Universidade Estadual de Campinas

Instituto de Biologia

Identificação e Caracterização de Genes Expressos em Resposta ao Estresse por Baixa Temperatura em Cana-de-açúcar

Fábio Tebaldi Silveira Nogueira

Tese apresentada no Instituto de Biologia da Universidade Estadual de Campinas, para obtenção do título de Doutor em Genética e Biologia Molecular, área de concentração Genética Vegetal e Melhoramento.

Prof. Orientador: Dr. Paulo Arruda

Campinas
2004

Data da defesa: 17/08/2004

Banca Examinadora

Prof. Dr. Paulo Arruda

Prof. Dr. Ladaslav Sodek

Prof. Dr. Jörg Kobarg

Prof. Dr. Marcelo Menossi Teixeira

Profa. Dra. Gláucia Mendes Souza

Prof. Dr. Jiří Borecký

Profa. Dra. Maria Inês Tiraboschi Ferro

O cientista não estuda a natureza porque ela é útil; estuda-a porque se delicia com ela, e se delicia com ela porque ela é bela. Se a natureza não fosse bela, não valeria a pena conhecê-la e, se não valesse a pena conhecê-la, não valeria a pena viver.

(autor desconhecido)

Aos meus pais, Rogério e Maria Inêz.
Às minhas irmãs, Giovana, Marcela e Luci.
Às minhas avós, Ireni e Tereza.
Aos meus avôs Juca e Ciro (*in memorian*)

Pelo amor, carinho e todo o apoio incondicional em todos os momentos de minha vida, só me resta agradecer...

À minha amada esposa Alessandra,

Por me ensinar que a vida pode ser realmente extraordinária...

AGRADECIMENTOS

A Deus por minha existência.

Ao orientador, Prof. Dr. Paulo Arruda, pela orientação, amizade e por ter acreditado no meu trabalho desde o início.

Aos doutores Ladaslav Sodek, Jörg Kobarg, Jiří Borecký, Maria Inês Tiraboschi Ferro, Gláucia Mendes Souza, membros da banca examinadora.

Aos membros doutores da pré-banca Adriana Capella, Márcio de Castro Silva Filho e Michel Vincentz pelas críticas, revisões e sugestões.

Ao Dr. Eugênio C. Ulian (Copersucar, Piracicaba) pela amizade, sugestões e pelo material vegetal utilizado no trabalho.

Ao Dr. Marcelo Menossi Teixeira pela amizade e co-orientação nos trabalhos com os arranjos de DNA.

À Dra. Anete pela compreensão e por me ajudar nos momentos cruciais da defesa de tese.

Às meninas da secretaria da pós, Lourdes e Nilvana por quebrarem todos os meus “galhos” nas etapas finais.

Às meninas da secretaria, Tânia, Sandra e Fabiana por resolverem as questões burocráticas (leia-se: chatas) e monetárias (leia-se: liberação de verba científica).

Ao Dr. Laurent Grivet (Syngenta, França) pelas críticas e sugestões do trabalho.

Ao Dr. Nilson Zanchin (LNLS, Campinas) pelo auxílio nas análises microscópicas.

Ao grande amigo Vicente, pelas discussões, conselhos, convivência, amizade duradoura e paciência, até quando nem eu mesmo me agüentava.

Às minhas fiéis “desorientandas”, Kívia e Paula, pela amizade e por “quebrar vários galhos” no laboratório.

Aos amigos Almir “cabeça”, Geraldo, Jiri e Paulo “xilol”, pela grande amizade e pelas longas “discussões científicas” regadas à cerveja. Vou sentir saudades...

Aos atuais e antigos amigos do laboratório genoma de plantas, Ana, Susan, Adriana, Heidi, Fabiana, Daniela “Danizinha”, Natália, Thaís (as “genometes”), Rodrigo

“the flash”, Danilo, Germano e Márcio pela paciência que tiveram comigo e pelo companheirismo demonstrado desde o início.

Aos amigos do laboratório genoma funcional de plantas Jú, Renato, Sandra, Jorge, Melina, Rodrigo, Dudu, pelos divertidos momentos que passamos juntos, pela troca de idéias e pela amizade.

Aos Sartori, minha segunda família, por seu grande afeto e amizade que fizeram de São Paulo minha segunda casa.

Aos ex-CBMEGuianos Luciane, Isabel e Fábio, pelas discussões, conselhos e amizade.

Aos novos amigos da Alellyx pela força e compreensão na etapa final.

Aos funcionários do CBMEG e do instituto de biologia pela colaboração.

À FAPESP pelo apoio financeiro.

Enfim, a todos que direta ou indiretamente contribuíram para realização deste trabalho.

ÍNDICE

LISTA DE ABREVIAÇÕES

AAB/ABA	ácido abscísico
ACC	ácido 1-aminociclopropano-1-carboxílico
APRT	gene codificador da enzima adenosina fosforribosil transferase
AOx	proteína oxidase alternativa
AOx	gene codificador da proteína oxidase alternativa
At	*Arabidopsis thaliana*
[α-^{32}dCTP]	desoxirribonucleotídeo que contém fósforo 32
[α-^{33}dCTP]	desoxirribonucleotídeo que contém fósforo 33
Bn	*Brassica napus*
cDNA	DNA complementar
cors	*cold-responsive genes* (genes que respondem a baixa temperatura)
CV	coeficiente de variação
DNA	ácido desoxirribonucléico
EROS	espécies reativas de oxigênio
EST	etiqueta de seqüência expressa
GFP	*green fluorescent protein*
GUS	gene codificador da enzima β-glucorinidase de *Escherichia coli*
GUS	enzima β-glucorinidase de *Escherichia coli*
H_2O_2	peróxido de hidrogênio
IP3	inositol 1,4,5-trifosfato
Kb	kilobase(s)
kDa	kilodalton(s)
$\Delta\mu_H+$	gradiente eletroquímico de prótons
MACF	*mitochondrial anion carrier family* (família de proteínas carreadoras mitocôndriais)
METS	*mitochondrial energy transfer signature* (assinatura de transferência de energia)
MP	máxima parsimônia
NAC	proteína que contém domínio NAC
NAC	gene que codifica proteína que contém domínio NAC
NLS	*nuclear localization signal* (sinal de localização nuclear)
ORF	*open reading frame* (seqüência aberta de leitura)
Os	*Oryza sativa*
PCR	reação da polimerase em cadeia
pb	par(es) de base(s)
p/v	peso por volume
PEG	Polietilenoglicol
Pi	fósforo inorgânico
pI	ponto isoelétrico
PUMP	proteína desacopladora mitocondrial de plantas
PUMP	gene codificador de proteína desacopladora mitocondrial de plantas
R	coeficiente de correlação
RNA	ácido ribonucléico
RNAm/mRNA	RNA mensageiro
RT-PCR	*reverse transcriptase* PCR
SAGE	*serial analysis of gene expression* (análise serial da expressão gênica)
SD	desvio-padrão
Ss	*Saccharum sp.*

RESUMO

Estresses ambientais representam um dos principais fatores limitantes para a produtividade agrícola. Sendo assim, fazem-se importantes estudos que têm por objetivo analisar ao nível molecular os efeitos causados por esses estresses. Neste trabalho, analisamos o perfil de expressão gênica em plântulas de cana-de-açúcar submetidas a baixas temperaturas. Inicialmente, membranas de alta densidade (DNA *macroarrays*) contendo 1536 ESTs (Etiquetas de Seqüências Expressas) do projeto SUCEST (*Sugarcane EST Project*) foram hibridadas com sondas de cDNAs obtidas a partir de RNA total de tecido foliar de plântulas de cana-de-açúcar crescidas *in vitro* e submetidas a 4 ^{o}C por 3, 6, 12, 24 e 48 h. Como controle, foram utilizadas plântulas *in vitro* crescidas a 26 ^{o}C. O experimento foi repetido duas vezes, utilizando-se RNA total de novas plântulas tratadas e controle. Somente ESTs que apresentaram padrão de indução ou repressão semelhantes nos dois experimentos foram considerados para posterior análise. Desta forma, os resultados permitiram a identificação de 59 ESTs que foram induzidos ou reprimidos pelo estresse abiótico, dentre os quais alguns ESTs sem homólogos no *GenBank*. Alguns genes selecionados tiveram seu perfil de expressão confirmado por *RNA-blot*. Utilizando o banco de dados do SUCEST e seqüências de proteínas oriundas do NCBI, foi gerado um banco de dados contendo 33 proteínas de cana-de-açúcar relacionadas ao frio.

Dentre essas proteínas foi identificada proteína similar a PUMP (*Plant Uncoupling Mitochondrial Protein*), a qual está envolvida em mecanismos de redução da produção de espécies reativas de oxigênio durante estresses abióticos. Posteriormente, quatro outras proteínas pertencentes à família PUMP foram identificadas no banco de dados do SUCEST, além de quatro proteínas pertencentes à família AOx (*alternative oxidase*). Análises filogenéticas corroboraram a hipótese da existência de novos membros de ambas as famílias de proteínas em cana-de-açúcar (*Saccharum sp. PUMPs* e *AOxs*, ou seja, *SsPUMPs* e *SsAOxs*, respectivamente) e em *Arabidopsis* (*AtPUMPs* e *AtAOxs*). Análise do perfil de expressão desses genes sugere que *PUMPs* e *AOxs* são expressos em órgãos diferentes em cana-de-açúcar e em *Arabidopsis*. Adicionalmente, resultados de *RNA-blot* sugeriram que dois membros da família de PUMPs (*SsPUMP4* e *SsPUMP5*) e um membro da família AOx (*SsAOx1c*) foram induzidos por baixa temperatura (4 ^{o}C), enquanto que, em *Arabidopsis*, dois membros da família PUMP (*AtPUMP4* e *AtPUMP5*) e um membro da família AOx (*AtAOx1a*) foram também induzidos pelo estresse abiótico, entretanto, com perfis diferentes.

Interessantemente, o gene *AtAOx1d* foi reprimido após exposição das plantas a baixas temperaturas.

Análises *in silico* demonstraram a existência de 26 membros da superfamília de proteínas NAC em cana-de-açúcar, incluindo um gene identificado em estudos de expressão gênica, denominado *SsNAC23* (*Saccharum* sp. *NAC23*). Este gene foi completamente seqüenciado e posteriormente analisado. Resultados de *RNA-blot* sugerem que o gene *SsNAC23* é regulado por baixa temperatura (4 °C), mas não por 12 °C, além de herbivoria (*Diatraea saccharalis*) e estresse hídrico (Polietilenoglicol, PEG 6000, 20%). Ensaios de localização subcelular demonstraram que a proteína SsNAC23 localizou-se preferencialmente no núcleo de células epidérmicas de cebola (*Allium cepa*). A construção de um modelo molecular do domínio NAC da proteína SsNAC23 permitiu a análise de determinantes estruturais possivelmente envolvidos em ligação a DNA, sugerindo seu papel como fator de transcrição. Alinhamento desses determinantes de outras proteínas contendo domínio NAC revelou alta conservação de importantes aminoácidos envolvidos no contato entre proteína e molécula de DNA. Em conjunto, nossos resultados sugerem que a proteína SsNAC23 é um fator de transcrição envolvido em mecanismos de resposta a estresses bióticos e abióticos.

ABSTRACT

Environmental stresses represent one of the main factors limiting crop productivity, being very important studies that aim to analyze, at a molecular level, the effects caused by these stresses. In the present project, we analyzed the sugarcane gene expression profiles in response to low temperature. Initially, high-density filter arrays containing 1536 ESTs (Expressed Sequence Tags) from SUCEST (Sugarcane EST Project) were hybridized with P^{33}-labeled cDNA probes obtained from total RNA of sugarcane plantlets growing at 4 °C during 3, 6, 12, 24, and 48 h. As a control, plantlets growing at 26 °C were utilized. The experiment was repeated twice, being total RNA extracted from novel cold-treated and untreated plantlet sets. Only those ESTs, which presented similar expression profiles in both experiments, were taken into consideration for a further analysis. The results allowed identifying 59 ESTs that were induced or repressed by chilling stress. Selected genes had their expression profiles confirmed by RNA-blot analysis. Using SUCEST database and protein sequences from NCBI, a database containing 33 sugarcane proteins related to cold temperatures was generated.

Among those proteins, we identified a PUMP (Plant Uncoupling Mitochondrial Protein) gene family member. This protein was reported to be involved in oxidative stress responses. Further, in SUCEST database, four additional proteins belonging to PUMP and four proteins belonging to AOx family (alternative oxidase) were identified. Phylogenetic analyses confirmed the existence of novel members of both protein families in sugarcane (*Saccharum sp. PUMPs* and *AOxs*: *SsPUMPs* and *SsAOxs*) and *Arabidopsis* (*AtPUMPs* and *AtAOxs*). Gene expression profile analyses suggested that *PUMPs* and *AOxs* are expressed in different plant organs. RNA-blot results suggested that two members of sugarcane PUMP gene family (*SsPUMP4* and *SsPUMP5*) and one member of AOx family (*SsAOx1c*) were induced by low temperature (4 °C). *Arabidopsis PUMP4* and *PUMP5* orthologs and one member of AOx family (*AtAOx1a*) were also induced by cold stress. Interestingly, *AtAOx1d* was down-regulated after plantlets were exposed to low temperature.

In silico analyses demonstrated the existence of 26 members of sugarcane NAC domain protein family, including the cold-inducible *SsNAC23* (*Saccharum* sp. *NAC23*). This gene was completely sequenced and analyzed. RNA-blot results suggested that *SsNAC23* is modulated by extreme low temperatures (4 °C), but not at moderate ones (12 °C). *SsNAC23* transcripts also accumulated in response to herbivory (*Diatraea saccharalis*) and PEG-induced water stress. Nuclear localization essays demonstrated that SsNAC23 was targeted to nucleus of

onion cells (*Allium cepa*). A molecular model of the NAC domain of SsNAC23 allowed analyzing the structural determinants putatively involved in DNA-binding, evidencing the SsNAC23 transcriptional activator nature. Alignment of other NAC domain proteins revealed high structural determinants conservation in different plant species. In summary, our results indicated that the SsNAC23 is a transcription factor involved in biotic and abiotic stress response.

INTRODUÇÃO

A cana-de-açúcar é uma espécie vegetal de origem tropical/subtropical (Daniels e Roach, 1987), a qual o Brasil cultiva desde o seu descobrimento (Carvalho, 1993). Dentre os problemas que afetam a cultura, destaca-se a sensibilidade ao estresse por baixa temperatura (Tai e Lentini, 1998). Muitas plantas desenvolvem tolerância ao frio (*chilling* e/ou *freezing*) quando expostas previamente a temperaturas baixas (na faixa de 0 a 15 °C), sendo este processo denominado de aclimatação ao frio, o qual envolve modificações na expressão gênica. Enquanto espécies vegetais de clima temperado possuem habilidade de sobrevivência em temperaturas de congelamento (*freezing*; Thomashow, 1998), espécies de clima tropical/subtropical possuem pouca e nenhuma capacidade de aclimatação a temperaturas de *chilling* (temperaturas baixas acima de 0 °C) e ao congelamento (temperaturas abaixo de 0 °C), respectivamente. Entretanto, nestas espécies, os mecanismos moleculares de percepção, sinalização molecular e modificação do metabolismo em resposta ao estresse por frio são pouco estudados. Desta forma, um maior entendimento desses mecanismos pode acarretar o desenvolvimento de metodologias capazes de aumentar a tolerância dessas espécies vegetais ao frio, permitindo seu plantio em áreas agrícolas ainda não utilizadas devido aos extremos de temperatura. Neste contexto, a presente tese de doutoramento teve como principal objetivo avaliar o perfil de expressão de cana-de-açúcar em resposta ao estresse por baixa temperatura.

Para melhor explicitação dos assuntos abordados nesta tese, será apresentada uma síntese da bibliografia utilizada para sua confecção, seguida dos objetivos, capítulos (artigos) – em que são apresentados os resultados do trabalho–, bem como a discussão e conclusões gerais.

1. Cana-de-açúcar: origem, genética e importância econômica

A cana-de-açúcar (*Saccharum* sp.) é uma espécie alógama que pertence à família *Gramineae* (*Poaceae*), sendo do gênero *Saccharum*. Essa espécie vegetal é originária do Sudeste Asiático, na grande região central da Nova Guiné e Indonésia (Daniels e Roach, 1987). As variedades cultivadas atuais de cana-de-açúcar são resultado de hibridação interespecífica envolvendo *Saccharum officinarum* L. (2*n*=70-122). Os cruzamentos ocorrem com maior freqüência com *S. espontaneum* L. (2*n*=36-128), algumas vezes com *S. barberi* (2*n*=60-140) ou *S. sinense* (2*n*=104-128) e raramente com *S. robustum* (2*n*=60-170). Devido à sua origem

multiespécie, a cana-de-açúcar possui um dos mais complexos genomas vegetais, apresentando um variável nível de ploidia (geralmente de 5X a 14X, em que X = 5, 6, 8, 10, 12 ou 14) (Ingelbrecht et al., 1999). Desta forma, as variedades cultivadas de cana-de-açúcar são heterozigotas, contendo vários alelos diferentes por lócus. Essa redundância genômica poderia ter um efeito "tamponante" para as mutações, além de facilitar o aparecimento de novos genes durante os eventos de duplicação gênica (Ma et al., 2003).

Apesar do mapeamento genético estar sendo realizado (Grivet e Arruda, 2001), a seleção por intercruzamento visando à obtenção de variedades de cana-de-açúcar mais produtivas, resistentes a pragas e doenças e adaptadas a ambientes diversos, é um processo bastante lento, cuja finalização pode levar de 12 a 15 anos. Este quadro, aliado à complexidade de seu genoma, dificulta a aplicação de técnicas convencionais de melhoramento genético, tornando a cana-de-açúcar uma excelente candidata ao melhoramento por meio da engenharia genética. A identificação de genes responsáveis por qualidades agronomicamente desejáveis e sua posterior manipulação por meio de técnicas de biologia molecular podem proporcionar a obtenção de variedades bem sucedidas, reduzindo drasticamente as perdas na agricultura, além de permitir o aproveitamento de solos até então não utilizáveis. Além disso, devido aos altos níveis de similaridade genética entre gramíneas (Guimarães et al., 1997), os estudos realizados em cana-de-açúcar podem ser extrapolados para outras gramíneas cultivadas.

A cana-de-açúcar é considerada uma das 10 mais importantes culturas mundiais, contribuindo com cerca de 60% a 70% de todo o açúcar produzido no globo (Ingelbrecht et al., 1999) e com uma das maiores produções de biomassa entre as espécies cultivadas (Ma et al., 2003). Desde os tempos do Brasil colônia até os dias atuais, a cultura de cana-de-açúcar tem sido uma grande fonte de riquezas para a economia brasileira (Saciloto, 2003). O Brasil é o maior produtor mundial de cana-de-açúcar com uma safra estimada em 338 milhões de toneladas em 2001, equivalente a 27% da produção mundial (http://sucest.lad.ic.unicamp.br). A cultura é cultivada em cerca de cinco milhões de hectares, com concentração nas regiões Centro-Sul e Nordeste (Carvalho, 1993). Alagoas e Pernambuco, além de Minas Gerais e Rio de Janeiro, são importantes produtores, contribuindo em conjunto com quase 30% da safra brasileira. Entretanto, o estado de São Paulo concentra a maioria das lavouras de cana-de-açúcar. Há cerca de 50 mil produtores e 308 unidades de processamento industrial, produzindo 17,7 milhões de toneladas de açúcar e 13,7 milhões de m^3 de etanol por ano (Saciloto, 2003).

No Brasil, a indústria da cana-de-açúcar mantém o maior sistema de produção de energia comercial de biomassa do mundo, por meio do etanol (substituindo cerca de 20% da gasolina) e do uso do bagaço (equivalente a 11 milhões de toneladas de óleo), o qual pode ser

uma fonte alternativa para a resolução da questão energética. Além disso, a indústria da cana-de-açúcar contribui com a produção de plásticos biodegradáveis, açúcares não calóricos e compostos químicos de interesse farmacêutico. A agroindústria canavieira é responsável por milhões de empregos e geração de divisas para o país através da exportação de açúcar e álcool e pelo aproveitamento racional da biomassa vegetal, gerando energia elétrica e produtos para a indústria e para a agropecuária (Raizer, 1998). Um de seus produtos, o bagaço, é capaz de gerar excedentes de cerca de três quilowatts/tonelada de energia elétrica. Entretanto, com a utilização de parte da palha e de tecnologias mais eficientes, este potencial poderia ser elevado para cinco quilowatts, aumentando ainda mais a importância econômica da cultura (Saciloto, 2003).

A cana-de-açúcar obteve um grande sucesso de produção por meio do uso de variedades advindas do melhoramento genético clássico desenvolvido por centros de pesquisa e estações experimentais, tais como o Instituto Agronômico de Campinas (IAC) e a Copersucar. Na Copersucar, com a utilização dessas variedades, o faturamento bruto anual é de R$ 3,7 milhões (safra 2001/2002). As usinas e destilarias associadas à Copersucar processam anualmente um volume aproximado de 60 milhões de toneladas de cana, sendo a produção da ordem de 3,6 milhões de toneladas de açúcar e 2,4 bilhões de litros de álcool (safra 2001/2002). Esses volumes permitem à cooperativa uma participação de, aproximadamente, 21% do mercado brasileiro de açúcar e álcool (Saciloto, 2003).

2. Bases fisiológicas e moleculares da resposta aos estresses ambientais

2.1. Estresses ambientais

Estresses ambientais ou abióticos representam um dos principais fatores limitantes para a produtividade agrícola no mundo inteiro. Esses estresses não só afetam as áreas cultivadas hoje, mas também representam uma barreira significativa para a introdução de espécies cultiváveis em áreas que ainda não são utilizadas para a agricultura (Cherry, 1994).

Sendo o termo estresse recorrente neste estudo, faz-se necessário defini-lo. Estresse é geralmente entendido como qualquer fator que promove uma situação desvantajosa para o organismo. Em muitos casos, o nível do estresse é medido em termos de sobrevivência do organismo, acúmulo de biomassa, e, no caso de plantas, alterações em processos relacionados diretamente com o crescimento vegetal, tais como a absorção de CO_2 e de minerais (Taiz e Zeiger, 1998). Além disso, o conceito de estresse está intimamente relacionado a outro

conceito, o de tolerância ao estresse, o qual é definido como a capacidade de aclimatação da planta às condições desfavoráveis (Taiz e Zeiger, 1998). Como predito por Taiz e Zeiger (1998), o processo de aclimatação distingue-se do de adaptação, sendo este último resultado de processos de seleção contínuos através de várias gerações de uma espécie vegetal. A aclimatação, por sua vez, envolve geralmente alterações transientes da expressão gênica e produção de compostos específicos durante a exposição do organismo à (s) condição (ões) desfavorável (is). Entretanto, uma determinada condição ambiental pode ser considerada desfavorável para uma espécie vegetal e não para outra. Um exemplo poderia ser a comparação de temperaturas ideais de crescimento entre duas espécies vegetais. Enquanto ervilha (*Pisum sativum*) desenvolve-se melhor em temperaturas de até 20 °C, soja (*Glycine max*) cresce em temperaturas de 30 °C ou superiores, sem, entretanto, apresentar sintomas de estresse por altas temperaturas (Taiz e Zeiger, 1998).

Como as plantas são organismos que não podem se locomover, elas necessitam aclimatar-se às condições impostas de clima, solo e temperatura. Sendo assim, não é de se surpreender que seus mecanismos fisiológicos e moleculares sejam extremamente complexos (Keegstra e Thomashow, 2002). As alterações ambientais atuam tanto no metabolismo primário quanto no secundário dos organismos vegetais. O processo fotossintético, um exemplo do metabolismo primário, constantemente adapta-se às diferentes condições ambientais impostas, tais como o excesso de luminosidade e temperatura (Keegstra e Thomashow, 2002). Por outro lado, o metabolismo secundário é também largamente afetado pelas modificações do ambiente. A biossíntese de flavonóides e monolignóis, por exemplo, é profundamente modificada em resposta a estresses bióticos e abióticos (Keegstra e Thomashow, 2002). Finalmente, os fatores ambientais afetam o crescimento vegetal não somente no nosso planeta, mas também no espaço. Ferl et al. (2002) realizaram experimentos na Estação Espacial Internacional americana (*International Space Station, ISS*) utilizando plantas de *Arabidopsis thaliana* e observaram que o ambiente espacial causou alterações no desenvolvimento vegetal, incluindo modificações na regulação da expressão gênica.

Os primeiros passos para o processo de aclimatação são a percepção do estresse e o posterior “disparo” das respostas fisiológicas e moleculares (Knight e knight, 2001). As respostas fisiológicas, de um modo geral, incluem o fechamento estomático, redução da fotossíntese, redução na absorção de nutrientes, produção de compostos específicos, etc. As respostas moleculares, por sua vez, envolvem a alteração de genes específicos, bem como catabolismo e anabolismo de proteínas (Guy, 1990; Holmberg e Bülow, 1998). Muitos dos

compostos e proteínas produzidos podem atuar direta ou indiretamente no aumento de tolerância ao estresse ambiental, como apresentado na Tabela 1.

Tabela 1. A complexidade da aclimatação aos estresses ambientais (adaptada de Cushman e Bohnert, 2000)

Classe dos compostos/proteínas	Exemplos	Possível modo de ação
Osmoprotetores	Prolina, glicina betaína, manitol, sorbitol, sacarose, trealose, frutanas.	Ajustamento osmótico, proteção contra espécies reativas de oxigênio.
Protetores contra espécies reativas de oxigênio	Enzimas (catalase, superóxido dismutase, ascorbato peroxidase, glutationa S-transferase, glutationa peroxidase, oxidase alternativa, proteína desacopladora mitocondrial); Compostos não-enzimáticos (ascorbato, flavonas, carotenóides, antocianinas).	Detoxificação celular.
Proteínas de estresse	*Late embryogenesis abundant* (*LEA*), aquaporinas.	Estabilização de proteínas, estabilização da membrana celular, redução do estresse hídrico.
Heat shock proteins (*HSP*)	Várias *HSPs* presentes em diferentes compartimentos intracelulares.	Prevenção do dobramento incorreto de proteínas, modulação traducional.
Transportadores de prótons/íons	Transportadores de K^+, H^+/ATPases, Na^+/H^+ antiporters.	Estabelecimento do gradiente de prótons, remoção de íons tóxicos do citoplasma e de organelas.
Proteínas de metabolismo específico	Desaturases de ácidos graxos	Aumento da fluidez da membrana plasmática.
Componentes da sinalização celular	*Hisdine kinases*, *MAP kinases*, *Ca^{2+}-dependent protein kinases*,	Transdução de sinais celulares.

	phosphatases, *Ca^{2+}-sensors (SOS3), inositol kinases.*	
Fatores de transcrição	Proteínas pertencentes a várias famílias gênicas (*EREBP/AP2*, *zinc finger*, *Myb*)	Ativação da transcrição gênica.
Reguladores de crescimento	Ácido abscísico, citocininas, brassinoesteróides, poliaminas, etileno, etc.	Alteração do metabolismo e da expressão gênica.

2.2. Estresse por baixa temperatura

Um dos principais estresses abióticos que afetam o desenvolvimento e produtividade de plantas cultiváveis é aquele relacionado com extremos de temperaturas (Iba, 2002). A faixa de temperatura na qual as plantas estão expostas pode variar consideravelmente tanto temporal como espacialmente (Iba, 2002). Cada espécie vegetal possui uma temperatura ótima de crescimento; conseqüentemente, sua distribuição geográfica é determinada pela faixa de temperatura na qual a espécie é capaz de sobreviver. Enquanto plantas nativas de regiões equatoriais toleram temperaturas de até 60 °C, espécies vegetais desenvolvendo-se em regiões temperadas ou mesmo frias podem tolerar temperaturas em torno de -30 °C durante a estação de inverno (Thomashow, 1999).

Com exceção de algumas poucas espécies vegetais que possuem a capacidade de gerar calor metabolicamente (Saltveit e Morris, 1990), a maioria das plantas desenvolveu estratégias para evitar os extremos sazonais de temperatura que ocorrem em seu habitat natural. O estresse por baixas temperaturas ou por frio é um dos fatores limitantes da atividade agrícola mundial, promovendo redução na produção de alimentos em diversas regiões no globo (Pearce, 1999). Desta forma, fazem-se necessários estudos cujo objetivo é entender os processos fisiológicos e moleculares da resposta a esse estresse ambiental.

O estresse por frio pode ser dividido em duas categorias: *chilling* (temperaturas na faixa de 0 a 15 °C, ou seja, sem a formação de cristais de gelo intercelulares) e *freezing* (temperaturas de congelamento). A primeira afeta diretamente o funcionamento celular, enquanto a segunda atua indiretamente, danificando as células por desidratação (Pearce, 1999). No texto, de um modo geral, esse estresse será tratado como estresse por frio, englobando ambas as categorias.

Baixas temperaturas causam alterações fisiológicas, sendo estas divididas em alterações primárias e secundárias. Alteração primária é definida como a percepção do estresse, que desencadeia disfunções na planta, as quais são rapidamente revertidas se o organismo vegetal retornar às condições favoráveis de crescimento. Alterações secundárias, por sua vez, ocorrem em conseqüência das modificações primárias e podem ou não ser reversíveis (Saltveit e Morris, 1990). Desta forma, vários processos biológicos podem ser afetados quando as plantas estão expostas a baixas temperaturas, levando à perda de vigor e redução na razão de crescimento na ausência de outros sintomas visíveis de injúrias. Dentre as principais injúrias causadas por esse estresse, podemos citar necrose, redução da fotossíntese e da absorção de água, além de alterações na integridade da membrana plasmática, ocasionando perda de fluidos celulares para o apoplasto (Saltveit e Morris, 1990). Outro importante sintoma de estresse por frio, o amarelecimento foliar, pode ocorrer na presença de luz como conseqüência da fotooxidação, a qual aumenta a geração de espécies reativas de oxigênio (EROS) em compartimentos intracelulares. Além desses compostos funcionarem como sinalizadores secundários, modificando o metabolismo e a expressão gênica durante a exposição ao estresse (Prasad et al., 1994), as injúrias causadas pelo frio ocorrem, pelo menos em parte, em decorrência do aumento da produção de EROS que, por sua vez, atuam nas membranas e nos fotossistemas. Apesar dos cloroplastos serem considerados os principais locais de produção de EROS, principalmente em tecidos foliares (Richter et al., 1990), evidências experimentais sugerem que as mitocôndrias são uma importante fonte de espécies reativas de oxigênio em diversos tecidos de plantas sensíveis ao estresse por baixas temperaturas (Purvis et al., 1995). Finalmente, plantas expostas a baixas temperaturas podem tornar-se susceptíveis ao ataque de patógenos, tais como fungos e vírus (Gaudet, 1994; Hull, 2002). Com isso, a análise da resposta fisiológica e molecular ao estresse por frio mostra-se extremamente complexa, visto que poderia estar relacionada a outros fatores que não ao estresse *per si*.

Tipicamente, plantas de origem tropical e subtropical são susceptíveis às injúrias causadas pelo frio, incluindo milho (*Zea mays*), arroz (*Oryza sativa*), tomate (*Lycopersicum esculentum*), algodão (*Gossipium hirsuntum*) e cana-de-açúcar (Taiz e Zeiger, 1998). Os efeitos causados pela exposição a baixas temperaturas em cana-de-açúcar (Fig.1) dependem de diversos fatores, tais como o tipo de estresse (*chilling* ou *freezing*), estádio de desenvolvimento vegetal, variedade, duração do estresse, entre outros (Tai e Lentini, 1998; Du et al., 1999). A exposição de sementes de cana a temperaturas positivas baixas (*chilling*) por cerca de três semanas resulta na redução do número, comprimento e peso das hastes, bem como da

produção de açúcar. Prolongada exposição da planta a temperaturas de congelamento (*freezing*) pode eliminar ramos novos ou mesmo aqueles que estão se desenvolvendo durante a rebrota (Tai e Lentini, 1998).

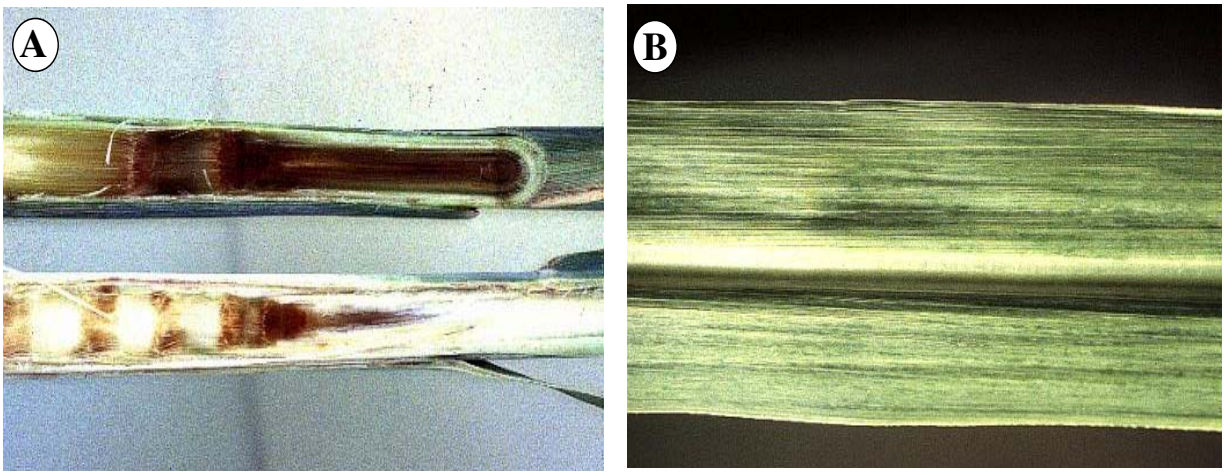


Figura 1. Danos causados aos tecidos da haste (**A**) e foliares (**B**) de cana-de-açúcar após exposição a temperaturas de congelamento (Fonte: Isakeit, T., TAEX, Weslaco, 1996).

Todavia, mesmo em espécies tropicais e subtropicais, a tolerância ao frio pode aumentar quando as plantas são previamente aclimatadas, ou seja, expostas por um período determinado a temperaturas baixas (na faixa de 2 °C a 15 °C), já que esse processo desencadeia mudanças na expressão de genes específicos, os quais promovem alterações metabólicas nas células (Guy, 1990; Thomashow, 1998). Entretanto, nestas espécies, os mecanismos moleculares de percepção, sinalização molecular e modificação do metabolismo em resposta ao estresse por frio são pouco estudados.

Desde 1988, genes induzidos por baixas temperaturas (coletivamente denominados aqui como *cold responsive* - *cors*) têm sido clonados de várias espécies vegetais, incluindo *Medicago falcata*, *Arabidopsis thaliana*, *Triticum aestivum*, *Solarium sogarandinum*, dentre outras (Mohapatra et al., 1988; Kurkela e Borg-Franck, 1990; Ouellet et al., 1993; Rorat et al., 1997). Os genes *cors* estão envolvidos em diferentes fases da aclimatação ao frio, as quais se estendem desde a percepção do estresse até a biossíntese de proteínas que participam diretamente dos mecanismos de tolerância (Pearce, 1999; Xiong et al., 2002). Baixas temperaturas causam alteração na fluidez da membrana plasmática (Murata e Los, 1997), podendo ser “percebida” por sensores primários que promovem a formação de uma cascata

primária de sinais intracelulares, que, por sua vez, resulta na modificação da expressão de genes específicos, os quais podem estar direta ou indiretamente relacionados com mecanismos de adaptação do metabolismo e/ou com a tolerância ao estresse abiótico. Finalmente, os sinais secundários (isto é, reguladores de crescimento e/ou moléculas sinalizadoras secundárias) podem iniciar outra (s) cascata (s) de eventos de sinalização, diferindo da sinalização primária temporal e espacialmente. Além disso, a sinalização secundária pode ser compartilhada por diferentes estímulos ambientais, gerando pontos de conexão com outros estresses abióticos (Xiong et al., 2002).

A percepção e sinalização em resposta ao estresse por frio estão resumidas na Figura 2. Um dos sensores primários bastante conhecido do estresse por frio são as proteínas formadoras dos canais de cálcio presentes na membrana plasmática (Knight, 2000). Durante o estresse, ocorre o influxo transiente de cálcio na forma iônica (Ca^{2+}) da parede celular para o citoplasma, o qual facilita a geração de moléculas sinalizadoras secundárias (Fig. 2; van der Luit et al., 1999). O envolvimento de Ca^{2+} na sinalização em resposta ao frio é bastante conhecido e foi inferido por meio de experimentos que utilizaram bloqueadores de canais de cálcio e agentes quelantes, os quais impedem o processo de aclimatação ao frio (Monroy et al., 1993). Moléculas sinalizadoras secundárias, tais como o inositol 1,4,5-trifosfato (*IP3*), induzem a liberação de Ca^{2+} de organelas intracelulares, levando à formação de uma nova cascata de sinais (Sanders et al., 1999; Fig. 2). EROS, por sua vez, atuam como sinalizadores secundários modificando a expressão gênica sem, entretanto, alterar a concentração intracelular de Ca^{2+} (Xiong et al., 2002; Fig. 2). Finalmente, o Ca^{2+} intracelular pode regular todas as etapas anteriores via retroalimentação (Fig. 2).

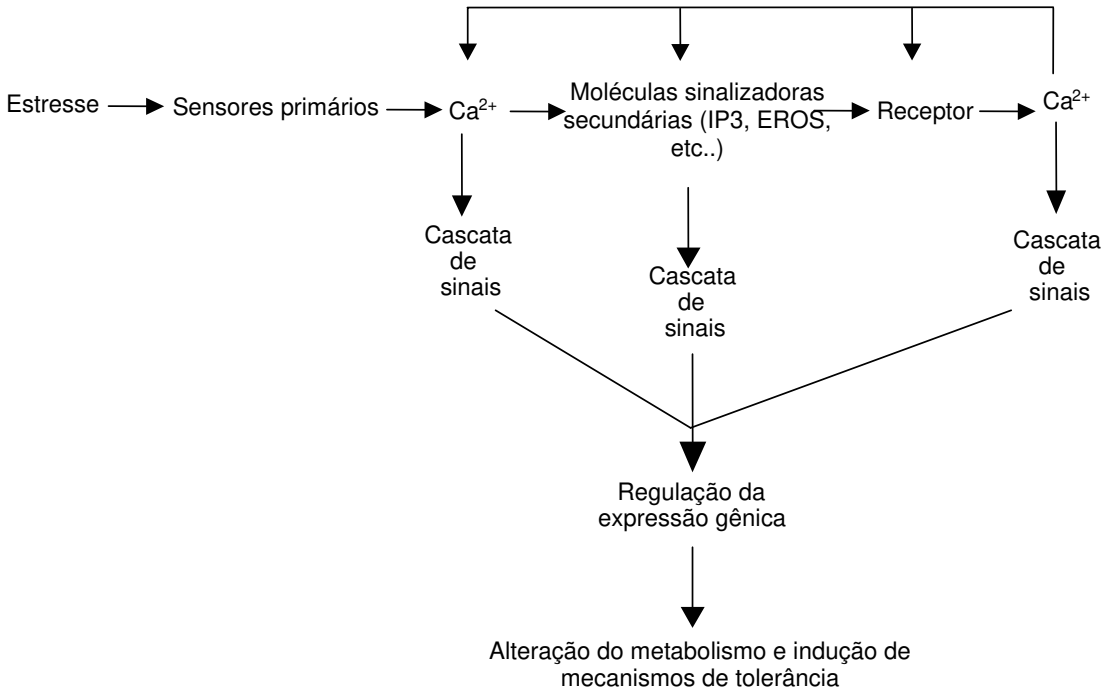


Figura 2. Esquema resumido da percepção do estresse por frio: sinalização primária e secundária (modificado de Xiong et al., 2002).

Aumentos transientes na concentração de Ca^{2+} citoplasmático induzem a expressão de genes tais como o *NpCaM-1*, que codifica para uma isoforma de calmodulina de *Nicotiana plumbaginifolia* (van der Luit et al., 1999), além de ativar proteínas do tipo *Ca^{2+}-dependent protein kinases* (CDPKs; Martin e Busconi, 2001). A superexpressão do gene *OsCDPK7* em plantas transgênicas de arroz resultou no aumento de tolerância ao estresse por baixas temperaturas (Saijo et al., 2000), sugerindo que essas proteínas desempenham um importante papel no desenvolvimento de mecanismos de tolerância a esse tipo de estresse. Além das CDPKs, outras fosfoproteínas são moduladas por estresse por frio, tais como as proteínas MAPKs. Em leveduras e animais, as MAPKs ou *mitogen-activated protein kinases* participam de vias metabólicas responsáveis pela produção de osmoprotetores, como o glicerol, e de compostos antioxidantes (Xiong et al., 2002). Apesar de ainda não comprovado, é possível que em plantas existam vias metabólicas similares, visto que alguns estresses osmóticos ativam diferentes MAPKs (Ichimura et al., 2000). Em alfafa (*M. falcata*), por exemplo, a expressão do gene *MMK4* foi altamente induzida por baixas temperaturas e déficit hídrico. Embora não se

tenha observado aumento na quantidade de proteína sintetizada, a atividade da proteína MMK4 foi aumentada em plantas submetidas a esses estresses (Jonak et al., 1996). Wen et al. (2002) analisaram a expressão de dois genes que codificam proteínas do tipo MAPK de arroz, *OsMEK1* e *OsMAP1*. Interessantemente, ambos os genes foram induzidos em resposta ao estresse por frio mais ameno (12 °C) e não por 4 °C, sugerindo que, em arroz, existem duas vias de sinalização em resposta a baixas temperaturas. Analogamente, Provat et al. (2003) identificaram diferenças qualitativas e quantitativas na expressão gênica em plantas de *Arabidopsis thaliana* expostas a 13 °C e 4 °C. Os autores sugerem que essas diferenças podem ocorrer devido ao fato de que o tratamento de 13 °C por 48 horas não foi suficiente para disparar os mecanismos de tolerância que ocorrem normalmente quando as plantas são expostas a 4 °C (temperatura comumente utilizada nos tratamentos de estresse por baixas temperaturas; Thomashow, 1999).

Além das fosfoproteínas citadas, a resposta ao estresse por baixas temperaturas pode ser sinalizada via proteínas do tipo RPK *(receptor protein kinase*). Essas proteínas são importantes componentes no processo de sinalização em resposta aos estresses ambientais (Viswanathan e Zhu, 2002). Em *A. thaliana*, a proteína RPK1 contém domínio extracelular com seqüências repetidas ricas no aminoácido leucina, além de um domínio *kinase* citoplasmático, sendo, então, um provável membro da família RPK (Hong et al., 1997). Além disso, o gene *RPK1* é rapidamente induzido em resposta ao frio, estresse salino e estresse hídrico (Hong et al., 1997).

Estresse por baixas temperaturas promove o acúmulo de EROS, tais como superóxidos, peróxido de hidrogênio e radicais hidroxilas (Sato et al., 2001). EROS podem funcionar tanto como moléculas sinalizadoras secundárias, induzindo mecanismos de tolerância, quanto como agentes causadores de injúrias e morte celular (Prasad et al., 1994). Enquanto as conecções entre as vias de transdução de sinais mediadas por EROS e estresses abióticos estão apenas começando a ser elucidadas (Xiong e Zhu, 2001), o envolvimento de espécies reativas de oxigênio em vias de sinalização em resposta ao ataque de patógenos é um processo bem documentado na literatura (Lamb e Dixon, 1997). Em *Arabidopsis thaliana*, a proteína AtMPK6, uma MAPK, ativada por estresse por frio e/ou estresse osmótico, é também regulada por estresse oxidativo (Yuasa et al., 2001), demonstrando que a transdução dos sinais moleculares em resposta ao estresse por baixas temperaturas pode ocorrer via sinalização por EROS.

Outras moléculas sinalizadoras secundárias, tais como o ácido abscísico (AAB) e o etileno, parecem desempenhar um importante papel na habilidade de espécies vegetais em tolerar estresses abióticos, incluindo o estresse por baixas temperaturas (Skriver & Mundy,

1990). Além desses reguladores de crescimento, alguns autores sugerem que o ácido salicílico e o ácido jasmônico também podem estar envolvidos em mecanismos de tolerância ao frio (Xiong et al., 2002); entretanto, o papel desses fitohormônios em resposta a esse tipo de estresse não foi ainda completamente elucidado.

O etileno tem sido associado ao aumento de tolerância a baixas temperaturas em plantas sensíveis, tais como feijão (*Phaseolus vulgaris* L.) e tomate (Field, 1984; Ciardi et al., 1997). Recentemente, foi relatado o envolvimento do etileno na regulação da atividade de proteínas anticongelantes em folhas de centeio (*Secale cereale*). Yu et al. (2001) demonstraram que plantas tratadas com *ethephon* (ácido 2-cloroetilfosfônico) ou com ACC (ácido 1-aminociclopropano-1-carboxílico), ambos precursores de etileno, acumularam proteínas anticongelantes, mesmo quando não expostas a temperaturas de congelamento (*freezing*). Desta forma, é possível que, em plantas submetidas ao estresse por baixas temperaturas, ocorra maior produção de etileno, que, por sua vez, funciona como sinalizador para alteração da expressão de genes responsáveis pelo processo de tolerância (Yu et al., 2001).

O AAB regula muitos genes induzidos por baixas temperaturas. Yamaguchi-Shimozaki e Shinozaki (1994) propõem que estresses por frio e por desidratação promovem a produção do fitorregulador, o qual induz a expressão de vários genes *cors*. Seqüências promotoras de alguns desses genes contêm o elemento *cis* ABRE (Stockinger et al., 1997; Shinozaki e Yamaguchi-Shimozaki, 2000). Vários fatores de transcrição do tipo *bZIP* (*ABF/AREB*), regulados por AAB, ligam-se ao elemento ABRE (PyACGTGGC), aumentando a expressão de genes que contêm esse motivo em seus promotores (Choi et al., 2000). Além disso, em *A. thaliana*, três genes que codificam proteínas do tipo *C2H2 zinc finger* (*AZF1*, *AZF3* e *STZ*) são regulados por baixas temperaturas e por AAB (Sakamoto et al., 2000). A superexpressão do gene *SCOF1* (coficador de proteína pertencente à família *C2H2 zinc finger*) em *A. thaliana* promoveu a expressão constitutiva de vários genes *cors*, aumentando a tolerância ao estresse por congelamento (Kim et al., 2001). Os autores também demonstraram que a proteína SCOF1 aumentou a atividade de ligação ao DNA da proteína SGBF1, a qual é um fator de transcrição do tipo *bZIP* induzido por frio e AAB (Hong et al., 1995). Xiong et al. (2001) identificaram mutantes de *Arabidopsis thaliana* (*los5*) que apresentaram deficiência na produção de AAB e reduzida tolerância ao estresse por frio. Análises genéticas comprovaram que o lócus *LOS5/ABA3* codifica proteína do tipo *molybdenum cofactor* (*Moco*) *sulfurase*, a qual está envolvida na última etapa da biossíntese desse fitorregulador em plantas (Xiong et al., 2001). A expressão do gene *LOS5/ABA3* é induzida em resposta ao estresse por baixas temperaturas, comprovando sua importância para o processo de aclimatação ao estresse. Todas as

informações supracitadas denunciam a importância do ácido abscísico na resposta e no processo de aclimatação de plantas a baixas temperaturas. Entretanto, através de estudos que utilizaram mutantes de *A. thaliana*, vários grupos têm demonstrado que alguns genes induzidos por baixas temperaturas são regulados por via (s) independente (s) de AAB (Thomashow, 1999; Shinozaki e Yamaguchi-Shimozaki, 2000). Gilmour e Thomashow (1991) observaram que os transcritos de alguns genes *cors* (*COR78*, *COR47* e *COR6.6*) se acumulavam em mutantes *aba1* (deficientes na produção de AAB) de *Arabidopsis thaliana*, quando induzidos por frio. Portanto, os autores sugerem que esses genes são induzidos por uma via independente de AAB, demonstrando a existência de diferentes rotas metabólicas induzidas por esse estresse ambiental.

