-2 cells. Plant Phys. 128:201–211.
- Terol J, Domingo C, Talón M. (2006) The GH3 family in plants: genome wide analysis in rice and evolutionary history based on EST analysis. Gene 371:279–290.
- Vandepoele K, Casneuf T, Van de Peer Y. (2006) Identification of novel regulatory modules in dicotyledonous plants using expression data and comparative genomics. Genome Biol. 7:R103.
- Vettore AL *et al.* (2003) Analysis and functional annotation of an expressed sequence tag collection for tropical crop sugarcane. Genome Res. 13:2725-2735.

- Volff JN. (2010) Tame affairs: domesticated transposase and domestic pigs. EMBO Rep. 11:241-242.
- Yamaguchi-Shinozaki K, Shinozaki K. (2005) Organization of cis-acting regulatory elements in osmotic- and cold-stress responsive promoters. Trends Plant Sci.10:88-94.
- Yamashita D, Sano Y, Adachi Y, Okamoto Y, Osada H, Takahashi T, Yamaguchi T, Osumi T, Hirose F. (2007) hDREF regulates cell proliferation and expression of ribosomal protein genes. Mol. Cell Biol. 27:2003-2013.
- Yesilirmak F, Sayers Z. (2009) Heterelogous expression of plant genes. International Journal of Plant Genomics 2009:1-16.
- Yu Z, Wright SI, Bureau TE. (2000) *Mutator*-like elements in Arabidopsis thaliana: structure, diversity and evolution. Genetics 156:2019-2031.
- Xiao YL, Redman JC, Monaghan EL, Zhuang J, Underwood BA., Moskal WA, Wang W, Wu HC, Town CD. (2010) High throughput generation of promoter reporter (GFP) transgenic lines of low expressing genes in Arabidopsis and analysis of their expression patterns. Plant Methods 6:1-13.
- Zhang J, Madden TL. (1997) PowerBLAST: A new network BLAST application for interactive or automated sequence analysis and annotation. Genome Res. 7:649-656.
- Walbot V. (2000) Saturation mutagenesis using maize transposons. Curr. Opin. Plant Biol. 3:103-107.
- Wolters H, Jurgens G. (2009) Survival of the flexible: hormonal growth control and adaptation in plant development. Nat. Rev. Genet. 10:305-317.