

UNIVERSIDADE FEDERAL DE SÃO CARLOS
Centro de Ciências Biológicas e da Saúde
Departamento de Genética e Evolução
Laboratório de Biologia Molecular

MARIANA CARDOSO MIGUEL

**PRODUÇÃO RECOMBINANTE E CARACTERIZAÇÃO DE DUAS
CISTATINAS DE CANA-DE-AÇÚCAR**

SÃO CARLOS

2014

UNIVERSIDADE FEDERAL DE SÃO CARLOS
Centro de Ciências Biológicas e da Saúde
Departamento de Genética e Evolução
Laboratório de Biologia Molecular

MARIANA CARDOSO MIGUEL

**PRODUÇÃO RECOMBINANTE E CARACTERIZAÇÃO DE DUAS
CISTATINAS DE CANA-DE-AÇÚCAR**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular do Centro de Ciências Biológicas e da Saúde da Universidade Federal de São Carlos para obtenção do título de Mestre em Genética Evolutiva e Biologia Molecular.

Orientador: Prof. Dr. Flávio Henrique Silva

Coorientador: Dr. Luciano Takeshi Kishi

SÃO CARLOS

2014

**Ficha catalográfica elaborada pelo DePT da
Biblioteca Comunitária da UFSCar**

M636pr Miguel, Mariana Cardoso.
Produção recombinante e caracterização de duas
cistatinas de cana-de-açúcar / Mariana Cardoso Miguel. --
São Carlos : UFSCar, 2014.
68 f.

Dissertação (Mestrado) -- Universidade Federal de São
Carlos, 2014.

1. Biologia molecular. 2. Cistatina. 3. Fitocistatina. 4.
Inibidores enzimáticos. 5. Cisteíno peptidase. 6. Cana-de-
açúcar. I. Título.

CDD: 574.88 (20ª)

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA EVOLUTIVA E BIOLOGIA
MOLECULAR

**“Produção recombinante e caracterização de duas
cistatinas de cana-de-açúcar”**

Dissertação de Mestrado de

MARIANA CARDOSO MIGUEL

Banca Examinadora

Prof. Dr. Flavio Henrique Silva



.....

Prof. Dr. Daniel Scherer de Moura



.....

Prof. Dr. Otávio Henrique Thiemann



.....

SÃO CARLOS – SP
2014

*Aos meus pais,
Edson Luiz e Maristela,
à minha irmã Mariângela,
e ao meu noivo Anderson Luiz,
pelo suporte, dedicação,
e amor sem limites.*

*"Nessa estrada não nos cabe conhecer ou ver o que virá,
O fim dela ninguém sabe bem ao certo onde vai dar,
Vamos todos numa linda passarela
De uma aquarela que um dia enfim,
Descolorirá."*

(Toquinho)

AGRADECIMENTOS

Agradeço a Deus, por me agraciar com saúde, disposição e paciência para que concluísse esse trabalho.

Agradeço imensamente à minha família, pelo amor incondicional e dedicação. Aos meus pais, Edson e Maristela, por serem meus primeiros orientadores, por me ensinarem, por meio dos próprios exemplos, que devemos ser o melhor que pudermos. À minha querida irmã Mariângela, e melhor amiga desde que nasci, sempre disposta a me amparar e incentivar nos momentos de dificuldade, e comemorar as minhas conquistas. À Josefa, que me ama como uma mãe. Ao meu noivo, Anderson Luiz, que teve paciência e muito amor para me dar suporte, principalmente no final da dissertação. Amo vocês.

Ao meu orientador Prof. Dr. Flávio Henrique Silva, pela amizade e orientação. Por ter me apoiado, não apenas nos momentos difíceis durante a pesquisa, mas também em momentos difíceis em minha vida pessoal. E pelos ensinamentos que levarei para a vida.

Ao meu querido amigo e coorientador Dr. Luciano Takeshi Kishi, pelos ensinamentos, incentivo, dedicação e amizade sem limites.

À Prof^a. Dra. Andrea Soares da Costa Fuentes, sempre disposta a ajudar, e compartilhar todo o seu conhecimento.