De fato, seqüências promotoras de vários genes *cors* contêm, além do elemento *cis* ABRE, o elemento DRE/CRT (CCGAC; Thomashow, 1999). Fatores de transcrição do tipo *DREB* (*dehydration-responsive element binding*) ligam-se ao elemento DRE, induzindo a expressão de alguns genes *cors* (Thomashow, 1999). As proteínas *DREBs* contêm domínio de ligação ao DNA tipo *APETALA2/ethylene-responsive element binding* e são divididas em duas famílias: DREB1 (DREB1A ou CBF3, DREB1B ou CBF1 e DREB1C ou CBF2) e DREB2 (DREB2A e DREB2B) (Liu et al., 1998). A expressão de genes *DREB1* é especificamente induzida por frio, independentemente de AAB, enquanto que a expressão dos genes *DREB2* pode ser induzida por estresse salino e/ou estresse hídrico (Viswanathan e Zhu, 2002). Plantas de *A. thaliana* que superexpressam os genes *DREB1* adquiriram tolerância constitutiva ao *freezing*, demonstrando a importância desses fatores de transcrição para a aclimatação ao frio (Jaglo-Ottosen et al., 1998; Liu et al., 1998; Kasuga et al., 1999). Membros ortólogos da família DREB1 foram identificados em espécies vegetais sensíveis ao estresse por baixa temperatura, tais como o tomate (Jaglo et al., 2001). Além disso, Hsieh et al. (2002) obtiveram plantas de tomate transgênicas que superexpressam o gene *CBF1*, as quais apresentaram maior nível de tolerância ao estresse por frio. Além disso, os autores observaram que, em plantas transgênicas, ocorreram redução na produção de H_2O_2 e aumento da atividade da enzima catalase, quando comparado aos níveis obtidos com as plantas selvagens. Portanto, é possível que a expressão heteróloga do gene *DREB1B* tenha resultado na indução de genes envolvidos na tolerância ao estresse oxidativo, o qual está intimamente associado ao estresse por baixas temperaturas (Sato et al., 2001). Contudo, as plantas de tomate transgênicas não foram capazes de sobreviver a temperaturas abaixo de 0 °C (*freezing*), sugerindo a não existência em tomate de genes *cors* ortólogos que são especificamente envolvidos nos processos de tolerância a essa categoria de estresse (Hsieh et al., 2002).

Recentemente, novos elementos têm sido adicionados à via de resposta ao frio mediada pelos genes *DREBs* em *Arabidopsis thaliana*. Stockinger et al. (2001) demonstraram que a ativação transcricional dos genes *cors*, mediada por proteínas *DREBs*, depende de modificações na estrutura da cromatina, sendo estas realizadas por proteínas do complexo *histone acetyltransferase*. Além disso, Chinnusamy et al. (2003) identificaram o gene *ICE1* (*inducer of CBF expression 1*), que codifica uma proteína pertencente à família *MYC-like bHLH* de fatores de transcrição. Esse gene é expresso constitutivamente em condições favoráveis de crescimento e a proteína liga-se especificamente a seqüências *MYC*, presentes nos promotores dos genes *DREBs*. A superexpressão do gene *ICE1* aumentou a expressão dos genes *DREBs* e também a tolerância de plantas transgênicas de *A. thaliana* submetidas ao estresse por frio (Chinnusamy et al., 2003).

Além das proteínas que regulam positivamente as cascatas de sinais em resposta ao estresse por frio, outras podem regular negativamente, reduzindo a expressão de importantes genes *cors*. Um exemplo interessante é a proteína HOS1, que contém motivo do tipo *RING finger*. Plantas de *A. thaliana* que possuem mutação no lócus *hos1* exibem aumento da expressão de vários genes *cors*, incluindo os *DREBs* (Lee et al., 2001). Os autores sugerem que a proteína HOS1 atua como um importante regulador negativo na resposta a esse tipo de estresse. Contudo, o real papel dessa proteína não foi ainda elucidado. Como muitas proteínas do tipo *RING finger* controlam a degradação de fatores de transcrição (Osterlund et al., 2000), Lee et al. (2001) especulam que a proteína HOS1 poderia controlar o *turnover* do ativador transcricional *ICE1*, afetando toda a via de resposta ao frio mediada pelos genes *DREBs* em *A. thaliana*.

Os genes *cors* também participam de vários processos celulares visando à aclimatação da planta à condição de estresse por frio. Entretanto, muitos desses genes, além de participarem de diferentes processos fisiológicos durante o desenvolvimento vegetal, não são expressos exclusivamente em resposta ao frio (Pearce, 1999). Estudos têm demonstrado que os genes expressos em baixas temperaturas são induzidos também em resposta à desidratação causada por vários fatores, tais como: congelamento extracelular, seca, ou maturação de sementes (Dhinsa e Monroy, 1994; Yamaguchi-Shimozaki e Shinozaki, 1994). Vários desses genes codificam proteínas altamente hidrofílicas com propriedades similares às proteínas LEA (*late embryogenesis abundant*, Dhinsa e Monroy, 1994), OSM (osmotinas, Zhu et al., 1993) e à família de proteínas KIN (Rorat et al., 1997). Outro exemplo são as proteínas ngLTPs (*non-specific lipid transfer proteins*), que são reguladas por frio e seca (White et al., 1994; Plant et al., 1997). Essas proteínas parecem estar envolvidas na formação da cutícula e

essa estrutura, por sua vez, pode reduzir a perda de água via transpiração durante a exposição da planta ao estresse (Pearce, 1999).

Outros genes também podem responder ao estresse por frio, principalmente aqueles relacionados à redução da produção de espécies reativas de oxigênio, as quais são produzidas em grandes quantidades quando plantas são sujeitas a estresses bióticos e/ou abióticos (Xiong et al., 2002). Duas proteínas distintas, a oxidase alternativa (AOx) e a proteína desacopladora mitocondrial de plantas (PUMP), recentemente têm sido implicadas na redução da produção de EROS (Kowaltowski et al., 1998; Maxwell et al., 1999).

Durante os estresses ambientais, podem ocorrer alterações no sistema de transporte de elétrons mitocondrial, promovendo um aumento na produção de EROS. A proteína AOx, presente em plantas como um sistema alternativo de respiração, pode reduzir a produção de EROS via ubiquinol (Purvis, 1997). As proteínas AOxs são classificadas em duas subfamílias: AOx1 e AOx2, ambas codificadas por genes nucleares (Considine et al., 2002). Membros da família AOx1, presentes em di e em monocotiledôneas, são mais largamente estudados, sendo diferencialmente regulados em diversos órgãos e estádios de desenvolvimento, além de responderem a estresses bióticos e abióticos. AOx2, por outro lado, possui expressão gênica constitutiva e está presente somente em dicotiledôneas (Considine et al., 2002).

Recentemente, uma nova proteína foi descrita como participante da manutenção do fluxo de elétrons na mitocôndria, a proteína desacopladora (denominada UCP em outros organismos e PUMP em plantas) (Vercesi et al., 1995; Laloi et al., 1997). Esta proteína, identificada em diversos reinos (Watanabe e Hirai, 2002), parece ser responsável pela dissipação do gradiente de prótons do espaço intermembrana, sem haver, entretanto, a produção de ATP pelo complexo ATP sintase (Vercesi et al., 1995). PUMPs/UCPs contêm três “assinaturas” da superfamília de proteínas carreadoras mitocondriais e seis domínios transmembrana (Borecký et al., 2001; Watanabe e Hirai, 2002). Vários genes foram caracterizados em diferentes espécies vegetais (Laloi et al., 1997; Maia et al., 1998; Nantes et al., 1999; Watanabe e Hirai, 2002), sem, contudo, a real função desta proteína em plantas ter sido elucidada. No entanto, estudos independentes têm demonstrado que, além da expressão gênica ser induzida por baixas temperaturas, PUMPs possuem a capacidade de reduzir a produção de EROS durante estresses ambientais (Vercesi et al., 1995; Maia et al., 1998; Kowaltowski et al., 1998; Pastore et al., 2000).

Apesar do avanço nos estudos de identificação e caracterização gênica, muitos genes *cors* possuem funções ainda desconhecidas. HVA22 é uma pequena família multigênica em *Arabidopsis thaliana*, na qual alguns de seus membros respondem ao estresse por frio e

também a outros estresses, tais como o estresse hídrico e estresse salino (Chen et al., 2002). Apesar de um homólogo aos genes *AtHVA22* em levedura (*Saccharomyces cerevisiae*) estar envolvido no sistema de transporte de vesículas intracelulares, em plantas essa função não foi ainda demonstrada (Chen et al., 2002). Finalmente, novos genes que respondem ao estresse por frio foram identificados por meio da análise de expressão gênica em larga escala, em diferentes espécies vegetais (Seki et al., 2002; Rabbani et al., 2003). Dentre eles, destacam-se as proteínas específicas de plantas que possuem um domínio altamente conservado, definido como NAC. Este domínio foi nomeado com base nas primeiras proteínas identificadas em *A. thaliana*: NAM, ATAF1 e 2 e CUC2 (Aida et al., 1997). Membros dessa superfamília em *Arabidopsis thaliana* foram identificados e caracterizados como pertencentes a uma nova família de fatores de transcrição (Ooka et al., 2003). Por exemplo, o gene *NAC1*, está envolvido na formação de raízes laterais por meio da regulação do gene *TIR1*(Xie et al., 2000). Outros genes foram também identificados em várias espécies vegetais (Souer et al., 1996, Ruiz-Medrano et al., 1999, Xie et al., 1999, KiKuchi et al., 2000). Collinge e Boller (2001) analisaram a expressão diferencial de uma proteína NAC sob diferentes estresses, neste caso em resposta à infecção por *Phytophthora infestans* e em resposta a ferimentos no tecido foliar de batata (*Solanum commersonii*). Os autores observaram que os genes de batata e *Arabidopsis thaliana*, pertencentes à subfamília ATAF (KiKuchi et al., 2000), responderam semelhantemente ao estresse por ferimento, demonstrando que os membros dessa subfamília podem compartilhar um conservado papel na resposta a estresses.

A maioria das pesquisas tem focado no estudo de genes com expressão induzida quando plantas estão expostas ao frio. Porém, muitos genes também apresentam expressão reprimida em resposta a baixas temperaturas. Fowler e Thomashow (2002) identificaram vários genes reprimidos por esse estresse, incluindo aqueles envolvidos em vias de regulação da transcrição gênica, sinalização molecular, fotossíntese e defesa contra patógenos. Os autores sugerem que tanto a indução quanto a repressão da expressão gênica são mecanismos moleculares essenciais para o processo de aclimatação.

Apesar de um grande número de genes envolvidos na resposta ao estresse por frio ter sido identificado em *A. thaliana* e em outras espécies de clima temperado, poucos trabalhos avaliaram a expressão gênica de espécies tropicais e subtropicais. Conforme observado por Rabbani et al. (2003), muitos genes *cors* de arroz induzidos ou reprimidos por frio foram diferentes daqueles em *A. thaliana*. Portanto, os conhecimentos gerados pela identificação e caracterização de novos genes advindos de outras espécies vegetais não somente são importantes para o total entendimento de como as plantas percebem as variações ambientais,

mas também possuem importantes aplicações práticas, permitindo o desenvolvimento de novas estratégias para aumentar a produtividade de espécies vegetais de importância agronômica, inclusive plantas de climas tropical e subtropical, como a cana-de-açúcar.

3. Genômica funcional: projetos EST, SAGE e arranjos de DNA

Com o advento da era genômica, a identificação de genes tornou-se um processo mais dinâmico, capaz de gerar um vasto volume de informação em um curto período de tempo (Grivet e Arruda, 2001). Vários projetos cujo objetivo é estudar o transcriptoma, ou seja, a população de RNAs transcrita de um determinado organismo, vêm sendo conduzidos em diferentes espécies vegetais (Ewing et al., 1999; White et al., 2000; Dong et al., 2003; Ma et al., 2003), denominados projetos EST (*Expressed Sequence Tag* ou Etiqueta de Seqüências Expressas). Esses projetos são uma poderosa ferramenta para identificar genes expressos em determinados tecidos e/ou tipos celulares de interesse. Nos projetos EST, bancos de dados contendo pequenas seqüências de DNA são gerados a partir do seqüenciamento de moléculas de cDNA, sintetizadas das populações de mRNA (RNA mensageiro), empregando iniciadores específicos do vetor (plasmídio), o qual foi utilizado no processo de clonagem gênica. Essas seqüências são usadas no procedimento de montagem dos *contigs* ou *clusters* que, na maioria das vezes, possuem *ORFs* (*open reading frames*) representando a região codificadora de diversos genes (Telles et al., 2001). Desta forma, a tradução destas *ORFs* fornece os primeiros indícios da função da proteína codificada por um determinado clone de cDNA. ESTs produzem informações biológicas de centenas de genes de um organismo, além de permitirem a identificação de diferentes isoformas de transcritos (Andrews et al., 2000) e o mapeamento gênico (Schuler, 1997; Wu et al., 2002). Outro aspecto importante dos ESTs é o acesso a informações sobre os genes expressos em organismos que contêm um genoma complexo (Vettore et al., 2001), tais como a cana-de-açúcar. No Brasil, concluiu-se em outubro de 2000 o projeto EST da cana-de-açúcar ("SUCEST - *Sugarcane EST Project*"), cujo objetivo foi gerar um banco de dados contendo cerca de 250.000 ESTs, obtidos de diferentes tecidos/orgãos vegetais (Vettore et al., 2003).

A grande quantidade de ESTs seqüenciados a partir de diferentes bibliotecas de cDNA pode fornecer uma estimativa da abundância relativa de transcritos de genes de interesse em diferentes tecidos/órgãos vegetais e também em diversas condições biológicas (Audic e Claverie, 1997). Esse "*northern* digital", aliado a metodologias experimentais, possibilita a

identificação e análise de novos genes, os quais podem ser utilizados em programas de melhoramento genético via biotecnologia.

Outra metodologia utilizada para avaliar a expressão gênica via seqüenciamento é a SAGE (*serial analysis of gene expression*). O método baseia-se no isolamento de pequenas seqüências de cDNA (denominadas *tags*), obtidas a partir de moléculas individuais de mRNAs, e sua posterior concatenação em longas moléculas de DNA. Essas longas moléculas de DNA são clonadas e seqüenciadas, permitindo assim a análise qualitativa e quantitativa dos transcritos celulares (Velculescu et al., 1995). Recentemente, Lee e Lee (2003) utilizaram essa metodologia para avaliar a expressão gênica em grãos de pólen de *Arabidopsis thaliana* submetidos ao estresse por frio.

Juntamente com os projetos EST, surgiram novas técnicas capazes de avaliar a expressão de milhares de genes simultaneamente. Dois exemplos dessas técnicas são o *DNA micro* e *macroarray* (micro e macroarranjos de DNA). Esses arranjos de DNA podem ser gerados utilizando-se cDNAs inteiros, ESTs, produtos de PCR ou mesmo oligonucleotídeos; os microarranjos são fixados a superfícies de lâminas de vidro ou *chips* de silício, enquanto os macroarranjos são fixados a membranas de náilon (Schena et al., 1995; Schena et al., 1998; Desprez et al., 1998; Schummer et al., 1999; Freeman et al., 2000).

Os microarranjos de DNA são largamente utilizados por grupos de pesquisa no mundo inteiro. Nessa metodologia, as seqüências de DNA são hibridadas simultaneamente a sondas de cDNA marcadas com dois fluoróforos (*Cy-3* e *Cy-5*), permitindo a análise comparativa direta da expressão gênica (Seki et al., 2002). Como a resposta aos estresses ambientais envolve a regulação de um grande número de genes (Seki et al., 2002), vários grupos têm utilizado a metodologia de microarranjos de DNA para avaliar o perfil de expressão gênica em espécies vegetais (Reymond et al., 1999; Kawasaki et al., 2001; Fowler e Thomashow, 2002; Seki et al., 2002; Provart et al., 2003; Rabbani et al., 2003).

Os macroarranjos de DNA são uma alternativa para os *microarrays*. Sua principal vantagem é que essa técnica não necessita de equipamentos especiais para hibridação e podem ser construídos tanto automatica como manualmente. Uma desvantagem é não permitir uma alta densidade de genes, quando comparados aos microarranjos de DNA (Freeman et al., 2000). Vários relatos na literatura demonstraram o uso dessa tecnologia com sucesso (Zhao et al., 1995; Desprez et al., 1998; Schummer et al., 1999; Carson et al., 2002).

Desta forma, no presente trabalho de doutoramento, empregaram-se a técnica de macroarranjos de DNA e a mineração do banco de dados de ESTs com o objetivo de avaliar a expressão e identificar novos genes em tecidos de cana-de-açúcar, os quais são

diferencialmente regulados em condições de baixas temperaturas. Quatro artigos científicos, dois publicados, um em fase de submissão e outro em fase final de preparação, descrevem os estudos e as análises desenvolvidas durante o projeto. O primeiro artigo (capítulo I) trata de uma revisão sobre a técnica de *macroarrays* e alguns exemplos de sua aplicação. O segundo (capítulo II) discorre sobre a análise do perfil de expressão gênica em tecido foliar de cana-de-açúcar sob estresse por baixas temperaturas. O terceiro artigo (capítulo III) trata da identificação e análise do perfil de expressão gênica, em diferentes tecidos e sob estresse por frio, de proteínas da família PUMP e da família AOx, em cana-de-açúcar e em *A. thaliana*. Finalmente, o quarto artigo (capítulo IV) descreve a caracterização de um fator de transcrição identificado em nossos *macroarrays*, o qual pode ser um componente importante para a via de sinalização molecular em resposta a estresses.

4. Objetivos

Dada a relevância do entendimento da resposta molecular ao estresse por baixa temperatura em plantas de origem tropical/subtropical, essa tese de doutoramento foi realizada considerando-se os seguintes objetivos:

- Avaliar o perfil de expressão gênica no tecido foliar de plântulas de cana-de-açúcar submetidas ao estresse abiótico, utilizando a técnica de macroarranjos de DNA;

- Caracterizar alguns genes de cana-de-açúcar induzidos por estresse de baixa temperatura e especular seu possível papel em resposta a essa condição ambiental.

Capítulo I

Genoma funcional: uso de arranjos de DNA em náilon para a análise de expressão gênica em larga escala

Juliana M. Felix, Rodrigo D. Drummond**, Fábio T.S. Nogueira**, Vicente E. de Rosa Júnior, Renato A. Jorge, Paulo Arruda, Marcelo Menossi

Biotecnologia Ciência & Desenvolvimento 24, 60-67 (2002).

Pesquisa

GENOMA FUNCIONAL

Uso de arranjos de DNA em náilon para a análise da expressão gênica em larga escala

Juliana de Maria Felix
Mestranda em Biologia Celular e Estrutural (UNICAMP)
felix@unicamp.br

Rodrigo Duarte Drummond
Mestrando em Genética e Biologia Molecular (UNICAMP)
rduarte@unicamp.br

Fábio Tebaldi Silveira Nogueira
Doutorando em Genética e Biologia Molecular (UNICAMP)
tebaldi@unicamp.br

Vicente Eugenio de Rosa Junior
Doutorando em Genética e Biologia Molecular (UNICAMP)
vicentej@unicamp.br

Renato Atílio Jorge
Professor Depto. de Físico-Química
Instituto de Química - (UNICAMP)
rjorge@iqm.unicamp.br

Paulo Arruda
Professor Doutor – Depto. de Genética e Evolução
Pesquisador do Centro de Biologia Molecular e Engenharia Genética (CBMEG-UNICAMP)
parruda@unicamp.br

Marcelo Menossi
Professor Doutor Depto. Genética e Evolução
Pesquisador do Centro de Biologia Molecular e Engenharia Genética (CBMEG-UNICAMP)
menossi@unicamp.br

Os organismos respondem a diferentes estímulos ambientais por meio de alterações no proteoma, conjunto de proteínas das células. Esse processo pode permitir a adaptação a condições adversas, como, por exemplo, choque térmico, exposição a drogas, ataque de patógenos, etc. O fluxo da informação gênica do DNA nos cromossomos (genoma) até o proteoma, é intermediado pelo conjunto das moléculas de RNA (transcritoma). Assim, a concentração relativa dos transcritos de um determinado gene em uma célula é um indicativo do quanto esse gene está sendo expresso, isto é, do quanto a célula está investindo do seu maquinário bioquímico para produzir a proteína codificada pelo gene.

Em vista disso, foram desenvolvidas metodologias visando a medir a concentração relativa dos transcritos dos genes em células e tecidos. Até pouco tempo atrás a análise da expressão gênica era feita com metodologias que avaliavam poucos genes de cada vez: Northern blot, dot blots, RT-PCR, entre outros. O crescimento exponencial do número de novos genes descobertos nos projetos genoma e o desenvolvimento dos arranjos de DNA propiciaram uma nova abordagem nos estudos da regulação gênica, tornando possível o monitoramento dos níveis de transcritos de um grande número de genes simultaneamente. O presente artigo enfoca a tecnologia dos arranjos de DNA em náilon, que empregam medologias rotineiras, robustas e de fácil implementação na maioria dos laboratórios de biologia molecular, sem a necessidade de investimentos vultosos e longos períodos de padronização.

Os arranjos de DNA são suportes sólidos, comumente vidro ou náilon, aos quais estão fixadas, de forma ordenada, seqüências completas ou parciais de genes. A Figura 1 mostra as etapas de uma análise de expressão gênica em larga escala empregando arranjos em náilon. O DNA dos genes é depositado em membranas réplicas, contendo os mesmos genes nas mesmas posições, em densidades que variam de 10 a 1000 genes/cm^2. A partir do RNA das células em estudo são produzidas sondas de cDNA via transcrição reversa na presença de um nucleotídeo radioativo, que permite sua detecção posterior. As sondas são hibridadas contra os arranjos e quanto maior a expressão de um gene em uma determinada condição, maior será o número de moléculas de RNAm, sendo maior o número de cDNAs desse gene na sonda sintetizada. Conseqüentemente, maior será o valor de intensidade do sinal derivado da sonda hibridada na região do arranjo que contém a seqüência desse gene.

Caso o exemplo esquematizado na Figura 1 fosse feito com arranjos em vidro, o procedimento seria muito similar. No entanto, uma das duas sondas seria sintetizada na presença de outro nucleotídeo modificado, não radioativo, mas fluorescente. A outra sonda seria sintetizada com um segundo nucleotídeo que fluoresce num comprimento de onda distinto do utilizado na primeira sonda. Dessa forma, ambas as sondas poderiam ser hibridadas num mesmo arranjo, o que é uma clara vantagem. A detecção da intensidade das duas sondas hibridadas no arranjo é

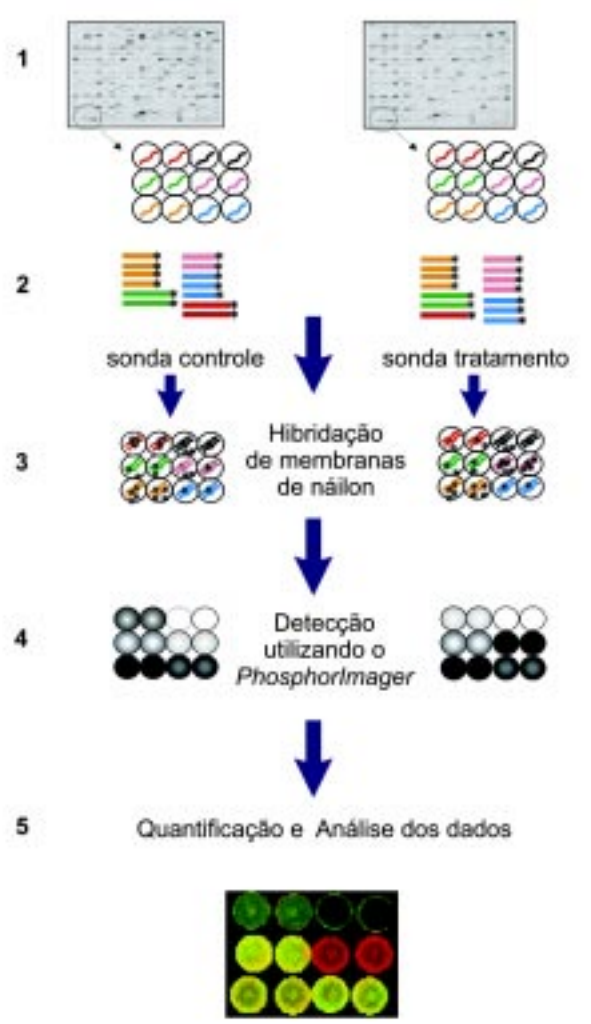


Figura 1:
Representação esquemática de um experimento utilizando macroarranjos
(1). O DNA plasmidial contendo os ESTs do projeto SUCEST é fixado nas membranas de náilon, em duplicata (cada cor representa um gene distinto); (2). A partir das amostras de RNA extraídas de diferentes tecidos ou tratamentos, é feita uma transcrição reversa em presença de α33-dCTP, produzindo as sondas de cDNA; (3). As membranas são hibridadas com as sondas de cDNA radioativas; (4). A radioatividade emitida de cada spot é detectada utilizando-se um phosphorimager ; (5). Os dados são comparados utilizando-se programas específicos para identificação do perfil de expressão de cada gene

realizada com um equipamento especial. A densidade de genes em arranjos de vidro pode chegar a ser uma ordem de magnitude maior. Para uma revisão sobre os arranjos em vidro, veja Freeman et al., 2000. Apesar dessas vantagens, a hibridação de arranjos em vidro requer investimentos muito maiores que os necessários para arranjos em náilon. Além disso, é um consenso que a padronização dessa metodologia é muito mais complexa.

Convém ressaltar que existem diversos termos empregados para descrever os arranjos de DNA: glass arrays, DNA chips, biochips e microarray por um lado, os quais, geralmente, refletem arranjos em vidro, e nylon array, filter arrays, high density membranes e macroarray, que se referem a arranjos em náilon. Neste artigo utilizaremos os termos arranjos em náilon e arranjos em vidro, uma vez que o suporte empregado é determinante para a seleção das metodologias de deposição do DNA no arranjo, síntese de sonda, hibridação e detecção dos níveis de expressão gênica.

Os arranjos em náilon têm sido empregados com sucesso desde o trabalho de Lennon & Lehrach (Lennon & Lehrach, 1991), os quais identificaram genes diferencialmente expressos em diferentes tratamentos. A partir daí, vários estudos vêm sendo realizados empregando-se essa metodologia, como, por exemplo, a análise da expressão gênica em 2.505 genes no cérebro humano (Zhao et al., 1995), a identificação de genes expressos diferencialmente em função da presença ou ausência de luz em Arabidopsis (Desprez et al., 1998), a análise de 21.500 genes de ovário para identificação de genes superexpressos em carcinomas (Schummer et al., 1999), o monitoramento do perfil de expressão gênica em raízes de tomateiro durante suplementação com nitrogênio (Wang et al., 2001), entre outros. O primeiro relato na literatura sobre o uso de arranjos em vidro foi publicado pelo grupo de Patrick Brown, da universidade de Stanford (Wenzl et al., 2001). A partir daí, vários estudos foram feitos utilizando-se essa metodologia (para mais detalhes veja revisão de Passos et al., 1999 e de Freeman et al., 2000).

O nosso grupo tem estudado o perfil de expressão gênica da cana-de-açúcar em resposta a estresses abióticos empregando macroarranjos (macroarrays), que são arranjos em náilon com média densidade de genes (ao redor de 20 genes/cm^2). Os genes utilizados são provenientes de bibliotecas de cDNA do projeto genoma da cana-de-açúcar, SUCEST (http://sucest.lad.ic.unicamp.br).

Análises genômicas empregando macroarranjos

Os arranjos em náilon são uma opção interessante aos arranjos em vidro, tanto pelo custo como pela facilidade de implementação. A seguir estão descritas as diversas etapas que realizamos em nosso laboratório para implementação dessa metodologia. Para maiores detalhes, veja link na Tabela 1.

1. Confecção dos macroarranjos

As amostras de DNA são preparadas a partir de bactérias contendo os plasmídeos de bibliotecas de EST, empregando lise alcalina em placas de 96 poços, prática comum nos laboratórios de seqüenciamento em larga escala. O DNA é fixado às membranas manualmente, com o auxílio de um replicador composto por 96 pinos (V&P Scientific, EUA, http://www.vp-scientific.com), que possibilita a transferência simultânea de ~0,1µL de DNA de cada posição de uma placa de 96 poços, equivalente a 10 ng de DNA em cada spot.

Com esse procedimento, o DNA de dezesseis placas (1.536 amostras de DNA) pode ser transferido para uma membrana de 12 x 8 cm (Fig. 1.1). Normalmente cada EST é depositado em duplicata, para maior confiabilidade da análise, com o qual os macroar-

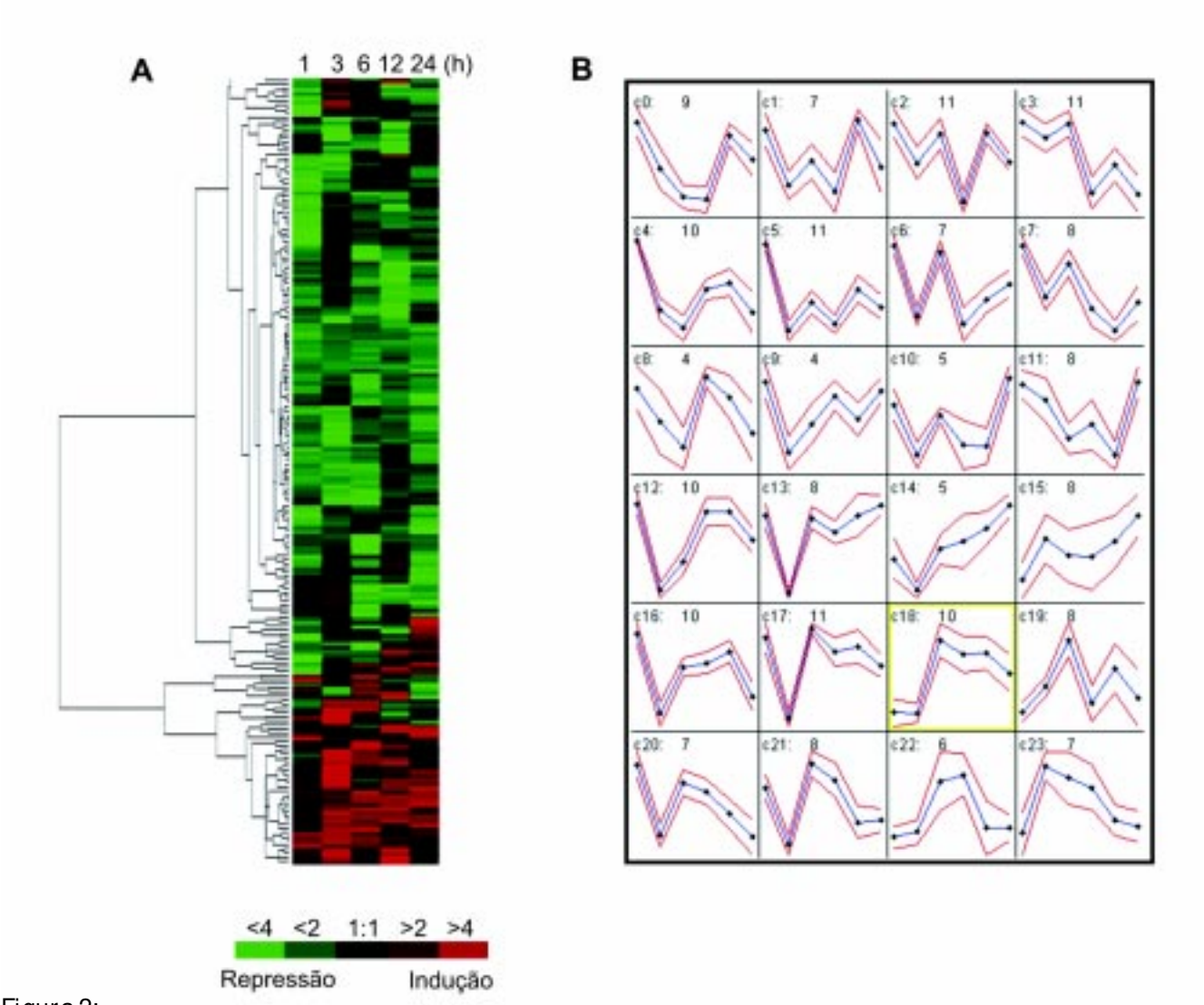


Figura 2:
Formas de visualização dos dados de macroarranjos
(A). Dendrograma mostrando os padrões de expressão ao longo do tempo em resposta a metil jasmonato. Cada linha representa um gene e as colunas representam os tratamentos. A expressão de cada gene é calculada a partir da razão entre o sinal no tratamento e no controle, sendo representada com diferentes cores, de acordo com a referência mostrada na parte inferior da figura. (B). Agrupamentos obtidos com o algorítmo de Self Organizing Maps (SOM). Os genes foram agrupados em 24 conjuntos (c0 a c23) e o número de genes em cada grupo está indicado à direita do nome do cluster. A linha central de cada gráfico representa a média do sinal de expressão dos genes do grupo nos seis pontos experimentais (controle, 1, 3, 6, 12 e 24h). As linhas vermelhas representam o desvio padrão no grupo para cada valor de expressão

ranjos produzidos com o replicador manual contém 768 ESTs (aproximadamente 8 ESTs/cm^2). A confecção de membranas de maior densidade pode ser conseguida com o auxílio de robôs, podendo-se obter uma densidade de genes dez vezes maior. O BCCCenter (http://www.bcccenter.fcav.unesp.br), que gerencia os clones do projeto SUCEST, tem previsão de disponibilizar para a comunidade científica membranas de 22,2 x 22,2 cm contendo 27.648 ESTs em duplicata.

2. Síntese de sonda e hibridação das membranas de náilon

As sondas são produzidas a partir de RNA total através de reação de transcrição reversa, na presença α33-dCTP (Fig. 1.2). Os cDNAs marcados radioativamente são hibridados contra as membranas de alta densidade (Fig. 1.3) em fornos de hibridação comuns, empregando um protocolo muito similar ao de Southern blot, técnica robusta e de uso corrente na maioria dos laboratórios de biologia molecular.

3. Obtenção das imagens

Após as lavagens para eliminar a hibridação inespecífica, a membrana de náilon é posta em contato com uma placa composta por material sensível à radioatividade (imaging plate). Após 96 h, a placa é "lida" em um aparelho do tipo Phosphorimager que a armazena na forma de uma imagem digitalizada (Fig. 1.4). Esse aparelho é o

Tabela1. Relação de programas utilizados para análise de perfil de expressão gênica

Softwares para Banco de Dados	
ARGUS	Brigham and Women's Hospital and Harvard Medical School (http://vessels.bwh.harvard.edu/software/argus/default.htm)
GeneX	Banco de dados aberto para expressão gênica (http://genex.ncgr.org/)
MAExplorer	Padrões de DataMining e Expressão Gênica (http://www.lecb.ncifcrf.gov/MAExplorer/)
MAXDSQL	Manchester University, baseado em requerimentos do MIAME-EBI (http://www.bioinf.man.ac.uk/microarray/maxd/maxdSQL/)
MGED	Grupo de Banco de dados de expressão gênica em Microarrays (http://www.mged.org/)

Exemplos de Banco de Dados para Expressão Gênica	
Microarray Centre	Centro de Microarray, The Ontario Cancer Institute (http://www.oci.utoronto.ca/services/microarray/index.html)
NIH	Banco de dados de expressão gênica, Molecular Pharmacology of Cancer (http://discover.nci.nih.gov/nature2000/naturemain.html)
SMD	Banco de Dados de Microarrays, Stanford University (http://genome-www4.stanford.edu/MicroArray/SMD/index.html)
yMGV	Visão global sobre Microarray de levedura (http://transcriptome.ens.fr/ymgv/)

Softwares para Análises	
ArrayViewer	Visualização e análise de dados de Microarrays – gratuito (http://www.tigr.org/softlab/)
BRB ArrayTools	Pacote integrado para visualização e análise estatística de dados de expressão gênica em Microarrays; Richard Simon & Amy Peng, National Cancer Institute, USA – gratuito (http://linus.nci.nih.gov/BRB-ArrayTools.html)
Cluster	Software de clusterização; Michael Eisen, Eisen Lab, University of California, Berkeley - gratuito (http://rana.lbl.gov/EisenSoftware.htm)
Expresssion Profiler	Ferramenta para clusterização, análise e visualização de expressão gênica. Análises on line – gratuito (http://ep.ebi.ac.uk/)
Multi Experiment Viewer	Aplicação Java que permite análise de dados de Microarrays para identificar padrões de expressão e genes expressos diferencialmente – gratuito. (http://www.tigr.org/softlab/)
SAM	Análise de Significância de Microarrays – gratuito (http://www-stat.stanford.edu/~tibs/SAM/)
SOM	Aplicação de Self Organizing Maps; Gavin Sherlock, Stanford University – gratuito (http://genome-ww.stanford.edu/~sherlock/SOMviewer.html)
Spotfinder	Detecta e quantifica spots – gratuito (http://www.tigr.org/softlab/)
Treeview	Visualização e clusterização de dados provindos do software Cluster; Michael Eisen, Eisen Lab, University of California, Berkeley - gratuito (http://rana.lbl.gov/EisenSoftware.htm)
X-Cluster	Software de clusterização; Gavin Sherlock, Stanford University – gratuito (http://genome-www.stanford.edu/~sherlock/SOMviewer.html)

Protocolos para arranjos de DNA	
LGF	Home page do laboratório de Genoma Funcional, CBMEG, UNICAMP, na qual estão disponibilizados protocolos completos para arranjos de DNA (http://cafe.cbmeg.unicamp.br)

equipamento mais custoso da metodologia (preço ao redor de U$ 40.000). A imagem é então quantificada utilizando-se programas específicos (veja Tabela 1) que geram tabelas de dados nas quais cada spot (região da membrana contendo um gene) recebe um valor numérico de acordo com a sua intensidade na imagem. A inspeção visual em filmes de raio-X também é possível, porém indicada apenas em ensaios qualitativos, uma vez que a quantificação das intensidades dos spots fica muito prejudicada.

4. Análise de dados

Cada gene está representado por dois spots no macroarranjo e a média desses dois dados é considerada a expressão do gene no macroarranjo, na condição experimental testada. Com a tecnologia disponível hoje para os arranjos em náilon e vidro, em boas condições experimentais, geralmente toma-se como genes induzidos aqueles que apresentaram um sinal pelo menos duas vezes maior no tratamento comparado ao controle, enquanto os reprimidos possuem uma intensidade de sinal no tratamento menor que a metade do controle. No entanto, essa premissa pode não ser verdadeira, sendo necessário avaliar a variabilidade dos dados obtidos nos diversos tratamentos.

As metodologias para análise estatística dos dados de arranjos de DNA estão em franco desenvolvimento,

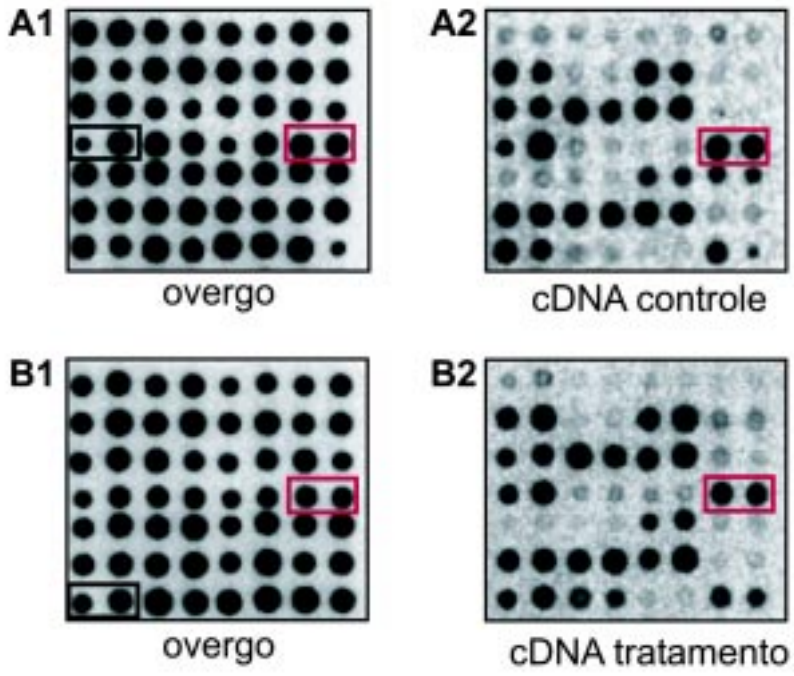


Figura 3:
Variações no sinal detectado devido a diferentes quantidades de DNA fixadas no spot
Macroarranjos réplicas contendo genes em duplicata foram hibridados com a sonda overgo (A1 e B1) e, em seguida, com sondas de cDNA de células do controle (A2) e de um tratamento (B2). Em A1 e B1, os retângulos em negro mostram spots réplicas com diferentes massas de DNA depositadas na mesma membrana, enquanto que os retângulos em vermelho mostram spots com variações entre membranas. Em A2 e B2, os retângulos em vermelho mostram os sinais das duas sondas de cDNA. Observa-se uma menor intensidade na sonda do tratamento. Essa aparente repressão do gene está relacionada com a diferença de massa nos spots

existindo diversas estratégias que atribuem níveis de confiança às diferenças observadas entre tratamentos. Em tais análises, é necessário obter o logaritmo dos valores de expressão, já que esses dados não apresentam distribuição normal.

O volume de dados gerados em nossos experimentos, com aproximadamente, 3.000 genes e 5 a 6 tratamentos, ainda é passível de ser analisado com planilhas de cálculo, como o MS Excel. Nessas análises, calculamos os logaritmos das razões entre a expressão de cada gene em um tratamento e a expressão observada no controle. Como é esperado que a maioria dos genes não altere sua expressão, a média dessas razões deve ser próxima a 1 (com o que a média dos logaritmos das razões deve ser próxima a zero). O indicador da variabilidade experimental é o desvio padrão dos logaritmos das razões de todos os genes de um macroarranjo. Em cada experimento, consideramos induzidos os genes cujo logaritmo da razão seja superior à média de todos os genes somada a 1,65 vezes o desvio-padrão (o inverso vale para os genes reprimidos). Esse fator multiplicador do desvio é escolhido com base na suposição de que os valores de logaritmos apresentam distribuição normal, de modo que a chance de um gene aparecer como falso positivo é menor que 5%. Por fim, selecionamos aqueles genes que foram induzidos (ou reprimidos) em duas repetições do experimento. Com esse critério, a chance de um gene ser erroneamente identificado como sendo expresso diferencialmente em função do tratamento é de $(5\%)^2$, o que garante um nível de significância de mais de 99% nos resultados.

Uma vez selecionados os genes que variam a expressão, são feitas outras análises que envolvem programas baseados em algoritmos de agrupamento (tabela 1). Esses programas identificam grupos de genes com um padrão de expressão semelhante nos tratamentos avaliados, permitindo a visualização dos dados através de dendrogramas e gráficos. A Figura 2 mostra duas formas de agrupamento e visualização de dados de expressão gênica em plântulas de Arabidopsis thaliana expostas ao hormônio metil jasmonato ao longo de um período de 24 h (dados calculados a partir dos resultados de Sasaki et al., 2001). Através dessa análise, podemos identificar genes ligados a um mesmo processo fisiológico ou por via metabólica. Assim, é possível inferir o papel bioquímico de genes desconhecidos, caso existam no seu grupamento genes cuja função já tenha sido descrita.

Uma outra forma de representar os genes induzidos e reprimidos pode ser obtida sobrepondo-se duas imagens digitalizadas (tratamento e controle), e gerar, assim, uma única imagem falsamente colorida (Fig. 1.5). Nessas imagens, similares às imagens oriundas dos experimentos de arranjos em vidro, que utilizam dois fluoróforos, os spots em vermelho representam genes induzidos, os spots em verde representam genes reprimidos, os spots em amarelo representam genes que não alteraram sua expressão em resposta ao tratamento, enquanto que os spots em negro representam genes com níveis de expressão próximos a zero.

Banco de dados: O destino dos dados de expressão gênica

O grande volume de informação gerado pelos projetos de análise de transcriptomas tem tornado cada vez mais complexo o armazenamento e a análise dos dados. Para contornar tal dificuldade, devem ser implementados, bancos de dados, que disponibilizem, de modo confiável, os dados e ferramentas de análise. Em muitos casos, esses bancos são abertos, o que aumenta ainda mais a aplicabilidade da pesquisa.

Alguns programas de bancos de dados estão disponíveis gratuitamen-

te (Tabela 1). No momento, nossa equipe está-se empenhando em implementar o banco GeneX (Mangalam et al., 2001). O GeneX é um banco de dados com código fonte aberto, que possibilita sua adaptação às necessidades específicas de cada grupo de pesquisa, além de possuir uma interface gráfica tipo web que pode ser acessada remotamente, facilitando a utilização de ferramentas de análises e a interação entre diversos grupos.

Padronização da metodologia

Diversos experimentos de padronização foram realizados visando à implementação da técnica de macroarranjos de DNA em nosso laboratório. Vários fatores podem afetar a sensibilidade, a reprodutibilidade e a confiabilidade da técnica, destacando-se: as variações na massa de DNA presente nos spots fixados na membrana; as flutuações que ocorrem durante as etapas de transcrição da sonda, hibridação e lavagem das membranas; além das variações no background (Schuchhardt et al., 2000; Schummer et al., 1999). Uma vez que esses fatores dificultam a implementação da técnica, nós detalharemos a seguir alguns dos procedimentos adotados em nosso grupo:

1. Variação na quantidade de DNA depositada nas membranas

O sinal obtido nas hibridações de ácidos nucléicos é proporcional não só à massa de transcrito presente na sonda, mas também à massa de DNA depositada na membrana. Tal como foi mostrado na Figura 1, para comparar diferentes tratamentos, é necessário empregar macroarranjos réplicas. Assim, existe o risco de considerar variações de sinal, que refletem apenas as diferentes massas de DNA nos spots nas membranas réplicas, não relacionadas com diferenças de expressão entre os tratamentos avaliados. Uma forma de se estimar flutuações na massa de DNA em cada spot é a hibridação das membranas de náilon com uma sonda que reconhece uma região específica do plasmídeo (chamada overgo), comum a todos os ESTs analisados (Perret, 1998). Uma vez que todos os spots competem igualmente pela sonda overgo, a hibridação entre a sonda e o spot é proporcional à quantidade de DNA nele depositada. Ao hibridar os macroarranjos com uma sonda que reconhece o gene da ampicilina do plasmídeo, nós observamos variações nos sinais entre spots de uma mesma membrana e entre membranas réplicas, que, teoricamente, deveriam ter a mesma quantidade de DNA (Fig. 3A e 3C). Usualmente, o dados de spots réplicas (com o mesmo EST) que apresentaram uma variação de mais de duas vezes entre seus respectivos sinais ou entre membranas não são considerados nas análises (J. Amselem, INRA – Versailles, com. pessoal).

Uma estratégia para diminuir a variação no sinal entre spots réplicas é a transferência do DNA em etapas, fixando-o mais de uma vez em cada spot (L. Reis, Instituto Ludwig, com. pessoal). Ao fixar uma só vez, caso ocorra alguma imperfeição na coleta do DNA ou na sua deposição na membrana, não há como corrigir. Ao fixar mais de uma vez, a probabilidade de que o erro aconteça duas vezes no mesmo spot é menor. Esse procedimento foi avaliado em um experimento no qual um macroarranjo contendo 768 ESTs em duplicata, depositados com o replicador manual, foi hibridado com a sonda overgo. Nós observamos que a variabilidade na quantidade de DNA depositada foi reduzida em 50% quando o DNA é transferido em duas etapas, sendo que a transferência em três etapas não causou maior redução (dados não mostrados). Aplicando duas transferências, cerca de 98% do total de spots réplicas tiveram suas razões de intensidade de sinal entre 0,5 - 2,0, resultado muito similar ao obtido com arranjos comerciais.

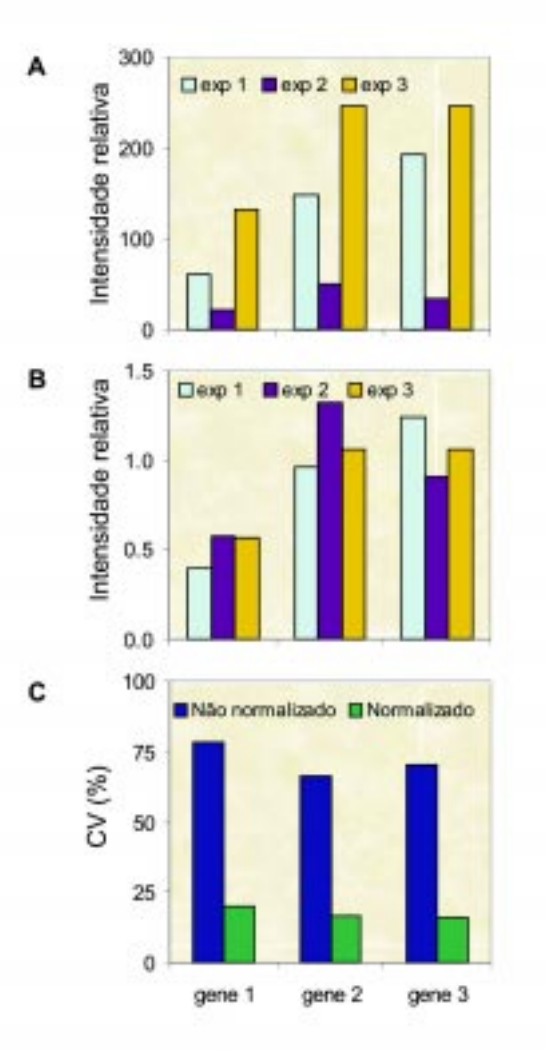


Figura 4:
Correção de variações do sinal entre experimentos empregando a normalização com a mediana de todos os genes
Três membranas réplicas foram hibridadas com sondas de cDNA obtidas a partir da mesma amostra de RNA de folha de cana-de-açúcar. Os sinais de 3 genes são mostrados. (A). Sinal dos três genes observados em três membranas réplicas hibridadas com sondas de cDNA, sem normalização. (B). Sinal observado em (A), após normalização utilizando as medianas de todos os genes contidos em cada membrana. (C). Coeficiente de variação (CV) dos genes antes e após a normalização

2. Subtração do background

Após a captação da imagem e a quantificação dos sinais obtidos para cada spot, é feita a subtração do valor referente à emissão de fundo (ba-

ckground), não específica, presente em toda membrana. Essa subtração procura remover o nível de "ruído", permitindo fazer a comparação entre os sinais específicos de cada spot. Existem várias maneiras de determinar o background, como a inserção de spots contendo plasmídeos sem inserto de cDNA. Porém, existem variações do nível de background ao longo da membrana, podendo afetar mais alguns spots que outros. Dessa forma, a subtração do background local, ao redor de cada spot é a maneira mais confiável para a obtenção dos dados. A maioria dos programas de quantificação dos sinais de arranjos de DNA já subtrai esse valor, gerando tabelas que contêm os valores locais de background, os valores brutos para cada spot e esses valores subtraídos do background, que serão os utilizados nas próximas etapas de normalização.

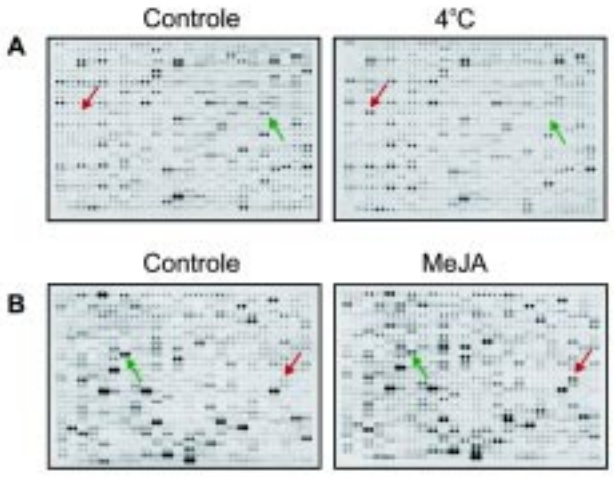


Figura 5:
Identificação de genes envolvidos na resposta a estresses abióticos utilizando macroarranjos
Macroarranjos contendo 768 ESTs em duplicata foram hibridados com sondas de folhas jovens de cana-de-açúcar em condições de controle e tratados com MeJa (A) e controle, e expostas a 4°C (B). As setas indicam os genes induzidos (seta vermelha) e reprimidos (seta verde)

3. Normalização do sinal causado por flutuações experimentais

A eficiência do processo de hibridação é influenciada por diversos parâmetros experimentais, principalmente no que se diz respeito a qualidade e quantidade de RNA utilizado na síntese de cDNA (Schuchhardt et al., 2000). Assim, os sinais de todos os genes obtidos da hibridação com uma determinada sonda podem ser maiores que aqueles obtidos com uma segunda sonda. Um parâmetro empregado para reduzir essa variação é a mediana do sinal de todos os genes da membrana, pois ela reflete as flutuações de hibridação, a síntese de sonda, etc. Nessa estratégia, divide-se a intensidade de sinal de cada spot pela mediana do sinal de todos os spots daquela membrana (Schummer et al., 1999). Uma alternativa à mediana é a média de todos os valores; no entanto, como a mediana é o valor central do sinal de todos os spots, ela é menos sensível à presença de valores extremos de intensidade.

Para avaliar essa estratégia de normalização, três macroarranjos idênticos foram hibridados individualmente com três sondas de cDNAs obtidas de reações de transcrição reversa independentes, porém da mesma amostra de RNA total de cana-de-açúcar. O sinal para cada gene deveria ser similar entre os três macroarranjos, porém observamos uma flutuação no sinal detectado (Fig. 4A). Essa flutuação é um provável reflexo das diferenças no processo de hibridação ou, alternativamente, na síntese de cDNA. Houve uma menor flutuação no sinal normalizado com suas respectivas medianas (Fig. 4B), com conseqüente redução do coeficiente de variação (Fig. 4C). Esses resultados comprovam que, em uma determinada membrana, a divisão do sinal de cada gene pela mediana de todos os genes reduz a flutuação nos dados observados, aumentando a confiabilidade da técnica. Porém, essa estratégia não se aplica a tratamentos nos quais a maioria dos genes tem diferentes níveis de expressão. Nesse caso, uma alternativa é identificar genes que não variam entre os tratamentos e usá-los como controle.

4. Sensibilidade para detecção de transcritos pouco abundantes

A maioria dos genes apresenta baixos níveis de expressão, com cerca de 10 moléculas de mRNA por célula, enquanto que os genes intermediários e abundantes (aprox. 300 e 12.000 moléculas/célula, respectivamente) são uma pequena fração do transcritoma (Huang et al., 1999). Assim, é fundamental que o método de análise de expressão tenha sensibilidade suficiente para detectar genes pouco expressos. Os macroarranjos construídos em nosso laboratório apresentaram um bom nível de sensibilidade, suficiente para detectar genes com expressão muito reduzida, que correspondem a 0.004% dos transcritos na célula (Felix et al., manuscrito em preparação).

Identificação de genes envolvidos na resposta a estresses abióticos, utilizando a técnica de macroarranjos

Após a padronização e normalização dessa metodologia, nosso grupo vem-se empenhando na identificação de padrões de expressão gênica em cana-de-açúcar em resposta a estresses abióticos, como baixas temperaturas e a hormônios como o elicitor metiljasmonato e o ácido abscísico.

Os estresses abióticos representam um dos principais fatores limitantes para a produtividade agrícola no mundo inteiro, além de representarem uma barreira para a introdução de espécies cultiváveis em áreas que ainda não são utilizadas para agricultura (Cherry, 1994). Como exemplo desses estresses, podemos destacar o decorrente de baixas temperaturas. Diversas plantas desenvolvem tolerância ao frio quando expostas a baixas temperaturas, processo denominado aclimatação ao frio. Embora já se conheçam alguns genes cuja expressão é alterada nesse processo, existem ainda muitas dúvi-

das quanto às rotas de sinalização envolvidas na resposta ao frio.

Sendo assim, com o intuito de descobrir novos genes induzidos por baixas temperaturas, macroarranjos de DNA contendo 1.536 ESTs de cana-de-açúcar foram hibridados com sondas de cDNA sintetizadas a partir de 30 µg de RNA total de plantas expostas ou não a 4 °C, durante vários períodos (Fig. 5A). Um total de 60 genes com expressão diferencial foram identificados. Dentre eles, destacam-se genes que participam da fixação de CO_2, dobramento correto de proteínas, metabolismo antioxidante, entre outros (Nogueira et al., manuscrito em preparação).

Por outro lado, as plantas estão sujeitas a outros estresses, tais como a radiação ultravioleta, vento e granizo e ao ataque de patógenos e pragas que acarretam em severos danos celulares. Sob tais condições, as plantas promovem a biossíntese de ácido jasmônico (JA) e seu metil éster volátil, metiljasmonato (MeJA). Esses reguladores de crescimento participam de diversos processos fisiológicos como, por exemplo, senescência, mecanopercepção, morfogênese, defesa a patógenos e pragas, entre outras (Wasternack & Parthier 1997; Glazebrook, 1999).

Para avaliar os efeitos do MeJA na expressão gênica em plântulas de cana-de-açúcar, macroarranjos contendo 1.536 ESTs oriundos do projeto SUCEST foram hibridados com sondas de cDNA sintetizadas a partir de RNA total extraído de tecidos foliares de cana-de-açúcar após diversos períodos de tratamento com MeJA (Fig. 5B). Os resultados permitiram a identificação de 37 genes induzidos e 44 reprimidos, entre os quais se encontram genes que participam da biossíntese do ácido jasmônico como a desaturase, três fatores de transcrição que participam da transdução de sinais, e as enzimas catalase e superóxido dismutase, que participam de estresse oxidativo, além de genes sem similaridade no GenBank. Esses resultados podem ajudar na compreensão dos efeitos produzidos por esse potente octadecanóide, além de revelar novos genes que podem ser importantes para a fisiologia e/ou defesa da planta a estresses.

Perspectivas e Conclusões

A tecnologia dos arranjos de DNA têm permitido um enfoque genômico para elucidar a regulação gênica dos mais diversos processos fisiológicos, colocando-nos diante de uma nova era na ciência. Em particular, os arranjos de DNA em náilon tal como descritos neste trabalho são uma alternativa robusta e pouco dispendiosa para análises de transcritomas.

Agradecimentos:

Gostaríamos de agradecer aos técnicos do projeto SUCEST e especialmente a Renato V. dos Santos pela ajuda na bioinformática; e à Cooperativa de Produtores de Açúcar e Álcool do Estado de São Paulo (Copersucar) pelo cultivo e manutenção das culturas de cana-de-açúcar. O trabalho em nosso grupo é financiado pela Fapesp, Comunidade Econômica Européia, CNPq/PADCT e Capes.

Referências Bibliográficas

Cherry JH (1994). Biochemical and cellular mechanisms of stress tolerance in plants. 604 pp Springer-Verlag

Desprez T, Amselem J, Caboche M, Hofte H (1998) Differential gene expression in Arabidopsis monitored using cDNA arrays. Plant Journal 14: 643-652

Freeman WM, Robertson DJ, Vrana KE (2000) Fundamentals of DNA hybridization arrays for gene expression analysis. Biotechniques 29: 1042-1055

Glazebrook J (1999) Genes controlling expression of defense responses in Arabidopsis. Current Opinion in Plant Biology 2: 280-286

Huang GM, Ng WL, Farkas J, He L, Liang HA, Gordon D, Yu J, Hood L (1999) Prostate cancer expression profiling by cDNA sequencing analysis. Genomics 59: 178-186

Lennon GG e Lehrach H (1991) Hybridization analyses of arrayed cDNA Libraries. Trends in Genetics 7: 314-317

Mangalam H, Stewart J, Zhou J, Schlauch K, Waugh M, Chen G, Farmer AD, Colello G, Weller JW (2001) GeneX: An Open Source gene expression database and integrated tool set. Ibm Systems Journal 40: 552-569

Passos GAS, Nguyen C, Jordan B (1999) Projeto Transcriptoma, Análise da expressão gênica em larga escala usando DNA-arrays. Biotecnologia Ciência & Desenvolvimento 12: 34-37

Perret E, Ferrán EA, Marinx O, Liauzun P, Dumont X, Fournier J, Kaghad M, Ferrara P, and Caput D.(1998) Improved differential screening approach to analyze transcription variations in organized cDNA libraries. Gene 208 , 103-115.

Sasaki Y, Asamizu E, Shibata D, Nakamura Y, Kaneko T, Awai K, Amagai M, Kuwata C, Tsugane T, Masuda T, Shimada H, Takamiya X, Ohta H, Tabata S (2001) Monitoring of methyl jasmonate-responsive genes in Arabidopsis by cDNA macroarray: Self-activation of jasmonic acid biosynthesis and crosstalk with other phytohormone signaling pathways. Dna Research 8: 153-161.

Schuchhardt J, Beule D, Malik A, Wolski E, Eickhoff H, Lehrach H, and Herzel H. (2000) Normalization strategies for cDNA microarrays. Nucleic Acids Research 28, e47.

Schummer M, Ng VLV, Baumgarner RE, Nelson PS, Schummer B, Bednarski DW, Hassell L, Baldwin RL, Karlan BY, Hood L (1999) Comparative hybridization of an array of 21 500 ovarian cDNAs for the discovery of genes overexpressed in ovarian carcinomas. Gene 238: 375-385

Wang YH, Garvin DF, Kochian LV (2001) Nitrate-induced genes in tomato roots. Array analysis reveals novel genes that may play a role in nitrogen nutrition. Plant Physiology 127: 345-359

Wasternack C, Parthier B (1997) Jasmonate signalled plant gene expression. Trends in Plant Science 2: 302-307

Wenzl P, Patino GM, Chaves AL, Mayer JE, RAO IM (2001) The high level of aluminum resistance in signalgrass is not associated with known mechanisms of external aluminum detoxification in root apices. Plant Physiology 125: 1473-1484

Zhao Nd, Hashida H, Takahashi N, Misumi Y, Sakaki Y (1995) High-density cDNA filter analysis - a novel-approach for large-scale, quantitative-analysis of gene-expression. Gene 156: 207-213.

Capítulo II

RNA expression profiles and data mining of sugarcane response to low temperature

Fábio T.S. Nogueira, Vicente E. De Rosa Jr., Marcelo Menossi , Eugênio C. Ulian, Paulo Arruda

Plant Physiology 132, 1811–1824 (2003).

RNA Expression Profiles and Data Mining of Sugarcane Response to Low Temperature[1]

Fábio T.S. Nogueira, Vicente E. De Rosa Jr., Marcelo Menossi, Eugênio C. Ulian, and Paulo Arruda*

Centro de Biologia Molecular e Engenharia Genética, Universidade Estadual de Campinas, 13083–970, Campinas, SP, Brazil (F.T.S.N., V.E.D.R., M.M., P.A.); Departamento de Genética e Evolução, Instituto de Biologia, Universidade Estadual de Campinas, 13083–970, Campinas, SP, Brazil (F.T.S.N., V.E.D.R., M.M., P.A.); and Centro de Tecnologia Copersucar, 13400–970, Piracicaba, SP, Brazil (E.C.U.)

Tropical and subtropical plants are generally sensitive to cold and can show appreciable variation in their response to cold stress when exposed to low positive temperatures. Using nylon filter arrays, we analyzed the expression profile of 1,536 expressed sequence tags (ESTs) of sugarcane (*Saccharum* sp. cv SP80-3280) exposed to cold for 3 to 48 h. Thirty-four cold-inducible ESTs were identified, of which 20 were novel cold-responsive genes that had not previously been reported as being cold inducible, including *cellulose synthase*, ABI3-interacting protein 2, a negative transcription regulator, *phosphate transporter*, and others, as well as several unknown genes. In addition, 25 ESTs were identified as being down-regulated during cold exposure. Using a database of cold-regulated proteins reported for other plants, we searched for homologs in the sugarcane EST project database (SUCEST), which contains 263,000 ESTs. Thirty-three homologous putative cold-regulated proteins were identified in the SUCEST database. On the basis of the expression profiles of the cold-inducible genes and the data-mining results, we propose a molecular model for the sugarcane response to low temperature.

Cold is one of the most important environmental stresses affecting plant growth and crop productivity. Chilling (low temperatures above 0°C) and freezing (temperatures below 0°C inducing extracellular ice formation) limit the geographical distribution and growing season of many crops and cause significant crop losses (Xin and Browse, 2000). However, chilling and freezing stresses differ from each other in that the former involves a direct effect of low temperature on cells, whereas freezing often acts indirectly, damaging cells by dehydration (Pearce, 1999).

Plants vary considerably in their ability to survive under chilling and freezing temperatures. At one extreme, some herbaceous plants from temperate regions can survive under freezing temperatures ranging from −5°C to −30°C. At the other extreme, plants from tropical and subtropical regions have virtually little or no capacity to survive even the slightest freezing (Thomashow, 2001). In addition, even species considered to be sensitive to chilling can show substantial variation in their response to colder temperatures. Cold (chilling and freezing) tolerance often increases if plants are first hardened by exposure to a period of acclimation at low positive temperatures (Thomashow, 2001). Cold acclimation is associated with biochemical and physiological changes that include alterations in carbohydrate metabolism, membrane lipid composition, phenylpropanoid content, respiration, photosynthesis, and oxidative stress defenses (Allen and Ort, 2001). In general, plants exposed to low temperatures show two basic responses: the adjustment of metabolism to kinetic constraints imposed by low temperatures, and the induction of tolerance to chilling and freezing via the activation of specific genes. These alterations are frequently caused by changes in gene expression (Guy et al., 1985), as shown by the isolation and characterization of several cold-regulated (*COR*) genes from different plant species (Thomashow, 1998).

Cold-induced genes can also be induced by water stress and, in several cases drought- and cold-inducible genes are also induced by the phytohormone abscisic acid (ABA). Dehydration caused by water stress or cold appears to trigger ABA biosynthesis, which in turn induces the expression of several genes (Liu et al., 1998). However, an ABA-independent pathway of gene induction by low temperature and water stress has been described (Yamaguchi-Shinozaki and Shinozaki, 1994; Gilmour et al., 1998; Liu et al., 1998). Direct evidence of an ABA-independent pathway came from the identification of three cold-induced transcription factors (*CBF1*, *CBF2*, and *CBF3*) from Arabidopsis that encode proteins containing AP2 DNA-binding motifs (Medina et al., 1999). These transcription factors are not induced by exogenous ABA, suggesting that they participate in a cold-induced ABA-independent

[1] This research was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (grant no. 98/12250–0 to P.A.). F.T.S.N. and V.E.R.J. were recipients of fellowships from Fundação de Amparo à Pesquisa do Estado de São Paulo. E.C.U. was the recipient of a fellowship from the Copersucar Technology Center, Universidade Estadual de Campinas.

* Corresponding author; e-mail parruda@unicamp.br; fax 55–19–3788–1089.

Article, publication date, and citation information can be found at www.plantphysiol.org/cgi/doi/10.1104/pp.102.017483.

pathway (Yamaguchi-Shinozaki and Shinozaki, 1994; Seki et al., 2001). These factors bind to the dehydration-responsive element/C-repeat (DRE/CRT) cis-acting element present in the promoter of drought-, high salt-, and cold-inducible genes (Liu et al., 1998). Constitutive overexpression of *CBF1* and *CBF3* genes in nonacclimated transgenic Arabidopsis plants induces the expression of several cold-inducible genes, thereby increasing the tolerance of the plant to freezing (Gilmour et al., 2000).

Sugarcane (*Saccharum* sp.) is generally considered as a cold-sensitive plant (Tai and Lentini, 1998). The magnitude of cold (chilling and freezing) damage is dependent on the severity and duration of the low temperature, cultivar resistance to post-freezing deterioration, and time lapse and temperature fluctuations between the freeze event and harvest (Tai and Lentini, 1998). However, field observations have shown that the sensitivity of sugarcane to cold varies among varieties. Du et al. (1999) demonstrated that some subtropical hybrid species are more cold tolerant than tropical species. Thus, the identification of sugarcane cold-responsive genes could be an important resource in breeding programs to detect genetic differences among hybrids and to obtain transgenic plants with improved tolerance to cold and its related stress.

The rapid advance of genome-scale sequencing has led to the development of methods for analyzing transcript abundance in a large set of genes involved in abiotic stress responses (Perret et al., 1998; Schena et al., 1998; Schummer et al., 1999; Freeman et al., 2000; Seki et al., 2001; Fowler and Thomashow, 2002). Expressed sequence tags (ESTs) are a rich source for gene discovery in several organisms. EST databases can be used for large-scale data mining of genes involved in specific pathways and, in association with techniques for expression profile analysis, can be helpful in analyzing the global response of tissues or whole organisms under biotic or abiotic stress and in the discovery of novel genes (Cushman and Bohnert, 2000). In this work, we describe the use of sugarcane ESTs from the Sugarcane EST Genome Project (SUCEST; http://sucest.lad.ic.unicamp.br) to construct high-density filter arrays containing random ESTs to identify cold-responsive genes. Using the expression profile data generated and extensive data mining in the SUCEST database, we constructed a putative model for global sugarcane gene expression under cold exposure.

RESULTS

Construction of Sugarcane EST Macroarrays and Data Analysis

Using SUCEST cDNA clones and a hand-held tool with a 96-pin printhead (V&P Scientific, San Diego, CA), we constructed sets of two high-density filters, each filter containing 768 random EST targets, thereby totaling 1,536 ESTs. The macroarray nomenclature used was that established for filter-based methods in which the target is the DNA spotted onto the filter surface, and the probe is the labeled DNA that is hybridized to the surface-bound DNA (Rose, 2000).

To decrease variation in the amount of DNA among spots and filters, each EST clone was spotted twice at the same position on all filters. This procedure reduced the coefficient of variation (cv) among spots by 50% (J.M. Felix, R. Drummond, F.T.S. Nogueira, R.A. Jorge, P. Arruda, and M. Menossi, unpublished data). Each EST was spotted at two positions on the filters to assess the reproducibility of spotting.

Before cDNA probe hybridization, the high-density filters were hybridized with a probe corresponding to the plasmid vector (see "Materials and Methods"), and the signal intensity was measured. The median value for all spot intensities of each filter was determined and then we estimated the cv of these median values to assess fluctuation in the amount of DNA among replicate filters. Only replicate filters with cv values lower than 10% were used for subsequent analysis. In addition, to reduce the variation among replicate filters caused by differences in the experimental conditions, the average of all signal intensities obtained with the cDNA probe from each filter was set to 1 (Schummer et al., 1999). To assess the signal variation between replicate spots, we estimated the ratio between the signal intensities of the two replicate spots representing each EST spotted onto filters. Around 98% of ESTs had ratios ranging from 0.5 to 2.0, indicating that the DNA variation between replicate spots was less than 2-fold for most ESTs (data not shown). This finding agreed with data for nylon filter arrays spotted with a manual device (Schummer et al., 1997). Finally, 12 spots representing the empty plasmid vector were used in each filter as a negative control to assess nonspecific hybridization. These spots consistently produced a hybridization signal above background, probably because of some degree of similarity between sugarcane cDNA and the vector sequences. Thus, only those sugarcane ESTs displaying signal intensities higher than the average signal intensity of the negative control plus two SD for three or more arrays were considered for further data analysis.

Identification and Functional Classification of Cold-Responsive Sugarcane ESTs

To determine the threshold for changes in gene expression that could be attributable to cold treatment, we used the strategy reported by Friddle et al. (2000) in which the $\log_2$ of the expression ratios (i.e. ratios between the normalized signal intensity of each time point for treated and control samples of each EST) were normally distributed and centered on

a ratio of 0 (i.e. the average of the expression ratios was 1). The average and the SD of the expression ratios ($\log_2$ transformed) for each cold-treatment time point were estimated from two independent experiments. Comparison of the ratio of signal intensities between two independent, untreated control probes yielded an *R* value of around 0.92, indicating a good correlation between replicate experiments. Assuming that experimental variation was nonspecific for any particular EST, we selected as the threshold of significance an expression ratio at least 1.65 SD above or below the average expression ratio for each cold-treatment time point in two replicate experiments ($P < 0.0025$). Moreover, among the previously selected ESTs, we only considered those displaying at least a 2-fold induction or repression at each interval of cold treatment relative to the control. A representative filter comparing control and cold-treated plantlets samples (48-h time point) is shown in Figure 1. Fifty-nine high quality sugarcane ESTs in the SUCEST database showing significantly altered expression during cold treatment were identified, of which 34 were up-regulated by cold treatment. Scatter plots of the expression ratios of cold-treated versus control plantlets are shown in Figure 2. Ratios below 1.0 were inverted and multiplied by −1 to aid data interpretation (Girke et al., 2000). In most ESTs, expression was unchanged by cold treatment (Fig. 2). The expression of a few ESTs was altered after 6 h of cold treatment, whereas the highest induction or repression occurred after 24 h of cold exposure. Figure 3 shows the expression profiling of a randomly sampled subset of six cold-inducible sugarcane ESTs from two independent experiments. Although the absolute -fold induction values were not identical between biological samples, the expression profiles were similar, corroborating the reproducibility of our array data.

The usefulness of our arrays for screening *COR* genes was demonstrated by the identification of several cold-inducible genes that had already been reported for other plants (Tables I and II). The putative relevant biological functions of all cold-responsive sugarcane ESTs identified are shown in Table II. The sugarcane cold-inducible ESTs were distributed in three classes (Table I). The first class contained 14 ESTs with homologs in other organisms in which they represent drought and cold-inducible genes. These included xanthine dehydrogenase (*XDH*), ocs-element binding factor 1 (*OCSBF-1*), pyruvate orthophosphate dikinase (*PPDK*), superoxide dismutase (*SOD*), NADP-dependent malic enzyme (*NADP-ME*), a putative sugar transporter, *polyubiquitin*, and NAC genes (Table II). The second class consisted of seven ESTs not previously described as being induced by cold stress. Among them were an EST encoding ABI3-interacting protein 2 (AIP2) involved in development (Kurup et al., 2000) and an EST encoding cellulose synthase (Table II). The third class contained 13 ESTs encoding unknown proteins with no hits in the GenBank nr database (Table II).

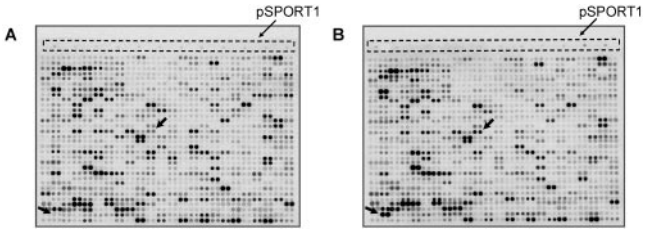


Figure 1. Examples of macroarray filters used to analyze gene expression under cold acclimation. Filters containing 768 random sugarcane cDNAs in duplicate were probed with [^{33}P]cDNA reverse-transcribed from total RNA of control (A) and cold-treated (48 h at 4°C; B) sugarcane plantlets. Nor-specific hybridization was monitored using 12 spots representing the empty pSPORT1 vector (dashed rectangles). The signals were detected in a phosphorimager analyzer. The arrows indicate examples of cold-inducible sugarcane ESTs.

Among the total cold-responsive ESTs, 25 (47%) were down-regulated by cold exposure, 13 of which encode proteins presenting no hit in the GenBank nr database (Table II). Twelve ESTs encode proteins with a wide range of functions, including transcription, signaling (receptor-like protein kinases), amino acid metabolism (acetohydroxyacid synthase and Asn synthetase), defense (pathogenesis-related protein) development (NAM-like protein), and water status (aquaporin). These results suggest that several metabolic processes, including perception of stress signals and regulation of gene expression, were repressed during cold stress.

To estimate the relative contribution of cold-inducible genes from each SUCEST library used in the array experiments, we first calculated the normalized number of cold-inducible ESTs (i.e. the ratio between the number of cold-inducible ESTs from each library and the total ESTs spotted onto filters for each SUCEST library) and then estimated the percentage of cold-inducible ESTs from each cDNA library (Fig. 4). Interestingly, the library HR1 (Vettore et al., 2001) accounted for most of the cold-inducible ESTs. This library was constructed from bacterium-infected sugarcane tissues (Vettore et al., 2001) and could be a good source for the discovery of biotic and abiotic stress-related genes.

RNA-Blot Analysis

To validate the macroarray data, we did blot analysis using total RNA from a new set of cold-treated and untreated plantlets. Five cDNA clones representing *polyubiquitin*, *OsNAC6*, and three novel cold-inducible genes were analyzed. Figure 5 compares the gene expression profiling obtained using macroarrays and RNA gel blots. Although the absolute -fold induction values of the blots were not identical to those on the array, there was a high consistency between the two data sets.

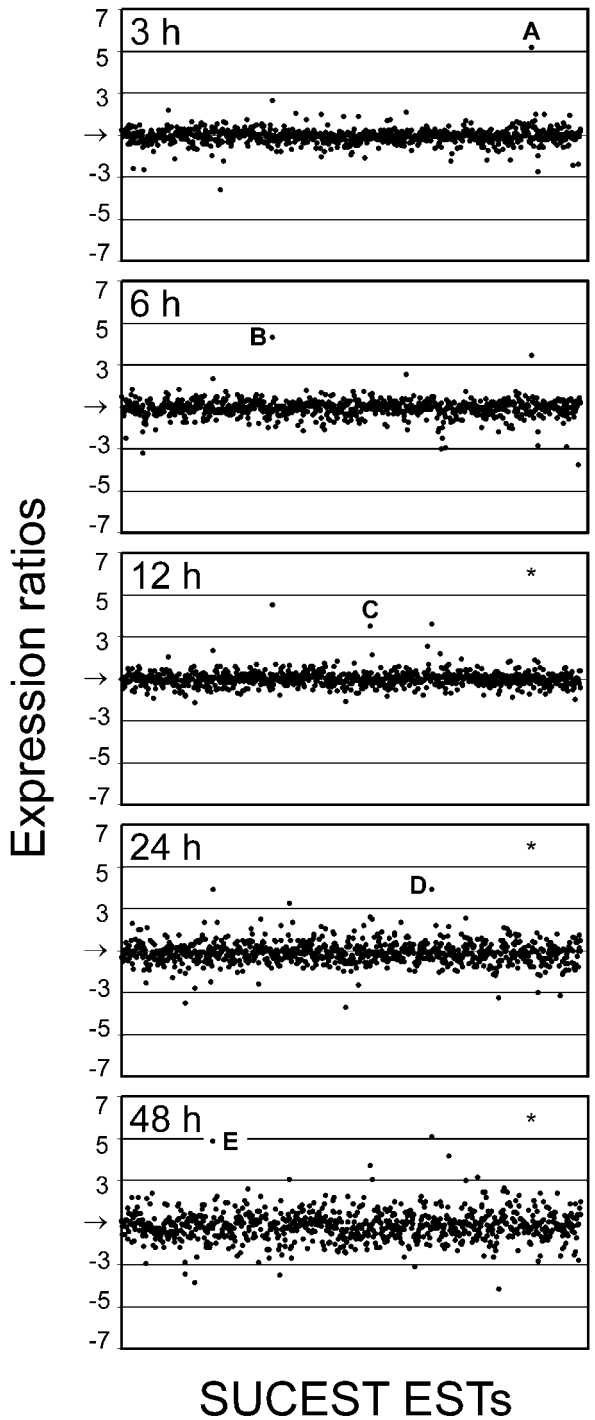


Figure 2. Scatter plots of the expression ratios of all ESTs analyzed. The average of the normalized relative intensities of each EST was used to calculate the expression ratios. The ratios that were higher with cold treatment (3, 6, 12, 24, and 48 h at 4°C) are plotted upward, and those that were higher in the controls are plotted downward. The lateral arrows represent ratios of approximately 1. Letters on the scatter plots indicate examples of cold-inducible sugarcane ESTs. A, EST encoding to no hit protein; B, *OsNAC6* gene; C, EST encoding to unknown protein; D, *OsNAC4* gene and E, *cellulose synthase* gene. Asterisks in the 12, 24, and 48 h panels indicate expression ratios that surpassed the ratio scale. Only one replicate macroarray experiment is represented in this figure.

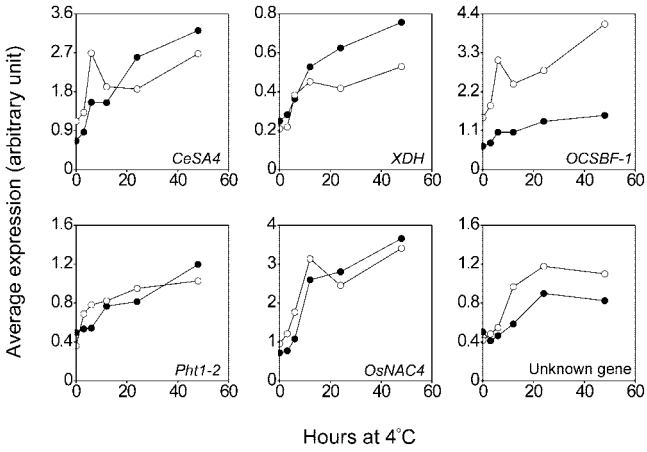


Figure 3. Examples of the expression profile of cold-inducible sugarcane ESTs. The average expression was calculated from normalized relative intensities of each time point of each EST from two independent experiments (white and black circles). *CeSA4*, Cellulose synthase 4; *Pht1-2*, phosphate transporter; and *OsNAC4*, rice NAC gene 4.

Data Mining and Domain Analysis

To complement the expression profiling data of cold-inducible genes obtained with macroarrays and to provide a global view of the up-regulation of gene expression in sugarcane during cold exposure, we undertook extensive data mining in the SUCEST databank to find homologs of cold-inducible genes reported in other plant species. Initially, we created a protein sequence database containing 250 cold-inducible genes reported in the literature and present in the GenBank database (http://www.ncbi.nlm.nih.gov). With the tBLASTN algorithm, these protein sequences were used as drivers to identify putative assembled sequences (combined set of contigs and singlets representing different transcripts from the SUCEST database). The criteria used to select the SUCEST-assembled sequences were the E value and percentage of protein coverage. Assembled sequences with an E value $\leq 10^{-20}$ and a protein sequence coverage greater than 70% were considered to represent putative cold-inducible gene homologs. Thirty-three SUCEST-assembled sequences encoding proteins similar to COR proteins described for other plants were identified (Table III). Interestingly, one

Table I. *Cold-inducible sugarcane ESTs identified by macroarray expression profiling*

Summary	No. of ESTs	% ESTs
Total cold-inducible ESTs	34	100
Cold- and drought-inducible ESTs[a]	14	41
ESTs unrelated to cold[b]	7	21
Novel cold-inducible ESTs[c]	13	38

[a]ESTs homologous to previously described cold- and drought-inducible genes. [b]ESTs homologous to previously described genes, but not induced by cold. [c]ESTs encoding proteins that have not been described previously.

Table II. *Average expression ratios and sequence similarities of cold-responsive sugarcane ESTs*

Clone Identification	BLAST Hit[a]	E Value	Description[b]	Ratios[c]				
				3 h	6 h	12 h	24 h	48 h
Up-regulated								
			Proteolysis					
SCEPCL6029G10	AAC16012	7E − 86	Polyubiquitin protein	1.1	1.4	2.1	2.0	3.4
SCCCHR1002E03	S28426	1E − 133	Polyubiquitin protein	1.0	1.2	1.8	2.5	3.0
SCCCAD1002H02	CAA66667	1E − 131	Polyubiquitin protein	0.8	1.1	1.7	2.0	3.0
SCCCHR1002A02	CAA66667	1E − 119	Polyubiquitin protein	0.9	1.2	1.6	1.6	2.8
SCEPCL6028G03	S17435	0.0	Polyubiquitin protein	1.0	1.2	1.6	1.7	3.0
			Plant development					
SCMCCL6027D02	BAA89800	1E − 141	OsNAC6 protein	2.2	3.3	4.0	6.2	8.5
SCCCHR1004F11	BAA89798	1E − 52	OsNAC4 protein	1.2	1.7	3.4	3.2	4.3
SCRLCL6032B05	BAB09485	3E − 16	NAM protein	1.0	1.0	2.0	2.1	2.6
			Antioxidant metabolism					
SCCCHR1004E11	T10235	0.0	Xanthine dehydrogenase	1.1	1.6	2.1	2.2	2.8
SCCCLR1068H03	P93407	1E − 42	Cu/ZnSOD protein	2.0	1.5	1.1	2.2	2.0
			Transcription regulation					
SCEPCL6021E11	P24068	3E − 57	OCSBF-1 protein	1.2	1.8	1.6	2.0	2.6
SCEPCL6028F01	BAA97498	2E − 21	Negative transc regulator	1.2	1.2	1.1	1.5	2.2
SCCCHR1003E12	CAB75509	2E − 26	AIP2 protein	2.1	1.4	1.1	1.0	0.9
			CO_2 fixation					
SCEPCL6023E08	CAA06247	0.0	PPDK protein	1.4	1.8	1.8	2.1	1.3
SCEPCL6023H04	AAK91502	0.0	NADP-ME	1.6	2.2	2.1	1.2	1.1
			Cell wall metabolism					
SCMCCL6027C02	AAF89964	0.0	Cellulose synthase-4	1.3	2.4	2.0	2.8	3.6
SCCCAD1003B03	T04331	2E − 75	se-wap41 protein	1.3	1.8	2.1	1.9	3.2
			Protein metabolism					
SCCCHR1003G02	AJ309824	0.0	25S ribosomal gene	1.3	1.2	1.4	2.8	3.2
SCCCAD1001H05	AJ309824	0.0	25S ribosomal gene	1.5	1.7	2.3	3.5	4.0
			Transporter					
SCCCAD1001G06	AAK25880	5E − 47	Putative sugar transporter	1.1	1.2	2.0	1.9	2.5
SCCCHR1003H07	AAM14593	0.0	Pht1-2	1.5	1.6	1.9	2.1	2.6
			Unknown/unclassified					
SCEPCL6021A10	T04466	1E − 94	Unknown protein	1.2	1.9	1.8	2.1	3.1
SCMCCL6027G02	AAD20139	3E − 54	Unknown protein	1.2	1.3	1.5	1.8	3.4
SCCCHR1003D12	T45625	5E − 30	Unknown protein	2.1	1.2	1.0	1.0	0.9
SCCCHR1004E10	AAD32824	3E − 55	Unknown protein	1.0	1.6	2.9	2.4	3.9
SCEPCL6028F02	AAD20391	2E − 45	Unknown protein	1.1	1.2	1.1	1.6	2.4
SCCCAD1003C07	BAB56055	4E − 19	Unknown protein	0.8	1.4	1.7	2.2	2.5
SCCCHR1003A12	NP_567065	3E − 13	Unknown protein	1.4	2.1	2.6	2.8	2.7
SCCCAD1002F02	NP_564419	3E − 12	Unknown protein	1.0	1.1	1.7	2.3	2.1
SCCCLR1068G12	BAB07950	4E − 24	Unknown protein	1.3	1.2	2.1	1.8	1.7
SCCCHR1003H01	–	–	No protein match	4.0	4.1	8.7	15.3	29.9
SCEPCL6028G02	–	–	No protein match	0.8	1.1	1.3	1.9	2.6
SCEPCL6029H09	–	–	No protein match	0.8	1.2	2.1	1.4	2.4
SCUTCL6036C05	–	–	No protein match	1.5	0.9	1.3	0.9	3.1
Down-regulated								
			Plant development					
SCRLCL6032B11	AAN64996	1E − 24	Putative NAM protein	0.7	0.4	0.7	0.8	0.6
			Transcription regulation					
SCRLCL6032F11	BAC55608	0.0	Scarecrow-like protein	0.7	0.5	0.7	0.6	0.5
			Signal transduction					
SCEPCL6028A06	AAK43512	1E − 141	Putative receptor kinase	1.2	0.4	1.1	0.9	1.1
SCRLCL6032H12	AAL67082	0.0	Putative receptor kinase	0.7	0.5	0.7	0.7	0.6

(*Table continues*)

CBF transcription factor was found among these 33 SUCEST putative genes (Table III). Recently, Jaglo et al. (2001) identified components of the CBF regulon (DRE/CRT-containing genes that are induced by CBF transcription factors) in other plant species. Thus, sugarcane may have novel genes in the CBF regulon with roles in mediating the responses to cold. In Arabidopsis, proteins encoded by the CBF regulon protect cells against freezing and other stress associated with dehydration (Thomashow, 2001).

Table II. *Continues*

Clone Identification	BLAST Hit[a]	E Value	Description[b]	Ratios[c]				
				3 h	6 h	12 h	24 h	48 h
			Amino acid metabolism					
SCCCHR1001G10	CAA45117	0.0	Acetohydroxyacid synthase	0.9	0.8	1.0	0.5	0.4
SCCCLR1068G05	AAK49456	1E − 159	Asn synthetase	0.7	0.5	0.9	0.8	0.6
			Water channel					
SCCCHR1004G12	AAK26758	1E − 163	Aquaporin	0.7	0.6	0.6	0.3	0.2
			Defense					
SCRLCL6032B09	T05694	6E − 67	Pathogenesis-related protein	1.2	0.4	1.0	1.4	1.0
			Heat shock response					
SCEPCL6023F09	CAA77978	0.0	HSP82	0.9	0.8	0.6	0.6	0.3
SCAGCL6016D10	CAA77978	0.0	HSP82	0.8	0.7	0.6	0.6	0.3
			Lipid metabolism					
SCRLCL6032B08	NP_178516	0.0	Acyl-CoA synthetase	0.7	0.5	1.0	0.9	0.6
			Protein metabolism					
SCCCHR1001C11	P49210	1E − 95	Ribosomal protein L9	0.9	0.7	0.7	0.4	0.4
			Unknown/unclassified					
SCEPCL6023A10	BAA92204	4E − 52	Unknown protein	0.9	0.9	0.5	0.5	0.5
SCEPCL6029A12	AAL36075	1E − 155	Unknown protein	0.4	0.3	1.2	0.5	0.4
SCCCHR1001B11	BAB64285	5E − 13	Unknown protein	0.7	0.5	0.4	0.2	0.2
SCCCHR1004B05	BAB64285	5E − 13	Unknown protein	0.8	0.5	0.4	0.3	0.2
SCCCAD1002A11	–	–	No protein match	0.6	0.6	0.4	0.2	0.3
SCEPCL6023A04	–	–	No protein match	1.3	0.3	1.4	1.3	1.2
SCAGCL6016A10	–	–	No protein match	0.4	0.9	0.9	0.7	0.6
SCUTCL6036B02	–	–	No protein match	0.9	1.0	0.5	0.5	0.4
SCCCHR1001B05	–	–	No protein match	0.6	0.4	0.4	0.2	0.2
SCRLCL6032D11	–	–	No protein match	0.6	0.4	0.7	0.6	0.4
SCRLCL6032E11	–	–	No protein match	0.6	0.4	0.7	0.8	0.5
SCRLCL6032G07	–	–	No protein match	0.8	0.4	0.8	0.9	0.6
SCRLCL6032E10	–	–	No protein match	0.7	0.5	0.8	0.7	0.5

[a]BLAST hit of each sequence was obtained using the BLASTX algorithm, except for the ribosomal genes, which were also analyzed by the BLASTN algorithm. [b]Description indicates the putative functions of the gene products expected from similarity sequences. [c]Each value represents the average of the expression ratios between the normalized relative intensity of each interval of cold treatment and the control (0-h exposure to cold) from two independent experiments.

To identify putative conserved domains, the inserts of all cDNAs encoding for unknown/unclassified cold-inducible proteins were completely sequenced. Most of the domains identified were related to proteins involved in the regulation of gene expression and signal transduction (Table IV). It is possible that these unknown proteins belong to novel cold-response pathways and cold/drought-tolerance. Finally, Figure 6 shows a possible panel of the up-regulation of gene expression during sugarcane cold adaptation, based on the cold-inducible gene expression profiling data and the data mining results described above.

DISCUSSION

Cold-Responsive Genes in Sugarcane

We employed high-density filters to assess the expression profile of sugarcane ESTs when plants were submitted to cold treatment for up to 48 h. Eleven of the cold-inducible ESTs found in our experiments represented genes reported to be induced by cold and drought in other plants. Acclimation to cold induces several biochemical and physiological alterations in the cellular machinery and probably improves plant tolerance to cold and other cold-related stress (Guy et al., 1985; Allen and Ort, 2001). Cold- and drought-related stresses induce a set of common genes (Liu et al., 1998). In this study, we identified five cold-inducible sugarcane ESTs encoding for polyubiquitin proteins (Table II). These proteins occur in all eukaryotes either as a free monomer or covalently linked to a variety of other proteins. One of their major functions is in tagging proteins for selective degradation by the 26S proteasome (O'Mahony and Oliver, 1999). As far as we know, there are no reports showing the induction of *polyubiquitin* gene expression by chilling or freezing in plants. However, O'Mahony and Oliver (1999) isolated a polyubiquitin cDNA involved in vegetative desiccation. Accordingly, it is possible that the late cold-induced expression of sugarcane *polyubiqitin* genes (Table II; Fig. 5) could be related to recovery from water stress caused by exposure to chilling temperatures.

One sugarcane EST encoding a putative XDH was significantly induced after 12 h of cold exposure (Table II). XDH is an NAD-dependent dehydrogenase

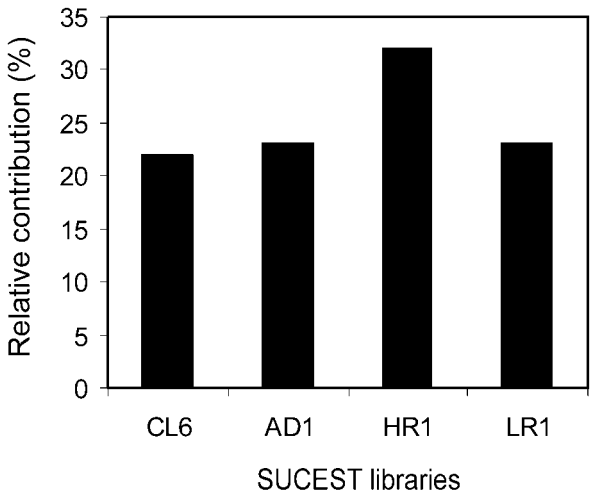


Figure 4. Relative contribution of the SUCEST libraries used in the macroarray experiments to identify cold-inducible genes. The values represent the percentage of normalized cold-inducible genes identified from each cDNA library (see "Results") in relation to the total number of ESTs identified as cold inducible in our arrays. CL6, Heat- and cold-treated and untreated callus; AD1, sugarcane plantlets infected with *Acetobacter diazotroficans*; HR1, plantlets infected with *Herbasperillum rubrisubalbicans*; and LR1, leaf row tissue.

that catalyzes the two final reactions of purine catabolism (Xu et al., 1996). This enzyme was induced by cold in fruitfly (*Drosophila melanogaster*; Duncker et al., 1995). Because chilling induces oxidative stress (Sato et al., 2001), it is possible that the sugarcane XDH is induced in response to oxidative stress generated by cold exposure. Another protein related to antioxidant metabolism is copper/zinc SOD (Cu/ZnSOD). This enzyme belongs to the group of metalloenzymes that protect cells from superoxide radicals by catalyzing the dismutation of the superoxide radical to molecular O_2 and H_2O_2 (Wu et al., 1999). A sugarcane EST encoding Cu/ZnSOD was induced after 24 h of cold exposure (Table II). The induction of two antioxidant genes suggests an increase in oxidative stress because cold-acclimating conditions do not result in the cessation of photosynthesis such that superoxide radicals were still being generated (Wu et al., 1999).

Kiyosue et al. (1998) observed that dehydration and chilling stress (4°C) induced a cDNA clone (ERD6) from Arabidopsis encoding a putative sugar transporter protein belonging to a multigene family (Williams et al., 2000). We also identified a sugarcane EST encoding a putative sugar transporter protein that was induced after 12 to 48 h of cold exposure (Table II). Sugar redistribution may be used as an energy source to protect cells against stress conditions (Kiyosue et al., 1998; Williams et al., 2000).

We also identified a cold-inducible inorganic phosphate transporter protein (Pht1-2). The role of nutrients such as nitrogen and phosphorus on cold hardiness has received attention because low temperatures inhibit photosynthesis and consequently reduce inorganic phosphate availability (Hurry et al., 2000). The induction of a sugarcane phosphate transporter gene may indicate a readjustment of the cellular P_i status and the recovery of photosynthetic carbon metabolism, thereby re-establishing the ability to produce Suc.