Às minhas grandes amigas e companheiras Ana Lúgia Buzolin e Danyelle Toyama, por terem compartilhado, não apenas as contas e despesas de um apartamento, mas também todos os momentos difíceis e alegres. Vocês sempre serão parte de minha família. À minha querida "chiquita" Adelita Carolina Santiago, por estar sempre torcendo por mim e ser uma amiga maravilhosa, dentro ou fora do laboratório. Amo vocês, meninas.

Ao meu amigo urso Wesley, pela amizade, e por me receber em São Carlos com muito carinho.

Ao nosso novo colega de trabalho, Célio Dias, por estar sempre feliz e disposto a ajudar, e ter colaborado com o meu trabalho. Aos amigos Darlan e Brunão, por todo incentivo e ajuda. À querida e fofa Marina, que chegou recentemente ao grupo e é sempre muito eficiente em ajudar. Obrigada, meninos, sempre torcerei pelo sucesso de vocês.

Aos colegas e amigos de trabalho, que permanecem no LBM ou que nos deixam saudades: Aliandra, Augusto, Fernando, Fran, Beth, Bruna, Daniela, Danilo, Kesser, Leo, Lucas, Gabriel, Rafael, Rodrigo, Priscila, Vanessa e Taíse. Por serem pessoas tão queridas, pelo bom humor, e por ajudarem de uma forma ou de outra. Que sorte eu tive de fazer parte desse grupo!

Aos meus amigos de Limeira, e aos que fiz estes anos em São Carlos.

À minha cachorrinha Lisa, pelos momentos de paz.

À secretária Ivanildes, por toda a ajuda com as documentações necessárias.

Aos professores e colegas do Departamento de Genética e Evolução, pela convivência.

Ao PPGGEv, pela oportunidade.

À CAPES pela concessão da bolsa e pelo auxílio financeiro que possibilitou a execução deste trabalho.

E a todos que direta ou indiretamente ajudaram na conclusão deste trabalho.

RESUMO

Cistatinas são inibidores reversíveis de cisteíno-peptidases. As cistatinas encontradas em plantas são denominadas fitocistatinas, e constituem uma subfamília independente da superfamília das cistatinas. Alguns estudos têm relatado importantes efeitos pleiotrópicos para cistatinas recombinantes expressas em plantas transgênicas, incluindo principalmente, fenótipos com tolerância ao ataque de artrópodes herbívoros e patógenos, e contra estresses bióticos e abióticos. Ademais, a cistatina recombinante de cana-de-açúcar, CaneCPI-4, apresentou potencial para inibir o desenvolvimento de células de melanoma. Dessa maneira, o estudo e conhecimento sobre as fitocistatinas, tornam-se interessantes do ponto de vista agrícola e na saúde. O projeto genoma da cana-de-açúcar (SUCEST) possibilitou a identificação de 25 possíveis cistatinas nesta planta, que foram reunidas em 4 grupos, por meio de análises filogenéticas. Nesse trabalho propomos uma reanálise das prováveis cistatinas encontradas no banco de dados do SUCEST. Além disso, descrevemos a expressão heteróloga, purificação e caracterização de duas novas cistatinas de cana-de-açúcar, CaneCPI-5 e CaneCPI-6, as quais apresentaram diferenças na ação inibitória contra a catepsina B humana. Enquanto a proteína CaneCPI-6 não foi capaz de inibir esta enzima de forma eficiente ($K_i = 1,83 \mu\text{M}$), a proteína CaneCPI-5 apresentou um bom poder inibitório contra a mesma enzima ($K_i = 6,87 \text{ nM}$). A cistatina CaneCPI-5 foi analisada também contra a catepsina L recombinante do inseto *Sphenophorus levis* (*rSl-CathL*), e apresentou alto poder inibitório contra essa enzima ($K_i = 0,059 \text{ nM}$). Por fim, ambas as proteínas, CaneCPI-5 e CaneCPI-6, mostraram-se termoestáveis quando mantidas à 100°C durante 30 minutos.

Palavras-chave: cistatinas, fitocistatinas, inibidores, cisteíno peptidase, cana-de-açúcar

ABSTRACT

Cystatins are reversible inhibitors of cysteine peptidases. The cystatins found in plants are called phytocystatins, and represent an independent subfamily of the cystatins superfamily. Some studies have reported significant pleiotropic effects for recombinant cystatins expressed in transgenic plants, notably including tolerance phenotypes against attack of herbivorous arthropods and pathogens, and against abiotic and biotic stresses. Besides, the recombinant sugarcane cystatin, CaneCPI-4, showed potential to inhibit development of melanoma cells. Thus, the study and knowledge about phytocystatins, become interesting from agricultural and medicinal point of view. The sugarcane genome project (SUCEST) allowed the identification of about 25 putative cystatins in this plant, which were gathered in 4 groups, by phylogenetic analysis. In this study, we propose a new classification for the putative cystatins found in the SUCEST database. Furthermore, we describe the heterologous expression, purification and characterization of two novel sugarcane cystatins, CaneCPI-5 and CaneCPI-6, which showed different inhibitory activities against human cathepsin B. While protein CaneCPI-6 was not able to inhibit this enzyme efficiently ($K_i = 1,83 \mu\text{M}$), the protein CaneCPI-5 showed a good inhibitory capacity against the same enzyme ($K_i = 6,87 \text{ nM}$). The CaneCPI-5 cystatin was also analyzed against recombinant cathepsin L from the beetle *Sphenophorus levis* (*rSl-CathL*), and showed a good inhibitory capacity against this enzyme ($K_i = 0,059 \text{ nM}$). Finally, both of proteins, CaneCPI-5 and CaneCPI-6, proved to be thermostable when kept at 100°C for 30 minutes.

Key words: cystatins, phytocystatins, inhibitors, cysteine peptidase, sugarcane

LISTA DE FIGURAS

Figura 1: (A) Visão frontal e lateral do modelo estrutural da Oyzacistatina. (B) Interação entre cistatina e cisteíno peptidase alvo (Fonte: Modificado de BENCHABANE <i>et al.</i> , 2010 e ALVAREZ-FERNANDEZ e ABRAHAMSON, 2006, respectivamente).....	17
Figura 2: Possível esquema evolutivo para cistatinas de plantas. (Fonte: Modificado de BENCHABANE <i>et al.</i> , 2010)	20
Figura 3: Representação esquemática do plasmídeo pTZ57R/T utilizado como vetor de propagação	26
Figura 4: A: Representação esquemática do plasmídeo pET-28a utilizado como vetor de expressão. B: Esquema da construção dos plasmídeos pET-28a contendo os genes de CaneCPI-5 e CaneCPI-6	28
Figura 5: Gráfico extraído do programa Blast2GO, representando as espécies que apresentaram maior similaridade com cistatinas de açúcar cana-de-açúcar.	33
Figura 6: Sequência de nucleotídeos traduzidos do cDNA que codifica CaneCPI-5	35
Figura 7: Sequência de nucleotídeos traduzidos do cDNA que codifica CaneCPI-6	36
Figura 8: Perfil de hidrofobicidade das cistatinas CaneCPI-5 e CaneCPI-6	38
Figura 9: Peptídeos sinais identificados nas sequências das cistatinas CaneCPI-5 e CaneCPI-6	39
Figura 10: Alinhamento entre as sequências das proteínas CaneCPI-5 e CaneCPI-6	40
Figura 11: Alinhamento da cistatina CaneCPI-5 com cistatinas de outras espécies vegetais	41
Figura 12: Alinhamento da cistatina CaneCPI-6 com cistatinas de outras espécies vegetais	41
Figura 13: Alinhamento das cistatinas de cana-de-açúcar	43