The recovery of photosynthetic carbon metabolism during cold stress may be achieved by increasing the

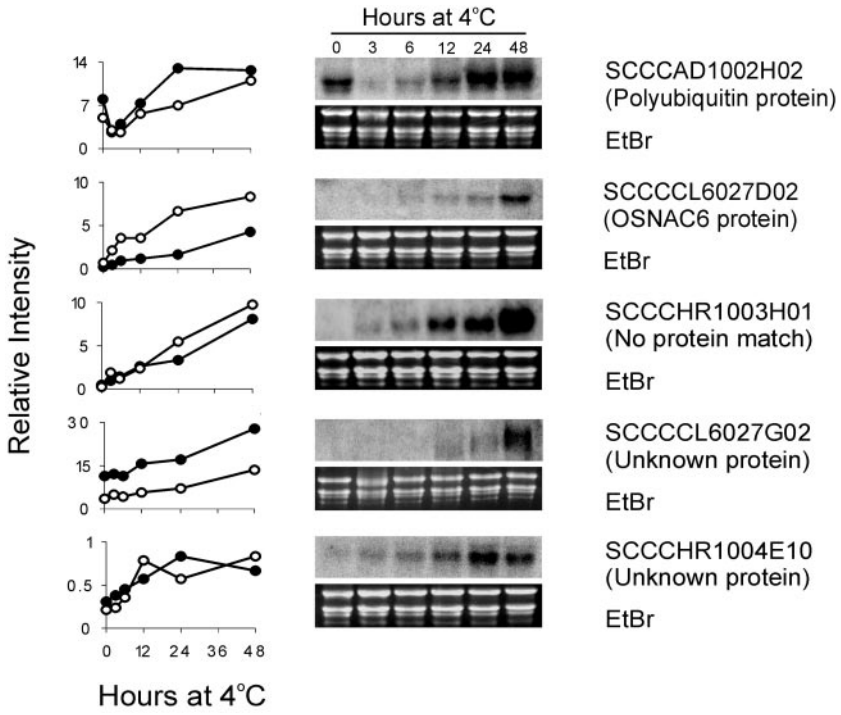


Figure 5. Comparison between EST macroarray and RNA-blot analysis for cold-inducible sugarcane ESTs. In RNA gel blots, each lane was loaded with 10 μg of total RNA isolated from plantlets grown at 26°C (0) and plantlets grown at 4°C for 3, 6, 12, 24, and 48 h. The graphs show the induction kinetics observed in the macroarrays (white circles) and RNA blots (black circles).

Table III. *Cold-regulated sugarcane homolog proteins identified by data mining the SUCEST database*

CHS, chalcone synthase; CLP, chitinase-like protein; P5CS, Δ (1) pyrroline-5-carboxylase synthetase; PLC1, phosphatidyl-specific phospholipase C protein; TLP, thaumatin-like protein; GLP, glucanase-like protein; ADH, alcohol dehydrogenase; GolS, galactinol synthase.

Assembled Sequences Identification[a]	Protein Similarity	Accession No.	Coverage[b]	E Value	Identity[d]	Similarity[d]	References
			%		%	%	
SCEPRZ1008C06	OsCDPK7	BAB16888	88	0.0	90	93	Saijo et al. (2000)
SCQSHR1023B08	CBF1	NP_567721	95	4E − 25	34	46	Stockinger et al. (2001)
SCBGLR1095B10	GCN5	O22929	80	4E − 34	57	72	Stockinger et al. (2001)
SCCCRZ1C01B03	ADA2b	AAK31320	99	1E − 137	49	59	Stockinger et al. (2001)
SCCCRZ1001D02	14-3-3	T04153	100	1E − 136	96	96	Jarillo et al. (1994)
SCRLLR1059D11	LIP15	S58692	100	6E − 48	70	74	Kusano et al. (1995)
SCCCRZ2001H12	NpCaM-1	P13565	100	1E − 80	100	100	van der Luit et al. (1999)
SCVPLR2027D06	CORTMC-AP3	CAA09867	100	3E − 37	52	62	Baldi et al. (1999)
SCCCLR1001F04	WCOR410b	T06802	100	2E − 27	33	39	Danyluk et al. (1998)
SCUTLR2015A11	WCOR413	AAG13395	96	2E − 89	82	86	Allard et al. (1998)
SCEQRT1028G10	CHS	AAD41878	99	0.0	96	96	Berberich et al. (1997)
SCJLRT1020F06	CLP	383024	99	1E − 143	75	80	Yu et al. (1999)
SCCCRZ1002G07	SUS1[c]	AAA68209	100	0.0	98	98	Déjardin et al. (1999)
SCEZRZ1016F07	α-Amylase	AAF63239	100	1E − 179	68	79	Wegrzyn et al. (2000)
SCACLR2014H05	β-Amylase	CAB58423	82	1E − 114	84	89	Seki et al. (2001)
SCCCRZ2C01H04	Ferritin	P29036	95	1E − 108	75	76	Seki et al. (2001)
SCQGLR1062E12	Glyoxalase I	BAA36759	100	1E − 149	88	92	Seki et al. (2001)
SCEPAM2011H12	atPUMP	CAA11757	96	1E − 127	73	79	Maia et al. (1998)
SCBFLR1026E05	V-ATPase	P49087	73	0.0	96	96	Carystinos et al. (1995)
SCCCRZ2C03E07	V-PPase[c]	S72527	94	0.0	96	97	Carystinos et al. (1995)
SCQGLR1085F11	DHN2	T14819	92	2E − 20	60	60	Zhu et al. (2000)
SCSGLR1045D05	HVA22	A48892	80	7E − 46	77	90	Shen et al. (2001)
SCCCLR1072E07	APX[c]	T03595	100	1E − 132	91	94	Zhang et al. (1997)
SCJLLR1054F05	HSC70	P24067	97	0.0	90	90	Li et al. (1999)
SCJFLR1073H12	P5CS	O04226	100	0.0	75	83	Gilmour et al. (2000)
SCSBHR1052C05	PLC1	AAK01711	75	0.0	71	77	Hirayama et al. (1995)
SCCCCL3003G05	60S Ribos. L13	BAA92738	94	2E − 95	86	89	Saez-Vasquez et al. (2000)
SCVPRT2073B04	TLP	444344	85	7E − 72	83	88	Yu et al. (1999)
SCJLRT1023E06	GLP	AF230109	91	3E − 75	49	65	Yu et al. (1999)
SCVPLR2005E09	ScRS7[c]	AAD26256	100	3E − 97	91	96	Berberich et al. (2000)
SCEZLR1031F04	GBF-1	T02084	97	1E − 162	78	80	de Vetten et al. (1995)
SCCCST2003C12	ADH	AAC34997	100	0.0	92	92	de Bruxelles et al. (1996)
SCCCLR1075C10	GolS	AAD26116	90	1E − 144	77	85	Pearce (1999)

[a]SUCEST-assembled sequence homologous to cold-regulated proteins reported for other plant species. [b]SUCEST-assembled sequence size relative to the total amino acid sequence of the driver protein. [c]Sugarcane protein homologs not been induced or repressed by cold treatment in our arrays. SUS1, Sucrose synthase 1; V-PPase, H^+-translocating pyrophosphatase; APX, ascorbate peroxidase; ScRS7, ribosomal protein S7. [d]The identity and similarity percentages were obtained using the BLOSUM62 matrix (http://www.ncbi.nlm.nih.gov).

activities of CO_2 fixation-related proteins such as PPDK and NADP-ME (Hurry et al., 2000). PPDK is an important enzyme in C_4 plant photosynthesis and is induced by ABA and stress associated with abiotic factors such as drying, cold, high salt, and mannitol treatment (Moons et al., 1998). NADP-ME is a major decarboxylating enzyme in NADP-ME-type C_4 species and has been reported to be induced by chilling stress (Fushimi et al., 1994). Sugarcane ESTs encoding PPDK and NADP-ME proteins were induced by cold exposition in our arrays (Table II), suggesting a possible maintenance of photosynthesis, even at low temperatures.

Members of the NAC protein family contain a highly conserved amino acid sequence in the N-terminal region known as the NAC domain (Kikuchi et al., 2000). The proteins of this family are specific to plant genomes and play important roles in plant growth, development, and senescence (Souer et al., 1996; Kikuchi et al., 2000). Recently, new members of the NAC family have been shown to interact with virus proteins and are involved in protein-protein interactions (Xie et al., 1999; Ren et al., 2000). Some NAC proteins also act as transcriptional regulators (Kikuchi et al., 2000). We identified three sugarcane ESTs encoding putative NAC proteins that were induced by cold exposure. Sequence analysis of these ESTs suggested that each one encodes a different member of the NAC family, with two of them encoding proteins similar to rice (*Oryza sativa*) NAC and the other one encoding a homolog of the Arabidopsis NAM protein (Table II). Interestingly, the ex-

Table IV. *Conserved domains of cold-inducible sugarcane putative and no hit proteins*

Clone Identification	Domain Description	Accession No.[a]	E Value	Similarity	Putative Function[b]
				%	
SCCCHR1004E10	SPX	pfam03105	2E − 22	48	G-protein-associated signal transduction; may also function as a phosphate sensor
SCEPCL6021A10	RING	smart00184	1E − 03	58	Protein-protein interactions
SCMCCL6027G02	DUF81	pfam01925	8E − 06	48	Unknown
SCEPCL6028F02	mTERF	pfam02536	2E − 17	57	DNA-binding motif (contains Leu zippers)
SCUTCL6036C05	e3-binding	pfam29152	4.6E − 08	40	Protein-protein interactions
SCCCHR1003H01	LEA 3	pfam03242	1.6E − 05	44	Desiccation tolerance
SCCCHR1003D12	Glycosylhydrolase	PF00332	5.1E − 23	76	Endo-1,3-β-glucanase-like protein

[a]Accession of each domain sequence was obtained by using the RPS-BLAST algorithm and the Pfam and SMART databases. [b]Putative functions expected based on similar domains described in SMART and Pfam domain family databases.

pression profiles of these ESTs differed from each other (Table II).

An EST encoding a protein similar to the OCSBF-1 was also induced by cold exposure (Table II). This protein was similar to the LIP19 of rice and LIP15 of maize (*Zea mays*), both of which are bZIP proteins up-regulated by exogenous ABA and low temperature (Singh et al., 1990; Aguan et al., 1991; Kusano et al., 1995). OCSBF-1 binds ocs cis-acting elements, G-box motifs, and ABA-responsive element (ABRE) in plant gene promoters (Kusano et al., 1995; Kim et al., 2001). Thus, sugarcane may have an ABA-dependent path-

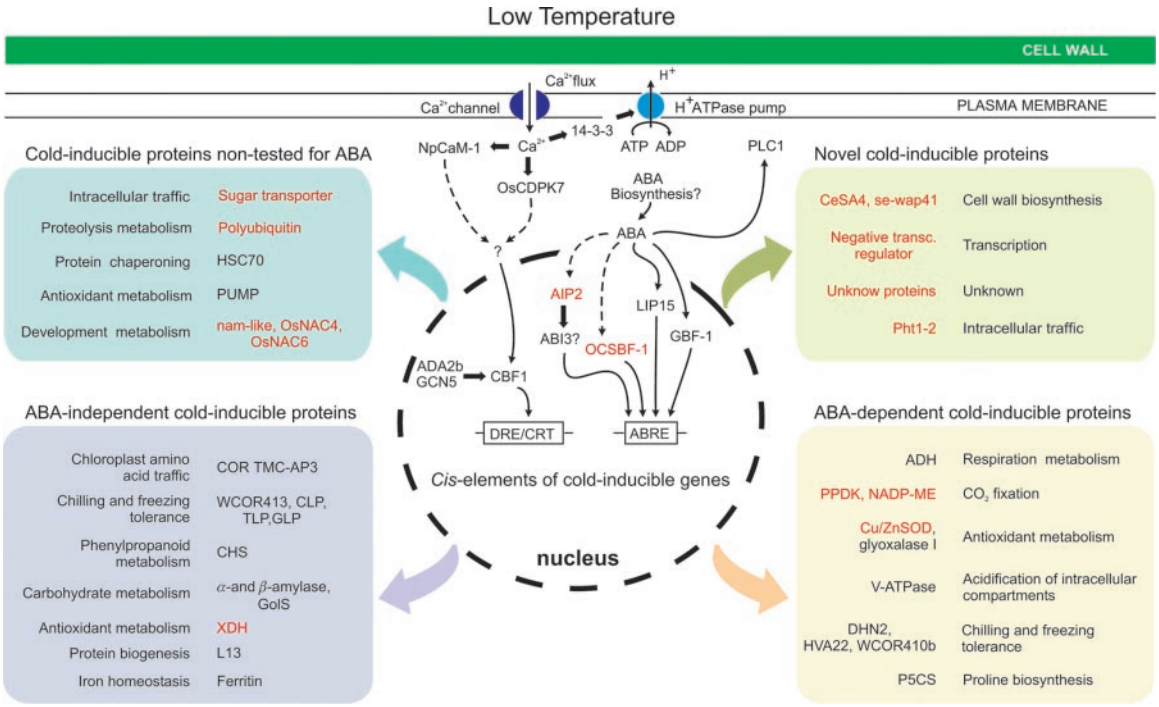


Figure 6. Hypothetical model of the sugarcane responses to low temperature. Colored rectangles contain pathways and protein names. The proteins identified by data mining are shown in black, whereas those in red were identified by macroarray expression profiling. Thin solid and dashed arrows represent gene induction (confirmed and putative, respectively) and large solid arrows represent protein interactions. OsCDPK7, rice Ca^{2+}-dependent protein kinase 7; NpCaM-1, calmodulin protein; 14-3-3, 14-3-3 protein; ADA2b, transcriptional adaptor; GCN5, HAT-like protein; CBF1, transcription factor containing AP2 DNA-binding motif; GBF-1, ABI3, LIP15, and OCSBF-1, bZIP transcription factors; AIP2,ABI3, interacting protein 2; COR TMC-AP3, chloroplastic amino acid selective channel protein; WCOR413, cold acclimation protein homolog F18B3.110; TLP, thaumatin-like protein; CLP, chitinase-like protein; GLP, glucanase-like protein; CHS, chalcone synthase; GolS, galactinol synthase; L13, ribosomal protein; Pht1-2, phosphate transporter; V-ATPase, vacuolar ATPase; HSC70, heat shock protein; PUMP, plant uncoupling mitochondrial protein; ADH, alcohol dehydrogenase; CeSA4, cellulose synthase 4; se-wap41, Golgi-associated protein; DHN2 and WCOR410b, dehydrin-like proteins; HVA22, ABA-inducible protein; P5CS, Δ (1) pyrroline-5-carboxylase synthetase; PLC1, phosphatidyl-specific phospholipase C protein; NAM-like, OsNAC4, and OsNAC6, NAC proteins. The cis-elements present in cold-inducible gene promoters are indicated by the rectangles: DRE/CRT and ABRE.

way triggered by low temperature. However, the presence of sugarcane protein similar to CBF1 (found in the data mining) also suggests the existence of an ABA-independent pathway (Table III; Fig. 6). These pathways may cross-talk during chilling stress.

Further evidence for the existence of an ABA-dependent cold-inducible pathway was the identification of a sugarcane EST encoding a protein similar to AIP2 from Arabidopsis. Significant induction of this gene was observed after 3 h of cold exposure (Table II). *AIP2* encodes a C_3HC_4-type zinc finger protein that interacts with ABA-INSENSITIVE3 (ABI3) protein during seed development in Arabidopsis (Kurup et al., 2000). ABI3 also appears to modulate low-temperature-induced freezing tolerance (Tamminen et al., 2001). Although no SUCEST-assembled sequence similar to ABI3 protein was identified, we cannot exclude the possibility that the sugarcane genome encodes this protein.

A sugarcane EST encoding a protein similar to an Arabidopsis general negative transcription regulator was up-regulated after 48 h of cold exposure (Table II). To our knowledge, this is the first time that the induction of a negative transcription regulator has been associated with low temperatures. This protein may be involved in down-regulating specific genes in response to chilling stress.

Finally, we identified two cold-inducible sugarcane ESTs encoding proteins similar to cellulose synthase 4 (CeSA4) and se-wap41. CeSA4 is involved in the production of cellulose, the major component of all higher plant cell walls (Richmond, 2000). Se-wap41 is a reversible glycosylated polypeptide associated with the Golgi complex that is involved in the synthesis of xyloglucan and other hemicelluloses (Dhugga et al., 1997). Nevertheless, the role of this putative reinforcement of the sugarcane leaf cell wall by cellulose and other components is not evident in chilling stress, because it has been suggested to function as a barrier against ice propagation during freezing stress (Pearce, 1999).

The repression of gene expression may also be an important component of the adaptation to low temperatures. To our knowledge, most of the ESTs found in our experiments represented genes that have not previously been reported as being down-regulated by chilling or freezing stresses. However, an aquaporin gene (*rwc1*) was identified as being chilling-repressed in rice leaves (Li et al., 2000). The authors suggested that this gene is involved in responses to water stress induced by chilling. It is possible that reduced transcript levels of sugarcane aquaporin (Table II) keep a suitable status of intracellular water during chilling stress.

Cold-Inducible Sugarcane Genes Mined in SUCEST Database and a Putative Model for Sugarcane Response to Low Temperature

A database containing the sequences of cold-inducible genes identified in different plant species was used as a driver to identify putative homologs in the SUCEST database. Thirty-three sugarcane-assembled sequences were identified, of which four were also found in our arrays (Table III). These genes are involved in many different functions including signaling pathways, transcriptional regulation, and other metabolic processes. The information associated with these proteins together with that provided by our arrays was used to develop a model for sugarcane responses to cold (Fig. 6).

A transient Ca^{2+} influx through plasma membrane Ca^{2+} channels or Ca^{2+} release from vacuole occurs during the initiation of cold acclimation (Monroy and Dhindsa, 1995; Knight et al., 1996; Knight and Knight, 2000). In addition, some Ca^{2+}-dependent proteins may be induced by low temperatures in plants (Jarillo et al., 1994; Monroy and Dhindsa, 1995; van der Luit et al., 1999). In our proposed model for sugarcane responses to cold adaptation (Fig. 6), the transient Ca^{2+} influx could be sufficient to activate Ca^{2+}-dependent multifunctional proteins such as 14-3-3, Ca^{2+}-dependent protein kinase (OsCDPK7), and calmodulin (NpCaM-1).

One SUCEST-assembled sequence showed similarity to CBF1 protein (Table III), including the nearly identical signature sequences (PPK/RPAGRxKFx-ETRHP and DSAWL) surrounding the AP2/EREBP domain, characteristic of CBF proteins (Jaglo et al., 2001). Gilmour et al. (1998) suggested that *CBF1* encodes the AP2-like transcriptional activator that binds to the CRT/DRE regulatory element present in the promoter of many *COR* genes and stimulates their transcription. On the basis of these observations, we speculate that sugarcane has cold-inducible genes that are members of the CBF regulon. While we were preparing this paper, Fowler and Thomashow (2002) reported that cold acclimation induced the expression of novel members of the CBF regulon in Arabidopsis. Among them, there are genes encoding for a putative sugar transporter and a galactinol synthase protein that were up-regulated by cold in our arrays and were also found in the SUCEST data mining (Tables II and III).

Another recent fundamental advance in understanding cold-associated transcriptional control mechanisms was the discovery of Arabidopsis histone acetyltransferase (HAT)-containing adapter complexes, which are recruited to promoters by transcriptional factors. These proteins can stimulate transcription (Stockinger et al., 2001). These complexes consist of the proteins GCN5 and ADA (Grant et al., 1997). The GCN5 protein has HAT activity, whereas the ADA protein may function as a transcriptional adapter for some activation domains (Marcus et al., 1994; Stockinger et al., 2001). We identified SUCEST-assembled sequences similar to Arabidopsis GCN5 and ADA2b proteins (Table III). These results provide preliminary evidence for the existence of HAT-

containing adapter complexes in sugarcane. The existence of this complex in plants was believed to be restricted to Arabidopsis (Stockinger et al., 2001). As in Arabidopsis (Stockinger et al., 2001), sugarcane GCN5 and ADA2b proteins may interact physically with each other and with CBF1, thereby stimulating the transcription of CBF1 target genes (Fig. 6).

We found assembled sequences encoding bZIP transcription factors similar to the G-box-binding factor 1 (GBF-1) and LIP15 (de Vetten and Ferl, 1995; Kusano et al., 1995; Kim et al., 2001), which bind to the ABRE cis-element present in cold- and ABA-induced gene promoters. Kim et al. (2001) observed that the level of *SGBF-1* transcripts in soybean (*Glycine max*) was up-regulated by cold and ABA. On the basis of these data, the LIP15 and GBF-1 proteins in sugarcane may be induced by cold treatment and may bind to the ABRE motif, thus enhancing the expression of cold-inducible genes through a cold-induced ABA-dependent pathway (Fig. 6).

A sugarcane-assembled sequence homologous to phosphatidyl-specific phospholipase C protein from Arabidopsis (atPLC-1; Hirayama et al., 1995) was also identified in the SUCEST database (Table III). This enzyme, which is induced by low temperature and ABA, hydrolyzes phosphatidylinositol 4,5-biphosphate into the 2-s messengers 1,4,5-triphosphate and 1,2-diacylglycerol (Hirayama et al., 1995).

Mining of the SUCEST database allowed us to identify several other proteins known to be induced by cold exposure (Table III). For instance, Δ (1) pyrroline-5-carboxylase synthetase, galactinol synthase, and α- and β-amylases may contribute to solute accumulation (Pearce, 1999; Gilmour et al., 2000; Seki et al., 2001). Other proteins, such as plant uncoupling mitochondrial protein, which can decrease the generation of reactive oxygen species during chilling stress (Maia et al., 1998), were also identified.

The cold anaerobic conditions caused by waterlogging and cold-induced increase of endogenous ABA can up-regulate alcohol dehydrogenase expression (de Bruxelles et al., 1996). In addition, other pathways, such as phenylpropanoid biosynthesis, may be involved in reactive oxygen species scavenging (Berberich and Kusano, 1997). We found an assembled sequence encoding a protein similar to chalcone synthase that catalyzes the first step in the branch pathway of phenylpropanoid synthesis specific for the formation of flavonoid products, which can function as antioxidant compounds (Grace and Logan, 2000). We also found an assembled sequence encoding a protein similar to the chloroplast amino acid-selective channel protein (COR TMC-AP3) from barley (*Hordeum vulgare*), whose gene expression is up-regulated by cold (Baldi et al., 1999).

A putative up-regulation of HSC70 mRNA expression in sugarcane during chilling stress may be required to sustain high levels of this heat shock protein that would stabilize some proteins compromised at low, nonfreezing temperatures (Li et al., 1999). Sugarcane protein biogenesis can also be altered during cold exposure through induction of a 60S ribosomal L13-like protein (Sáez-Vásquez et al., 2000).

Finally, we found ESTs encoding proteins that have been shown to be directly involved in chilling and freezing tolerance, including WCOR410b (Danyluk et al., 1998), WCOR413 (Allard et al., 1998), dehydrin 2 (DHN2; Zhu et al., 2000), barley ABA-inducible protein (HVA22; Shen et al., 2001), thaumatin-like protein, glucanase-like protein, and chitinase-like protein (Yu and Griffith, 1999). Our computer analyses suggest that sugarcane has two putative dehydrin-like proteins (WCOR410b and DHN2). These proteins could stabilize macromolecules and/or protect membranes against chilling damage (Pearce, 1999). Moreover, thaumatin-like protein, glucanase-like protein, and chitinase like-protein are examples of pathogenesis-related proteins with antifreeze activity, which can have direct effects on the stability of cellular membranes (Hincha et al., 1997; Pearce, 1999). A β-1,3-glucanase, for instance, has been reported as a protecting factor of thylakoid membranes against freeze-thaw effects (Hincha et al., 1997), as well as a fish antifreeze protein type I, which inhibited leakage across membranes during chilling to nonfreezing temperatures (Tomczak et al., 2002). It is possible that sugarcane putative antifreeze proteins can confer cellular membrane protection, reducing chilling injury.

MATERIALS AND METHODS

Plant Growth and Cold Treatment

Sugarcane (*Saccharum* sp. cv SP80-3280) plantlets were propagated axenically in vitro by excising the shoot apex of 2-month-old sugarcane plants kept in a greenhouse and culturing them in 5 mL of Murashige and Skoog medium (Murashige and Skoog, 1962) containing 150 mg citric acid L^{-1}. After 1 month, the shoots were transferred to 50 mL of the same medium supplemented with 0.2 mg 6-benzylaminopurine L^{-1} and 0.1 mg kinetin L^{-1}. Proliferating plantlets were subcultured every 3 weeks. Plantlets were then transferred to rooting medium in which 6-benzylaminopurine was substituted by 0.2 mg indole butyric acid L^{-1}. All plantlets were kept in a growth chamber at 26°C on a 16-h/8-h day/night cycle with a photon flux density of 70 μE m^{-2} s^{-1}. After 4 weeks, when the roots were fully developed, the plantlets were transferred to a growth chamber under the same photoperiod conditions but at 4°C. Control plantlets were maintained at 26°C. The leaves of control and cold-treated plantlets were harvested after 0, 3, 6, 12, 24, and 48 h of treatment. Six plantlets were used for each time point.

High-Density Filter Arrays and Probe Preparation

Sixteen 96-well plates containing EST plasmid clones were randomly sampled from the following sugarcane cDNA libraries: heat- and cold-treated and untreated callus (CL6), sugarcane plantlets infected with *H. rubrisubalbicans* (HR1) or *A. diazotroficans* (AD1), and leaf row tissue (LR1; Vettore et al., 2001). The plasmid DNA was denatured in 0.2 N NaOH for 15 min at 37°C and then spotted onto Hybond-N filters (Amersham Biosciences, Piscataway, NJ) with a hand-held 96-pin printhead tool (V&P Scientific, San Diego, CA). This tool typically deposited 0.1 μL of DNA solution, which corresponded to approximately 5 ng of DNA. The set of sixteen 96-well plates was spotted on two nylon filters (85 × 125 mm) in a 4 × 4 array configuration (768 ESTs per filter). Additionally, 12 spots

containing DNA of the empty plasmid vector pSPORT1 (Invitrogen, Carlsbad, CA) were applied to the filters as a negative hybridization control. The DNA was fixed to the filters by baking at 80°C for 2 h, and the filters were then stored at room temperature until used for hybridizations.

Total RNA was isolated from the leaves of treated and untreated sugarcane plantlets using Trizol Reagent (Invitrogen) according to the manufacturer's instructions. Probes were produced as described by Schummer et al. (1999) with slight modifications. In brief, 30 μg of total RNA was reverse transcribed with Superscript II (Invitrogen) using an oligo-dT18V (3 μM) primer, with 3,000 Ci mmol^{-1} [α-^{33}P]dCTP and unlabeled dATP, dGTP, and dTTP (1 mM each) for 20 min at 42°C. Unlabeled dCTP was then added to a final concentration of 1 mM, and the reaction continued for another 40 min. The cDNA probes were purified by using ProbeQuant G-50 microcolumns according to the manufacturer's instructions (Amersham Biosciences). Variations in the amount of DNA in the spots were estimated by hybridizing the filters with an oligonucleotide probe that recognized the sequence of the *Amp*^{r} gene of the pSPORT1 vector. This probe was synthesized with the primers 5′-GTGGTCCTGCAACTTTATCCGC-3′ and 5′-TAGACTGGATGGAGGCGGATAA-3′ in the presence of [α-^{33}P]dCTP, according to the protocol described by McPherson (2000; http://www.tree.caltech.edu/protocols/overgo.html).

Macroarray Analysis

Filters were initially hybridized with the oligo vector probe for 16 h at 58°C. Further details of the hybridization procedures can be obtained from the Web site cited in the previous paragraph. After hybridization and washing, the filters were exposed to imaging plates for 96 h and then scanned in a phosphorimager FLA3000-G (Fujifilm, Tokyo). The oligo vector probe was removed from the filters by boiling in 0.1% (w/v) SDS solution, with the efficiency of probe removal being monitored by phosphorimager scanning. After stripping, the filters were hybridized with cDNA probes for 18 h at 42°C as described by Schummer et al. (1997) and then sequentially washed in the following solutions: 0.2× SSC with 0.5% (w/v) SDS for 20 min at room temperature, 0.5× SSC with 0.5% (w/v) SDS for 20 min at 65°C (twice), and 0.2× SSC with 0.5% (w/v) SDS for 20 min at 65°C. After the last wash, the filters were sealed with plastic film, immediately exposed to imaging plates for 96 h, and scanned as described above. All signals were quantified by using Array Vision software (Imaging Research, St. Catherines, ON, Canada). The grids were predefined and manually adjusted to obtain optimal spot recognition, and the spots were then quantified individually. The local background was subtracted automatically from each spot, and the intensity data were rearranged into Microsoft Excel (Microsoft, Redmond, WA) files for further analysis. The two signal intensity values of duplicated DNA spots were averaged and used to calculate the expression ratios between cold-treated and control (untreated) samples.

RNA-Blot Analysis

Ten micrograms of total RNA was electrophoresed in a 1% (w/v) agarose gel containing formaldehyde and transferred to a Hybond-N+ filter (Amersham Biosciences) as described by Sambrook et al. (1989). The filters were hybridized with the cDNA inserts of cold-inducible sugarcane ESTs labeled with [α-^{32}P]dCTP, and hybridization was done at 42°C (Sambrook et al., 1989). The blots were then washed at high stringency and exposed to imaging plates. Digitized images of the RNA-blot hybridization signals were quantified using the Image Gauge software (Fujifilm).

Bioinformatics

The sequences of 250 cold-related plant proteins obtained from the National Center for Biotechnology Information were analyzed for similarity against the 43,141 assembled sequences of the SUCEST database (Telles and da Silva, 2001) using the tBLASTN algorithm. Unknown protein domain analysis was done using the RPS-BLAST algorithm, Pfam (Bateman et al., 2000), and SMART (Schultz et al., 2000) databases.

Distribution of Materials

Upon request, all novel materials described in this publication will be made available in a timely manner for noncommercial research purposes, subject to the requisite permission from any third-party owners of all parts of the material. Obtaining any permissions will be the responsibility of the requestor.

Accession Numbers

The sequence data described in this paper have been submitted to GenBank under accession numbers BU102492 to BU103710. The array data described in this manuscript have been submitted to Gene Expression Omnibus under accession numbers GPL210 (platform), GSM2431 to GSM2442 (samples), and GSE83 and GSE84 (series).

ACKNOWLEDGMENTS

We thank A.S. Zanca, S. Meire, D. Althmann, A.L. Beraldo, and R. Drummond for excellent technical assistance, R.V. dos Santos for help in constructing the COR protein sequence database, and Dr. L. Grivet for critical reading of the manuscript.

Received November 8, 2002; returned for revision January 27, 2003; accepted April 24, 2003.

LITERATURE CITED

Aguan K, Sugawara K, Suzuki N, Kusano T (1991) Isolation of genes for low-temperature-induced proteins in rice by a simple subtractive method. Plant Cell Physiol **32:** 1285–1289

Allard F, Houde M, Kr M, Ivanov A, Huner NPA, Sarhan F (1998) Betaine improves freezing tolerance in wheat. Plant Cell Physiol **39:** 1194–1202

Allen DJ, Ort DR (2001) Impacts of chilling temperatures on photosynthesis in warm-climate plants. Trends Plant Sci **6:** 36–42

Baldi P, Grossi M, Pecchioni N, Valè G, Cattivelli L (1999) High expression level of a gene coding for a chloroplastic amino acid selective channel protein is correlated to cold acclimation in cereals. Plant Mol Biol **41:** 233–243

Bateman A, Birney E, Durbin R, Eddy SR, Howe KL, Sonnhammer ELL (2000) The Pfam protein families database. Nucleic Acids Res **28:** 263–266

Berberich T, Kusano T (1997) Cycloheximide induces a subset of low temperature-inducible genes in maize. Mol Gen Genet **254:** 275–283

Berberich T, Uebeler M, Feierabend J (2000) cDNA cloning of cytoplasmic ribosomal protein S7 of winter rye (*Secale cereale*) and its expression in low-temperature-treated leaves. Biochim Biophys Acta **1492:** 276–279

Carystinos GD, MacDonald HR, Monroy AF, Dhindsa RS, Poole RJ (1995) Vacuolar H^+-translocating pyrophosphatase is induced by anoxia or chilling in seedlings of rice. Plant Physiol **108:** 641–649

Cushman JC, Bohnert HJ (2000) Genomic approaches to plant stress tolerance. Curr Opin Plant Biol **3:** 117–124

Danyluk J, Perron A, Houde M, Limin A, Fowler B, Benhamou N, Sarhan F (1998) Accumulation of an acidic dehydrin in the vicinity of the plasma membrane during cold acclimation of wheat. Plant Cell **10:** 623–638

de Bruxelles GL, Peacock WJ, Dennis ES, Dolferus R (1996) Abscisic acid induces the alcohol dehydrogenase gene in Arabidopsis. Plant Physiol **111:** 381–391

Déjardin A, Sokolov LN, Kleczkowski LA (1999) Sugar/osmoticum levels modulate differential abscisic acid-independent expression of two stress-responsive sucrose synthase genes in *Arabidopsis*. Biochem J **344:** 503–509

de Vetten NC, Ferl RJ (1995) Characterization of a maize G-box binding factor that is induced by hypoxia. Plant J **7:** 589–601

Dhugga KS, Tiwari SC, Ray PM (1997) A reversibly glycosylated polypeptide (RGP1) possibly involved in plant cell wall synthesis: purification, gene cloning, and trans-Golgi localization. Proc Natl Acad Sci USA **94:** 7679–7684

Du Y-C, Nose A, Wasano K (1999) Thermal characteristics of C4 photosynthetic enzymes from leaves of three sugarcane species differing in cold sensitivity. Plant Cell Physiol **40:** 298–304

Duncker BP, Chen C-P, Davies PL, Walker VK (1995) Antifreeze protein does not confer cold tolerance to transgenic *Drosophila melanogaster*. Cryobiology **32:** 521–527

Fowler S, Thomashow MF (2002) Arabidopsis transcriptome profiling indicates that multiple regulatory pathways are activated during cold

acclimation in addition to the CBF cold response pathway. Plant Cell **14:** 1675–1690

Freeman WM, Robertson DJ, Vrana KE (2000) Fundamentals of DNA hybridization arrays for gene expression analysis. Biotechnology **29:** 1042–1055

Friddle CJ, Koga T, Rubin EM, Bristow J (2000) Expression profiling reveals distinct sets of genes altered during induction and regression of cardiac hypertrophy. Proc Natl Acad Sci USA **97:** 6745–6750

Fushimi T, Umeda M, Shimazaki T, Kato A, Toriyama K, Uchimiya H (1994) Nucleotide sequence of a rice cDNA similar to a maize NADP-dependent malic enzyme. Plant Mol Biol **24:** 965–967

Gilmour SJ, Sebolt AM, Salazar MP, Everard JD, Thomashow MF (2000) Overexpression of the Arabidopsis CBF3 transcriptional activator mimics multiple biochemical changes associated with cold acclimation. Plant Physiol **124:** 1854–1865

Gilmour SJ, Zarka DG, Stockinger EJ, Salazar MP, Houghton JM, Thomashow MF (1998) Low temperature regulation of the Arabidopsis CBF family of AP2 transcriptional activators as an early step in cold-inducible *COR* gene expression. Plant J **16:** 433–442

Girke T, Todd J, Ruuska S, White J, Benning C, Ohlrogge J (2000) Microarray analysis of developing Arabidopsis seeds. Plant Physiol **124:** 1570–1581

Grace SC, Logan BA (2000) Energy dissipation and radical scavenging by the plant phenylpropanoid pathway. Philos Trans R Soc Lond B Biol Sci **355:** 1499–1510

Grant PA, Duggan L, Cote J, Roberts SM, Brownell JE, Candau R, Ohba R, Owen-Hughes T, Allis CD, Winston F et al. (1997) Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. Genes Dev **11:** 1640–1650

Guy CL, Niemi KJ, Brambi R (1985) Altered gene expression during cold acclimation of spinach. Proc Natl Acad Sci USA **82:** 3673–3677

Hincha DK, Meins F Jr, Schmitt JM (1997) β-1,3-Glucanase is cryoprotective in vitro and is accumulated in leaves during cold acclimation. Plant Physiol **114:** 1077–1083

Hirayama T, Ohto C, Mizoguchi T, Shinozaki K (1995) A gene encoding a phosphatidyl-specific phospholipase C is induced by dehydration and salt stress in *Arabidopsis thaliana*. Proc Natl Acad Sci USA **92:** 3903–3907

Hurry V, Strand A, Furbank R, Stitt M (2000) The role of inorganic phosphate in the development of freezing tolerance and the acclimatization of photosynthesis to low temperature is revealed by the *pho* mutants of *Arabidopsis thaliana*. Plant J **24:** 383–396

Jaglo KR, Kleff S, Amundsen KL, Zhang X, Haake V, Zhang JZ, Deits T, Thomashow MF (2001) Components of the Arabidopsis C-repeat/dehydration-responsive element binding factor cold-response pathway are conserved in *Brassica napus* and other plant species. Plant Physiol **127:** 910–917

Jarillo JA, Capel J, Leyva A, Martinez-Zapater JM, Salinas J (1994) Two related low-temperature-inducible genes of Arabidopsis encode proteins showing high homology to 14-3-3 proteins, a family of putative kinase regulators. Plant Mol Biol **25:** 693–704

Kikuchi K, Ueguchi-Tanaka M, Yoshida KT, Nagato Y, Matsusoka M, Hirano H-Y (2000) Molecular analysis of the *NAC* gene family in rice. Mol Gen Genet **262:** 1047–1051

Kim JC, Lee SH, Cheong YH, Yoo C-M, Lee SI, Chun HJ, Yun D-J, Hong JC, Lee SY, Lim CO et al. (2001) A novel cold-inducible zinc finger protein from soybean, SCOF-1, enhances cold tolerance in transgenic plants. Plant J **25:** 247–259

Kiyosue T, Abe H, Yamaguchi-Shinozaki K, Shinozaki K (1998) ERD6, a cDNA clone for an early dehydration-induced gene of Arabidopsis, encodes a putative sugar transporter. Biochim Biophys Acta **1370:** 187–191

Knight H, Knight MR (2000) Imaging spatial and cellular characteristics of low temperature calcium signature after cold acclimation in Arabidopsis. J Exp Bot **51:** 1679–1686

Knight H, Trewavas AJ, Knight MR (1996) Cold calcium signaling in *Arabidopsis* involves two cellular pools and a change in calcium signature after acclimation. Plant Cell **8:** 489–503

Kurup S, Jones HD, Holdsworth MJ (2000) Interactions of the developmental regulator ABI3 with proteins identified from developing Arabidopsis seeds. Plant J **21:** 143–155

Kusano T, Berberich T, Harada M, Suzuki N, Sugawara K (1995) A maize DNA-binding factor with a bZIP motif is induced by low temperature. Mol Gen Genet **248:** 507–517

Li L-G, Li S-F, Tao Y, Kitagawa Y (2000) Molecular cloning of a novel water channel from rice: its products expression in *Xenopus* oocytes and involvement in chilling tolerance. Plant Sci **154:** 43–51

Li Q-B, Haskell DW, Guy CL (1999) Coordinate and non-coordinate expression of the stress 70 family and other molecular chaperones at high and low temperature in spinach and tomato. Plant Mol Biol **39:** 21–34

Liu Q, Kasuga M, Sakuma Y, Abe H, Miura S, Yamaguchi-Shinozaki K, Shinozaki K (1998) Two transcriptions factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in Arabidopsis. Plant Cell **10:** 1391–1406

Maia IG, Benedetti CE, Leite A, Turcinelli SR, Vercesi AE, Arruda P (1998) AtPUMP: an Arabidopsis gene encoding a plant uncoupling mitochondrial protein. FEBS Lett **429:** 403–406

Marcus GA, Silverman N, Berger SL, Horiuchi J, Guarente L (1994) Functional similarity and physical association between GCN5 and ADA2: putative transcriptional adaptors. EMBO J **13:** 4807–4815

McPherson JD (2000) Overgo probing of high-density filters. http://www.tree.caltech.edu/protocols/overgo.html

Medina J, Bargues M, Terol J, Pérez-Alonso M, Salinas J (1999) The *Arabidopsis CBF* gene family is composed of three genes encoding AP2 domain-containing proteins whose expression is regulated by low temperature but not by abscisic acid or dehydration. Plant Physiol **119:** 463–469

Monroy AF, Dhindsa RS (1995) Low-temperature signal transduction: induction of cold acclimation-specific genes of alfalfa by calcium at 25°C. Plant Cell **7:** 321–331

Moons A, Valcke R, Van Montagu M (1998) Low-oxygen stress and water deficit induce cytosolic pyruvate orthophosphate dikinase (PPDK) expression in roots of rice, a C3 plant. Plant J **15:** 89–98

Murashige T, Skoog F (1962) A revised medium for rapid growth and biossays with tobacco tissue culture. Physiol Plant **15:** 471–497

O'Mahony PJ, Oliver MJ (1999) The involvement of ubiquitin in vegetative desiccation tolerance. Plant Mol Biol **41:** 657–667

Pearce RS (1999) Molecular analysis of acclimation to cold. Plant Growth Regul **29:** 47–76

Perret E, Ferrán EA, Marinx O, Liauzun P, Dumont X, Fournier J, Kaghad M, Ferrara P, Caput D (1998) Improved differential screening approach to analyze transcription variations in organized cDNA libraries. Gene **208:** 103–115

Ren T, Qu F, Morris TJ (2000) *HRT* gene function requires interaction between a NAC protein and viral capsid protein to confer resistance to turnip crinkle virus. Plant Cell **12:** 1917–1925

Richmond T (2000) Higher plant cellulose synthases. Genome Biol **1:** 3001.1–3001.6

Rose D (2000) Microfluidic technologies and instrumentation for printing DNA microarrays. *In* M Schena, eds, Microarray Biochip Technology, Ed 1. Eaton Publishing, BioTechniques Books Division, Natick, MA, pp 19–37

Sáez-Vásquez J, Gallois P, Delseny M (2000) Accumulation and nuclear targeting of BnC24, a *Brassica napus* ribosomal protein corresponding to a mRNA accumulating in response to cold treatment. Plant Sci **156:** 35–46

Saijo Y, Hata S, Kyozuka J, Shimamoto K, Izui K (2000) Over-expression of a single Ca^{2+}-dependent protein kinase confers both cold and salt/drought tolerance on rice plants. Plant J **23:** 319–327

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual, Ed 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Sato Y, Murakami T, Funatsuki H, Matsuba S, Saruyama H, Tanida M (2001) Heat shock-mediated APX gene expression and protection against chilling injury in rice seedlings. J Exp Bot **51:** 1679–1686

Schena M, Heller RA, Theriault TP, Konrad K, Lachenmeier E, Davis RW (1998) Microarrays: biotechnology's discovery platform for functional genomics. Trends Biotechnol **16:** 301–306

Schultz J, Copley RR, Doerks T, Ponting CP, Bork P (2000) SMART: a web-based tool for the study of genetically mobile domains. Nucleic Acids Res **28:** 231–234

Schummer M, Ng W-L, Bumgarner RE, Nelson PS, Schummer B, Bednarski DW, Hassell L, Baldwin RI, Karlan BY, Hood L (1999) Comparative

hybridization of an array of 21500 ovarian cDNAs for the discovery of genes overexpressed in ovarian carcinomas. Gene **238:** 375–385

Schummer M, Ng W-L, Nelson PS, Bumgarner RB, Hood L (1997) A simple high-performance DNA arraying device for comparative expression analysis of a large number of genes. Biotechnology **23:** 1087–1092

Seki M, Narusaka M, Abe H, Kasuga M, Yamaguchi-Shinozaki K, Carninci P, Hayashizaki Y, Shinozaki K (2001) Monitoring the expression pattern of 1,300 Arabidopsis genes under drought and cold stress by using a full-length cDNA microarray. Plant Cell **13:** 61–72

Shen Q, Chen CN, Brands A, Pan SM, Ho TH (2001) The stress- and abscisic acid-induced barley gene HVA22: developmental regulation and homologues in diverse organisms. Plant Mol Biol **45:** 327–340

Singh K, Dennis ES, Ellis JG, Llewellyn DJ, Tokuhisa JG, Wahleithner JA, Peacock WJ (1990) OCSBF-1, a maize ocs enhancer binding factor: isolation and expression during development. Plant Cell **2:** 891–903

Souer E, van Houwelingen A, Kloos D, Mol J, Koes R (1996) The *no apical meristem* gene of petunia is required for pattern formation in embryos and flowers and is expressed at meristem and primordia boundaries. Cell **85:** 159–170

Stockinger EJ, Mao Y, Regier MK, Triezenberg SJ, Thomashow MF (2001) Transcriptional adaptor and histone acetyltransferase proteins in Arabidopsis and their interactions with CBF1, a transcriptional activator involved in cold-regulated gene expression. Nucleic Acids Res **29:** 1524–1533

Tai PYP, Lentini RS (1998) Freeze damage of Florida sugarcane. *In* DL Anderson, eds, Sugarcane Handbook, Ed 1. Florida Cooperative Extension Service, University of Florida, Gainesville, FL, pp 1–3

Tamminen I, Makela P, Heino P, Palva ET (2001) Ectopic expression of ABI3 gene enhances freezing tolerance in response to abscisic acid and low temperature in *Arabidopsis thaliana*. Plant J **25:** 1–8

Telles GP, da Silva FR (2001) Trimming and clustering sugarcane ESTs. Gen Mol Biol **24:** 17–23

Thomashow MF (1998) Role of cold-responsive genes in plant freezing tolerance. Plant Physiol **118:** 1–7

Thomashow MF (2001) So what's new in the field of plant cold acclimation? Lots! Plant Physiol **125:** 89–93

Tomczak MM, Hincha DK, Estrada SD, Wolkers WF, Crowe LM, Feeney RE, Tablin F, Crowe JH (2002) A mechanism for stabilization of membranes at low temperatures by an antifreeze protein. Biophys J **82:** 874–881

van der Luit A, Olivari C, Haley A, Knight MR, Trewavas AJ (1999) Distinct calcium signaling pathways regulate calmodulin gene expression in tobacco. Plant Physiol **121:** 705–714

Vettore A, da Silva FR, Kemper EL, Arruda P (2001) The libraries that made SUCEST. Gen Mol Biol **24:** 1–7

Wegrzyn T, Reilly K, Cipriani G, Murphy P, Newcomb R, Gardner R, MacRae E (2000) A novel alpha-amylase gene is transiently up-regulated during low temperature exposure in apple fruit. Eur J Biochem **267:** 1313–1322

Williams LE, Lemoine R, Sauer N (2000) Sugar transporters in higher plants: a diversity of roles and complex regulation. Trends Plant Sci **5:** 283–290

Wu G, Wilen RW, Robertson AJ, Gusta LV (1999) Isolation, chromosomal localization, and differential expression of mitochondrial manganese superoxide dismutase and chloroplastic copper/zinc superoxide dismutase genes in wheat. Plant Physiol **120:** 513–520

Xie Q, Sanz-Burgos AP, Guo H, Gárcia JA, Gutiérrez C (1999) GRAB proteins, novel members of the NAC domain family, isolated by their interaction with a geminivirus protein. Plant Mol Biol **39:** 647–656

Xin Z, Browse J (2000) Cold comfort farm: the acclimation of plants to freezing temperatures. Plant Cell Environ **23:** 893–902

Xu P, Huecksteadt TP, Hoidal JR (1996) Molecular cloning and characterization of the human xanthine dehydrogenase gene (XDH). Genomics **34:** 173–180

Yamaguchi-Shinozaki K, Shinozaki K (1994) A novel *cis*-acting element in an Arabidopsis gene is involved in responsiveness to drought, low-temperature, or high-salt stress. Plant Cell **6:** 251–264

Yu XM, Griffith M (1999) Antifreeze proteins in winter rye leaves form oligomeric complexes. Plant Physiol **119:** 1361–1370

Zhang H, Wang J, Nickel U, Allen RD, Goodman HM (1997) Cloning and expression of an Arabidopsis gene encoding a putative peroxisomal ascorbate peroxidase. Plant Mol Biol **34:** 967–971

Zhu B, Choi DW, Fenton R, Close TJ (2000) Expression of the barley dehydrin multigene family and the development of freezing tolerance. Mol Gen Genet **264:** 145–153

Capítulo III

Plant energy-dissipating systems: depicting the genomic structure and expression profiles of uncoupling protein and alternative oxidase gene families in monocots and dicots

Jiří Borecký, **Fábio T.S. Nogueira**, Ivan G. Maia, Felipe R. da Silva, Aníbal E.Vercesi, Paulo Arruda

Plant Energy-Dissipating Systems: Depicting the Genomic Structure and Expression Profiles of Uncoupling Protein and Alternative Oxidase Gene Families in Monocots and Dicots

Jiří Borecký[‡,§,1], Fábio T.S. Nogueira[§,‖,1], Ivan G. Maia[§,2], Felipe R. da Silva[§], Aníbal E.Vercesi[‡], and Paulo Arruda[§,‖,*]

[‡] Departamento de Patologia Clínica (NMCE), FCM, Universidade Estadual de Campinas (UNICAMP), 13083-970, Campinas, SP, Brazil

[§] Centro de Biologia Molecular e Engenharia Genética, Universidade Estadual de Campinas (UNICAMP), 13083-970, Campinas, SP, Brazil

[‖] Departamento de Genética e Evolução, Instituto de Biologia, Universidade Estadual de Campinas (UNICAMP), 13083-970, Campinas, SP, Brazil

Keywords: plant uncoupling protein, alternative oxidase, gene family, chilling stress, *Arabidopsis thaliana,* sugarcane

FOOTNOTES

[1] These authors contributed equally to this work.

[2] Present address: Instituto de Biociências, Departamento de Genética, Universidade Estadual Paulista (UNESP), 18618-000, Botucatu, SP, Brazil. E-mail: igmaia@ibb.unesp.br

* Corresponding author
Centro de Biologia Molecular e Engenharia Genética
Universidade Estadual de Campinas (UNICAMP),
CP 6109, 13083-970, Campinas, SP, Brazil
E-mail: parruda@unicamp.br

Running title: Plant uncoupling protein and alternative oxidase gene families

Abstract

Plant mitochondria contain an alternative oxidase (AOx) and an uncoupling protein (PUMP), both involving in energy metabolism tuning, and limiting of reactive oxygen species production. However, why plants need these two energy-dissipating systems? Here we describe a catalogue of PUMP and AOx multigene families, using dicot *Arabidopsis thaliana* and monocot sugarcane (*Saccharum* sp.) models. Four novel *AtPUMPs* were identified in *Arabidopsis* and five *SsPUMPs* and four *SsAOxs* in sugarcane. Tissue-enriched expression of *PUMPs* was more ubiquitous than of *AOxs* in both plants. Distinct expression patterns among gene family members were observed between monocots and dicots and during chilling stress, suggesting different transcriptional regulation of each energy-dissipating system member in plants.

1. Introduction

The proton electrochemical gradient ($\Delta\mu_H+$) energy-dissipating pathway involving plant uncoupling mitochondrial protein (PUMP) and the redox energy-dissipating pathway involving an alternative ubiquinol cyanide-resistant oxidase (AOx) lead to the same final effect - a decrease in the oxidative phosphorylation efficiency and an increase in heat production. PUMPs are probably involved in optimizing $\Delta\mu_H+$ across the inner mitochondrial membrane, while AOxs function predominantly in thermogenic processes [1-2]. Although both systems are present in the plant mitochondria, PUMPs and AOxs do not appear to work at their maximal rates simultaneously, since increasing concentrations of free fatty acids (FFA) block the activity of AOx (K_i ~4 μM) while activating PUMP (K_m ~10 μM; [3]). Studies analyzing ripening tomato [3-4] or mango fruit [5] have shown that AOx and PUMP may work sequentially.

Members of the uncoupling protein family display in general similar biochemical properties [6-15]. At present, five members in animals (UCP1-5 [16-20]) and two in plants (AtPUMP1-2 [21-22]) compose this family. Furthermore, two discrete gene families encoding alternative oxidases were described in plants. The AOx1 type family is supposedly induced by different types of stress and is present in both monocot and dicot plants. The AOx2 type family is usually expressed in a constitutive manner in dicots, but is absent in all monocots examined [23]. This molecular distinction suggests a divergence in AOx across plant families and may even have implications for its physiological role in different plant species. The properties and physiological roles of different PUMP and AOx isoforms have been investigated. Evidences for a role in the regulation of energy metabolism in mitochondria [24-26] in order to avoid an extremely high $\Delta\mu_H+$, which can lead to the excessive production of reactive oxygen species (ROS) have been provided [26-28].

Here we characterize a probably complete gene family of uncoupling proteins in plants, which consists of six members, *PUMP1-6*, found in the whole genome of the dicotyledonous *Arabidopsis thaliana* or in a database of 237,954 ESTs of monocotyledonous sugarcane (SUCEST) [29]. We also identified four members of the AOx1 gene family in sugarcane. The expression profiles of *PUMPs* and

AOxs in specific organs and during chilling stress were also examined. The gene regulation of these two energy-dissipating systems is discussed.

2. Materials and methods

2.1. Plant growth and chilling stress treatment

Sugarcane plantlets (*Saccharum* sp. cultivar SP80-3280), propagated axenically *in vitro* [30], were kept at 26°C on a 16 h/8 h day/night cycle with a photon flux density of 70 μE m^{-2} s^{-1}. Three-month-old plantlets were transferred to 4°C under the same photoperiod conditions. Control plantlets were maintained at 26°C. The leaves of control and chilling-treated plantlets were harvested after 0, 6, 12, 24, and 48 h of treatment (six plantlets per time point). The expression patterns of the *SsPUMP* and *SsAOx* genes were examined using one-month-old sugarcane plantlets cultivated in a greenhouse. Floral organs were obtained from field grown sugarcane plants.

Seeds of *Arabidopsis thaliana* ecotype Columbia (Col-0) were surface-sterilized and plated in Petri dishes containing MS medium [31]. Twenty one-day-old seedlings were incubated either at 4°C or at room temperature (22°C) and the aboveground organs were harvested after 0, 3, 6, 12, 24, and 48 h of treatment (60 seedlings per time point). To analyze the expression profiles of *AtPUMPs* and *AtAOxs* in different plant organs, seeds were germinated under controlled environment conditions: 70 μE m^{-2} s^{-1}, day/night temperature of 22°C, and a photoperiod 16 h. Tissues samples were obtained from ten-week-old plants.

2.2. Identification and in silico analysis of sequences encoding sugarcane and Arabidopsis PUMPs and AOxs

The *Arabidopsis thaliana* genome database (TIGR, http://www.tigr.org/tdb/ath1/htmls/ index.html) was searched for sequences homologous to *UCP*/*PUMP* or *AOx* using the tBLASTn algorithm [32] and UCP1-5, AtPUMP1-2, AtAOx1a-1c, and AtAOx2 as drivers. The same approach, including the novel *Arabidopsis* homologues, was used to identify SASs (sugarcane assembled sequences;

http://sucest.lad.ic.unicamp.br/public [29]) sharing similarities to *PUMPs* and *AOxs*. Sugarcane *PUMP* and *AOx* sequences were also identified by keyword searches using a provisional functional assignment generated by automated annotation, and the resulting hits were validated against existing homologues in public databases (NCBI, http://www.ncbi.nlm.nih.gov).

To obtain the full-length cDNA sequences that represented both sugarcane *PUMP* and *AOx* genes, the longest EST of each identified SAS was resequenced using an ABI PRISM 3700 DNA sequencer and sequence data were analyzed using the PHRED-PHRAP package (http://www.phrap.org/).

The full-length cDNAs representing the *AtPUMP* and *AtAOx* genes were obtained by RT-PCR, using clone-specific forward and reverse (in parentheses) primers for each gene: *AtPUMP1*, 5'-GCTTTAGCCGTAATCGTCG-3' (5'-GTTGCTCTCATTCCCTCTGC-3'), *AtPUMP2*, 5'-CAATGGCGGATTTCAAAC-3' (5'-CTAGGGATCTGAGAATCAATCG-3'), *AtPUMP3*, 5'-ATGGAGCGGAGCCGAGTG-3' (5'-CGCGTTAATCAGAAACTGATGC-3'), *AtPUMP4*, 5'-ATGGGAGTCAAAAGTTTCG-3' (5'-TCAAAAATCTCGAAGCAGC-3'), *AtPUMP5*, 5'-CCAGTGAGATCAGCTCCAATTC-3' (5'-CGCCACCATCATTCATCTTC-3'), *AtPUMP6*, 5'-TGTCTTCTTCCATTCCAAC-3' (5'-TGAAGAAATATGGGAATCTC-3'), *AtAOx1a*, 5'-ATGATGATAACTCGCGGTGGAGC-3' (5'-GCAACATTCAAAGAAAGCCGAATC-3'), *AtAOx1b*, 5'-ATGATGATGAGTCGTCGCTATG-3' (5'-CCCATTAAAGCCCATTTAGG-3'), *AtAOx1c*, 5'-CAAATCTCCCTTGAATCCG-3' (5'-GCTCTTCTGATTCAGTGATATCC-3'), *AtAOx1d*, 5'-CCCAACTGTTGTTACTCATG-3' (5'-CACAGCTTTGTGACTTTGTC-3'), and *AtAOx2*, 5'-ATGGGTATGAGTTCTGCATCG-3' (5'-TTAGTGATAACCAATCGGAGCTG-3'). Reverse transcription was done on 5 μg of total RNA extracted from *Arabidopsis* seedlings using Superscript II RNase H^- reverse transcriptase (Invitrogen, USA) under the recommended conditions. PCR was carried out using the following conditions: 94°C for 3 min and 35 cycles of 94°C for 1 min, 52 to 60°C for 45 s, and 72°C for 1 min and 30 s. The single PCR products were cloned into the pGEM-T EASY vector (Promega, USA) and resequenced.

The amino acid sequences deduced from the cDNAs of the *PUMP* candidates mined in the databases were initially aligned with the corresponding protein sequences of all known UCPs/PUMPs using the CLUSTALX program (weight matrix Gonnet 250) [33]. The aligned sequences were grouped and subsequently aligned with sequences representative of known UCPs/PUMPs, together with malate/2-oxoglutarate carriers (M2OMs), and the dicarboxylate carriers (DICs) as the phylogenetically closest mitochondrial anion carrier family (MACF) members and corresponding phylogenetic trees (1000 bootstraps) were inferred using the MEGA2 program [34]. Sequences (accession numbers in parentheses) used for phylogenetic analyses: UCP1_BOVIN (P10861), UCP1_HUMAN (P25874), UCP1_MESAU (P04575), UCP1_MOUSE (P12242), UCP1_RABIT (P14271), UCP1_RAT (P04633), DgUCP1 (AAM49148), PsUCP1 (AAG33983), UCP2_BRARE (Q9W720), UCP2_CANFA (Q9N2J1), UCP2_CYPCA (Q9W725), UCP2_HUMAN (P55851), UCP2_MOUSE (P70406), UCP2_PIG (O97562), UCP2_RAT (P56500), PmUCP2 (AAL92117), PsUCP2 (AAG33984), UCP3_BOVIN (O77792), UCP3_CANFA (Q9N2I9), UCP3_HUMAN (P55916), UCP3_MOUSE (P56501), UCP3_PIG (O97649), UCP3_RAT (P56499), PsUCP3 (AAG33985), UCP4_HUMAN (O95847), MmUCP4 (BAC66453), RnUCP4a (CAC20898), RnUCP4b (CAC20899), RnUCP4c (CAC20900), UCP5_MOUSE (Q9Z2B2), MmUCP5 (NP_035528), RnUCP5 (NP_445953), UCP5_HUMAN (O95258), CeUCP (NP_505414), EmUCP (AAK16829), HmUCP (BAC06495), UCP-gallus (AAL35325.2), MgUCP (AAL28138), XlUCPput (AAH44682), UCPhom-yeast (S25357), AtPUMP1 (CAA11757), AtPUMP2 (NP_568894), AtPUMP3 (F7A19_22), AtPUMP4 (F22K18_230), AtPUMP5 (F14M13_10), AtPUMP6 (T5E8_270), GmPUMP1a (AAL68562), GmPUMP1b (AAL68563), MiPUMP1 (AAK70939), LePUMP (AAL82482), OsPUMP1 (BAB40657), OsPUMP2 (BAB40658), SfUCPa (BAA92172), SfUCPb (BAA92173), TaPUMP1a (BAB16385), TaPUMP1b (BAB16384), ZmPUMP (AAL87666), DIC_HUMAN (Q9UBX3), DIC_MOUSE (Q9QZD8), RnDCput (NP_596909), CeM2OM (NP_509133), NcDCput (XP_327953), M2OM_MOUSE (Q9CR62), M2OM_RAT (P97700), M2OM-pig (AAD01440), OaM2OM (AAF44754), M2OM_HUMAN (Q02978), M2OM_BOVIN (P22292), NtDTC (CAC84545), AtM2OM (NP_197477), PlaFalM2OM (CAD51134), PlaYoeM2OM (EAA21506), PmM2OM2 (S65042), StM2OM (CAA68164),

SsM2OM (SCRURT2010A08.g), DNC_HUMAN (Q9HC21), ODC1_YEAST (Q03028), and ODC2_YEAST (Q99297).

Further, the complete amino acid sequences sharing homology with known UCPs, PUMPs, M2OMs, and DICs were analyzed using the pattern prediction program package MEME-MAST, version 3.0 (http://www.sdsc.edu/MEME) [35]. Finally, the identified sequences of sugarcane and *Arabidopsis* were checked for i) the number of a 10-amino acid sequence motif known as the mitochondrial energy transfer signature (METS; Prosite PDOC00189) and ii) the presence of these METS in variants specific for uncoupling proteins [2].

Similar *in silico* analysis was done for AOx candidates. Representative fungal and plant mitochondrial AOx and plastidic AOxs sequences were retrieved from the NCBI database. AOx sequences (accession numbers in parentheses) used for phylogenetic analyses: AcaAOx (AAD29681), AfuAOx (AAL87459), AniAOx (AAN39883), AOX_ASPNG (O74180), AtAOx1a (Q39219), AtAOx1b (O23913), AtAOx1c (O22048), AtAOx1d (NP_564395), AtAOxput (NP_567658), AOX2_ARATH (O22049), BfAOx (CAD42731), BgAOx (AAL56983), CaAOx0a (AAF21993), CaAOx0b (AAC98914), CneAOx (AAM22475), CrAOx0a (AAG33633), CrAOx0b (T07947), CroAOx (BAA23803), CsAOx2a (AAP35170), CsAOx2b (EAA32850), Cw80AOx (BAA23725), DdAOx (BAB82989), EniAOx (BAA93615), GspAOx (AAN39884), AOX_HANAN (Q00912), LeAOx1a (AAK58482), LeAOx1b (AAK58483), LeAOx1c (AAP92756), LeAOx2 (AAP92755), AOX1_MANIN (Q40294), MfAOx (AAL24516), MgAOx (AAG49588), MiAOx1a (AAK70935), MiAOx1b (AAK70936), MiAOx1c (AAK70937), AOX_NEUCR (Q01355), OsAOx1a (BAA28772), OsAOx1b (BAA28771), OsAOx1c (BAB71944), OsAOxPut (AAL76179), PaAOx (AAK58849), PapAOx (CAE11918), PtPtAOx1a (CAB64356), PtPtAOx1b (CAB72441), AOX1_SAUGU (P22185), AOX1_SOYBN (Q07185), AOX2_SOYBN (Q41266), AOX3_SOYBN (O03376), StAOx1 (2208475A), TaAOx1a (BAB88646), TaAOx1b (BAB88645), TaAOx1c (BAB88646), AOX1_TOBAC (Q41224), AOX2_TOBAC (Q40578), AOX_TRYBB (Q26710), VinAOx0a (AAK61349), VinAOx0b (AF279690), VuAOx2a (CAC42836),

ZmAOX1 (AAL27795), ZmAOX2 (AAL27796), ZmAOX3 (AAL27797), VuAOx2b (CAD12835), YlAOx (CAD21442), and LeAOxPlastid (AAG02286).

2.3. Digital mRNA expression profiling

The digital expression profiling of the *SsPUMP* and *SsAOx* transcripts among sugarcane tissues was analyzed as an estimate of their relative abundance calculated as a count of the number of ESTs in the given library pool normalized by the total number of ESTs in each SUCEST library pool (pooled from 25 different EST libraries prepared from mRNA isolated from different tissues [29]). The library pools are listed as following (total number of ESTs from each pool in parentheses): IL – plants inoculated with endophytic bacteria (24,430), M – meristem (39,116), F – floral organs (52,430), L – leaves (19,676), R – roots (25,302), S – seeds (17,106), and ST – stem (49,785).

The *Arabidopsis* ESTs were retrieved from libraries described in the *Arabidopsis* Gene Index database at The Institute for Genomic Research (TIGR; http://www.tigr.org/tdb/tgi/agi). Libraries containing more than 9,000 ESTs and completely described for tissues or organs were used to further digital mRNA expression profiling analysis. All subtractive and normalized libraries were removed from the analysis and selected EST libraries were pooled into five library pools. The library pools are listed (total number of ESTs from each pool in parentheses): A – aboveground organs (12,985), R – roots (17,573), F – floral organs (9,120), G – green siliques (12,589), and S – seeds (10,800). The sequences representing *AtPUMP* and *AtAOx* transcripts were used as drivers against TIGR *Arabidopsis* gene index and corresponding tentative consensus (TC) sequences (227,742 TCs as of April 24, 2003) [36] were analyzed in order to estimate the relative abundance of each gene in the selected library pools.

Nuclear-encoded subunits of ATP-synthase (subunit γ, accession D88374 or SAS SCJFLR1013C03.g) and two respiratory chain representatives, Complex I (51 kDa subunit, accession NM_120938 or SAS SCCCCL3003D05.b) and Complex IV (subunit Vb, accession NM_112434 or SAS SCSFLR2031H06.g) were included in the analysis as mitochondrial nuclear gene controls.

2.4. RNA gel blot analysis

Total RNA was isolated from chilling-treated and untreated sugarcane and *Arabidopsis* tissues using Trizol reagent (GibcoBRL, USA) according to the manufacturer's instructions. Twenty micrograms of total RNA were electrophoresed in a 1% (w/v) agarose gel containing formaldehyde and transferred to a Hybond-N^+ filter (Amersham Pharmacia Biotech, USA). The filters were hybridized with full-length cDNA fragments of *SsPUMPs*, *SsAOxs*, *AtPUMPs*, and *AtAOxs* labeled with α-^{32}P dCTP. Hybridization was performed as described [37]. Digitalized images of the RNA-blot hybridization signals were analyzed using the Image Gauge software (Fujifilm, Japan).

2.5. Semi-quantitative RT-PCR

Semi-quantitative RT-PCR was used to detect the accumulation of *AtPUMP4*, *AtPUMP5*, *AtAOx1b*, and *AtAOx1d* transcripts. cDNA was synthesized from 500 ng (*AtAOx1b-1d* and *AtPUMP5*) and 1 μg (*AtPUMP4*) of total RNA from leaves, floral organs, and roots using Superscript II RNase H^- RT (Invitrogen, USA). Primers for the constitutive gene encoding adenine phosphoribosyltransferase (*APRT*) [38] were used as an internal control to normalize the quantity of total RNA used in each sample. The PCR was done with a 12.5-fold dilution of the first-strand cDNA under the following conditions: 94°C for 3 min and 27 cycles of 94°C for 1 min, 52 to 60°C for 45 s, and 72°C for 1 min and 30 s. The experiment was repeated twice.

3. Results

3.1. PUMP gene family

Homology-based searches (tBLASTn algorithm; E-value $<10^{-40}$) in the *Arabidopsis* genome database yielded 21 genomic sequences (BAC, TAC, and PL clones) comprising *PUMP*-like ORFs. Alignment of the protein sequences deduced from these clones with known UCPs and PUMPs, together with M2OMs and DICs, led to the identification of six *Arabidopsis* genes highly similar to known

UCPs/PUMPs. Using the same search that included all the six *Arabidopsis* sequences as drivers, additional 28 *PUMP*-like sugarcane sequences representing EST clusters produced by the assembly of 237,954 ESTs [29], were found in the SUCEST database (data not shown). Phylogenetic analysis of the six *Arabidopsis* and five sugarcane PUMPs along with the UCPs/PUMPs plus M2OM and DIC representatives revealed a complex group that consisted exclusively of uncoupling proteins of all types (Fig. 1A, groups I-VI), distinct from the branch containing M2OMs and DICs (Fig. 1A, group VII).

The *PUMP* genes identified within the *Arabidopsis* genome are summarized in the Table 1. *AtPUMP1* and *AtPUMP2* (*AtUCP2*) were characterized previously [21,39]. The sequence of *AtPUMP3* was previously submitted to GenBank as a "putative mitochondrial uncoupling protein". The sequences representing *AtPUMP4-6* were identified for the first time in this work. Mapping of these *AtPUMP* genes revealed that each gene is located on a different chromosome, except for *AtPUMP2* and *AtPUMP6*, which are on the opposite ends of chromosome 5 (Table 1). Figure 2A illustrates the genomic structure of *PUMP* genes in the *Arabidopsis* genome. Intron/exon composition of *AtPUMP1* and *AtPUMP2* were almost identical, with only first exon and first and sixth introns differing in their size [2]. Genomic structures of *AtPUMP3* and *AtPUMP6* were also similar. *AtPUMP4* and *AtPUMP5* genes do not have any intron. Intronless genes homologous to *AtPUMP4* and *AtPUMP5* were also detected in rice and maize (accession 8356.t03613 and accession AZM3_88176, respectively) by searches in TIGR databases (http://www.tigr.org/tdb).

The five sugarcane PUMP orthologs identified in this work were denominated *SsPUMP1-5* (*Saccharum* sp. *PUMP1-5*; Table 2). Comparison of the protein sequences of SsPUMPs with PUMPs from other plants revealed a new specific insertion of four alanines in the proximity of the first energy transfer protein signature (data not shown), which may represent a putative monocotyledonous PUMP-specific motif. This motif is also present in PUMPs from maize [40], wheat [41], and rice [42] but absent in PUMPs from dicotyledonous plants. In addition, genes encoding putative new M2OMs were also detected in the *Arabidopsis* and sugarcane databases (named here as *AtM2OM* and *SsM2OM*, respectively), and were phylogenetically much closer to known M2OMs than PUMP4-6 (Fig. 1A).

All members of the MACF possess the METS P-x-[DE]-x-[LIVAT]-[RK]-x-[LRH]-[LIVMFY] [2]. The known uncoupling proteins have three copies of this signature, while other MACF members possess only one or two copies (Table 3). The three copies of the METS signature found in all of the new PUMPs displayed specific variants for the PUMPs already identified [2], while the M2OM sequences lacked the second signature copy (Table 3). The novel PUMP members were also screened for the presence of proposed four UCP-specific signatures [43]. PUMP3 from both *Arabidopsis* and sugarcane showed all three signatures with the exception of one substitution (a nonpolar amino acid for Tyr) in the middle of the second signature (Table 3). PUMP4-6 differed slightly; they lacked the negatively charged amino acid (Asp or Glu) in the middle of the first signature and had Gly at this position, and the terminal Phe/Leu/Ile was substituted for His (Table 3). In the second signature, there were three substitutions of a nonpolar amino acid for Thr (positions 1, 6, and 11). In the third signature, a negatively charged amino acid was substituted for Gln (position 13) and the first Pro was changed to Leu (Table 3). All of the new PUMPs had a complete fourth signature, found in all UCPs/PUMPs but UCP5 [43]. Moreover, the MEME-MAST prediction of additional conserved motifs revealed a M2OM/DIC-specific motif [AG]PM[TV][VM][LM]T[FLW]IFL[EM]Q[LMI][NRQ]K located at the C-end of the molecule. This motif was not detected in any of the already known or new UCPs/PUMPs. These analyses provided strong evidence that the novel sequences identified in *Arabidopsis* and sugarcane corresponded to true PUMPs and not to any other MACF member.

3.2. Alternative oxidase gene family

The five *Arabidopsis* AOx protein sequences (Table 1, [44-45]) were used as drivers to identify *AOx* orthologs in the SUCEST database. Eleven SASs found as homologous to AOx were reblasted to the GenBank non-redundant database and six of them showed high identity levels (E-value $<10^{-40}$) to known AOxs (*SsAOxs* for *Saccharum* sp. *AOxs*, Table 2). Phylogenetic analysis (Fig. 1B) showed that four SsAOxs (Table 2) together with AtAOx1a-1c were grouped within AOx1 type groups (Fig. 1B, group I and II) but not with AOx2 (Fig. 1B, group III), in agreement with recently published results [23] showing

that monocots lack *AOx2* genes. Interestingly, AtAOx1d together with LeAOx1b and PtPtAOx1b formed a new branch, the third AOx type – AOx3 (Fig. 1B, group IV). Analysis of the genomic organization of the members of *AtAOx* gene family revealed a similar intron/exon distribution pattern for *AtAOx1* type, all of which present four exons that differ in size among each family member. *AtAOx2* gene structure consists of five exons (Fig. 2B).

3.3. Tissue-enriched expression profiling of PUMP and AOx genes

EST datasets have been used recently to extract information on gene expression levels [57]. The rationale of the so-called 'digital mRNA expression profiling' is that the number of EST sequences representing a given mRNA obtained from a cDNA library is proportional to the abundance of this mRNA species in the tissue originally used to make the library [58]. However, some precautions need to be implemented to prevent potentially misleading gene expression results. First, records of individual EST sequences were individually inspected and clone IDs were retrieved and displayed in a catalog, such that EST sequences obtained from opposite ends of the same cDNA clone were counted as one. Second, only non-normalized sugarcane and *Arabidopsis* EST collections were used for this analysis (see Materials and Methods). Therefore, the expression profiling of members of the *PUMP* and *AOx* gene families in several tissues and/or organs of sugarcane and *Arabidopsis* was assessed using the information available in the SUCEST and TIGR *Arabidopsis* EST databases. ESTs representing three respiratory chain genes (ATP-synthase, Complex I, and Complex IV subunits) in both species were found in the most cDNA pools analyzed, suggesting that these libraries are representative of the mitochondrial-nuclear transcriptome (Fig. 3). The results showed that *SsPUMP1* was expressed only in floral organs whereas *SsPUMP3* was found only in stems (Fig. 3A). *SsPUMP2* was expressed in most tissues, but preferentially in the meristem and roots. *SsPUMP4* and *SsPUMP5* were more abundant in stems, leaves, and plantlet inoculated with endophytic bacteria pools [29]; the latter *PUMPs* were expressed at lower levels in roots and seeds (Fig. 3A).

The sugarcane *AOx* genes were expressed at very low levels when compared with the expression of PUMP genes. In general, *AOx* gene expression was observed preferentially in nongreen tissues, in contrast to *SsPUMPs* (Fig. 3A). ESTs of *SsAOx1a* and *SsAOx1b* were found in roots, while *SsAOx1c* was present in the stem and meristem and *SsAOx1d* only in leaves (Fig. 3A).

AtPUMP1 was expressed in several tissues, in agreement with the previous results [39]. *AtPUMP2* was detected in roots and green siliques, whereas *AtPUMP3* was found only in roots (Fig. 3B). *AtPUMP4* showed high levels of expression in roots and seeds, whereas *AtPUMP5* was expressed predominantly in aboveground organs and roots, and less in green siliques. No *AtPUMP6* expression was detected in any tissue (Fig. 3B), suggesting that it could be a pseudogene. *AtAOx1a* expression was observed in aboveground organs and *AtAOx1c* was detected only in floral organs (Fig. 3B).

To validate our digital mRNA expression profiling analysis, we analyzed the expression of *SsPUMP2, SsPUMP4*, *SsAOx1a*, and *SsAOx1d* using RNA-blots. *SsPUMP2* was expressed preferentially in roots, while only traces of its transcript were found in floral organs, leaves, and stems (Fig. 3C). RNA-blots of *SsPUMP4* indicated high levels of expression in non-reproductive tissues, especially in leaves and roots. *SsAOx1a* was expressed at very low levels in all tissues except floral organs, whereas *SsAOx1d* was undetectable in all of the tissues analyzed (Fig. 3C). In general, the RNA-blots agreed with the results obtained by digital expression analysis.

Because of the low EST frequency of *AtAOxs* in the library pools analyzed (Fig. 3B), a semi-quantitative RT-PCR of *AtPUMP4*, *AtPUMP5*, *AtAOx1b*, and *AtAOx1d* was used to validate the digital expression analysis. By RT-PCR, *AtPUMP4-5* were ubiquitously expressed in leaves, floral organs, and roots at considerably high levels. *AtAOx1b* was detected only in floral organs at very low levels whereas *AtAOx1d* was expressed in all three organs, although the expression in roots was low (Fig. 3D). The discrepancies observed between 'digital mRNA expression profiling' and semi-quantitative RT-PCR might be result of different development stages of the tissues/organs used in both analyses, since these genes were reported to be temporally modulated [51]. Despite of these differences, as with sugarcane, the

Arabidopsis AOx genes showed extremely low overall levels of expression when compared with *AtPUMPs* in both analyses.

3.4. PUMP and AOx Expression Profiles in Response to Chilling Stress

RNA-blots revealed that *SsPUMP2* transcript accumulation was unaltered when sugarcane plantlets were submitted to chilling stress (4°C) during 48 h (Fig. 4A). *SsPUMP1* and *SsPUMP3* were undetectable during whole chilling treatment suggesting no induction by low temperature (data not shown). In contrast, *SsPUMP4* and *SsPUMP5* were strongly induced, with the transcript levels of both reaching a maximum after 48 h at 4°C (Fig. 4A). The level of *SsAOx1a* transcripts was unaffected by chilling stress, whereas *SsAOx1c* was slightly induced after 6 h at 4°C, and returned to basal levels after 12 h (Fig. 4A). In contrast, *SsAOx1b* and *SsAOx1d* transcripts were undetectable in plantlets exposed to cold (data not shown).

AtPUMP1 has been described as being chilling-inducible [39]. We therefore analyzed the transcript accumulation of the other *AtPUMPs* identified here. *AtPUMP2-3* were not induced by chilling treatment (Fig. 4B). In contrast, *AtPUMP4-5* were strongly induced after 3-h cold treatment, returning to basal levels after 12 h (Fig. 4B). Although their expression profiles had similar tendencies, the levels of transcript accumulation for both *AtPUMPs* were different (Fig. 4B). *AtPUMP6* transcripts were undetectable throughout the entire exposure to cold (data not shown).

Under chilling stress, the mRNA levels of *AtOx1a* increased with time reaching a maximum level after 48 h at 4°C (Fig. 4B). However, the accumulation of *AtAOx1b* and *-1c* transcripts was undetectable throughout the exposure to cold (data not shown). Surprisingly, *AtAOx1d*, a recently described member of the *Arabidopsis* AOx gene family [23], was clearly downregulated after exposure to low temperature (Fig. 4B), indicating a possibly different role for this *AOx* gene in the response to cold stress. Finally, *AtAOx2* was not responsive to chilling stress.

4. Discussion

4.1. Identification and Tissue-enriched Expression Profiles of PUMP and AOx Gene Families in Sugarcane and Arabidopsis

The roles of PUMP and AOx energy-dissipating systems in plants have been extensively discussed [3-4, 13, 46, 59]. Both systems can play a similar role in mitochondrial energy-linked processes, either through tissue-specific thermogenesis and/or the maintenance of tissue temperature, or by protecting plant cells against oxidative stress and/or tuning the $\Delta\mu_H+$ across inner mitochondrial membrane [1-2].

The results presented here indicate a new perspective for their physiological roles. A whole family of PUMP-encoding genes was identified in dicotyledonous *Arabidopsis thaliana* and in monocotyledonous sugarcane. For *Arabidopsis*, in addition to *AtPUMP1* [39], *AtPUMP2* [21], and a putative gene *AtUCP4* [47] classified here as *AtPUMP3*, three new members, *AtPUMP4*, *AtPUMP5*, and *AtPUMP6*, were identified (Table 1). The location of *AtPUMP1*, *AtPUMP2*, *AtPUMP4*, and *AtPUMP5* within or close to duplicated regions of *Arabidopsis* genome [48] and the gene structure similarities among pairs of the family members suggest these genes originated through gene duplication events (Fig. 2A). These observations is in keeping with the phylogenetic subfamilies defined in Figure 1A, corroborating the existence of three distinct groups of PUMPs as discussed below. Although *AtPUMP6* has one intron (Fig. 2A), it is phylogenetically close to *AtPUMP4-5* (Fig. 1A) and share higher nucleotide similarity with *AtPUMP4-5* than with other *AtPUMPs* (data not shown). We hypothesize that mechanisms of intron gain/loss occurred during the evolution of the *PUMP* gene family, as have been documented for other plant gene families [49].

The five *AtPUMP* orthologs were also identified in sugarcane and denominated as *SsPUMP1-5* (Table 2). The amino acid sequences deduced from the identified genes exhibited the structural features of uncoupling proteins (Table 3). The major PUMP feature, the presence of three METS [2], was conserved in all novel PUMP isoforms. In addition, PUMPs were clearly distinguished in specific branches of the phylogenetic trees when analyzed together with other members of the MACF family (Fig. 1A). In general, type 1 and 2 PUMPs from both *Arabidopsis* and sugarcane are homologous to the animal UCPs belonging

to types 1, 2, and 3 while type 3 PUMPs are similar to UCP4, and PUMPs of type 4, 5, and 6 formed a distinct PUMP group, analogous to the group formed by the UCP5 members (Fig. 1A).

Two *Arabidopsis* DNA clones (At4g24570 and At2g22500) corresponding to *AtPUMP4* and *AtPUMP5*, respectively (Table 1), were classified as putative DICs [50]. The results presented here showed clearly that these genes had the conserved PUMP-specific motifs (Table 3; [2]). In addition, the proposed UCP-specific motifs [43] were also conserved in a high degree. The two substitutions of the negatively charged amino acids for Gly or Gln do not contradict the proposed involvement of these signatures in fatty acid anion binding and translocation [43]. Indeed, the substitutions should increase PUMP affinity to FFA anions. Finally, the identification of an M2OM/DIC-specific motif that is also present in AtM2OM and SsM2OM but absent in the PUMPs described here, provides strong support that PUMP3-6 are true uncoupling proteins in nature.

Four members of *AOx1* family were mined in the sugarcane EST database (Table 1). Alignment of the protein sequences deduced from newly identified genes showed a primary protein structure pattern typical for all AOxs, *i.e.*, a variable N-terminal portion followed by a highly conserved region (data not shown). The phylogenetic analysis showed that four sugarcane (*SsAOx1a-1d*) and three *Arabidopsis* (*AtAOx1a-1c*) *AOx* genes actually belonged to the *AOx1* type. Another *Arabidopsis* gene belonged to the *AOx2* type that was absent in sugarcane (Fig. 1B). The results agree with the findings of Considine et al. [23], who postulated that monocotyledonous plants lack the genes encoding AOx2. Nonetheless, to our surprise, a recently described member of the *Arabidopsis AOx* family [23], AtAOx1d (Fig. 1B), initially classified into the *AOx1* type gene family, clustered with tomato AOx1b (LeAOx1b) and *P. tremula x P. tremuloides* AOx1b (PtPtAOx1b) into a discrete group (group IV, Fig. 1B). This new group was separate from the AOx1 (Fig. 1B, groups I and II) and AOx2 groups (group III, Fig. 1B). Thus, we propose the existence of a new AOx family, AOx3 type, probably present exclusively in dicotyledonous plants.

The existence of multiple members of the PUMP and AOx protein families suggests that the members of each mitochondrial energy-dissipating system are probably subject to different transcriptional regulation, and may be expressed in a cell-, tissue-, or organ-specific manner. *AtAOx* expression profiling

was only recently investigated [45], however, we performed our expression analysis of *AtAOxs* to be able to compare it with that of *AtPUMPs*. The results showed that *PUMP* members were generally expressed at higher levels in both plants when compared with *AOx* members (Fig. 3B and 3D). When seeking for specific gene expression in individual tissues/organs, *PUMP* and *AOx* genes displayed very different tissue-enriched expression patterns in both plant species (Figs. 3). By *in silico* analysis, there was no organ in which all *SsPUMP* members were expressed and in some tissues only two or three *SsPUMPs* were expressed in parallel (Fig. 3A and 3C). Similarity, it seems that *SsAOxs* were expressed differentially in the sugarcane tissues/organs (Fig. 3A and 3C). *In silico* and experimental results suggested that *AtPUMP4* and *AtPUMP5* are the most expressed *AtPUMP* genes, while all *AtAOxs* were expressed at extremely low levels (Figs. 3B and 3D). AOx activity in non-thermogenic plants was reported to be present in most tissues [51] but quite low. In fact, the low abundance of *AOx* transcripts in several soybean tissues should explain the lack of immunodetection of these proteins in mitochondria. Moreover, AOx activity is tissue-dependent and developmental stage-dependent and can vary with growth conditions [51].

4.2. Influence of Chilling Stress on PUMP and AOx Expression

Among the *SsAOx* genes, only *SsAOx1c* increased transiently the transcript level in response to cold stress. Chilling treatment did not induce *SsAOx1a*, whereas *SsAOx1b* and *-1d* were not detected at any time point during the experiment (Fig. 4A). Although no expression was detected for some *AOxs* (and *PUMPs* as well) in both cold-treated plants (data not shown), the possibility exists that they may be downregulated by low temperature. *Arabidopsis AOx1a* was upregulated during chilling stress after a 12 h exposure (Fig. 4B). This gene is induced by antimycin A, an inhibitor of Complex III in the mitochondrial respiratory chain [44]. Moreover, the overexpression of *AOx1a* in transgenic tobacco reduced the formation of mitochondrial ROS, suggesting a possible role for AOx in protecting plant cells against oxidative stress [52].

AOxs has been proposed to have a housekeeping function in respiratory metabolism and a protective function during stress [23]. Nevertheless, the unexpected downregulation of *AtAOx1d* by low

temperature (Fig. 4B) suggested an additional role for AOx proteins. AtAOx1d fell within a novel subgroup of AOx (Fig. 1B, group IV) and its very different response to chilling stress corroborates with a third type of dicot AOx (AOx3), as mentioned above. In *Arabidopsis*, a cold-tolerant plant species, acclimation to low temperatures leads to the modulation of gene expression, including up- and downregulation of many different genes [53]. These authors suggested that transferring *Arabidopsis* from a warm to low temperature produces "waves" of changes in its transcriptome. These "waves" may differentially regulate the gene family members, such as the AOx families.

Analysis of the expression profile of the *SsPUMP* genes during chilling stress revealed that *SsPUMP2* was not chilling-regulated (Fig. 4A). In contrast, a strong response was observed for *SsPUMP4* and *SsPUMP5* after a 12 h exposition to 4°C (Fig. 4A). Interestingly, the EST frequency of *SsPUMP4* and *SsPUMP5* was elevated in libraries prepared from sugarcane plantlets inoculated with endophytic bacteria [29], suggesting that these genes may respond to biotic and abiotic stress signaling. The *Arabidopsis* counterparts of *SsPUMP4* and *SsPUMP5* were also upregulated by exposure to low temperature, but with different transcript accumulation pattern (Fig. 4A). Analogously to other UCPs/PUMPs [39,54], these proteins may be involved in reducing ROS generation during chilling stress in monocot and dicot plants.

As shown in Figure 4, the expression profiles of PUMP compared to the AOx genes were quite different in sugarcane and *Arabidopsis*, suggesting that multiple regulatory pathways were likely to be involved in PUMP and AOx gene expression. Interestingly, computational inspection using PLACE [55] and PlantCARE [56] of the promoter region (within a 2-kb region upstream of the transcription initiation site) of *AtPUMP4*, *AtPUMP5*, and *AtAOx1a* indicated the presence of 2-3 copies of a CCGAC core sequence, which is recognized by the C-repeat/DRE-binding factor that is involved in cold responses [53].

Another interesting finding was that *SsPUMP4* and *-5* were long-term upregulated by chilling stress, while *SsAOx1c* was upregulated transiently (Fig. 4). In *Arabidopsis*, the opposite was observed (Fig. 4), *i.e*, *AtPUMP4* and *AtPUMP5* were induced rapidly and transiently and *AtAOx1a* was long-term upregulated. Sugarcane, a chilling-sensitive species, appears not to have a complete cold-regulated signal pathway, comparing to *Arabidopsis* and other temperate-climate species [30]. This finding suggests that,

in addition to the temporal regulation of *PUMPs* and *AOxs* within a plant species, there could also be species-dependent regulatory pathways in the response to chilling stress.

In summary, the results of this work support the hypothesis that the differential regulation of each member of *PUMP* and *AOx* gene families at the transcriptional level can be very adaptable in their responses to biotic and abiotic stresses involving oxidative stress.

ACKNOWLEDGMENTS

The authors thank A. S. Zanca, S. Meire, G. S. Ferraz, C. R. Brambile and A L. Beraldo for excellent technical assistance and L. Neubauer for providing the *AtPTR* specific primers.

References

[1] Day, D.A., Arron, G.P., and Laties, G.G. (1980) in: Biochemistry of Plants. (Davies, D.D. Ed), Ed. 1, Vol. 4. pp. 197-241, Academic Press Inc., New York.

[2] Borecký, J., Maia, I.G. and Arruda, P. (2001) Biosci. Rep. 21, 201-211.

[3] Sluse, F.E. and Jarmuszkiewicz, W. (2000) Br. J. Med. Biol. Res. 33, 259-268.

[4] Almeida, A.M., Navet, R., Jarmuszkiewicz, W., Vercesi, A.E., Sluse-Goffart, M.C. and Sluse, F.E. (2002). J. Bioenerg. Biomembr. 34, 487-498.

[5] Considine, M.J., Daley, D.O. and Whelan, J. (2001) Plant. Physiol. 126, 1619-1629.

[6] Almeida, A.M., Jarmuszkiewicz, W., Khomsi, H., Arruda, P., Vercesi, A.E. and Sluse, F.E. (1999) Plant Physiol. 119, 1-7.

[7] Costa, A.D.T., Nantes, I.L., Ježek, P., Leite, A., Arruda, P. and Vercesi, A.E. (1999) J. Bioenerg. Biomembr. 31, 527-533.

[8] Jarmuszkiewicz, W., Almeida, A.M., Sluse-Goffart, C., Sluse, F.E. and Vercesi, A.E. (1998) J. Biol. Chem. 273, 34882-34886.

[9] Ježek, P., Costa, A.D.T. and Vercesi, A.E. (1996) J. Biol. Chem. 271, 32743-32748.

[10] Ježek, P., Costa, A.D.T. and Vercesi, A.E. (1997) J. Biol. Chem. 272, 24272-24278.

[11] Kowaltowski, A.J., Costa, A.D.T. and Vercesi, A.E. (1998) FEBS Lett. 425, 213-216.

[12] Nantes, I.L., Fagian, M.M., Catisti, R., Arruda, P., Maia, I.G. and Vercesi, A.E. (1999) FEBS Lett. 457, 103-106.

[13] Sluse, F.E., Almeida, A.M., Jarmuszkiewicz, W. and Vercesi, A.E. (1998) FEBS Lett. 433, 237-240.

[14] Vercesi, A.E., Chaimovich, H. and Cuccovia, I.M. (1997) Recent Res. Devel. Plant Physiol. 1, 85-91.

[15] Vercesi, A.E., Ježek, P., Costa, A.D.T., Kowaltowski, A.J., Maia, I.G. and Arruda, P. (1998) in: Plant Mitochondria: From Gene to Function (Moller, I.M., Gardeström, P., Glimelius, K. and Glaser E. Eds), Ed 1. pp. 435-440, Backhuys Publishers, Leiden, The Netherlands.

[16] Ricquier, D. and Kader, J.-C. (1976) Biochem. Biophys. Res. Commun. 73, 577-583.

[17] Fleury, C., Neverova, M., Collins, S., Raimbault, S., Champigny, O., Levi-Meyrueis, C., Bouillaud, F., Seldin, M.F., Surwit, R.S., Ricquier, D. and Warden, C.H. (1997) Nat. Genet. 15, 269-272.

[18] Boss, O., Samec, S., Paoloni-Giacobino, A., Rossier, C., Dullo, A., Seydoux, J., Muzzin, P. and Giacobino, J.-P. (1997) FEBS Lett. 408: 39-42.

[19] Mao, W., Yu, X.X., Zhong, A., Li, W., Brush, J., Sherwood, S.W., Adams, S.H. and Pan, G. (1999) FEBS Lett. 443, 326-330.

[20] Sanchis, D., Fleury, C., Chomiki, N., Goubern, M., Huang, Q., Neverova, M., Gregoire, F., Easlick, J., Raimbault, S., Levi-Meyrueis, C. et al. (1998) J. Biol. Chem. 273, 34611-34615.

[21] Vercesi, A.E., Martins, I.S., Silva, M.A.P., Leite, H.M.F., Cuccovia, I.M. and Chaimovich, H. (1995) Nature 375, 24.

[22] Watanabe, A., Nakazono, M., Tsutsumi, N., Hirai, A. (1999) Plant Cell. Physiol. 40, 1160-1166.

[23] Considine, M.J., Holtzapffel, R.C., Day, D.A., Whelan, J. and Millar, A.H. (2002) Plant Physiol. 129, 949–953.

[24] Ježek, P. and Garlid, K.D. (1998) Int. J. Biochem. Cell Biol. 30, 1163-1168.

[25] Ricquier, D. and Bouillaud, F. (2000) Biochem. J. 345, 161-179.

[26] Skulachev, V.P. (1998) Biochem. Biophys. Acta Bioenerg. 1363, 100-124.

[27] Brandalise, M., Maia, I.G., Borecký, J., Vercesi, A.E. and Arruda, P. (2003) J. Bioenerg. Biomembr. 35, 203-209.

[28] Considine, M.J., Goodman, M., Echtay, K.S., Laloi, M., Whelan, J., Brand, M.D. and Sweetlove, L.J. (2003) J. Biol. Chem. 278, 22298-22302.

[29] Vettore, A.L., da Silva, F.R., Kemper, E.L., Souza, G.M., da Silva, A.M., Ferro, M.I.T., Henrique-Silva, F., Giglioti, E.A., Lemos, M.V.F., Coutinho, L.L. et al. (2003) Genome Res. 13: 2725-2735.

[30] Nogueira, F.T.S., De Rosa Jr., V.E., Menossi, M., Ulian, E.C. and Arruda, P. (2003) Plant Physiol. 132: 1811-1824.

[31] Murashige, T. and Skoog, F. (1962) Physiol. Plant. 15, 471-497.

[32] Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Nucleic Acids Res. 25, 3389-3402.

[33] Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997) Nucleic Acids Res. 25, 4876-4882.

[34] Kumar, S., Tamura, K., Jakobsen, I.B. and Nei, M. (2001) Bioinformatics 17, 1244-1245.

[35] Bailey, T.L. and Gribskov, M. (1998) Bioinformatics 14, 48-54.

[36] Quackenbush, J., Liang, F., Holt, I., Pertea, G. and Upton, J. (2000) Nucleic Acids Res. 28, 141-145.

[37] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular cloning: a laboratory manual, Ed 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

[38] Moffatt, B.A., McWhinnie, E.A., Agarwalb, S.K. and Schaff, D.A. (1994). Gene 143, I211-I216.

[39] Maia, I.G., Benedetti, C.E., Leite, A., Turcinelli, S.R., Vercesi, A.E. and Arruda, P. (1998) FEBS Lett. 429, 403-406.

[40] Brandalise, M., Maia, I.G., Borecký, J., Vercesi, A.E. and Arruda, P. (2003) Plant Sci. 165, 329-335.

[41] Murayama, S. and Handa, H. (2000) Mol. Gen. Genet. 264, 112-118.

[42] Watanabe, A. and Hirai, A. (2002) Planta 215, 90-100.

[43] Ježek, P. and Urbánková, E. (2000) IUBMB Life 49, 63-70.

[44] Saisho, D., Nambara, E., Naito, S., Tsutsumi, N., Hirai, A. and Nakazono, M. (1997) Plant Mol. Biol. 35, 585-596.

[45] Thirkettle-Watts, D., McCabe, T.C., Clifton, R., Moore, C., Finnegan, P.M., Day, D.A. and Whelan, J. (2003) Plant Physiol. 133, 1158-1169.

[46] Jarmuszkiewicz, W., Almeida, A.M., Vercesi, A.E., Sluse, F.E. and Sluse-Goffart, C.M. (2000) J. Biol. Chem. 275, 13315–13320.

[47] Hanák, P. and Ježek, P. (2001) FEBS Lett. 24771, 1-5.

[48] Blanc, G., Barakat, A., Guyot, R., Cooke, R. and Delseny, M. (2000) Plant Cell 12, 1093–1101.

[49] Frugoli, J.A., McPeek, M.A., Thomas, T.L. and McClung, C.R. (1998) Genetics 149, 355-365.

[50] Millar, A.H., and Heazlewood, J.L. (2003) Plant Physiol. 131, 443-453.

[51] Finnegan, P.M., Whelan, J., Millar, A.H., Zhang, Q., Smith, M.K., Wiskich, J.T. and Day, D.A. (1997) Plant Physiol. 114, 455-466.

[52] Maxwell, D.P., Wang, Y. and McIntosh, L. (1999) Proc. Natl. Acad. Sci. USA 96, 8271-8276.

[53] Fowler, S. and Thomashow, M.F. (2002) Plant Cell 14, 1675-1690.

[54] Laloi, M., Klein, M., Reismeier, J.W., Müller-Röber, B., Fleury, C., Bouillaud, F. and Ricquier, D. (1997) Nature 389, 135-136.