Figura 14: Alinhamento das cistatinas de cana-de-açúcar, CaneCPI-4 e CaneCPI-5	43
Figura 15: Árvore filogenética correspondentes às prováveis cistatinas de cana-de-açúcar	44
Figura 16: Alinhamento de cistatinas de cana-de-açúcar pertencentes a diferentes grupos.....	46
Figura 17: Amplificação das ORFs das cistatinas da cana-de-açúcar	48
Figura 18: PCR de colônia de <i>E. coli</i> DH5-α transformadas com o vetor pTZ57R/T com as ORFs de CaneCPI-5 e CaneCPI-6 clonadas.....	49
Figura 19: Reações de digestão dos vetores pTZ57R/T, contendo as ORFs de CaneCPI-5 e CaneCPI-6.	50
Figura 20: PCR de colônia de <i>E. coli</i> DH5-α transformadas com o vetor pET-28a com ORFs de CaneCPI-5 e CaneCPI-6 subclonadas.....	50
Figura 21: Expressão heteróloga e coleta periódica a cada 1 hora após a indução das cistatinas CaneCPI-5 e CaneCPI-6.	51
Figura 22: Expressão heteróloga e purificação das cistatinas CaneCPI-5 e CaneCPI-6.	52
Figura 23: Inibição da catepsina B humana pelas cistatinas recombinantes... 	54
Figura 24: Inibição da catepsina L de <i>Sphenophorus levis</i> (<i>rSl-CathL</i>) pela cistatina recombinante CaneCPI-5	57
Figura 25: Estabilidade térmica das cistatinas recombinantes de cana-de-açúcar.....	60

LISTA DE TABELAS

Tabela 1: Sequências dos <i>primers</i> utilizados para amplificar a região gênica das ORFs de CaneCPI-5 e CaneCPI-6.....	25
Tabela 2: Descrição das bibliotecas que compõem o SUCEST. (Fonte: Modificado de VETTORE <i>et al.</i>, 2001).	32
Tabela 3: Resumo das informações de cada clone de cDNA selecionados para estudo	34
Tabela 4: Tamanho das ORFs e massas moleculares teóricas das cistatinas.....	36
Tabela 5: Identidade de sequências de aminoácidos entre as cistatinas de cana-de-açúcar e cistatinas de outras espécies vegetais	40
Tabela 6: Rendimento das cistatinas de cana-de-açúcar	53
Tabela 7: Constantes de Inibição (K_i) das cistatinas CaneCPI-5 e CaneCPI-6 contra a catepsina B humana (1,8 nM).	55
Tabela 8: Comparação entre as constantes de inibição (K_i) obtidas neste estudo com outras cistatinas de cana-de-açúcar contra catepsina B humana.	55
Tabela 9: Comparação entre as constantes de inibição (K_i), obtidas neste estudo, com outras cistatinas de cana-de-açúcar, contra a enzima <i>rSL-CathL</i> (11 nM).	58

LISTA DE ABREVIATURAS E SIGLAS

Arg (A)	arginina
BLAST	<i>Basic Local Alignment Search Tool</i>
BSA	<i>bovine serum albumin</i>
CathB	humana catepsina B humana
cDNA	DNA complementar
cm	centímetros
CPI	<i>cysteine protease inhibitor</i>
Cys (C)	cisteína
DNA	ácido desoxirribonucléico
dNTP	desoxirribonucleotídeos trifosfatos
DTT	ditiotreitól
EST	<i>expressed sequence tags</i>
g	gramas
g	gravidade
Gln (E)	ácido glutâmico
Gly (G)	glicina
GO	<i>Gene Ontology</i>
h	hora
HCl	ácido clorídrico
His (H)	histidina
KCl	cloreto de potássio
kDa	quilodalton
KEGG	<i>Kyoto Encyclopedia of Genes and Genomes</i>
K _i app	constante de inibição aparente
K _i	constante de inibição
K _m	constante catalítica de Michaelis-Menten
L	litro
LB	Luria-Bertani
LBM	laboratório de biologia molecular
Leu (L)	leucina
M	molar
mg	miligramas
MgCl ₂	cloreto de magnésio
Min	minuto
mL	mililitros
mM	milimolar
Na ₂ HPO ₄	fosfato de sódio dibásico
NaCl	cloreto de sódio
NaH ₂ PO ₄	fosfato de sódio monobásico
NCBI	<i>National Center for Biotechnology Information</i>
Ng	nanogramas
Nm	nanômetros
NR	<i>non-redundant protein sequences</i>
ORF	fase aberta de leitura (<i>open reading frame</i>)
Pb	pares de bases
PCR	reação em cadeia da polimerase (<i>polymerase chain reaction</i>)
pH	potencial de hidrogênio
Phe (F)	fenil alanina