[55] Higo, K., Ugawa, Y., Iwamoto, M. and Korenaga, T. (1999) Nucleic Ac. Res. 27, 297-300.

[56] Lescot, M., Dehais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., Rouze, P. and Rombauts, S. (2002) Nucleic Ac. Res. 30, 325-327.

[57] Mekhedov S., de Ilarduya, O.M., and Ohlrogge, J. (2000) Plant Physiol. 122, 389-402.

[58] Audic, S., and Claverie, J.M. (1997) Genome Res. 7, 986-995.

[59] Hourton-Cabassa, C., Matos, A. R., Zachowski, A., Moreau, F. (2004) Plant Physiol. Biochem. 42, 283–290.

Table 1. *PUMP* and *AOx* homologues from *Arabidopsis* and their mapping to chromosomes.

Protein	CHR position[a]	CHR locus[b]	No. of ESTs	Highest identity (%)
PUMPs				
AtPUMP1	20656336-20658442	At3g54110	16	UCP2 (45%)
AtPUMP2	30436384-30434148	At5g58970	2	UCP3 (41%)
AtPUMP3	5488903-5487246	At1g14140	1	UCP4 (41%)
AtPUMP4	1782788-1783729	At4g24570	20	UCP3 (31%)
AtPUMP5	10640322-10641674	At2g22500	13	UCP2 (31%)
AtPUMP6	2950514-2949242	At5g09470	0	UCP3 (31%)
AOxs				
AtAOx1a	7904163-7905391	At3g22370	1	PtPtAOx1a (68%)
AtAOx1b	7906799-7908423	At3g22360	0	NtAOx1b (66%)
AtAOx1c	10231842-10230521	At3g27620	2	NtAOx1b (64%)
AtAOx1d	11669668-11666957	At1g32350	0	PtPtAOx1e (67%)
AtAOx2	25293532-25291541	At5g64210	0	VuAOx2a (64%)

[a]Chromosome position in base pairs.

[b]Chromosome locus based on MIPS (http://mips.gsf.de/proj/thal/proj/thal_overview.html).

Table 2. *PUMP* and *AOx* homologues from sugarcane.

Protein	SAS ID[a]	No. of ESTs	Highest identity (%)
PUMPs			
SsPUMP1	SCRFFL5034A04.g	1	UCP2 (47%)
SsPUMP2	SCEPAM2011H12.g	22	UCP3 (45%)
SsPUMP3	SCRLSB1040H12.g	1	UCP4 (48%)
SsPUMP4	SCCCRZ2C01E09.g	31	UCP3 (36%)
SsPUMP5	SCCCLR1076E12.g	37	UCP5 (42%)
AOxs			
SsAOx1a	SCACCL6008G11.g	5	TaAOx1a (77%)
SsAOx1b	SCRURT3064E04.b	2	OsAOx1b (39%)
SsAOx1c	SCCCCL3080G10.b	3	OsAOx1c (48%)
SsAOx1d	SCBGLR1095F09.g	1	OsAOx1c (30%)

[a]Assembled sequences from the SUCEST database [45].

Table 3. Specific variants of the energy transfer protein signature with an additional tripeptide in known UCPs/PUMPs compared with new PUMP genes.

Protein name	1st signature	2nd signature	3rd signature
UCP1[a]	PLDTAKVRL-QLQ	PTEVVKVR[LM]-QAQ	PVDVVKTRF-[IV]NS
UCP2[a]	PLDTAKVRL-QLQ	PTDVVKVRF-QAQ	PVDVVKTRY-MNS
UCP3[a]	PLDTAKVRL-QLQ	PTDVVKVRF-QAQ	PVDVVKTRY-MNS
UCP4[a]	PLDLTKTRL-QMQ	PTDLVKV••-QMQ[b]	PADVIKSRI-MNQ
BMCP1[a]	PVDLTKTRL-QVQ	PTDVLKIRM-QAQ	PVDVVRTRM-MNQ
AtPUMP1-2	PLDTAKVRL-QIQ	PTDLVKVRL-Q[AS]E	PVDV[VM]KSRM-MGD
AtPUMP3	PIDLTKTRM-QLH	PADLVKVRMQAD	PADVVKTRM-MNG
AtPUMP4	PLDLIKVRM-QLQ	PADVAMVRM-QAD	PVDVIKTRV-MNM
AtPUMP5	PLDLIKVRM-QLQ	PADVAMVRM-QAD	PVDVIKTRV-MNM
AtPUMP6	PLDLIKVRM-QLH	PADVAMVRM-QAD	PIDVVKTRM-MN•[b]
SsPUMP1	PLDTAKVRL-QLQ	PTDLVKVRL-QAD	PVDVVKSRM-MGD
SsPUMP2	PLDTAKVRL-QLQ	PTDLVKVRL-QAE	PVDVVKSRM-MGD
SsPUMP3	PLDAVKTRL-QLH	PADLMKVRM-QAD	PADVIKTRM-MNQ
SsPUMP4	PLDLIKVRM-QLQ	PADVAMVRM-QAD	PVDVVKTRM-MNM
SsPUMP5	PLDLIKVRM-QLQ	PADVAMVRM-QAD	PVDVVKTRM-MNM
M2OM[a]	PIDMVKVKI-QLG	-	PFDYVKTQI-QKM
AtM2OM	PIDMIKVRI-QLG	-	PFDYVKTQI-QKM
SsM2OM	PIDMVKVRI-QLG	-	PFDYVKTQI-QKM

[a]Signature variants are representative for proteins of all species.

[b]Filled circles represent possible amino acid deletions.

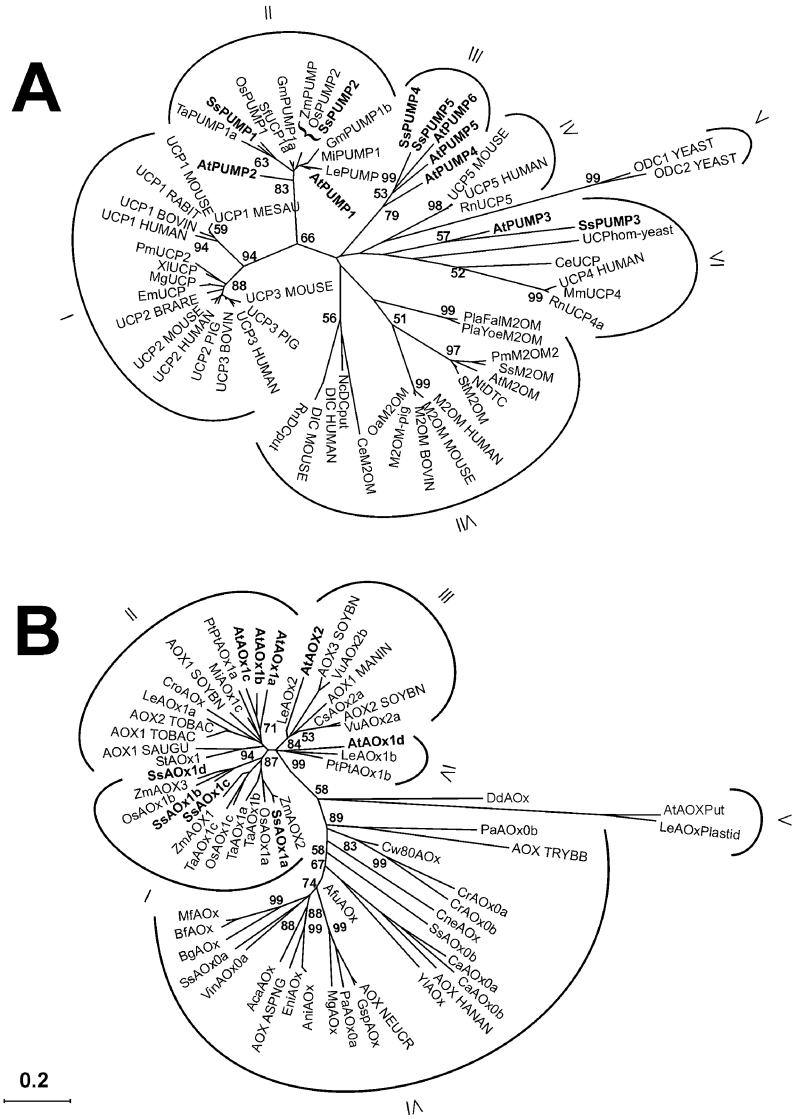


Figure 1. Unrooted phylograms of the amino acid sequences of uncoupling proteins (A) and alternative oxidases (B) of all types. The seven groups distinguished within the uncoupling protein tree are: I, UCP1-3; II, PUMP1-2; III, PUMP4-6; IV, UCP5; V, yeast ODC, VI, UCP4 and PUMP3; and VII, M2OM and DIC. The six groups discriminated within the alternative oxidase tree are: I, monocot AOx1; II, dicot AOx1; III, dicot AOx2; IV, possible dicot AOx3; V, plastidic dicot AOx; and VI, fungal AOx. Phylograms were generated as a consensus of 1000 bootstrap replicates by the neighbor joining method (the bootstrap values are indicated close the branch divisions). The scale bar indicates the relative amount of change along branches. The *Arabidopsis* and sugarcane PUMP and AOx family members are in bold. The accession numbers are given in Materials and Methods.

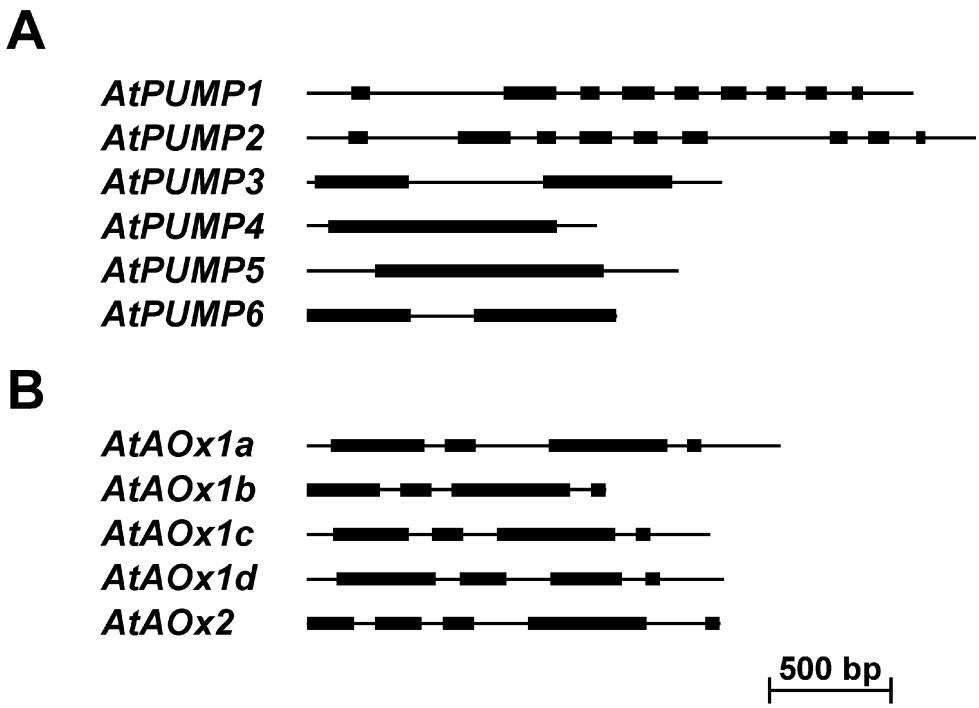


Figure 2. Gene structure of *Arabidopsis* PUMP (A) and AOx (B) genes. Filled boxes represent exons and lines represent introns and 5'- and 3'-unstranslated regions according to Locus data from The *Arabidopsis* Information Resource (TAIR). Bar corresponds to 500 bp of chromosomal DNA.

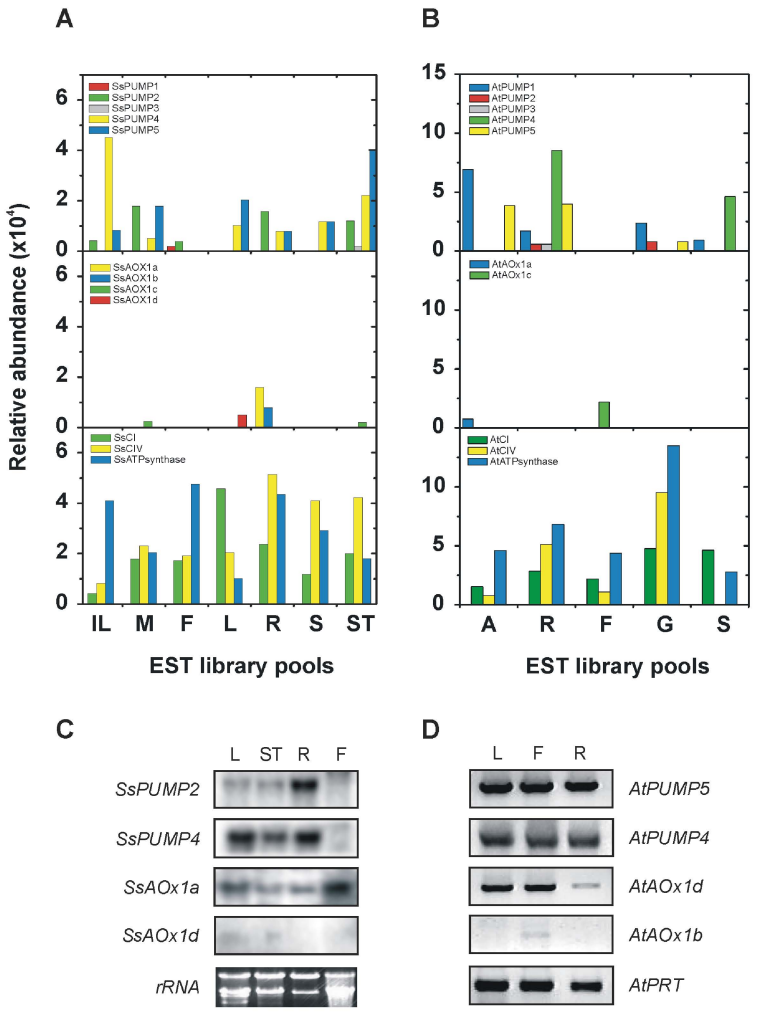


Figure 3. Relative abundance of PUMP- and AOx-representing ESTs in sugarcane (A) and *Arabidopsis* (B) tissue-specific library pools and transcript accumulation of *SsPUMPs* and *SsAOxs* (C) and of *AtPUMPs* and *AtAOxs* (D) in different plant organs. (A and B) The top panels show PUMP ESTs; the middle panels show AOx ESTs; and the bottom panels show ESTs of the nuclear-coded subunits of respiratory chain complexes I (CI) and IV (CIV), and of the ATP-synthase of both plant species. Sugarcane EST library pools: IL–plants inoculated with endophytic bacteria, M–meristem, F–floral organs, L– leaves, R–roots, S–seeds, and ST–stem. *Arabidopsis* EST library pools: A–aboveground organs, R–roots, F–floral organs, G–green siliques, and S–seeds. (C) Total RNA isolated from sugarcane organs was subjected to RNA-blot analysis using probes prepared from full-length cDNAs for *SsPUMP2*, *SsPUMP4*, *SsAOx1a*, and *SsAOx1d*. L– leaves, ST–stem, R–roots, and F–floral organs. (D) Total RNA isolated from *Arabidopsis* organs was subjected to semi-quantitative RT-PCR analysis using specific primers for *AtPUMP4*, *AtPUMP5*, *AtAOx1b*, and *AtAOx1d*. L– leaves, F–floral organs, and R–roots.

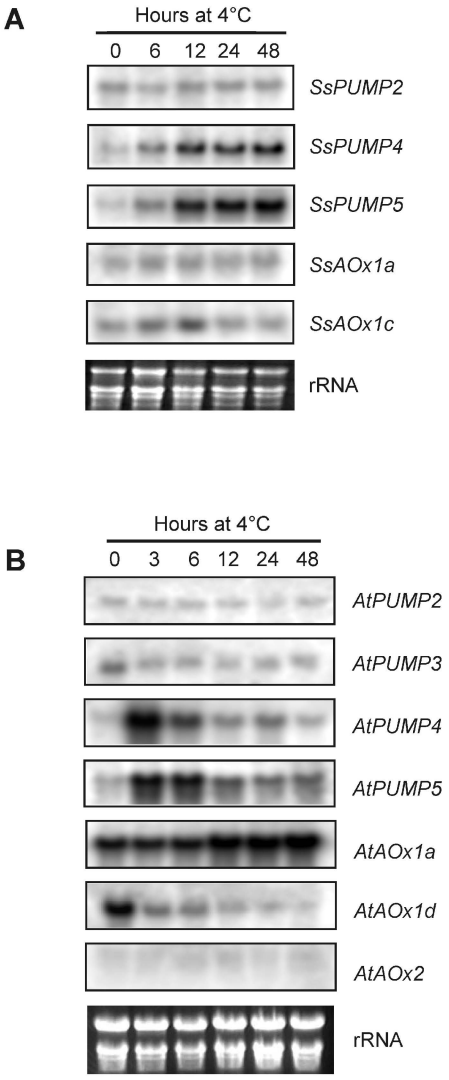


Figure 4. Chilling-regulated transcript accumulation of PUMPs and AOxs in sugarcane (Panel A) and *Arabidopsis* (Panel B). Total RNA from the leaves of chilling-treated and untreated plantlets colleted after 0, 3 (*Arabidopsis* only), 6, 12, 24, and 48 h at 4 °C was subjected to RNA-blot analysis using probes prepared from full-length cDNAs for *SsPUMPs* and *SsAOxs* (Panel A), or *AtPUMPs* and *AtAOxs* (Panel B).

Capítulo IV

SsNAC23, a NAC domain protein family member, is associated with cold, herbivory and water stress in sugarcane

Fábio T.S. Nogueira, Paulo Arruda

ABSTRACT

NAC domain gene family members are widely distributed in plant genomes and some of them are regulated by biotic and abiotic stresses. Using sugarcane EST database, SUCEST, we were able to identify 26 NAC domain gene family members, referred to collectively as *Saccharum sp*. *NACs* (*SsNACs*). Most of the *SsNACs* display a similar tissue/organ expression pattern. Among those *NACs*, a chilling-stress inducible gene designated *SsNAC23*, was characterized in details. SsNAC23 has strong similarity to the *Arabidopsis* ATAF1 and rice OsNAC6, which were phylogeneticaly classified into Dicot and Monocot ATAF subgroups, respectively. Molecular modeling of SsNAC23 3-D structure and its nuclear subcellular localization in transient expression experiments with onion epidermis evidenced its transcriptional activator nature. The expression of *SsNAC23* was strongly induced at 4 ^{o}C but not at 12 ^{o}C, suggesting that this gene is involved in the response to lower-range low temperature stress. *SsNAC23* was also induced by PEG-induced water stress and herbivory. The role of *SsNAC23* as a transcription factor in plants subjected to stress stimuli is discussed.

INTRODUCTION

Plants vary considerably in their ability to survive under stressful conditions. Stress-regulated genes play an important role in stress tolerance mechanisms when plants are exposed to hostile environments (Xiong et al., 2002). Chilling, for instance, is a common abiotic stress that unfavorably affects plant growth, especially those of tropical and subtropical origin, like sugarcane (Tai and Lentini, 1998). Molecular studies using cDNA arrays have identified genes that are upregulated by biotic and abiotic stresses (Maleck et al., 2000; Kawasaki et al., 2001; Seki et al., 2002; Fowler and Thomashow, 2002; Nogueira et al., 2003; Provart et al., 2003). Many of these genes encode proteins involved in signaling pathways, including mitogen-activated protein kinases (MAPKs), histidine kinases, Ca^{2+}-dependent protein kinases (CDPKs), SOS3 family of Ca^{2+} sensors, as well as specific transcription factors. Regarding transcription factors, the sequencing of whole plant genomes has revealed that they often belongs large gene families, as for example ERF, bZIP, WRKY (Singh et al., 2002) and MYB (Jin and Martin, 1999). Some members of these families are involved in the regulation of defense responses to biotic and environmental stress while others appear to coordinate plant developmental pathways (Jin and Martin, 1999; Chen and Singh, 1999; Robatzek and Somssich, 2001).

An interesting family of transcription factors is the NAC domain protein family (Kikuchi et al., 2000; Duval et al., 2002). The members of this family contain a highly conserved N-terminal region (Aida et al., 1997) while the carboxy terminus is highly variable in both length and amino acid sequence. The whole NAC domain protein family shows no sequence homology to any other proteins in the GenBank nr database (Xie et al., 2000).

Some *Arabidopsis thaliana* NAC domain genes has been identified and characterized. Sablowski and Meyerowitz (1998) demonstrated that *NAP* is up-regulated in flower organ primordia by the two MADs box containing-proteins, APETALA3 and PISTILLATA. *NAP* is expressed mainly underneath the inflorescence meristem as it develops sepals and at the base of stamen filaments. Other examples reported are *Arabidopsis NAC1* and *CUC* genes. *NAC1* mediates auxin signaling to promote lateral root

development (Xie et al., 2000). *CUC1* is required for complete penetrance of the cupped cotyledon phenotype (Takada et al., 2001) while *CUC3* seems to be involved in the establishment of boundaries that contain cells with low proliferation and/or differentiation rates (Vroemen et al., 2003). These evidences suggest that NAC domain genes could be involved in several plant development mechanisms.

NAC domain proteins have also been shown to be implicated in biotic and abiotic stress responses. A potato NAC domain gene, *StNAC*, and *Arabidopsis ATAF1* and *ATAF2* genes were shown to be induced by pathogen attack and wounding (Collinge and Boller, 2001). More recently, Hegedus et al. (2003) found that BnNAC domain gene family members were differentially regulated in response to biotic and abiotic stresses, and Ooka et al. (2003) proposed that members of the subfamilies ATAF and NAP could be involved in the response to stress stimuli. Moreover, Ren et al. (2000) demonstrated that the *Arabidopsis* NAC domain protein, TIP, interacts specifically with the capsid protein of turnip crinkle virus (TCV), suggesting that TIP may participate in the virus resistance. However, correlation between family classification and putative physiological NAC domain protein roles in stress responses has not been elucidated in details.

Two lines of evidence support the concept that NAC domain proteins are involved in transcriptional regulation. First, *A. thaliana* lines over-expressing *NAC1* exhibited upregulation of the auxin-responsive genes and a corresponding decrease was observed in anti-sense lines (Xie et al., 2000). Second, *Brassica napus* NAC proteins (BnNACs) were capable of activating gene expression in yeast and recognized an element within the CaMV 35S promoter sequence, which is found in some plant promoter sequences (Hegedus et al., 2003). Additionally, in rice (*Oryza sativa*) and *A. thaliana*, NAC domain proteins can constitute a superfamily and it has recently been proposed a first classification of these specific plant transcriptional activators (Ooka et al., 2003).

Here we report the identification of 26 non-redudant sugarcane genes encoding NAC domain proteins, including a cold-inducible sugarcane NAC domain gene, *SsNAC23*, (Nogueira et al., 2003). RNA-blot analysis showed that *SsNAC23* is strongly induced by cold, water stress and herbivory. Prediction of the protein three-dimensional structure and *in vivo* analysis evidenced the SsNAC23

transcription factor nature. The role of *SsNAC23* as a transcription factor in response to stress stimuli is discussed.

RESULTS

Identification of sugarcane NAC domain containing-sequences and phylogenetic analysis of stress-responsive NAC proteins

An ordered set of *Arabidopsis* and rice NAC domain protein sequences was used as drivers to identify and classify sugarcane NAC domain containing-sequences encoded by EST clusters produced by the assembly of 237,954 ESTs, which were referred as SASs (sugarcane assembled sequences, Vettore et al., 2003). Queries consisting of full-length protein sequence of each *Arabidopsis* (ANAC) and rice (ONAC) NAC domain protein families (Ooka et al., 2003) were used to screen the SUCEST sugarcane EST database. Candidate SASs were selected based upon e-value < 10^{-30} and the presence of at least four N-terminal conserved NAC subdomains (A to D, Kikuchi et al., 2000). Using this approach, we were able to identify 35 sugarcane NAC domain containing-SASs. Further alignment of all of these SASs resulted in 26 non-redundant sequences referred here as *SsNACs* (*Saccharum sp*. *NACs*, Table 1).

To assess the expression profiling of all *SsNACs* in several sugarcane tissue/organs, we used the information available in the SUCEST EST database. The ESTs from libraries constructed from different tissues and/or organs were retrieved from SUCEST database and grouped in library pools (Table 2). The relative abundance of ESTs was calculated as a count of the number of *SsNAC* ESTs in a given library pool normalized by the total number of ESTs in that pool (Fig. 1). Most of *SsNACs* were expressed in "roots" in a similar pattern, except for *SsNAC14* and *SsNAC23*, which showed higher expression in roots. Thirty percent (8/26) of *SsNAC*s displayed tissue/organ-preferred expression, while the remaining showed constitutive expression. This could be an indication that the most of sugarcane NAC domain proteins could play redundant roles in distinct plant organs.

Phylogenetic analysis of SsNACs was done using the neighbor-joining method and selected sequences (only those containing entire NAC domain sequence, comprising around 160 amino acids), besides *Arabidopsis*/rice NAC domain protein sequences (Ooka et al., 2003), as described in the Materials and Methods. Twenty SsNACs were included into 14 subgroups of NAC families (data not shown), according to classification described by Ooka et al. (2003). Among the identified SsNACs, SsNAC23 is encoded by SAS SCCCCL4007D12 that was reported as being regulated by chilling stress (Nogueira et al., 2003). SsNAC23 together with other three sugarcane NAC domain proteins (SsNAC8, SsNAC14 and SsNAC19) fell into ATAF, NAP, and OsNAC3 subgroups respectively (Fig. 2A). We analyzed in details these subfamilies and our phylogenetic analysis clearly indicated that these three subfamilies could be divided into five subgroups, including Dicot ATAF subgroup (I), Monocot ATAF subgroup (II), OsNAC3 subgroup (III), ANAC subgroup (IV), and NAP subgroup (V) (Fig. 2A). The subgroups were determined by selecting suitable branches on the basis of the bootstrap values. To estimate the reliability of the classification of these five subgroups, we also performed the phylogenetic analysis by maximum parsimony (MP) method (PROTPARS in Felsenstein's program package PHYLIP, version 3.5c). An MP tree was constructed using MEGA2 program (Kumar et al., 2001). To assess the reliability of such classification into five subgroups, the significance level of the MP analysis was examined by bootstrap testing with 1000 replicates and similar results were observed (data not shown). Moreover, biotic and abiotic stress-responsive *ANACs* and *ONACs* can be found in all subgroups analyzed, mainly in subgroups I and II, represented by Dicot and Monocot ATAF proteins (Fig. 2A). Wounding induced the expression of both *ATAF1* and *ATAF2* but *ATAF1* is also induced by cold and drought stress besides ABA treatment (Fig. 2B; Collinge and Boller, 2002; Seki et al., 2002). Other *Arabidopsis* NAC domain genes were reported as being inducible or repressed by biotic and abiotic stress, including the already described *AtNAM* and *NAP* (Fig. 2B; Provart et al., 2002) and BnNACs (Fig. 2B; Hegedus et al., 2003). Finally, among those proteins classified into Monocot ATAF subgroup, OsNAC6 and SsNAC23 are encoded by stress-inducible genes (Fig. 2B; Nogueira et al., 2003; Rabbani et al., 2003).

Additionally, a set of conserved motifs in the C-terminal (transcriptional activation regions; Kikuchi et al., 2000) of subgroups I and II which may define important functional sequences were identified (Fig. 3). Figure 3A illustrates the position of each C-terminal motif found in the representative proteins of each subgroup. Based on these domains, Dicot ATAF subgroup could be still divided into ATAF1- and ATAF2-like proteins, sharing transcriptional activation regions with serine-rich (M1) and acidic amino acid-rich (M2) motifs (Fig. 2B). Moreover, ATAF1-like proteins enclose a motif containing proline-rich regions (M4, Fig. 3B). The motifs found in the proteins belonging to Monocot ATAF subgroup are similar, except for motif M2 that is present only in OsNAC6 and SsNAC23. Motifs M1, M2, and M5 were conserved in both subgroups (Dicot and Monocot ATAF subgroups, Fig. 3). Altogether, these motifs can be considered as informative since they shared derived characteristics that were used to classify members of the Dicot and Monocot ATAF subgroups, being consistent with the phylogenetic analysis (Figure 2A).

Characterization of protein structure and cellular localization of SsNAC23

Because *SsNAC23* was the second sugarcane gene more expressed in response to chilling stress (Nogueira et al., 2003), we focused on characterizing in details this SsNAC protein. The longest cDNA clone belonging to SAS SCCCCL4007D12 (Table 1) was completely sequenced and analyzed as described in Materials and Methods. A 1.56-kb cDNA was identified and found to contain an 890-bp open reading frame (ORF). The putative protein encoded by this cDNA has 296 amino acids with features characteristic of known NAC domain proteins, including the five N-terminal subdomains A to E (Fig. 4A). This sugarcane NAC protein is predicted to have an estimated molecular mass of 34 kDa and a pI of 10.34. The putative sugarcane protein designed here as SsNAC23 (Table 1) is closely related to the OsNAC6 (Fig. 2A), a rice NAC domain protein already described (overall 82% identity and 83% similarity at the amino acid sequence level; Kikuchi et al., 2000). *In silico* analyses using PEST-FIND program identified a poor PEST sequence in SsNAC23 that is present into M1 motif (Fig. 3) that, in turn, is conserved among Dicot and Monocot ATAF subgroup members. Hegedus et al. (2003) found similar

PEST domain in ATAF1- and ATAF1-like proteins. The PEST domains are common in regulatory proteins that are rapidly targeted for degradation by the ubiquitin-26S proteosome pathway (Rechsteiner and Rogers, 1996).

As explained in Earnst et al. (2004), the basic residues found in the NAC domain sequence (Arg, Lis and His), which are important to DNA-binding function, are conserved in *Arabidopsis*. Describing the ANAC structure and their DNA binding capacity, the authors speculated that the extremities of the β-strands 4 and 5 connecting turn protrude from the face of the NAC dimer is rich in positive charged residues. This internal surface appears to be the DNA-contacting interface (Earnst et al., 2004). This assumption is in according with the biological data collected by Duval et al. (2002) in AtNAM deletion mutants in yeast. Using 15 NAC sequences from different plants, was constructed a sequence alignment that lets us to conclude that these basic residues are conserved not only in *Arabidopsis* sequences but also in other plants as well (Fig. 4A). The more conserved parts of the alignment coincide with the secondary elements described in ANAC structure (Fig 4A). The primary structure conservation observed in NAC domain lets us to hypothesize that overall folding of the secondary structure element might be conserved in all NAC proteins.

Therefore, molecular modeling was used to obtain the possible structure of SsNAC23 protein. The Ramachandran plot of the modeled SsNAC23 structure shows that no residues lie in the disallowed regions, and stereochemical parameters checked in Procheck (Laskowski et al., 1993), were inside or better than expected at 95% confidence level (data not shown). To avoid the targeting of the initial 3D structure to the obtained model (a natural consequence of the homology modeling procedure) we performed 100 p.s. of unrestrained molecular dynamics in the initial model of the SsNAC23 using GROMACS (Lindahl et al., 2001) and additional energy minimization to the final structure.

Final SsNAC23 model fold maintain the overall structure of twisted antiparalell β-sheet with pack against an N-terminal α-helix (as already described for ANAC structure, Earnst et al., 2004). The secondary structure of modeled SsNAC23 is almost equal to the ANAC, with minor differences. The

backbone superimposition of SsNAC23 model and ANAC structures showed backbone r.m.s.d. equal to 1.5 Å, indicating high conservation in the domain structure (Fig. 4B).

The SsNAC23 dimer structure maintains a correct fold configuration (Fig. 4C) with internal positive surface rich in Arg, Lys and His. It is important to remark the conservative position of the residues that allow the formation of the interdomain salt bridge (Arg19 and Glu26 in Fig. 4B) and the Glu67 residue, important for the coordination of the Arg19 position at the domain interface. All the basic residues in DNA-contacting surface conserved their position if compare the already described ANAC structure and the obtained structure of the SsNAC23 model (Fig. 4D), suggesting that obtained SsNAC23 model maintain all the important features that characterize DNA-contacting proteins.

The degenerate bipartite nuclear localization signal (NLS, IKKALVFYAGKAPKGEKTN), suggested by Greve et al. (2003), was identified in the SsNAC23 NAC domain sequence, being very conserved in the several NAC domain proteins among different plant species (Fig. 5A). This suggested that SsNAC23 could be targeted to nucleus. To determine this hypothesis, *SsNAC23* ORF was fused to the N-terminal of the *Escherichia coli GUS* gene using pBI221 vector as described in Material and Methods. Both recombinant DNA constructs encoding SsNAC23-GUS fusion and a GUS protein alone were introduced into onion epidermis cells by particle bombardment (Varagona et al., 1992). Most of SsNAC23-GUS fusion protein was targeted to the nucleus as visualized by the blue staining in the area of the nucleus. In contrast, all of the GUS protein was detected throughout the cell, indicating that the GUS protein could remain in the cytosol (Fig. 5B). These results agreed with data previously observed for *Arabidopsis* NAC1and ANAC proteins (Xie et al., 2000; Greve et al., 2003), suggesting NLS conservation in sequence and function among NAC domain proteins from different plant species.

***SsNAC23* gene expression analysis under chilling stress**

Because NAC domain genes have a high degree of sequence conservation in the 5´ coding region, an XbaI-digested 3´ DNA fragment of SsNAC23 cDNA was generated to be used as specific probe for all RNA-blot analyses. In general, RNA-blot results for tissue/organ-enriched expression confirmed the

findings of the *in silico* analysis (Fig. 1). *SsNAC23* was expressed in nonstressed one-month-old sugarcane plantlets but in different levels among different sugarcane organs. It was expressed at moderate levels in roots and young leaves, being detected only traces of *SsNAC23* mRNA in stems and non-detected in apical meristematic tissues and floral organs (Fig. 6A).

To examine the expression of *SsNAC23* in response to low temperature in detail, we used leaves from three-month-old sugarcane plantlets. RNA-blot analyses revealed that the accumulation of the *SsNAC23* mRNA was long-term induced in leaves during a 48-h treatment at 4 °C (Fig. 6B). The induction of *SsNAC23* was initiated within one hour and steadily increased afterward until 48 h subsequent to the 4 °C treatment. As sugarcane growth is strongly altered at suboptimal temperature (below 16 °C; Tai and Lentini, 1998), we evaluated the *SsNAC23* expression profile under moderate low temperature stress (12 °C). Interestingly, the level of *SsNAC23* mRNA was not induced or repressed in leaves of three-month-old sugarcane plantlets that were subjected to 12 °C treatment (Fig. 6B). Therefore, it could be concluded that *SsNAC23* is induced by 4 °C stress but not by 12°C. Similar data was observed for rice *lip19* gene encoding a bZIP-type DNA-binding protein that is inducible by low temperatures (5 °C; Aguan et al., 1993) and it was recently reported as being responsive to 4 °C treatment but not to 12 °C treatment (Wen et al., 2002). Moreover, because the *SsNAC23* mRNA level remained high as long as the plantlets were kept at the extremely positive low temperature (4 °C treatment), we analyzed the mRNA accumulation during plantlets acclimation to normal growth temperature (26 °C) after chilling treatment. When 48 h-cold treated sugarcane plantlets were returned to 26 °C, *SsNAC23* mRNA levels rapidly declined to basal levels within 3 h (Fig. 6C).

Wen et al. (2002) observed that a rice mitogen-activated protein kinase gene *MAPK* (*OsMAP1* or *OsMAPK5*) was induced by 12 °C but not 4 °C, suggesting the presence of different pathways in response to ranges of low temperature exposition in rice. These evidences, together with distinct *SsNAC23* regulation to 4 °C and 12 °C (Fig. 6B), let us to hypothesize that sugarcane might also has differential response to extreme (4 °C) and moderate (12 °C) chilling stress. Using OsMAPK5 sequence as a driver and tBLASTn algorithm, we identified in the SUCEST database (Vettore et al., 2003) the SAS

SCCCRZ2C04G10.g that encodes protein very similar to MAPK5a (94% identity at the amino acid sequence level; Xiong and Yang, 2003). The 1,128-bp ORF of this SAS encodes a putative protein (named here as *Saccharum sp.* MAPK, SsMAPK) that has 375 amino acid residues with an estimated molecular mass of 43 kDa and a pI of 5.5. Figure 7 shows that the transcript accumulation of the *SsMAPK* was substantially induced by 4 °C- and 12 °C-treatment but with distinct expression profiles. These results, together with the differential *SsNAC23* expression in response to ranges of low temperature exposition, suggested that there could be distinct signaling systems that perceive and transduce temperature signals in sugarcane leaves.

Effects of water stress, ABA, and herbivory on the expression of *SsNAC23*

It has been widely reported that some cold-inducible genes also respond to water deficit (Seki et al., 2002) and wounding (Xiong and Yang, 2003). Therefore, we examined the effects of water stress and herbivore on the expression of *SsNAC23*. Figure 8A shows that the expression of *SsNAC23* could also be transiently induced by water stress (simulated by PEG 6000 solution; Money, 1989; Jiang and Zhang, 2002), similarly to results observed for *SsNAC23* orthologues *ATAF1* and *OsNAC6* (Seki et al., 2002; Rabbani et al., 2003); however, the mRNA level induced by water stress was lower than that induced by low temperature (Fig. 8A). Interestingly, sugarcane *NAC23* did not show clear responsiveness to ABA application (data not shown). The data contrast with previously identified *SsNAC23* orthologue genes from *Arabidopsis* and rice, *ATAF1* and *OsNAC6*, respectively (Fig. 2A), which have been shown to be responsive to exogenous ABA (Seki et al., 2002; Rabbani et al., 2003). This discrepancy may be attributable to the difference in plant species used as well as stress treatment and plant growth conditions. *SsNAC23* transcripts also accumulated transiently after 4 h of tissue damage caused by herbivory (*Diatraea saccharalis* feeding), reaching the maximum level of transcript accumulation at 12 h (Fig. 8B). Altogether, these results suggest a complex gene regulation of *SsNAC23* during biotic and abiotic stress.

DISCUSSION

Sugarcane NAC domain protein family members

Many transcription factors belong to large gene families, being several members responsive to biotic and abiotic stresses (Chen et al., 2002). In this work we identified 26 members of sugarcane NAC domains proteins (SsNACs, Table 1) using *Arabidopsis* and rice protein sequences as drivers and SUCEST database (Vettore et al., 2003). As the sugarcane genome was not completely sequenced, it is possible that a presence of an expanded NAC domain gene family in sugarcane, arising from genome and gene duplication events during and after speciation. Moreover, Ooka et al. (2003) found 75 predicted NAC domain proteins in full-length cDNA data sets of *O. sativa* and 105 from the *A. thaliana* genome, suggesting that a great number of NAC domain protein members is predicted to be identified in different plant species.

As expected for transcription regulators, most of *SsNACs* were expressed at low levels in sugarcane tissue/organs (Table 1, Fig. 1). Seven *SsNACs* were more expressed by sugarcane genome, including *SsNAC23* that is the most expressed sugarcane NAC gene, being represented by 28 ESTs in the SUCEST database (Table 1). Digital mRNA expression profile showed that few *SsNACs* displayed organ/tissue-enriched expression (Fig. 1). Most of *SsNACs* have at least one EST from "root" pool, suggesting an important role of these proteins in that plant organ. Indeed, when *Arabidopsis NAC1* was expressed in antisense orientation, the root development was blocked (Xie et al., 2000). Moreover, NAC1 activated the expression the two downstream auxin-responsive genes, therefore, belonging to the auxin-signaling pathway that induces lateral root formation (Xie et al., 2000).

Phylogenetic analysis conducted here using protein sequences from four SsNACs and ATAF, NAP and OsNAC3 subfamilies members (Ooka et al., 2003), besides *B. napus* NAC domain proteins revealed that these three subfamilies indeed comprised five distinct subgroups (Fig 2A). The inclusion of two dicot and two monocot species allowed an improvement in the resolution of the evolutionary relationships among this set of proteins. It is interesting to notice that the ATAF subgroup has a clearly distinction between dicots and monocots, suggesting the existence of a common ancestral, while the

subgroups ANAC and OsNAC3 contains only dicot and monocot protein members, respectively (Fig. 2A). Nevertheless, additional information such as common motifs, expression pattern, or map position is required to support this orthology (Vincentz et al., 2003). MEME-based analysis of transcriptional activation regions revealed common motifs (M1, M2 and M5) among proteins belonging to Dicot and Monocot ATAF subgroups (Fig. 3), suggesting that these regions were conserved in parallel with NAC domain structures. Therefore, it is possible that both NAC subdomains and transcriptional activation regions are involved in determining similar functions in dicot and monocot plant species. Moreover, several members of ATAF subgroups I and II were reported to be regulated by biotic and abiotic stress (Fig. 2B). Hegedus et al. (2003) identified eight *B. napus* NAC domain genes differently regulated in response to wounding, insect feeding, cold stress, and dehydration. *ATAF2* is wouding-inducible whereas *ATAF1* is predicted to be induced by biotic and abiotic stress as observed for its orthologues *OsNAC6* and *SsNAC23* (Fig 2B; Collinge and Boller, 2002; Seki et al., 2002; Nogueira et al., 2003; Rabbani et al., 2003). These findings suggest that stress-related functions could be conserved among dicot and monocot members of ATAF subgroups. Moreover, based on the gross similarity of expression profiles, it is tempting to speculate that dicot and monocot ATAF members participate in the same physiological regulatory network during stress conditions in both plant species.

Stress-inducible *SsNAC23* encodes a putative DNA-binding protein that is targeted to the nucleus

SsNAC23 encodes a NAC domain protein sharing high homology (82% identity) with the rice OsNAC6 and shows extensive similarity with other NAC domain homologs identified in plants, including *Arabidopsis* ANAC (Fig. 4A). In the sequence alignment, the secondary structure elements described in ANAC structure (Ersnt et al., 2004) coincides with the more conserved parts of the alignment and positive charged residues that were conserved in all studied plants as well (Fig. 4A). To date, ANAC was the unique NAC domain protein that had its DNA-binding domain structure determined by X-ray (Ersnt et al., 2004). In our model study, a very similar structure of the NAC domain fold from SsNAC23 was observed (Fig. 4B). The conservation of the fold structure of the NAC domain, the position of the residues that

allow the formation of the dimer through salt bridge and the abundance of basic residues in the contacting DNA surface lets us to conclude that obtained SsNAC23 model maintain all the important features that characterize DNA-contacting proteins. Furthermore, the conservation of the β-sheet scaffold in the NAC primary sequences among several plants (Fig. 4A), together with the extensive similarity in structure between SsNAC23 and ANAC (Fig. 4B), suggested that these proteins can use β-sheet structures not only for DNA-binding but also for interacting with other proteins, as proposed for ANAC by Greve et al. (2003). In fact, Hegedus et al. (2003) observed that some BnNACs interacted with NAC domains alone of other ones and Greve et al. (2003) demonstrated that NAC domain is sufficient for the interaction between *Arabidopsis* RING-H2-type RHA2a protein and ANAC. Finally, SsNAC23 contains a degenerate bipartite nuclear localization signal that could direct reporter GUS fusions to the nucleus (Fig. 5B). Both *Arabidopsis* NAC1 (Xie et al., 2000) and ANAC (Greve et al., 2003) were also able to target GFP and GUS fusions to the nucleus, respectively, suggesting that predicted degenerate bipartite nuclear localization signal is highly conserved among NAC domain proteins (Fig. 5A). Altogether, these data provide strong evidences that SsNAC23 is a sugarcane transcriptional activator.

Interestingly, *SsNAC23* was inducible by 4 °C-chilling treatment but not by a moderate low temperature exposition for 2 d (12 °C, Fig. 6B), similar to *Oslip19* (Wen et al., 2002). In contrast, another sugarcane gene that might be involved in cold stress signaling (Asai et al., 2002), *SsMAPK*, was long-term inducible during 4 °C treatment but transiently during 12 °C treatment (Fig. 7B). This sugarcane *MAPK* shares high similarity with other known stress-responsive plant *MAPKs* (Fig. 7A, Wen et al., 2002; Xiong and Yang, 2003). The similar expression profiles observed for *SsMAPK* and *SsNAC23* in severe low temperature exposition (4 °C, Figs. 6B and 7B) suggested that they might participate of similar signal transduction pathways. On the other hand, it is possible that *SsMAPK* acts in pathways distinct those of *SsNAC23* when sugarcane plantlets are exposed to moderate chilling stress. The differential expression profiling of *SsMAPK* and *SsNAC23* in response to 4 °C and 12 °C suggests that sugarcane has mechanisms to distinguish the two different ranges of low temperatures. Similarly, analyzing the transcript accumulation of two *OsMAPKs* and the bZIP transcription factor *Oslip19*, Wen et al. (2002) proposed that

rice plants also have mechanisms to discriminate different ranges of low temperature. Additionally, Provart et al. (2003) observed quantitative and qualitative differences between 13 °C- and 4 °C-treatment expression profiles in *Arabidopsis*. A possible reason for the existence of common and distinct components during extreme and moderate chilling stress may be related to the fact that a 12 °C treatment for 2 d only includes responses to acclimate sugarcane metabolism to grow at low temperature.

SsNAC23 could be involved in dehydration-related responses since it was also inducible by polyethylene glycol-induced water stress (Fig. 8A). Furthermore, sugarcane plantlets subjected to *D. saccharalis* feeding for variable time showed high *SsNAC23* transcript accumulation after 12 h of larvae exposition (Fig. 8B). Despite of biotic and abiotic stress are very complex and possess many different yet related attributes (Xiong et al., 2002), our results suggest that there is a similar molecular mechanism among these stimuli. Zhang et al. (2004) suggested that insect feeding induces plant responses, which would take part in a jasmonic acid-independent pathway crosstalk with those related to abiotic stress. Therefore, it is possible that the sugarcane *NAC23* is involved in a pathway triggered by severe chilling (4 °C treatment), PEG, herbivory-induced water stress and, in turn, it could regulate the expression of unknown (or known) water-deficit stress-related genes through binding to a putative "NAC" box present in their promoters.

MATERIALS AND METHODS

Plant Growth and Treatments

Sugarcane plantlets (*Saccharum* sp. cv. SP80-3280), propagated axenically *in vitro* (Nogueira et al., 2003), were kept at 26 °C on a 16 h/8 h day/night cycle with a photon flux density of 70 $\mu E \cdot m^{-2} \cdot s^{-1}$. For comparing *SsNAC23* expression profile between 4 °C and 12 °C treatments, three-month-old plantlets were transferred to 4 °C or 12 °C under the same photoperiod conditions. Control plantlets were maintained at 26 °C for both treatments. The leaves of control and chilling-treated plantlets were harvested after 0, 1, 3, 6, 12, and 48 h of treatment (six plantlets per time point). In a similar manner, ABA

or PEG treatments were performed by transferring sugarcane plantlets to rooting medium (Nogueira et al., 2003) supplemented by 100 μM ABA solution or 20% (p/v) polyethylene glycol (PEG) 6000, a polymeric osmoticum that does not penetrate plan cells and is relatively non-toxic (Chazen and Neumann, 1994). Plantlets were maintained as described above. For ABA treatment, the sampling time was 0, 1, 3, 6, 12, 24 and 48 h while for PEG treatment, the leaf samples were collected after 0, 3, 6, 12, 24, and 48 h of treatment. The expression pattern of the *SsNAC23* in apical meristem tissues, leaves, roots, and stem was examined using one-month-old sugarcane plantlets (*Saccharum* sp. cv. SP80-3280) cultivated in a greenhouse. Flowers were obtained from sugarcane plants (*Saccharum* sp. cv. SP80-87432) grown at the Copersucar experimental station (http://sucest.lad.dcc.unicamp.br).

For the bioassay with sugarcane borer, third and fourth instar *Diatraea saccharalis* larvae were starved overnight prior to wounding experiments. Larvae were placed on mature, terminal leaf of three-month-old sugarcane plants (*Saccharum* sp. cv. SP80-3280). Tissue samples (closer to larvae feeding) were collected after 0, 6, 12 and 24 hours of larvae exposition. Larvae were then removed, and the plant tissues were immediately frozen in liquid nitrogen. Experiments were initiated between 7 and 8 AM and conducted in a greenhouse.