RNA	ácido ribonucléico
RPM	rotações por minuto
rSI-CathL	catepsina L recombinante do inseto <i>S. levis</i>
s	segundos
SDS-PAGE	eletroforese em gel de poliacrilamida contendo dodecil sulfato de sódio
SUCEST	<i>Sugarcane Expressed Tags Project</i>
Trp (W)	triptofano
U	unidade enzimática
UV	ultravioleta
Val (V)	valina
°C	graus Celsius
µg	microgramas
µL	microlitros
µM	micromolar

SUMÁRIO

1. INTRODUÇÃO	15
1.1. CISTEÍNO-PEPTIDASES.....	15
1.2. CISTATINAS E FITOCISTATINAS.....	16
1.3. CISTATINAS DE CANA-DE-AÇÚCAR.....	20
1.4. PROJETO GENOMA DA CANA-DE-AÇÚCAR (SUCEST).....	21
2. OBJETIVOS	23
3. METODOLOGIA	24
3.1. ANÁLISES IN SILICO DAS SEQUÊNCIAS DAS CISTATINAS DE CANA-DE-AÇÚCAR...	24
3.2. ANÁLISE FILOGENÉTICA.....	25
3.3. ISOLAMENTO DA REGIÃO CODANTE DAS ORFS QUE CODIFICAM AS DUAS NOVAS CISTATINAS DE CANA-DE-AÇÚCAR.....	25
3.4. SUBCLONAGEM E CONSTRUÇÃO DOS VETORES DE EXPRESSÃO.....	26
3.5. EXPRESSÃO RECOMBINANTE DAS CISTATINAS.....	29
3.6. PURIFICAÇÃO DAS CISTATINAS RECOMBINANTES.....	30
3.7. ENSAIOS DE INIBIÇÃO DAS CISTATINAS CONTRA CISTEÍNO-PEPTIDASES.....	31
4. RESULTADOS E DISCUSSÃO	32
4.1. ANÁLISES IN SILICO DAS SEQUÊNCIAS DAS CISTATINAS DE CANA-DE-AÇÚCAR...	33
4.2. ANÁLISE FILOGENÉTICA.....	44
4.3. AMPLIFICAÇÃO DAS REGIÕES CODANTES DAS ORFS QUE CODIFICAM AS CISTATINAS CANECPI-5 E CANECPI-6, E CONSTRUÇÃO DOS VETORES DE EXPRESSÃO.....	48
4.4. EXPRESSÃO HETERÓLOGA E PURIFICAÇÃO DAS CISTATINAS RECOMBINANTES	51
4.5. INIBIÇÃO DAS ATIVIDADES DAS CISTEÍNO-PEPTIDASES CATEPSINA B HUMANA E RSL-CATHL PELAS NOVAS CISTATINAS RECOMBINANTES.....	53
4.6. ANÁLISES DE TERMOESTABILIDADE DAS CISTATINAS RECOMBINANTES.....	60
5. CONCLUSÕES	61
6. CONSIDERAÇÕES FINAIS E PERSPECTIVAS	62
7. REFERÊNCIAS BIBLIOGRÁFICAS	63
ANEXO I. TABELA DE IDENTIFICAÇÃO DOS CLONES DO SUCEST COM OS RESPECTIVOS CÓDIGOS USADOS NA CONSTRUÇÃO DA ÁRVORE FILOGENÉTICA	68

1. INTRODUÇÃO

1.1. CISTEÍNO-PEPTIDASES

Enzimas proteolíticas, também conhecidas como peptidases, são proteínas que catalisam a hidrólise das ligações peptídicas (RAWLINGS; MORTON; BARRETT, 2007). Estas enzimas estão envolvidas em várias funções biológicas, participando de numerosos processos bioquímicos e fisiológicos, tais como, digestão de alimentos, coagulação sanguínea e fibrinólise, ativação de pró-enzimas, liberação de peptídeos fisiologicamente ativos, ativação do sistema complemento, processos de inflamação, entre outros (LEUNG-TOUNG *et al.*, 2002). As peptidases são produzidas por inúmeros organismos, incluindo animais, bactérias, leveduras e plantas (BUGG, 2004). Podem ser divididas em seis classes principais, de acordo com os resíduos de aminoácidos catalíticos presentes no sítio ativo: as serino-peptidases, que possuem no centro ativo um resíduo de serina envolvido no processo catalítico, as treonino-peptidases, que possuem como resíduo catalítico a treonina, as aspartil-peptidases, cuja atividade catalítica depende de um ou dois resíduos de ácido aspártico, as metalo-peptidases, que utilizam um íon metálico, geralmente o Zn^{2+} , em seu mecanismo de ação catalítica, as glutamil-peptidases utilizam a díade catalítica glutamato/glutamina e as cisteíno-peptidases possuem um resíduo de cisteína em seu sítio ativo (RAWLINGS *et al.*, 2014). As peptidases podem ser classificadas também de acordo com o mecanismo de ação, da similaridade de sequências de aminoácidos e das estruturas tridimensionais das proteínas, categoria de inibidores, e pH ótimo de atividade (BARRETT *et al.*, 2004).