Identification and *in silico* analysis of sequences encoding sugarcane NAC domain proteins

The *Arabidopsis thaliana* as well as *Oryza sativa* NAC domain proteins already described (Ooka et al., 2003) were used as drivers to identify and classify a set of non-redundant sugarcane assembled sequences (SASs) sharing similarities to NACs in the SUCEST database (Vettore et al., 2003). SASs with an e-value $\leq 10^{-30}$ were considered to represent putative sugarcane NAC domain-containing genes (*SsNACs*). Subsequently, the identified sequences were checked for the presence of conserved N-terminal NAC domain sequence using Pfam database (Bateman et al., 2000) and BLASTP algorithm. To identify gene redundancy among sequences retrieved from SUCEST database, the DNA sequences of all accepted SsNAC candidates were aligned and compared using the CLUSTALX program (Thompson et al., 1997).

Those sequences showing more than 95% of identity in the DNA sequence were discarded of further analysis.

Unrooted, neighbor-joined phylogenetic tree (1000 bootstraps) was inferred using the MEGA2 program (Kumar et al., 2001) and protein sequences corresponding to the NAC domain amino acids to ensure the most conserved alignment. Sequences of NAC domain and those corresponding to C-terminal region of *Arabidopsis*/rice NAP, ATAF and OsNAC3 protein family members, *Brassica napus* NAC domain proteins and orthologue SAS, were detected using the pattern prediction program package MEME-MAST, version 3.0 (Bailey and Gribskov, 1998, http://www.sdsc.edu/MEME). Peptide sequences common to proteins that are targeted for rapid turnover were identified with PEST-FIND (http://www.at.embnet.org/embnet/tools/bio/PESTfind).

Sequences from the following list (accession numbers in parentheses) were used for the alignment and phylogenetic analysis: OsNAC6 (AK068392), OsNAC5 (AK063399), OsNAC4 (AK073848), OsNAC3 (AK073667), ONAC010 (AK063406), ONAC002 (AK104712), NAP (CAA10955), Jasmonate-regulated NAC1 (JaNAC1, NP_188170), Jasmonate-regulated NAC2 (JaNAC2, AAM61076), AtNAM (AAD17314), ANAC (NP_175697), ATAF1 (NP_171677), ATAF2 (CAC35884), ANAC025 (NP_564771), ANAC032 (NP_177869), ANAC047 (NP_187057), ANAC081 (NP_680161), ANAC102 (NP_201184), BnNAC5-8 (AY245883), BnNAC5-7 (AY245882), BnNAC5-11 (AY245884), BnNAC5-1 (AY245881), BnNAC3 (AY245880), BnNAC18 (AY245886), BnNAC14 (AY245885), BnNAC1-1 (AY245879), SsNAC19 (CA140418), SsNAC14 (CA132543), and SsNAC8 (CA120159).

Identification of *SsNAC23* and *SsMAPK* full-length sequences

To obtain the full-length cDNA sequence representing *SsNAC23* and *SsMAPK*, the longest EST of each individual SAS was resequenced using an ABI PRISM 3700 DNA sequencer and M13 forward and reverse primers. Subsequently, the sequence data were analyzed using the PHRED-PHRAP package (http://www.phrap.org/). The identified full-length sequences of *SsNAC23* and *SsMAPK* were submitted to GenBank database under accession number CA094345 and CA150281, respectively.

RNA-Blot Analysis

RNA was isolated using the Trizol reagent (Invitrogen, USA) as described by manufacturer and 15 μg of total RNA were electrophoresed in a 1 % (w/v) agarose gel containing formaldehyde and transferred to a Hybond-N+ filter (Amersham Pharmacia Biotech., USA) as described by Sambrook et al. (1989). RNA blots were hybridized with 3'-fragment of *SsNAC23* or *SsMAPK* full-length cDNA, both labeled with α-^{32}P dCTP and hybridization was done at 42^{o}C for 16 h (Sambrook et al., 1989). The blots were then washed at high stringency and exposed to imaging plates. Each RNA-blot experiment was repeated twice. Later, the blots containing total RNA from sugarcane organs/tissues and different treatments were rehybridized with a α-^{32}P-labeled sugarcane EST encoding a beta-actin protein under same conditions.

Nuclear Localization Analysis of SsNAC23

The SsNAC23 coding region (890 pb) was amplified by PCR using internal primers (forward: 5'-GGATCCGCGATTGAGCGAGGGAGGAA-3' and reverse: 5'-GGATCCGAACGGCTTGCCCCAGTA CA-3'). The single PCR product was isolated from agarose gel using Concert^{TM} rapid gel extraction system (Invitrogen, USA), then subcloned into the pGEM-T EASY vector (Promega, USA) and resequenced. Further, the vector was digested at BamHI restriction site included at the 3'and 5'-ends and ligated into pBI221 plant expression vector (Clontech, USA) downstream of the CaMV 35S RNA promoter and fused in-frame upstream of the *β-glucuronidase* (*GUS*) reporter gene. The construct was checked by appropriated digestion and DNA sequencing.

A helium biolistic gene transformation system (Bio-Rad, USA) was used to transiently transform onion (*Allium cepa*) epidermis cells with the SsNAC23::GUS fusion protein construct. Onion cells were also transformed with pBI221 as a negative control. Inner epidermal single cell layers of the onions were peeled and placed on MS solid medium (Murashige and Skoog, 1962). Five milligrams of column (Qiagen, USA) purified and water dialyzed plasmid DNA was precipitated onto 1.6 mm gold particles

(Bio-Rad, USA) using 2.5 M $CaCl_2$ and 1 M spermidine (Sigma, USA). DNA coated particles were rinsed with absolute ethanol, resuspended in ethanol, and used for bombardment of the onion cells at 1300 psi. The bombarded tissue was incubated for approximately 18 h in the dark before histochemical analysis. SsNAC23::GUS fusion protein was detected in transformed cells by histochemically staining the bombarded onion epidermis with X-gluc (Jefferson et al., 1987). The position of the nucleus in each cell was determined by staining with 4', 6-diamidino-2-phenylindole (DAPI, Varagona et al., 1992). X-gluc staining was detected and photographed in transformed cells under a Nikon Eclipse E600 (Nikon, Japan) microscope with bright-field optics, and the DAPI stained nuclei of the same cells were visualized using fluorescence optics.

Determination of SsNAC23 DNA-binding NAC domain three-dimensional structure

The structure of the DNA-binding NAC of *Arabidopsis* (pdb 1ut7, Ersnt et al., 2004) has been used to predict the 3-D structure of the ssNAC23. We have taken advantage of the 68% sequence identity between ANAC (chain A) and SsNAC23 to map the SsNAC23 sequence in the three-dimensional structure described for chain A of ANAC in 1ut4 structure. The atomic coordinates of the 1ut7 pdb structure, solved at 1.7 Å resolution, was used as a template in comparative modeling by satisfaction of spatial restraints (Sali and Blundell, 1993) implemented in the program MODELLER 6v2 (Fiser et al., 2000). In SsNAC23 modeling procedure, we generated 25 different models of the protein structure. The quality of the predicted fold was evaluated in MODELLER using the score of the variable target function method (Fiser et al., 2000). The stereochemical quality of the five best scoring models was assessed by PROCHECK (Laskowski et al., 1993) at the same resolution as the 1ut7 structure. The final model was selected based on the overall stereochemical quality. Obtained model was subjected to relaxation and energy minimization using GROMACS (Lindahl et al., 2001).

To refine the molecular model of the SsNAC23 structure, additional energy minimization and equilibrating molecular dynamics simulations were carried out in the program GROMACS 3.14 (Lindahl et al., 2001) on a Linux workstation. Initial protein model was submitted to a steepest-decent (SD) energy

minimization (5000 steeps) to remove bad van der Waals contacts. For further relaxation of the obtained structure an unrestrained molecular dynamic was performed for 100 ps. in water (spc216 model) with Berendsen-type temperature (312 K) and pressure (1 atm) coupling in a 30 X 30 X 30 simulation cell (Lindahl et al., 2001; Pellegrini et al., 2001; Louie et al., 2002), implementing the PME method (Essman et al., 1995). An unrestrained multiple step conjugate-gradient (CG) minimization process was used (to 0.1 kJ mol $^{-1}$ nm $^{-1}$) to obtain the energy minimized structure of our protein. For the superimposition of target and modeled backbones, a combinatorial extension method (CE) was used (Shindyalov et al., 1998). Obtained ssNAC23 model is available for non-profit uses upon direct request to author.

ACKNOWLEDGMENTS

The authors thank A. S. Zanca, S. Meire, D. Althmann and A L. Beraldo for excellent technical assistance, Dr. N. Zanchin for helping in the microscopic analyses, and Dr. Lo Leggio for sharing the ANAC PDB data before publication.

REFERENCES

Aguan K, Sugawara K, Suzuki N, Kusano T (1993) Low-temperature-dependent expression of a rice gene encoding a protein with a leucine-zipper motif. Mol Gen Genet **240**:1-8.

Aida M., Ishida T, Fukaki H, Fujisawa H and Tasaka, M (1997) Genes involved in organ separation in *Arabidopsis*: an analysis of the cup-shaped cotyledon mutant. Plant Cell **9**: 841–857.

Bailey TL, Gribskov M (1998) Combining evidence using p-values: application to sequence homology searches. Bioinformatics **14**: 48-54.

Bateman A, Birney E, Durbin R, Eddy SR, Howe KL, Sonnhammer, ELL (2000) The Pfam protein families database. Nucleic Acids Res **28:** 263-266

Chazen O, Neumann, PM (1994) Hydraulic signals from the roots and rapid cell wall hardening in growing maize leaves, are primary responses to PEG induced water deficits. Plant Physiol **104**: 1385-1392.

Chen, W, Provart, NJ, Glazebrook, J, Katagiri, F, Chang, H-S, Eulgem, T, Mauch, F, Luan, S, Zou, G, Whitham, SA, Budworth, PR, Tao, Y, Xie, Z, Chen, X, Lam, S, Kreps, JA, Harper, JF, Si-Ammour, A, Mauch-Mani, B, Heinlein, M, Kobayashi, K, Hohn, T, Dangl, JL, Wang, X, Zhu, T. (2002) Expression profile matrix of *Arabidopsis* transcription factor genes suggests their putative functions in response to environmental stresses. Plant Cell **14**: 559–574.

Collinge, M, Boller, T (2001) Differential induction of two potato genes, *Stprx2* and *StNAC*, in response to infection by *Phytophthora infestans* and to wounding. Plant Mol Biol **46**:521-529.

Duval, M, Hsieh, T-F, Kim, SY, Thomas TL (2002) Molecular characterization of AtNAM: a member of the *Arabidopsis* NAC domain superfamily. Plant Mol Biol **50**: 237-248.

Ernst HA, Olsen AN, Skriver K, Larsen A, Leggio LL (2004) Structure of the conserved domain of ANAC, a member of the NAC family of transcription factors. EMBO rep. **5**: 1-7.

Essman U, Perela L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) A smooth particle mesh Ewald method. J. Chem. Phys. **103**: 8577-8592.

Fiser A, Do RK, Sali A (2000) Modeling of loops in protein structures. Protein Sci. **9**: 1753-1773.

Fowler S, Thomashow MF (2002) *Arabidopsis* transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway. Plant Cell **14**: 1675-1690

Hegedus, D, Yu, M, Baldwin, D, Gruber, M, Sharpe, A, Parkin, I, Whitwill, S, Lydiate, D (2003) Molecular characterization of *Brassica napus* NAC domain transcriptional activators induced in response to biotic and abiotic stress. Plant Mol Biol **53**: 383–397.

Holmberg N, Bülow L (1998) Improving stress tolerance in plants by gene transfer. Trends Plant Sci **3**: 61-66.

Jefferson, RA, Kavanagh TA, Bevan, MW. (1987) GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J **20**: 3901-3907.

Jiang M, Zhang J. (2002) Water stress-induced abscisic acid accumulation triggers the increased generation of reactive oxygen species and up-regulates the activities of antioxidant enzymes in maize leaves. J Exp Bot **53**: 2401-2410.

Jin, H. and Martin, C (1999) Multifunctionality and diversity within the plant MYB-gene family. Plant Mol. Biol. **41**: 577–585.

Kawasaki S, Borchert C, Deyholos M, Wang H, Brazille S, Kawai K, Galbraith D, Bohnert HJ (2001) Gene expression profiles during the initial phase of salt stress in rice. Plant Cell **13**:889-905.

Kikuchi K, Ueguchi-Tanaka M, Yoshida KT, Nagato Y, Matsusoka M, Hirano H-Y (2000) Molecular analysis of the *NAC* gene family in rice. Mol Gen Genet **262**: 1047-1051

Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA2: molecular evolutionary genetics analysis software. Bioinformatics **17**: 1244-1245.

Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Cryst. **26**: 283-291.

Lindahl E, Hess B, van der Spoel D (2001) GROMACS 3.0: A Package for molecular simulation and trajectory analysis. J. Mol. Mod. **7**: 306-317.

Louie TM, Yang H, Karnchanaphanurach P, Xie XS, Xun L (2002) FAD is a preferred substrate and an inhibitor of *Escherichia coli* general NAD(P)H:flavin oxidoreductase. J Biol Chem. **277**: 39450-39455.

Maleck K, Levine A, Eulgem T, Morgan A, Schmid J, Lawton KA, Dangl JL, Dietrich, RA (2000) The transcriptome of *Arabidopsis thaliana* during systemic acquired resistence. Nature Gen **26**: 403-410.

Money, NP (1989) Osmotic pressure of aqueous polyethylene glycols: relationship between molecular weight and vapor pressure deficit. Plant Physiol **91**: 766-769.

Murashige T, Skoog F (1962) A revised medium for rapid growth and biossays with tobacco tissue culture. Physiol Plant **15**: 471-497.

Nogueira FTS, De Rosa Jr VE, Menossi M, Ulian EC, Arruda P (2003) RNA expression profiles and data mining of sugarcane response to low temperature. Plant Physiol **132**: 1811-1824.

Ooka, H, Satoh, K, Doi, K, Nagata, T, Otomo, Y, Murakami, K, Matsubara, K, Osato, N, Kawai, J, Carninci, P, Hayashizaki, Y, Suzuki, K, Kojima, K, Takahara, Y, Yamamoto, K, Kikuchi, S. (2003) Comprehensive Analysis of NAC Family Genes in *Oryza sativa* and *Arabidopsis thaliana*. DNA Res **10**: 239 –247.

Pellegrini M, Bremer AA, Ulfers AL, Boyd ND, Mierke DF (2001) Molecular characterization of the substance P*neurokinin-1 receptor complex: development of an experimentally based model. J Biol Chem. **276**: 22862-22867.

Provart, N.J., Gil, P., Chen, W., Han, B., Chang, H-S, Wang, X, Zhu (2003) Gene expression phenotypes of *Arabidopsis* associated with sensitivity to low temperatures. Plant Physiol **132**: 893-906.

Rabbani MA, Maruyama K, Abe H, Khan MA, Katsura K, Ito Y, Yoshiwara K, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2003) Monitoring expression profiles of rice genes under cold, drought, and high-salinity stresses and abscisic acid application using cDNA microarray and RNA gel-blot analyses. Plant Physiol **133**: 1755-1767.

Rechsteiner M, Rogers SW (1996) PEST sequences and regulation by proteolysis. Trends Biochem Sci. **21**: 267-271.

Ren T, Qu F, Morris TJ (2000) *HRT* gene function requires interaction between a NAC protein and viral capsid protein to confer resistance to turnip crinkle virus. Plant Cell **12**: 1917-1925

Robatzek, S. and Somssich, I.E (2001) A new member of the *Arabidopsis* WRKY transcription factor family, AtWRKY6, is associated with both senescence- and defence-related processes. Plant J. **28**: 123–133.

Sablowski, RWM, Meyerowitz, EM (1998) A homolog of *NO APICAL MERISTEM* is an immediate target of the floral homeotic genes *APETALA3/PISTILLATA*. Cell **92**: 93-103.

Sali A, Blundell TL.Comparative protein modeling by satisfaction of spatial restraints. 1993. J. Mol. Biol. **234**: 779-815.

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, Ed 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Seki M, Narusaka M, Ishida J, Nanjo T, Fujita M, Oono Y, Kamiya A, Nakajima M, Enju A, Sakurai T, Satou M, Akiyama K, Taji T, Yamaguchi-Shinozaki K, Carninci P, Kawai J, Hayashizaki Y, Shinozaki K. (2002) Monitoring the expression profiles of 7000 *Arabidopsis* genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. Plant J **31**:279-292.

Shindyalov IN, Bourne PE (1998) Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. Prot. Eng. **11**: 739-747.

Singh, K.B., Foley, R.C. and Onate-Sanchez, L (2002) Transcription factors in plant defense and stress responses. Curr. Opin. Plant Biol. **5**: 430–436.

Tai PYP, Lentini RS (1998) Freeze damage of Florida sugarcane. In DL Anderson, eds, Sugarcane handbook, Ed 1. Florida Cooperative Extension Service, University of Florida, Florida, pp. 1-3

Takada, S, Hibara, K-I, Ishida, T, Tasaka, M (2001) The *CUP-SHAPED COTYLEDON* gene of *Arabidopsis* regulates shoot apical meristem formation. Development 128: 1127-1135.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res **25**: 4876-4882.

Varagona, MJ, Schmidt, RJ, Raikhel, NV (1992) Nuclear localization signal (s) required for nuclear targeting of the maize regulatory protein Opaque-2. Plant Cell **4**: 1213-1227.

Vettore AL, da Silva FR, Kemper EL, Souza GM, da Silva AM, Ferro MIT, Henrique-Silva F, Giglioti EA, Lemos MVF, Coutinho LL et al. (2003) Analysis and functional annotation of an expressed sequence tag collection for tropical crop sugarcane. Genome Res **13**:2725-2735.

Vincentz M, Bandeira-Kobarg C, Gauer L, Schlögl P, Leite A (2003) Evolutionary pattern of angiosperm bZIP factors homologous to the maize Opaque2 regulatory protein. J Mol Evol **56**:105-116.

Vroemen CW, Mordhorst AP, Albrecht C, Kwaaitaal, MACJ, de Vries SC (2003) The *CUP-SHAPED COTYLEDON3* gene is required for boundary and shoot meristem formation in *Arabidopsis*. Plant Cell **15**: 1563-1577.

Xie, Q, Frugis, G, Colgan, D, Chua, NH (2000) *Arabidopsis* NAC1 transduces auxin signal downstream of TIR1 to promote lateral root development. Genes Dev **14**: 3024–3036.

Xiong L, Schumaker KS, Zhu JK (2002) Cell signaling during cold, drought, and salt stress. Plant Cell **14**: S165-S183.

Xiong L., Yang Y (2003) Disease resistance and abiotic stress tolerance in rice are inversely modulated by an abscisic acid-inducible mitogen-activated protein kinase. Plant Cell **15**: 745-759.

Wen JQ, Oono K, Imai R. 2002. Two novel mitogen-activated protein signaling components, OsMEK1 and OsMAP1, are involved in a moderate low-temperature signaling pathway in rice. Plant Physiol. **129**: 1880-1891.

Wicker, T., Stein, N., Albar, L., Feuillet, C., Schlagenhauf, E., Keller, B (2001) Analysis of a contiguous 211 kb sequence in diploid wheat (*Triticum monococcum* L.) reveals multiple mechanisms of genome evolution. Plant J. **26**: 307-316.

Zhang F, Zhu L, He G (2004) Differential gene expression in response to brown planthopper feeding in rice. J Plant Physiol **161**: 53-62.

Table 1. Summary of the SsNACs identified in the SUCEST database.

Name	SAS ID	Acc. No.[a]	EST No.	Highest identity[b]	Acc. No.[c]
SsNAC1	SCAGRT2039C10	CA137402	6	ONAC040	BAD09611
SsNAC2	SCBGLR1118D05	CA119152	15	ANAC014	NM_202225
SsNAC3	SCCCCL4009H05	CA094553	2	OsNAC7	AB028186
SsNAC4	SCCCCL5072D12	CA095553	3	ANAC033	NM_202449
SsNAC5	SCCCCL6005A08	CA096752	6	OsNAC8	BAA89802
SsNAC6	SCCCFL4092A07	CA235099	1	*Triticum* sp.GRAB2	AJ010830
SsNAC7	SCCCLB1024F12	CA111892	5	ANAC079	NM_120850
SsNAC8	SCCCLR1070C08	CA120159	19	OsNAC3	AB028182
SsNAC9	SCCCLR1C04H01	CA189789	18	ANAC083	NM_121321
SsNAC10	SCCCRZ1004G12	CA147133	3	ANAC036	NM_127259
SsNAC11	SCCCST3005F09	CA180484	4	ANAC073	NM_118992
SsNAC12	SCEPLB1042D02	CA112161	22	ANAC053	NM_111885
SsNAC13	SCEPRT2048G05	CA138286	3	ANAC021	NM_179486
SsNAC14	SCEQRT1024G08	CA132543	16	OsNAC5	AB028184
SsNAC15	SCEQRT1028D09	CA132861	1	PhNAM2	AAM34765
SsNAC16	SCEQRT1031E06	CA133129	11	ONAC007	AK062952
SsNAC17	SCEQRT2030D12	CA138754	6	ANAC041	NM_128908
SsNAC18	SCEZRT2019E05	CA140061	3	ANAC087	NM_203071
SsNAC19	SCEZRT2023G12	CA140418	8	*Triticum* sp. GRAB1	CAA09371
SsNAC20	SCJFRT1009A08	CA133624	2	ANAC100	NM_125536
SsNAC21	SCJLRT1015C08	CA135079	2	ANAC056	NM_112419
SsNAC22	SCJLRT1018B06	CA135405	1	*P. hybrida* NAC protein	AC007454
SsNAC23	SCCCCL4007D12	CA094345	28	OsNAC6 protein	AB028185
SsNAC24	SCMCLB2082D05	CA271774	1	*N. tabacum* NAC protein	AB021178
SsNAC25	SCMCRT2106G06	CA142777	1	ANAC021	NM_179486
SsNAC26	SCQGRT3042B03	CA260833	1	ANAC104	NM_125849

[a]Accession numbers refer to SAS containing entire NAC domain sequence.

[b]Highest identity based on Ooka et al. (2003) when applicable.

[c]Accession numbers refer to NAC domain proteins homologs to identified SsNACs.

Table 2. Summary of tissue/organ-specific library pools from the SUCEST database

EST library ID[a]	Tissue/organ-specific library pools[b]	Total no. of ESTs
CL6	Calli	5,518
AM1, AM2, LB1, LB2	Meristematic tissues	39,116
FL1, FL3, FL4, FL6, FL8	Floral organs	52,430
LV1, LR1, LR2	Leaves	19,676
RT1, RT2, RT3	Roots	25,302
ST1, ST3, SB1, RZ1, RZ2, RZ3	Stem	49,785

[a]Identification of each EST library within the SUCEST database (Vettore et al., 2003).

[b]EST libraries were grouped into seven library pools according to the source tissues from which the library was derived.

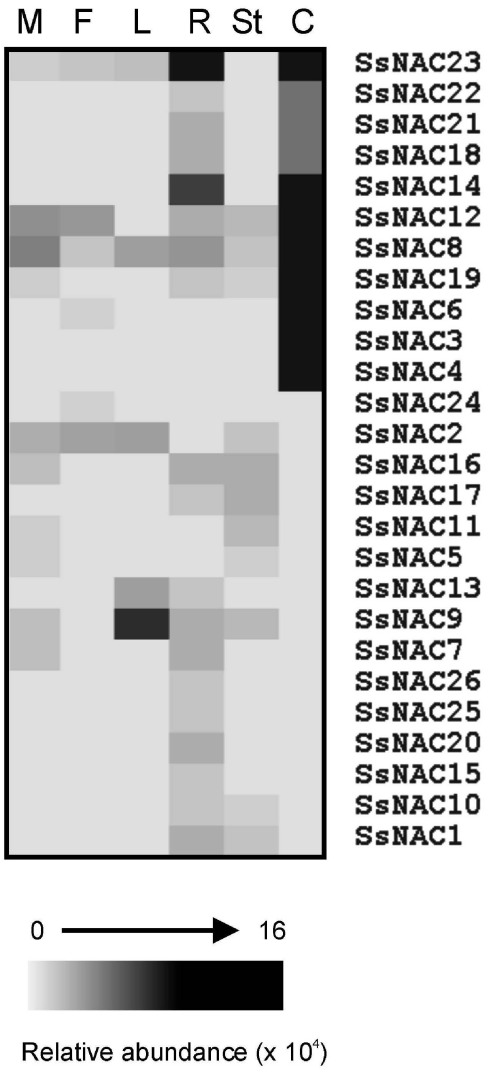


Figure 1. Relative abundance of SsNAC-representing ESTs in sugarcane tissue/organ-specific library pools. Each SsNAC is represented by a single row and each library pool is represented by a single column. The relative abundance of ESTs was calculated as a count of the number of SsNAC ESTs in the given library pool normalized by the total number of ESTs in that pool (Table 2). The normalized values were analyzed by CLUSTER program (Eisen et al., 1998) and the image generated by Tree View software downloaded from http://rana.lbl.gov. The SsNAC names are shown on the right and the color intensity scale representing the relative abundance is indicated at the bottom. C, calli; M, meristematic tissues; F, floral organs; L, leaves; R , roots; and St, stem.

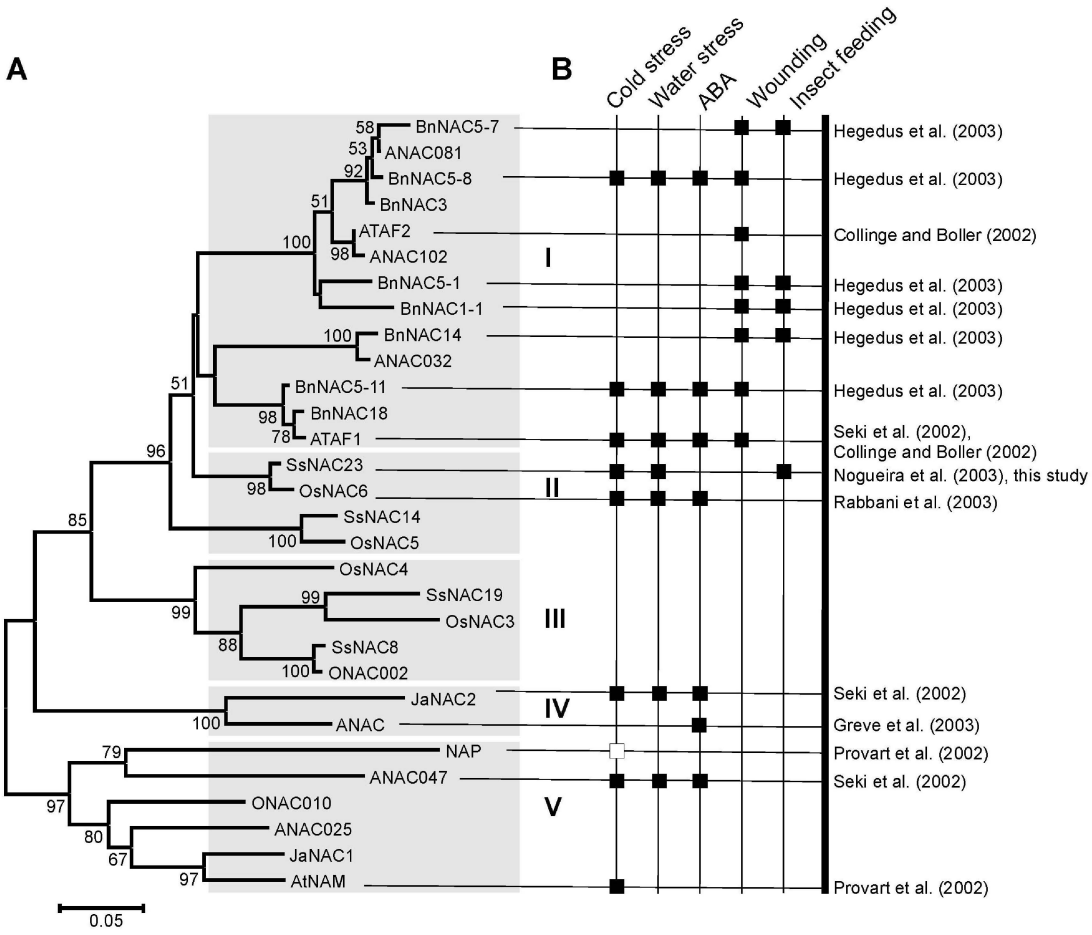
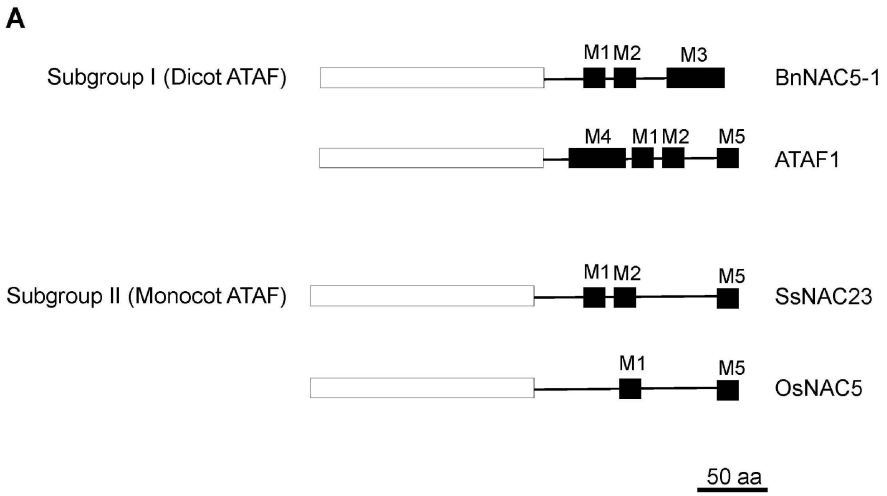


Figure 2. Comparison of deduced amino acid sequences and expression patterns of NAC domain genes. (A) Unrooted phylogram of *Arabidopsis*, *Brassica napus*, rice and sugarcane NAC domain proteins. The five subgroups distinguished within the ATAF, NAP and OsNAC3 subgroups (Ooka et al., 2003) are showed in gray rectangles. Phylogram was generated as a consensus of 1000 bootstrap replicates by the neighbor joining method (the bootstrap values are indicated close the branch divisions, when > 50%). The scale bar indicates the relative amount of change along branches. The accession numbers are given in the Material and Methods. (B) Comparison of expression profiles of *AtNACs, OsNACs, BnNACs*, and *SsNACs* and their respective references. Full squares represent an induction in the transcription level whereas empty squares represent downregulated mRNA levels in response to the stress/treatment under consideration.



B

M1: T[S/T][D/S]SS[C/G]S[E/S][H/Q]V[V/I/L]SP[D/E][F/V]

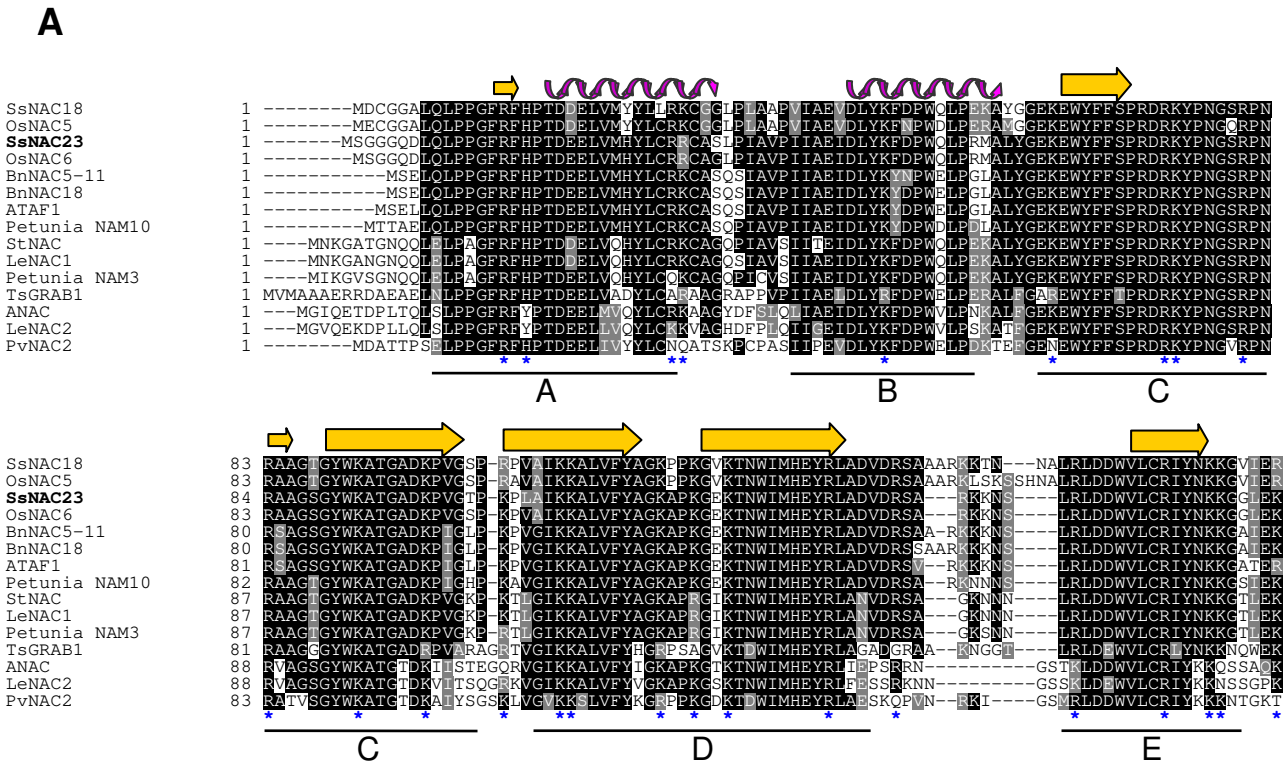
M2: [S/R]EVQS[E/Q]PK[W/I][G/K/S][E/D][W/L][E/S][D/R][A/T]

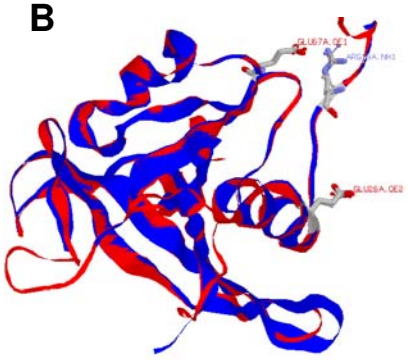
M3: LLQ[N/S]DAFVPQF[L/Q]YQ[S/P]D[F/Y][F/A/V][T/D/S][S/P]
F[Q/E]DP[F/P]EQKPFLNW[S/N]F[A/G]PQG

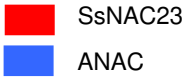
M4: VYGDE[V/I][V/M]EEKP[R/K][L/V][S/T]EM[G/V]MPPPP[V/Q]
[M/Q][P/T][N/S][D/E]F[V/A]YFDTSDSVPKLH

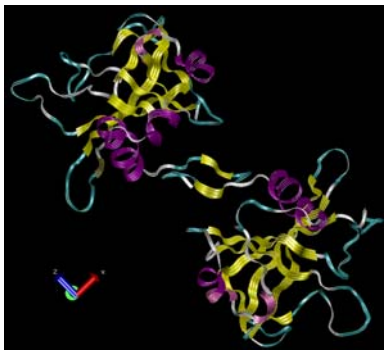
M5: [D/L/M][F/P][L/P]LQD[M/I][F/L]MY[W/M/N][G/M/Q][K/P][P/K][F/P]

Figure 3. Conserved motifs in transcriptional activation regions. (**A**) NAC domains are shown by white rectangles while motifs in transcriptional activation regions are shown by black rectangles. (**B**) Amino acid sequences of each motif found by MEME-MAST program (Bailey and Gribskov, 1998) are indicated.


A
SsNAC18
OsNAC5
SsNAC23
OsNAC6
BnNAC5-11
BnNAC18
ATAF1
Petunia NAM10
StNAC
LeNAC1
Petunia NAM3
TsGRAB1
ANAC
LeNAC2
PvNAC2
A
B
C
C
D
E




SsNAC23
ANAC



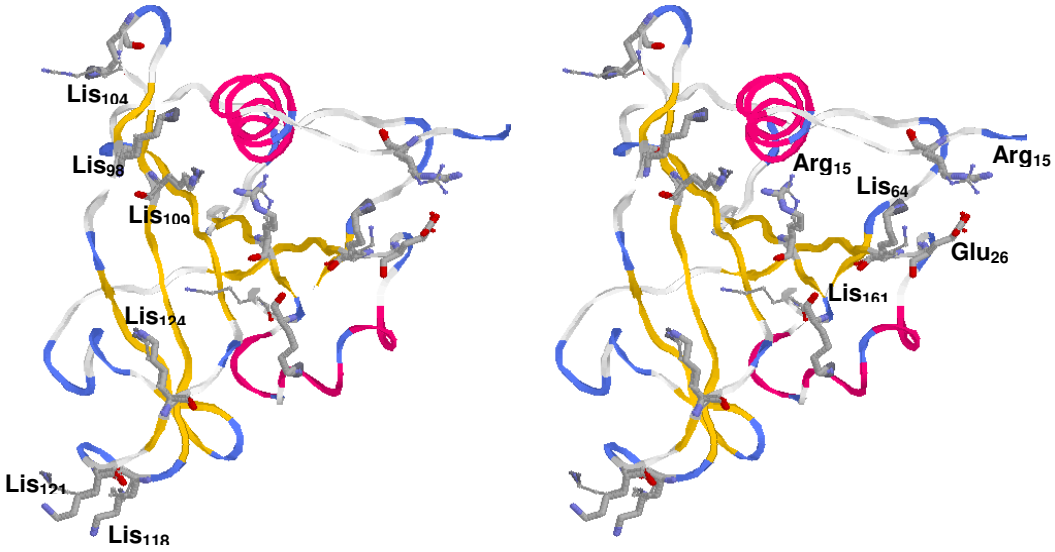
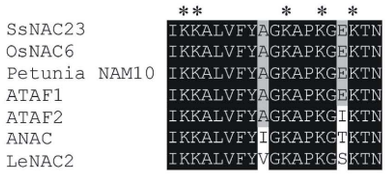

Lis104
Lis98
Lis109
Lis124
Lis121
Lis118
Arg15
Arg15
Lis64
Glu26
Lis161

Figure 4. NAC domain sequence and structure conservation of SsNAC23 protein. (A) Alignment of 15 NAC domain protein sequences from different plant species. Only amino acids belonging to NAC domain sequence were used in the alignment. Subdomains A to E are shown by lines below the sequences. Amino acid character: white on black, identical residues and white on light gray, conserved residues. Arg, Lys and His residues, conserved in these sequences, are marked by a blue star and the secondary elements described in ANAC structure (Ersnt et al., 2004) represented on the top of the alignment. The accession numbers for each protein sequence (in parentheses) are given in Material and Methods, except for Petunia NAM3 and NAM10 (AAM34766 and AAM34773, respectively), potato StNAC (CAC42087), tomato LeNAC1 and LeNAC2 (AAR88435 and AAF04915, respectively), wheat TsGRAB1 (CAA09371), and *Phaseolus* PvNAC2 (AAK84884). (B) Structural alignment of the obtained SsNAC23 model (in red) and ANAC structure (in blue) backbones. Conserved Arg19 and Glu26 residues responsible for interdomain salt bridges are represented in CPK (thick lines for SsNAC23 model and thin lines for ANAC). (C) Structure of the SsNAC23 backbone dimer model represented and colored according to secondary structure. (D) Stereo representation of the DNA-contacting basic surface of the SsNAC23 structure, colored according to the secondary structure. Conserved basic residues in the internal surface, important to DNA-binding function (Arg and Lys), were structurally aligned and represented in CPK (thick lines for SsNAC23 model and thin lines for ANAC).

A



B

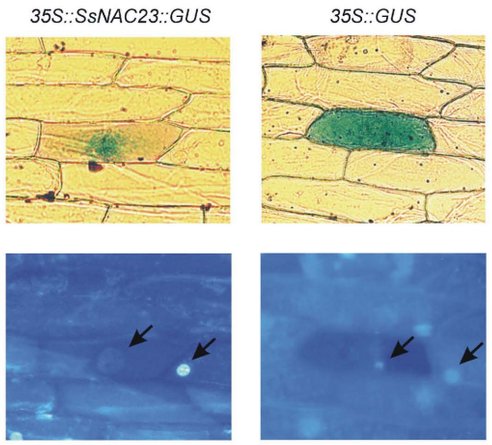


Figure 5. Nuclear localization of SsNAC23 in onion epidermis cells. (A) Alignment of degenerate bipartite nuclear localization motif using CLUSTALX program (Thompson et al., 1997) from selected NAC domain proteins. Accession numbers of SsNAC23, ATAF1, ATAF2, ANAC and OSNAC6 are specified in Material and Methods while those of Petunia NAM10 e *Lycopersicum esculentum* NAC domain protein 2 (LeNAC2) are described in Figure 4. Amino acid character: white on black, identical residues and white on light gray, conserved residues. Asterisks indicate the presence of the basic amino acid lysine (K). (B) Histochemical staining of GUS activity in cells 24 h after particle bombardment (upper panel) and DAPI staining of the same cells to identify nuclei (lower panel) were photographed with a Nikon Eclipse E600 (X 400). Arrows indicate the nuclei.

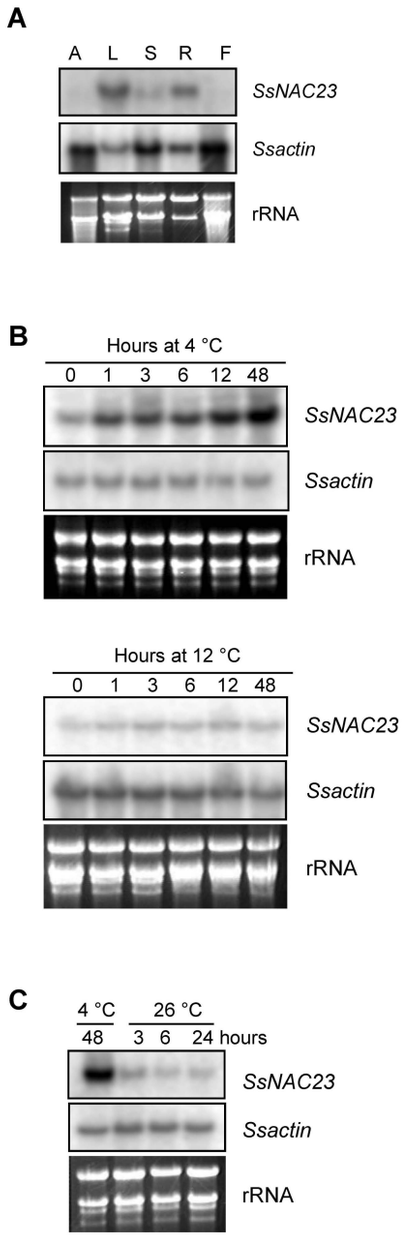


Figure 6. RNA-blot analysis of the *SsNAC23* expression. (A) Organ-enriched expression of *SsNAC23*. A, apical meristem; L, leaves; S, stem; R, roots; F, floral organs. (B) Time course of *SsNAC23* expression in response to low temperature treatments in sugarcane plantlets. Plantlets were subjected to 4 °C or 12 °C for the indicated periods (hours). (C) A set of sugarcane plantlets growing under 4 °C for 48 h were returned to 26 °C and kept for 3, 6, and 24 h, as indicated above the panel. Further, all filters were rehybridized with the ^{32}P-labeled sugarcane EST fragment of *actin* (*Ssactin*) as a control. The sugarcane EST clone (accession number Ca192497) encoding protein very similar to actin from *Triticum monococcum* (accession number AAK84080; Wicker et al., 2001).

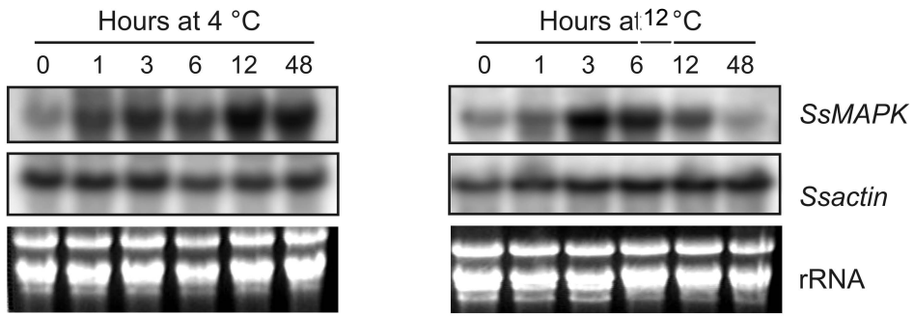


Figure 7. RNA-blot analysis of the *SsMAPK* expression. Three-month-old plantlets were exposed to 4 °C or 12 °C for the indicated periods (hours). RNA-blots were hybridized with the ^{32}P-labeled entire cDNA fragment of *SsMAPK*. Further, filters were rehybridized with the ^{32}P-labeled EST fragment of *Ssactin* as a control.

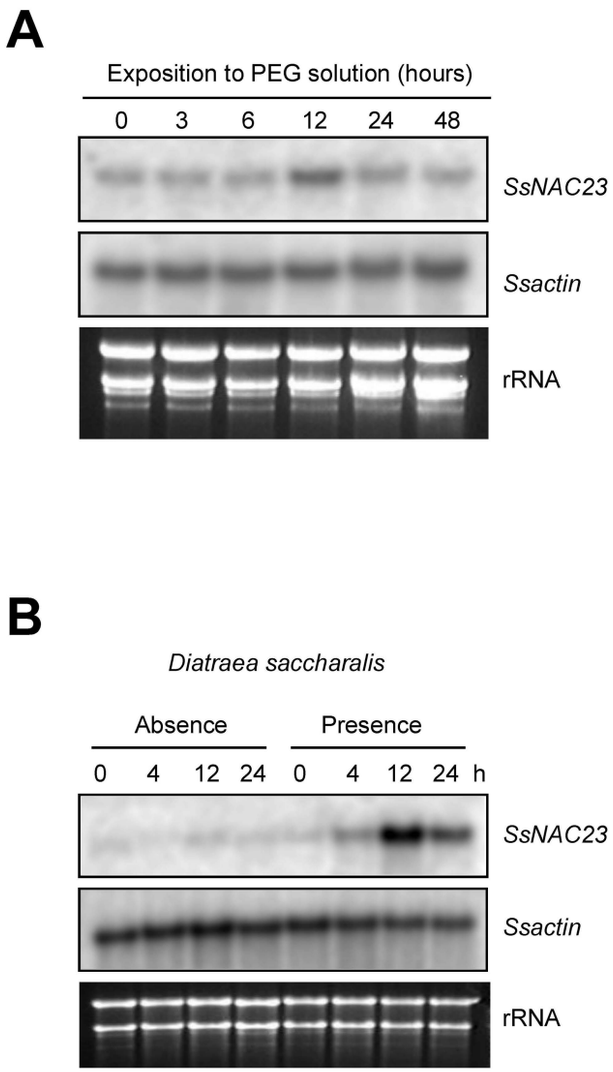


Figure 8. RNA-blot analysis of the *SsNAC23* expression in response to water stress and insect feeding. Plantlets were subjected to PEG-induced water stress (A) and sugarcane borer feeding (B) for the indicated periods (hours). Further, all filters were rehybridized with the ^{32}P-labeled EST fragment of *Ssactin* as a control.

DISCUSSÃO

Como cada um dos capítulos presentes no corpo da tese apresenta uma discussão dos seus respectivos resultados, esta seção constará de uma discussão geral em que se abordará a contribuição destes trabalhos no entendimento da análise do perfil de expressão de diferentes genes em cana-de-açúcar em resposta ao estresse por baixas temperaturas.

1. Análise do perfil de expressão gênica utilizando macroarranjos de DNA

Vários fatores, tais como variações na massa de DNA depositada na membrana, flutuações durante a síntese da sonda de cDNA e a hibridação e *background*, influenciam a reprodutibilidade dos resultados obtidos com os macroarranjos de DNA (Desprez et al., 1998; Perret et al., 1998; Schummer et al., 1999; Schuchhardt et al., 2000). A intensidade de sinal do *spot* detectada nas hibridações de ácidos nucléicos, por exemplo, é proporcional não só à massa do transcrito na sonda, mas também à massa de DNA depositada na membrana (Lara, 1995). Para estudar esse fator, foram realizados experimentos para avaliar o efeito das variações de massa de DNA fixada e o número de vezes em que a amostra é aplicada na membrana (capítulo I). Os resultados demonstraram que a estratégia de uso de sonda radioativa sintetizada a partir de seqüência específica complementar a uma região comum a todas as amostras de DNA (isto é, ao gene de resistência ao antibiótico ampicilina, *Amp*^{r}) permitiu o acompanhamento da variação de massa de DNA entre as membranas réplicas (Fig. 3, capítulo I). Como a metodologia de macroarranjos de DNA sugere o uso de membranas réplicas hibridadas com sondas de cDNA advindas de diferentes experimentos, a estimativa de variabilidade da massa de DNA fixada nessas membranas é uma etapa extremamente importante para evitar interpretações errôneas durante a análise do perfil de expressão gênica. Além disso, fixar cada amostra de DNA duas vezes numa mesma posição sobre a membrana de náilon reduziu o coeficiente de variação em cerca de 50% (dados não mostrados). Desta forma, ambas as estratégias permitiriam restringir possíveis imperfeições ocorridas durante a feitura das membranas, etapa crucial nos experimentos de *macroarrays*.

Como predito por Schuchhardt et al. (2000), a hibridação não-específica entre moléculas de ácidos nucléicos é uma típica fonte de erros que não deve ser ignorada. Como fixamos diretamente o DNA plasmidial sobre as membranas, é possível que ocorra pareamento inespecífico entre as seqüências de cDNA presentes na população de sondas e a seqüência do plasmídio utilizado na confecção das bibliotecas de cDNA do SUCEST (pSPORT1; Vettore et

al., 2001). Desta forma, foram fixadas nas membranas réplicas algumas amostras de DNA do vetor pSPORT sem inserto, o qual funcionou como um controle de hibridação inespecífica (capítulo II). Reymond et al. (2000) e Schenk et al. (2000) utilizaram clones-controle para análise de hibridação inespecífica e eliminaram *spots* em que os valores de intensidade obtidos com a sonda de cDNA foram menores do que ou iguais àqueles obtidos com os clones-controle. Estratégia similar foi adotada nas análises de nossos macroarranjos de DNA (capítulo II).

Além do exame dos fatores que podem promover variabilidade durante a preparação e condução de experimentos utilizando arranjos de DNA, os procedimentos de análise dos dados gerados são, do mesmo modo, extremamente importantes. Como apresentado no capítulo I, o uso do valor da mediana como fator de normalização, obtido a partir dos dados de intensidade de todos os *spots* de cada membrana réplica, mostrou-se bastante satisfatório, posto que o coeficiente de variação (CV) entre *spots* hibridados com mesma sonda de cDNA foi reduzido em cerca de três vezes quando comparado aos dados não normalizados (Fig. 4, capítulo I). Além disso, o valor médio de todos os *spots* de cada membrana também foi empregado com sucesso neste procedimento de normalização, como apresentado no capítulo II, demonstrando que tanto o valor da mediana como o valor da média podem ser utilizados no processo de normalização dos dados de intensidade de arranjos de DNA. Resultados similares foram obtidos por Schuchhardt et al. (2000).

Finalmente, parâmetros como a média e o desvio-padrão das razões de expressão (ou seja, a relação entre os valores de intensidade do sinal normalizados obtidos com a sonda tratamento e com a sonda controle) foram empregados para identificar os ESTs diferencialmente expressos nas condições especificadas (capítulos I e II). Friddle et al. (2000) utilizaram análise semelhante para identificar genes que apresentaram expressão variável durante hipertrofia cardíaca em camundongos. A principal premissa para o uso desta estratégia é a de que os valores em escala logarítima ajustam os dados o mais próximo a uma distribuição normal, dando maior confiabilidade aos resultados obtidos (Zar, 1999).

Entretanto, dependendo do tipo de *design* experimental, várias metodologias são empregadas para avaliar a expressão diferencial via arranjos de DNA. Por exemplo, Demura et al. (2001), estudando o perfil de expressão gênica em células de *Zinnia elegans* durante a xilogênese, avaliaram a alteração de expressão gênica baseada na diferença entre os valores máximo e mínimo de intensidade de cada *spot* durante todo o curso do experimento. Diferentemente, quando compararam a expressão gênica entre o endosperma e outros tecidos/órgãos de milho, Fernandes et al. (2002) utilizaram genes-controle (ou seja, genes de

expressão constitutiva em quase todos os tecidos/órgãos da planta) para identificar aqueles preferencialmente expressos no endosperma dentro do grupo de genes de interesse.

2. Resposta molecular da cana-de-açúcar ao estresse por baixa temperatura

Como apresentado no capítulo II, o perfil de expressão no tecido foliar de cana-de-açúcar (cv. SP80-3280) foi monitorado via *macroarrays* durante a exposição das plantas a baixas temperaturas (3 h - 48 h, 4 °C). Um pouco mais de 1.500 ESTs do projeto SUCEST foram analisados e, similarmente aos dados relatados para *Arabidopsis thaliana* (Seki et al., 2002; Fowler e Thomashow, 2002; Kreps et al., 2002; Provart et al., 2003) e arroz (Rabbani et al., 2003), várias classes de genes em cana-de-açúcar tiveram sua expressão alterada, incluindo fatores de transcrição, proteínas quinases, proteínas pertencentes a diferentes metabolismos, proteínas específicas de estresse, entre outras (Tabela 2, capítulo II). Fatores de transcrição do tipo *bZIP* e do tipo *NAC* estão entre os genes cuja expressão foi induzida por frio em cana-de-açúcar (capítulo II), em *A. thaliana* (Seki et al., 2002; Kreps et al., 2002; Provart et al., 2003) e em arroz (Rabbani et al., 2003), demonstrando que algumas vias de sinalização em resposta a esse estresse podem ser conservadas tanto em espécies sensíveis quanto em espécies tolerantes. Entretanto, após extensiva mineração do banco de dados do SUCEST, não foi identificado nenhum *cluster* que codifica proteína homóloga à ABI3 ou à ICE1, as quais são importantes componentes do processo de transdução de sinais durante exposição de plantas de *A. thaliana* a baixas temperaturas (Tamminen et al., 2001; Chinnusamy et al., 2003). Todavia, como o genoma da cana-de-açúcar não foi completamente seqüenciado, não se pode descartar a existência desses genes.

A expressão de genes codificadores de proteínas envolvidas em metabolismo primário (i.e., fixação de CO_2, biossíntese de proteínas, etc.), metabolismo secundário (biossíntese da parede celular), transporte de moléculas (transportadores de açúcar), metabolismo antioxidante e em respostas específicas de estresse (i.e., proteínas da família LEA) é alterada tanto em cana-de-açúcar, quanto em *Arabidopsis thaliana* (capítulo II, Seki et al., 2002; Provart et al., 2003). Esses dados sugerem que, embora mono e dicotiledôneas estejam separadas em termos evolutivos por milhões de anos, alterações metabólicas similares ocorrem quando ambas as espécies são submetidas a baixas temperaturas. Recentemente, Vincentz et al. (2004) demonstraram que cerca de 2/3 dos genes expressos em cana-de-açúcar codificam proteínas similares àquelas codificadas pelo genoma de *A. thaliana*.

Apesar de a hipótese de conservação do aparato molecular em cana-de-açúcar envolvido na resposta ao estresse por baixas temperaturas ter sido reforçada pela identificação de proteínas similares àquelas envolvidas no processo de aclimatação ao frio em várias espécies vegetais (Tabela 3, capítulo III), algumas proteínas detectadas no nosso estudo não foram ainda identificadas em outras plantas como sendo alteradas pelo estresse por frio. Um dos exemplos é a proteína poliubiquitina (Tabela 2, capítulo II). Utilizando *microarrays*, Kreps et al. (2002) observaram que vários genes que codificam poliubiquitinas não apresentaram expressão modificada durante a exposição de plantas de *A. thaliana* a baixas temperaturas. Em cana-de-açúcar, por outro lado, a expressão de genes que codificam diferentes poliubiquitinas foi induzida após 12 h de estresse (Tabela 2 e Fig. 5, capítulo II). Estas observações sugerem que, durante a exposição de plantas sensíveis a baixas temperaturas, ocorre aumento de degradação de proteínas, em razão de serem as poliubiquitinas responsáveis por "marcar" as proteínas que serão posteriormente degradadas pelo complexo 26S proteossoma (O'Mahony e Oliver, 1999). Além disso, Callis e Vierstra (2000) sugerem que o processo de proteólise é um importante componente da sinalização em resposta aos estresses abióticos. Outro interessante gene identificado em nossos macroarranjos de DNA, e não descrito na literatura como tendo sua expressão modificada por baixas temperaturas, codifica uma proteína homóloga (62% de similaridade) à proteína AIP2 de *Arabidopsis thaliana* (*ABI3-interacting protein 2*; Kurup et al., 2000). Esta proteína pertence à família de fatores de transcrição do tipo *C3HC4-type RING finger* e interage com a proteína ABI3 durante o desenvolvimento de tecidos embrionários em sementes de *A. thaliana* (Kurup et al., 2000). Entretanto, até o momento, não foi observada nenhuma conexão entre esse fator de transcrição e a resposta a estresses em outras espécies vegetais. Outros exemplos de genes não relatados na literatura, os quais apresentaram expressão alterada em resposta ao estresse por frio em cana-de-açúcar, estão descritos no capítulo II, sugerindo que algumas vias e metabolismos alterados durante a exposição de plantas a baixas temperaturas podem ser espécie-específicos.

A maioria dos genes (70%) que responderam ao tratamento de 4 °C em nossos experimentos apresentou gradual aumento ou diminuição da expressão durante todo o período de estresse, enquanto outros tiveram sua expressão modificada transientemente (Tabela 2, capítulo II). Resultados similares foram observados para o perfil de expressão gênica de plântulas de arroz e *Arabidopsis thaliana* submetidas ao mesmo tipo de estresse (Fowler e Thomashow, 2002; Rabbani et al., 2003), sugerindo que múltiplas vias regulatórias estão envolvidas no processo de aclimatação de plantas às condições de baixas temperaturas.

Finalmente, vários genes que codificam proteínas com funções desconhecidas (53% do total de genes, capítulo II) foram identificados nos macroarranjos de DNA. Análises *in silico* de domínios conservados nas proteínas codificadas por genes desconhecidos, os quais apresentaram expressão induzida por frio, demonstraram que essas proteínas podem estar envolvidas em processos que vão desde a transdução de sinais até os mecanismos de tolerância ao estresse. Futuras análises experimentais fornecerão informações adicionais sobre as possíveis funções desses genes e sua real importância para o processo de aclimatação da cana-de-açúcar ao estresse por frio.

De acordo com os dados apresentados no capítulo II, a cana-de-açúcar parece possuir grande parte do aparato molecular envolvido no processo de aclimatação a baixas temperaturas, o qual está presente em outras espécies vegetais, incluindo plantas de clima temperado, como *A. thaliana*. No entanto, a sensibilidade ao estresse observada em cana-de-açúcar (Tai e Lentini, 1998; Du et al., 1999) pode ser decorrente da ausência de genes que codificam proteínas-chave envolvidas tanto na transdução de sinais moleculares, quanto na ação direta de algumas proteínas na proteção contra injúrias especificamente causadas por esse tipo de estresse. Para exemplificar o último caso, não foi identificado no banco de dados do SUCEST nenhum *cluster* que codifica proteína similar às proteínas COR15a, COR47, COR78, entre outras. Essas proteínas podem funcionar como moléculas crioprotetoras de membranas de cloroplastos (Thomashow, 1999), protegendo contra as injúrias causadas pelo congelamento. A ausência desses genes no genoma de cana-de-açúcar e de outras espécies tropicais e subtropicais pode ser especulada com base no fato de que essas plantas desenvolvem-se em ambientes desprovidos de temperaturas negativas. Outra possibilidade para explicar a sensibilidade de cana-de-açúcar a baixas temperaturas é a regulação diferenciada de alguns genes quando comparada à regulação de seus homólogos em espécies de clima temperado. Jaglo et al. (2001) observaram que o gene *DREB1* de tomate apresentou expressão induzida por frio, mas com perfil desigual daquele observado para seus homólogos em plantas tolerantes a baixas temperaturas, tais como *A. thaliana*, canola, centeio e trigo.

3. Caracterização de genes regulados por baixas temperaturas em cana-de-açúcar

Durante a identificação de proteínas de cana-de-açúcar homólogas àquelas codificadas por genes cuja expressão é induzida por frio em outras espécies vegetais, foi detectado um *cluster* (SCEPAM2011H12, Tabela 3, capítulo II) que codifica proteína similar à proteína desacopladora mitocondrial de plantas (PUMP). Este resultado levantou a questão sobre a

existência de outros genes no genoma de cana-de-açúcar, os quais poderiam codificar outras isoformas de proteínas pertencentes à família PUMP. Em arroz (Watanabe e Hirai, 2002) e *Arabidopsis thaliana* (Maia et al., 1998; Watanabe et al., 1999), a identificação de diferentes genes *PUMPs* sugere a existência de família multigênica em plantas. Desta forma, utilizando seqüências de PUMPs de *Arabidopsis thaliana* e de arroz como *drivers*, além de outras espécies vegetais, identificamos cinco genes candidatos no genoma de cana-de-açúcar que codificam proteínas com características semelhantes às PUMPs já descritas (*SsPUMPs*, Tabela 2, capítulo III). As proteínas codificadas por esses genes foram classificadas, juntamente com proteínas de outras espécies vegetais, em 3 grupos (Fig.1, capítulo III), sugerindo a presença de diferentes subfamílias de PUMPs. Essa hipótese foi confirmada ao identificarmos quatro genes no genoma de *A. thaliana* (nomeados *AtPUMP3*, *AtPUMP4*, *AtPUMP5* e *AtPUMP6*), os quais codificam proteínas similares àquelas já descritas nessa espécie (*AtPUMP1* e *AtPUMP2*, Maia et al., 1998; Watanabe et al., 1999). A expressão dos genes *PUMPs* de ambas as espécies foi variável nos diferentes órgãos vegetais e também em resposta ao estresse por baixa temperatura (Figs. 3 e 4, capítulo III). Interessantemente, os genes ortólogos *PUMP4* e *PUMP5* apresentaram perfis de expressão contrastantes em resposta ao estresse por frio em ambas as espécies. Enquanto em cana-de-açúcar, os genes *SsPUMP4* e *SsPUMP5* apresentaram uma expressão gradual, aumentando com o tempo de exposição ao estresse, a expressão dos genes *AtPUMP4* e *AtPUMP5* de *A.thaliana* foi transientemente induzida por frio (Fig. 4, capítulo III). Como discutido na seção anterior, é possível que a regulação de genes *cors*, tais como *PUMPs*, seja diferenciada entre plantas de clima tropical/subtropical e plantas de clima temperado. Desta forma, *A. thaliana* é capaz de gerar mecanismos que diminuam rapidamente as injúrias causadas pelo estresse por frio, já que as proteínas PUMPs são implicadas em mecanismos de redução da produção de EROS durante estresses abióticos (Kowaltowski et al., 1998). Adicionalmente, experimentos utilizando *microarrays* indicaram que a expressão dos genes *AtPUMPs* é alterada por outros estresses abióticos, além de estresses bióticos (Kreps et al., 2002; Seki et al., 2002; Whitham et al., 2003).

Além das PUMPs, outra proteína mitocondrial está envolvida na resposta ao estresse por baixa temperatura: a proteína oxidase alternativa (AOx, Saisho et al., 1997). Apesar de não termos detectado a presença de *cluster* que codifica esta proteína em nossos estudos iniciais (Tabela 3, capítulo II), uma análise mais detalhada no banco de dados do SUCEST identificou quatro membros da subfamília AOx1 em cana-de-açúcar (*SsAOx1s*, Tabela 2, Fig. 1, capítulo III). Entretanto, o perfil de expressão dos genes *AOxs* em órgãos vegetais, tanto em cana-de-

açúcar quanto em *Arabidopsis thaliana*, é discrepante daquele observado para os genes *PUMPs* de ambas as espécies analisadas (Fig. 3, capítulo III).

Vários trabalhos têm analisado a regulação e produção dessas duas proteínas mitocondriais em diferentes órgãos vegetais. Sluse et al. (1998) observaram que o ácido linoleico (principal ativador da proteína PUMP) inibiu a respiração resistente a cianeto, a qual é realizada pela proteína AOx. Esses resultados sugerem que o aumento da concentração de ácidos graxos livres na mitocôndria pode afetar diferentemente ambos os sistemas de dissipação de energia, indicando que esses sistemas não operam juntos, mas seqüencialmente (Sluse e Jarmuszkiewicz,1999). De fato, durante o desenvolvimento e amadurecimento de frutos de tomate, Sluse e Jarmuszkiewicz, (1999) observaram que a atividade da proteína AOx foi maior na fase de desenvolvimento enquanto que a proteína PUMP apresentou maior atividade durante a fase de amadurecimento dos frutos, concomitante com o aumento da concentração de ácidos graxos livres no tecido vegetal.

Apesar de vários trabalhos focarem a análise bioquímica das proteínas PUMP e AOx, poucos têm avaliado a regulação gênica em diferentes tecidos/órgão vegetais e durante condições fisiológicas adversas. Neste contexto, os dados apresentados no capítulo III reforçam a hipótese de que esses dois mecanismos de dissipação de energia em plantas atuam em tecidos/órgãos distintos. Entretanto, em relação à resposta ao estresse por frio, a importância de um ou outro sistema de dissipação de energia não parece tão óbvia. Durante a exposição de plantas aos estresses bióticos ou abióticos, pode ocorrer o desequilíbrio entre o suprimento de substratos na forma reduzida (por exemplo, NADH) e a demanda de carbono e energia utilizados durante a biossíntese de compostos orgânicos, ambos acoplados pela cadeia transporte de elétrons. Tanto PUMP quanto AOx atuam corrigindo esse desequilíbrio, reduzindo a produção de EROS (Jarmuszkiewicz, 2001).

O capítulo IV apresentado no corpo da tese, *SsNAC23, a sugarcane member of the NAC domain protein family, is associated with cold, herbivory and osmotic stress*, compila os resultados de um trabalho desenvolvido a partir dos dados apresentados no Artigo1, no qual se observou que o segundo gene mais expresso em resposta ao estresse por frio em cana-de-açúcar (Tabela 2, capítulo II) codifica proteína similar àquelas pertencentes à família NAC (nomenclatura baseada nas primeiras proteínas identificadas NAM, ATAF1-2 e CUC2, Aida et al., 1997). A principal característica dessas proteínas é a presença de região amino-terminal altamente conservada de cerca de 150-160 aminoácidos, denominada de domínio NAC (Kikuchi et al., 2000). Além do gene identificado no capítulo II (denominado no capítulo IV como *SsNAC23*), outros 25 membros dessa família de proteínas foram identificados no genoma de

cana-de-açúcar (Tabela 1, capítulo IV), confirmando a hipótese da presença de um grande número de genes da família NAC em diferentes espécies vegetais (Ooka et al., 2003).

Tal como observado para oito membros dessa família em arroz (Kikuchi et al., 2000), a maioria dos genes *NAC* de cana-de-açúcar (*SsNACs*) foi expressa em diversos órgãos da planta (Fig. 1, capítulo IV), sugerindo que as proteínas SsNACs podem ter múltiplos ou mesmo redundantes papéis fisiológicos. Interessantemente, as proteínas que contém o domínio NAC são específicas de espécies vegetais, não sendo encontradas até o momento em nenhum outro organismo (Kikuchi et al., 2000; Ooka et al., 2003). Portanto, os genes *NACs* parecem ter se diversificado em associação com a evolução da arquitetura das espécies vegetais, sendo então importantes para seu correto desenvolvimento. Trabalhos futuros poderão elucidar o processo de diversificação e surgimento desses genes durante a evolução das plantas. Entretanto, como apresentado no capítulo IV e em outros trabalhos (Collinge and Boller, 2001; Hegedus et al., 2003), alguns membros da família de proteínas NAC também estão envolvidos na resposta aos estresses bióticos e abióticos. Análises filogenéticas e de perfis de expressão gênica demonstraram que genes *NACs* parálogos e ortólogos parecem ter funções conservadas em resposta aos diferentes estresses ambientais (Fig. 2, capítulo IV).

Trabalhos anteriores têm confirmado que proteínas NAC de *A. thaliana* (ANAC, AtNAM, NAC1; Xie et al., 2000; Duval et al., 2002; Earnst et al., 2004) possuem algumas características de fatores de transcrição, das quais a principal é a presença de uma estrutura protéica terciária ou quaternária capaz de interagir com regiões específicas da molécula de DNA (elementos em *cis*) e promover a transcrição gênica. Devido à alta porcentagem de identidade ao nível de aminoácidos com a proteína ANAC (>60%) e a disponibilidade de dados de sua estrutura em bancos de dados, foram gerados modelos do domínio NAC da proteína SsNAC23, como descrito no capítulo IV. Após o modelo com melhores parâmetros ter sido escolhido (Material e Métodos, capítulo IV), analisamos as regiões que supostamente poderiam interagir com moléculas de DNA. Interessantemente, todos os aminoácidos importantes para o contato (Earnst et al., 2004) são altamente conservados em todas as espécies vegetais analisadas (Fig. 4A, capítulo IV). Após a estrutura modelada do domínio NAC da proteína SsNAC3 ser sobreposta à estrutura do mesmo domínio da proteína ANAC (Earnst et al., 2004), detectamos que ambas são extremamente semelhantes (Fig. 4B, capítulo IV), sugerindo conservação de seqüência primária e de estrutura entre proteínas dessa família. Além disso, SsNAC23 localizou-se no núcleo de células epidérmicas de cebola (*Allium cepa*, Fig. 5, capítulo IV), reforçando a hipótese de que essa proteína possui características similares aos fatores de transcrição já descritos.

O perfil de expressão diferencial em resposta a duas faixas de temperatura testadas (4 °C e 12 °C), observado tanto para o gene *SsNAC23* quanto para o gene *SsMAPK* (Fig. 6 e 7, capítulo IV), indica que cana-de-açúcar possui distintos mecanismos de sinalização em resposta ao estresse por frio. Resultados similares foram observados para arroz (Wen et al., 2002) e *A. thaliana* (Provart et al., 2003). Um detalhamento maior sobre o possível papel dessa proteína nos mecanismos de sinalização celular, bem como a identificação de seus "genes-alvos", poderá auxiliar no entendimento da regulação gênica durante a resposta tanto a estresses abióticos, como a estresses bióticos.

Em resumo, os trabalhos apresentados como parte do corpo desta tese de doutoramento abrem novas possibilidades para a elucidação dos mecanismos envolvidos na resposta ao estresse por baixas temperaturas, não somente de cana-de-açúcar, mas também de outras espécies vegetais.

CONCLUSÕES

Com base nos resultados apresentados nesta tese de doutoramento, podemos concluir que:

I. Macroarranjos de DNA em membranas de *nylon* podem ser utilizados eficientemente e com boa reprodutibilidade para a análise do perfil de expressão de milhares de genes simultaneamente;

II. As estratégias de normalização testadas permitiram a redução da variabilidade nos dados gerados nos *macroarrays*, aumentando assim a confiabilidade da técnica;

III. O uso de sonda *overgo* permitiu a estimativa das flutuações na massa de DNA depositada nas membranas;

IV. Os resultados de *macroarray* e *data mining* sugerem que o genoma de cana-de-açúcar codifica vários genes homólogos àqueles induzidos ou reprimidos por frio em outras espécies vegetais, incluindo aquelas de clima temperado;

V. Múltiplas vias regulatórias parecem estar envolvidas na expressão gênica de *PUMPs* e *AOxs*;

VI. Além das diferenças observadas na regulação da expressão de *PUMPs* e *AOxs* dentro de uma mesma espécie, os resultados sugerem a existência de vias regulatórias espécie-dependentes;

VII. Os membros de cada sistema de dissipação de energia (*PUMP* e *AOx*) são diferencialmente regulados dependendo do tecido ou órgão vegetal. Deste modo, é possível que as espécies vegetais respondam de forma mais flexível às condições de estresse biótico e/ou abiótico;

VIII. Além do envolvimento de proteínas contendo domínio NAC no desenvolvimento vegetal, estas também podem fazer parte de mecanismos de respostas a estresses, tal como observado para a proteína SsNAC23;

IX. Proteínas pertencentes aos subgrupos ATAF, OsNAC3, ANAC e NAP compartilham não somente similaridade estrutural, mas também conservada função em resposta a diferentes estresses ambientais;

X. A expressão do gene *SsNAC23* foi bastante induzida quando plantas de cana-de-açúcar foram expostas a temperatura de 4 °C, porém não por estresse por frio moderado (12 °C), sugerindo o envolvimento dessa proteína em respostas a estresses mais extremos;

XI. A localização nuclear da proteína SsNAC23, juntamente com a conservação da estrutura terciária de seu domínio NAC, indica que esta possa ser um fator de transcrição, tal como observado para a proteína ANAC de *A. thaliana* (Greve et al., 2003).

REFERÊNCIAS BIBLIOGRÁFICAS

AIDA, M.; ISHIDA, T.; FUKAKI, H.; FUJISAWA, H.; TASAKA, M. Genes involved in organ separation in *Arabidopsis thaliana*: an analysis of the *cup-shaped cotyledon* mutant. **Plant Cell** v.9, p.841-857, 1997.

ANDREWS, J.; BOUFFARD, G.G.; CHEADLE, C.; LÜ, J.; BECKER, K.G.; OLIVER, B. Gene discovery using computational and microarray analysis of transcription in the *Drosophila melanogaster* testis. **Genome Res**. v.10, p.2030-2043, 2000.

AUDIC, S.; CLAVERIE, J.M. The significance of digital gene expression profiles. **Genome Res**. v.7, p.986-995, 1997.

CALLIS, J.; VIERSTRA, R.D. Protein degradation in signaling. **Curr Opin Plant Biol.** v.3, p.381-386, 2000.

CARSON, D.L.; HUCKETT, B.I.; BOTHA, F.C. Sugarcane ESTs differentially expressed in immature and maturing internodal tissue. **Plant Sci**. v.162, p.289-300, 2002.

CARVALHO, L.C.C. Perspectivas da cultura da cana-de-açúcar para a década de noventa. In: CAMARA, G.M.S; OLIVEIRA, E.A.M. (ed.). **Produção de cana-de-açúcar**. Piracicaba: FEALQ, 1993. p.1-17.

CHEN, C-N; CHU, C-C; ZENTELLA, R.; PAN, S-M; HO, T-H. D. *AtHVA22* gene family in *Arabidopsis thaliana*: phylogenetic relationship, ABA and stress regulation, and tissue-specific expression. **Plant Mol. Biol.** v.49, p.633-644, 2002.

CHERRY, J.H (ed.). **Biochemical and cellular mechanisms of stress tolerance in plants**. Berlin: Springer-Verlag, 1994.

CHINNUSAMY, V.; OHTA, M.; KANRAR, S.; LEE, B-H; HONG, X.; AGARWAL, M.; ZHU, J-K. ICE1: a regulator of cold-induced transcriptome and freezing tolerance in *Arabidopsis thaliana*. **Genes Develop**. v. 17, p.1043-1054, 2003.

CHOI, H.I.; HONG, J.H.; HA, J.O.; KANG, J.Y.; KIM, S.Y. ABFs, a family of ABA-responsive element binding factors. **J. Biol. Chem**. v.275, p.1723-1730.

CIARDI, J.A.; DEIKMAN, J.; ORZOLEK, M.D. Increased ethylene synthesis enhances chilling tolerance in tomato. **Physiol. Plant**. v.101, p.333-340, 1997.

CUSHMAN, J. C.; BOHNERT, H. J. Genomic approaches to plant stress tolerance. Cur. Op. **Plant Biol**. v.3, p.117-124, 2000.

DANIELS, J.; ROACH, B.T. Taxonomy and evolution. In: HEINZ, D.J. (ed.). **Sugarcane improvement though breeding**. Amsterdam: Elsevier, 1987. p.7-84.

DEMURA, T.; TASHIRO, G.; HORIGUCHI, G.; KISHIMOTO, N.; KUBO, M.; MATSUOKA, N.; MINAMI, A.; NAGATA-HIWATASHI, M.; NAKAMURA, K.; OKAMURA, Y.; SASSA, N.; SUZUKI, S.; YAZAKI, J.; KIKUCHI, S.; FUKUDA, H. Visualization by comprehensive microarray analysis of gene expression programs during transdifferentiation of mesophyll cells into xylem cells. **Proc. Natl. Acad. Sci. USA** v.24, p.15794-15799, 2002.

DESPREZ, T.; AMSELEM, J.; CABOCHE, M.; HÖFTE, H. Differential gene expression in *Arabidopsis thaliana* monitored using cDNA arrays. **Plant J.** v14, p.643-652, 1998.

DHINSA, R.S.; MONROY, A.F. Low temperature signal transduction, gene expression, and cold acclimation: multiple roles of low temperature. In: CHERRY, J.H. (ed.). **Biochemical and cellular mechanisms of stress tolerance in plants**. Berlin: Springer-Verlag, 1994.

DONG, Q.; ROY, L.; FREELING, M.; WALBOT, V.; BRENDEL, V. ZmDB, an integrated database for maize genome research. **Nuclei. Aci. Res**. v. 31, p.244–247, 2003.

DU, Y-C.; NOSE, A.; WASANO, K. Thermal characteristics of C4 photosynthetic enzymes from leaves of three sugarcane species differing in cold sensitivity. **Plant Cell Physiol**. v. 40, p.298–304,1999.

EWING, R.M.; KAHLA, A.B.; POIROT, O.; LOPEZ, F.; AUDIC, S.; CLAVERIE, J.M. Large-scale statistical analyses of rice ESTs reveal correlated patterns of gene expression. **Genome Res**. v.9, p.950-959, 1999.

FERL, R.; WHEELER, R.; LEVINE, H.G.; PAUL, A-L. Plants in space. **Cur. Opi. Plant Biol**. v.5, p.258-263, 2002.

FERNANDES, J.; BRENDEL, V.; GAI, X.; LAL, S.; CHANDLER, V.L.; ELUMALAI, R.P.; GALBRAITH, D.W.; PIERSON, E.A.; WALBOT, V. Comparison of RNA expression profiles based on maize expressed sequence tag frequency analysis and micro-array hybridization. **Plant Physiol**. v.128, p.896-910, 2002.

FIELD, R.J. The role of 1-aminocyclopropane-1-carboxylic acid in the control of low temperature induced ethylene production in leaf tissue of *Phaseolus vulgaris* L. **Ann. Bot.** v.54, p.61-67, 1984.

GAUDET, D.A. Progress towards understanding interaction between cold-hardiness and snow-mold resistance and development of resistant cultivars. **Can. J. Plant Pathol**. v.16, p.1615-1621.

GILMOUR, S.J.; THOMASHOW, M.F. Cold acclimation and cold-regulated gene expression in ABA mutants of *Arabidopsis thaliana thaliana*. **Plant Mol. Biol.** v.17, p.1233-1240, 1991.

GRIVET, L.; ARRUDA, P. Sugarcane genomics: depicting the complex genome of an important tropical crop. **Cur. Opi. Plant Biol**. v.5, p.122-127, 2001.

GUIMARÃES, C.T.; SILLS, G.R.; SOBRAL, B. W. S. Comparative mapping of Andropogoneae: *Saccharum* L. (sugarcane) and its relation to sorghum and maize. **Proc. Natl. Acad. Sci. USA** v. 94, p.14261–14266, 1997.

GUY, C.L. Cold acclimation and freezing stress tolerance: role of protein metabolism. **Annu. Rev. Plant Physiol. Plant Mol. Biol**. v.41, p.187-223, 1990.

HOLMBERG, N.; BÜLOW, L. Improving stress tolerance in plants by gene transfer. **Trends Plan Sci.** v.3, p.61-66, 1998.

HONG, J.C.; CHEONG, Y.H.; NAGAO, R.T.; BAHK, J.D.; KEY, J.L.; CHO, M.J. Isolation of two soybean G-box binding factors which interact with a G-box sequence of an auxin-responsive gene. **Plant J.** v.8, p.199-211, 1995.

HONG, S.W.; JON, J.H.; KWAK, J.M.; NAM, H.G. Identification of a receptor-like kinase gene rapidly induced by abscisic acid, dehydration, high salt, and cold treatments in *Arabidopsis thaliana thaliana*. **Plant Physiol**. v. 113, p.1203-1212, 1997.

HSIEH, T-H; LEE, J-T; YANG, P-T; CHIU, L-H; CHARNG, Y-Y; WANG, Y-C; CHAN, M-TS. Heterology expression of the *Arabidopsis thaliana c-repeat/dehydration response element binding factor 1* gene confers elevated tolerance to chilling and oxidative stresses in transgenic tomato. **Plant Physiol**. v.129, p.1086–1094, 2002.

HULL, R. **Matthews' Plant Virology**. San Diego: Academic Press, 2002.

IBA, K. Acclimate response to temperature stress in higher plants: approaches of gene engineering for temperature tolerance. **Annu. Rev. Plant Biol**. v.53, p.225-245, 2002.

ICHIMURA, K.; MIZOGUCHI, T.; YOSHIDA, R.; YUASA, T.; SHINOZAKI, K. Various abiotic stresses rapidly activate *Arabidopsis thaliana* MAP kinases ATMPK4 and ATMPK6. **Plant J**., v.24, p.655-665, 2000.

INGELBRECHT, I.L.; IRVINE, J.E.; MIRKOV. Posttranscriptional gene silencing in transgenic sugarcane. Dissection of homology-dependent virus resistance in a monocot that has a complex polyploid genome. **Plant Physiol.** v.119, p.1187-1197, 1999.

JAGLO, K.R.; KLEFT, S.; AMUNDSEN, K.L.; HAAKE, V.; ZHANG, J.Z., DEITS, T.; THOMASHOW, M.F. Components of the *Arabidopsis thaliana* C-repeat/dehydration-responsive element binding factor cold-response pathway are conserved in *Brassica napus* and other plant species**. Plant Physiol.** v.127, p.910-917, 2001.

JAGLO-OTTOSEN, K.R.; GILMOUR, S.J.; ZARKA, D.G.; SCHABENBERGER, O.; THOMASHOW, M.F. *Arabidopsis thaliana CBF1* overexpression induces *cor* genes and enhances freezing tolerance. **Science** v.280, p.104-106, 1998.

JARMUSZKIEWICZ, W. Uncoupling proteins in mitochondria of plants and some microorganisms. **Acta Bioch. Pol.** v.48, p145-155, 2001.

JONAK, C.;KIERGEL, S.; LIGTERINK, W.; BARKER, P.J.; HUSKISSON, N.S.; HIRT, H. Stress signaling in plants: a mitogen-activated protein kinase pathway is activated by cold and drought. **Proc. Natl. Acad. Sci. USA** v.93, p.11274-11279, 1996.

KASUGA, M.; LIU, Q.; MIURA, S.; YAMAGUCHI-SHIMOZAKI, K.; SHIMOZAKI, K. Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. **Nat. Biotech**. v.17, p.287-291, 1999.

KAWASAKI, S.; BORCHERT, C.; DEYHOLOS, M.; WANG, H.; BRAZILLE, S.; KAWAI, K.; GALBRAITH, D.; BOHNERT, H.J. Gene expression profiles during the initial phase of salt stress in rice. **Plant Cell** v.13, p.889-905, 2001.

KEEGSTRA, K.; THOMASHOW, M. Adapting physiology and metabolism to changes in the environment. **Cur. Opi. Plant Biol**. v.5, p.191-192, 2002.

KNIGHT, H. Calcium signaling during abiotic stress in plants. **Int. Rev. Cytol**. v. 195, p.269-325, 2000.

KNUTSON, R.M. Heat production and temperature regulation in eastern skunk cabbage. **Science**. v.186, p.746-747, 1974.

KREPS, J.A.; WU, Y.; CHANG, H-S; ZHU, T.; WANG, X.; HARPER, J. F. Transcriptome changes for *Arabidopsis thaliana* in response to salt, osmotic, and cold stress. **Plant Physiol**. v.130, p.2129–2141, 2002.

KURKELA, S.; BORG-FRANCK, M. Cloning and characterization of a cold- and ABA-inducible *Arabidopsis thaliana* gene. **Plant Mol. Biol.**, v.15, p.137-144, 1990.

LAMB, C.; DIXON, R.A. The oxidative burst in plant disease resistance. **Annu. Rev. Plant Physiol. Plant Mol. Biol.** v.48, p.251-275, 1997.

LARA, F.J.S. **Hibridação de ácidos nucléicos**. Ribeirão Preto: Sociedade Brasileira de Genética, 1995.

LEE, H.; XIONG, L.; GONG, Z.; ISHITANI, M.; STEVENSON, B.; ZHU, J-K. The *Arabidopsis thaliana* HOS1 gene negatively regulates cold signal transduction and encodes a RING finger protein the displays cold-regulated nucleo-cytoplasmic partitioning. **Genes Dev**. v.15, p. 912-924, 2001.

MA, H.; SCHULZE, S.; LEE, S.; YANG, M.; MIRKOV, E.; IRVINE, J.; MOORE, P.; PATERSON, A. An EST survey of the sugarcane transcriptome. **Theor. Ap. Gen**. v.108, p.851-863, 2003.

MARTIN, M.L.; BUSCONI, L. A rice membrane-bound calcium dependent protein kinase is activated in response to low temperature. **Plant Physiol**., v.125, p.1442-1449, 2001.

MOHAPATRA, S.S.; POOLE, R.J.; DHINDSA, R.S. Abscisic acid-induced gene expression in relation to freezing tolerance in alfalfa. **Plant Physiol.**, v.87, p.468-473, 1988.

MONROY, A.F.; SARHAN, F.; DHINDSA, R.S. Cold-induced changes in freezing tolerance, protein phosphorylation, and gene expression: Evidence for a role of calcium**. Plant Physiol**. v.102, p.1227-1235, 1993.

MURATA, N.; LOS, D.A. Membrane fluidity and temperature perception. **Plant Physiol**. v.115, p.875-879, 1997.

OHTA, T. Multigene families and the evolution of complexy. **J. Mol. Evol.** v.33, p34-41, 1991.

OSTERLUND, M.; HARDTKE, C.; WEI, N.; DENG, X-W. Targeted destabilization of HY5 during light regulated development of *Arabidopsis thaliana*. **Nature**, v.405, p.462-466, 2000.

OUELLET, F.; HOUDE, M.; SARHAN, F. Purification, characterization and cDNA cloning of 200 kA protein induced by cold acclimation in wheat. **Plant Cell Physiol.** v.34, p59-65, 1993.

PLANT, A.L.; COHEN, A.; MOSES, M.S.; BRAY, E.A. Nucleotide sequence and spatial expression pattern of a drought- and abscisic acid-induced gene of tomato. **Plant Physiol**. v.97, p900-906, 1997.

PRASAD, T.K.; ANDERSON, M.D.; MARTIN, B.A., STEWART, C.R. Evidence for chilling-induced oxidative stress in maize seedlings and a regulatory role for hydrogen peroxide. **Plant Cell** v.6, p65-74, 1994.

PURVIS, A.C.; SHEWFELT, R.L.; GEGOGEINE, J.W. Superoxide production by mitochondria isolated from green bell pepper fruit. **Physiol. Plant**. v.94, p743-749, 1995.

RAIZER, A.J. Interações genótipos x ambientes e estabilidade fenotípica em cana-de-açúcar no Estado de São Paulo. Piracicaba, 1998. 103p. Dissertação (Mestrado) – Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo.

REYMOND, P.; WEBER, H.; DAMOND, M.; FARMER, E.E. Differential gene expression in response to mechanical wounding and insect feeding in *Arabidopsis thaliana*. **Plant Cell** v.12, p.707-720, 2000.

RICHTER, M.; RÜHLE, A.; WILD, A. Studies on the mechanism of photosystem II photoinhibition. II. The involvement of toxic oxygen species. **Photosyn. Res**. v.24, p.237-243, 1990.

RORAT, T.; IRZYKOWSKI, W.; GRYGOROWICZ, W.J. Identification and expression of novel cold induced genes in potato (*Solarium sogarandinum*). **Plant Sci.**, v.124, p.69-78, 1997.

RUDD, S. Expressed sequence tags: alternative or complement to whole genome sequences? **Trends Plant Sci.** v.8, p.321-328, 2003.

RUIZ-MEDRANO, R.; XOCONOSTLE-CÁZARES, B.; LUCAS, W.J. Phloem long-distance transport of *CmNACP* mRNA: implications for supracellular regulation in plants. **Development** v.126, p.4405-4419, 1999.

SACILOTO, R.F.Z. Inserção do gene *pr5k* em cana-de-açúcar visando induzir resistência ao fungo da ferrugem *Puccinia melanocephala*. Piracicaba, 2003. 74p. Dissertação (Mestrado) – Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo.

SAKAMOTO, H.; ARAKI, T.; MESHI, T.; IWABUCHI, M. Expression of a subset of the *Arabidopsis thaliana* Cys2/His2-type zinc-finger protein gene family under water stress. **Gene** v. 248, p.23-32, 2000.

SALTVEIT, M.E.; MORRIS, L.L. Overview of chilling injury of horticultural crops. In: YANG, C.Y. (ed.) **Chilling Injury of Horticultural Crops**. Boca Raton: CRC Press, 1990, p.3-15.

SANDERS, D.; BROWNLEE, C.; HARPER, J.F. Communicating with calcium. **Plant Cell**, v.11, p.691-706, 1999.

SCHENA, M.; SHALON, D.; DAVIS, R.W.; BROWN, P.O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. **Science** v.270, p.467-470, 1995.

SCHENK, P.M.; KAZAN, K.; WILSON, I.; ANDERSON, J.P.; RICHMOND, T.; SOMERVILLE, S.C.; MANNERS, J.M. Coordinated plant defense responses in *Arabidopsis thaliana* revealed by microarray analysis. **Proc. Natl. Acad. Sci. EUA** v.97, p.11655-11660, 2000.

SCHULER, G. Pieces of the puzzle: expressed sequences tags and the catalog of human genes. **J. Mol. Med**. v.75, p.694-698, 1997.

SHINOZAKI, K.; YAMAGUCHI-SHIMOZAKI, K. Molecular response to dehydration and low temperatures: differences and cross-talk between two stress signaling pathways. **Curr. Opin. Plant Biol**. v.3, p.217-p223, 2000.

SKRIVER, K.; MUNDY, J. Gene expression in response to Abscisic acid and osmotic stress. **Plant Cell**, v.2, p.503-512, 1990.

SOUER. E.; VAN HOUWELINGEN, A.; KLOOS, D.; MOL, J.; KOES, R. The *no apical meristem* gene of petunia is required for pattern formation in embryos and flowers and is expressed at meristem and primordia boundaries. **Cell** v.85, p.159-170, 1996.

STOCKINGER, E.J.; GILMOUR, S.J.; THOMASHOW, M.F. *Arabidopsis thaliana thaliana* CBF1 encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a *cis*-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. **Proc. Natl. Acad. Sci. USA** v.94, p.1035-1040, 1997.

TAI, P.Y.P.; LENTINI, R.S. Freeze damage of Florida sugarcane. In: ANDERSON, D.L., (ed.) **Sugarcane Handbook** (Ed. 1). Gainesville: Florida Cooperative Extension Service, University of Florida, 1998, p.1–3.

TAIZ, L.; ZEIGER, E. (eds.) **Plant Physiology** (2nd ed.) Sinauer Associates, Sunderland, Massachusetts, 1998.

TELLES, G.P.; BRAGA, M.D.V; DIAS, Z.; LIN, T-L; QUITZAU, J.A.A.; DA SILVA, F.R.; MEIDANIS, J. Bioinformatics of the sugarcane EST project. **Gen. Mol. Biol.** v. 24, p.9-15, 2001.

THOMASHOW, M.F. Plan cold acclimation: freezing tolerance genes and regulatory mechanisms. **Annu. Rev. Plant Physiol. Plant Mol. Biol**. v.50, p571-599, 1999.

VETTORE, A.L.; DA SILVA, F.R.; KEMPER, E.L.; ARRUDA, P. The libraries that made SUCEST. **Gen. Mol. Biol.** v. 24, p.1-7, 2001.

VINCENTZ, M.; CARA, F.A.; OKURA, V.K.; DA SILVA, F.R.; PEDROSA, G.L.; HEMERLY, A.S.; CAPELLA, A.N.; MARINS, M.; FERREIRA, P.C.; FRANCA, S.C.; GRIVET, L.; VETTORE, A.L.; KEMPER, E.L.; BURNQUIST, W.L.; TARGON, M.L.; SIQUEIRA, W.J.; KURAMAE, E.E.; MARINO, C.L.; CAMARGO, L.E.; CARRER, H.; COUTINHO, L.L.; FURLAN, L.R.; LEMOS, M.V.; NUNES, L.R.; GOMES, S.L.; SANTELLI, R.V.; GOLDMAN, M.H.; BACCI, M. JR.; GIGLIOTI, E.A.; THIEMANN, O.H.; SILVA, F.H.; VAN SLUYS, M.A.; NOBREGA, F.G.; ARRUDA, P., MENCK, C.F. Evaluation of monocot and eudicot divergence using the sugarcane transcriptome. **Plant Physiol**. v.134, p.951-959, 2004.

VISWANATHAN, C.; ZHU, J-K. Molecular genetic analysis of cold-regulated gene transcription. **Phil. Trans. R. Soc. Lond. B** v.357, p.877-886, 2002.

XIE, Q.; SANZ-BURGOS, AP; GUO, H.; GÁRCIA, J.A.; GUTIÉRREZ, C. GRAB proteins, novel members of the NAC domain family, isolated by their interaction with a geminivirus protein. **Plant Mol Biol** 39: 647-656, 1999.

XIONG, L.; ZHU, J.K. Abiotic stress signal transduction in plants: molecular and genetics perspectives. **Physiol Plant**. v.112, p.152-166, 2001.

XIONG, L.; SCHUMAKER, K.S.; ZHU, J-K. Cell signaling during cold, drought, and salt stress. **Plant Cell**, S165-S183, 2002.

XIONG, L.; ISHITANI, M.; LEE, H.; ZHU, J-K. The *Arabidopsis thaliana* LOS5/ABA3 locus encodes a molybdenum cofactor sulfurase and modulates cold stress- and osmotic stress-responsive gene expression. **Plant Cell** v.13, p.2063-2083, 2001.

YU, X-M; GRIFFITH, M.; WISEMAN, S.B. Ethylene induces antifreeze activity in winter rye leaves. **Physiol Plant**. v.126, p.1232-1240, 2001.

YUASA, T.; ICHIMURA, K.; MIZOGUCHI, T.; SHINOZAKI, K. Oxidative stress activates AtMPK6, an *Arabidopsis thaliana* homologue of MAP kinase. **Plant Cell Physiol**. v. 42, p.1012-1016, 2001.

ZAR, J.H. (eds.) **Biostatistical analysis** (4th ed.) Simon & Schuster, Upper Saddle River, New Jersey, 1999.

ZHU, B., CHEN; T.H.H., LI, P.H. Expression of an ABA-responsive osmotin-like gene during the induction of freezing tolerance in *Solarium commersoni*. **Plant Mol. Biol.**, v.21, p.729-735, 1993.

WHITE, A.J.; DUNN, A.M.; BROWN, K.; HUGHES, M.A. Comparative analysis of genomic sequence and expression of a lipid transfer protein gene family in winter barley. **J. Exp. Bot.** v.281, p.1885-1892, 1994.

WHITE, J.A.; TODD, J.; NEWMAN, T.; FOCKS, N.; GIRKE, T.; DE ILARDUYA, O. M.; JAWORSKI, J.G.; OHLROGGE, J. B.; BENNING, C. A new set of *Arabidopsis thaliana* expressed sequence tags from developing seeds. The metabolic pathway from carbohydrates to seed oil**. Plant Physiol**. v. 124, p.1582–1594, 2000.

WHITHAM SA, QUAN S, CHANG HS, COOPER B, ESTES B, ZHU T, WANG X, HOU YM. Diverse RNA viruses elicit the expression of common sets of genes in susceptible *Arabidopsis thaliana thaliana* plants. **Plant J.** v.33, p271-283, 2003.

WU, J.; MAEHARA, T.; SHIMOKAWA, T.; YAMAMOTO, S.; HARADA, C.; TAKAZAKI, Y.; ONO, N.; MUKAI, Y.; KOIKE, K.; YAZAKI, J.; FUJII, F.; SHOMURA, A.; ANDO, T.; KONO, I.; WAKI, K.; YAMAMOTO, K.; YANO, M.; MATSUMOTO, T.; SASAKI, T. A comprehensive rice transcript map containing 6591 expressed sequence tag sites. **Plant Cell** v. 14, p525–535, 2002.