As cisteíno-peptidases compreendem um complexo grupo de enzimas de variadas propriedades físicas e bioquímicas. A atividade proteolítica de todas as cisteíno-peptidases depende de uma díade catalítica de cisteína e histidina. A ordem dos resíduos de cisteína e histidina (Cys/His ou His/Cys) na sequência linear diferem entre as famílias de cisteíno-peptidases, sugerindo que estas proteínas tiveram origens evolucionárias distintas (RAWLINGS e BARRETT, 1994). Algumas dessas peptidases podem apresentar ainda, um terceiro resíduo, de asparagina, para orientação do resíduo de histidina (BARRETT *et al.*, 2004).

As cisteíno-peptidases estão distribuídas em 105 famílias, e 10 clãs, de acordo com sua origem evolutiva (RAWLINGS *et al.*, 2014). Dentre as cisteíno-

peptidases, o membro melhor caracterizado é a papaína, pertencente à família C1, denominada família das papaínas (BUGG, 2004).

As cisteíno-peptidases já foram identificadas em vírus, bactérias, protozoários, plantas e mamíferos, e estão envolvidas em diversas funções em todos os organismos (OTTO e SCHIRMEISTER, 1997). Em organismos parasitas de animais e plantas, como bactérias, fungos, protozoários e vermes, as cisteíno-peptidases estão envolvidas em processos de patogenicidade, tais como, invasão celular e tecidual, imuno-evasão, por meio da degradação de proteínas envolvidas na defesa do hospedeiro, e obtenção de aminoácidos essenciais, para seu crescimento e desenvolvimento (VALUEVA *et al.*, 2004; SAJID e MCKERROW, 2002).

Cisteíno-peptidases de plantas são encontradas nos vacúolos, e atuam na mobilização de proteínas estocadas no endosperma durante a germinação das sementes (ARAI *et al.*, 2002); são também encontradas no meio extracelular, tais como as de papaia e figo (RAWLINGS e BARRETT, 1994).

Em invertebrados, como insetos e nematóides, as cisteíno-peptidases estão entre as enzimas digestivas (PERNAS *et al.*, 1998; PERNAS *et al.*, 2000). Em vertebrados, as cisteíno-peptidases estão envolvidas no processo de degradação de proteínas lisossomais, além de participarem de processos extracelulares como em algumas desordens metabólicas (RAWLINGS e BARRETT, 1994).

As cisteíno-peptidases mais importantes em mamíferos são as calpaínas citosólicas (cisteíno-peptidases ativadas por íons de cálcio), da família C2 (família da calpaína), e as catepsinas lisossomais (catepsinas B, C, H, K, L, M, N, S, T, V e W), da família C1 (família da papaína). As catepsinas lisossomais, pertencentes à família da papaína, exercem importantes funções em muitos processos fisiológicos, tais como, processamento proteolítico de pró-hormônios e pró-enzimas, degradação de proteínas e apresentação de antígenos, entre outros. Tais enzimas estão também envolvidas em várias condições patológicas como processos inflamatórios, doença de Alzheimer, osteoporose, invasão tumoral e metástase (OTTO e SCHIRMEISTER, 1997; ABRAHAMSON, 1993; VASILJEVA *et al.*, 2007).

1.2. CISTATINAS E FITOCISTATINAS

Cistatinas são inibidores reversíveis de cisteíno-peptidases e seu mecanismo de ação é baseado na inibição competitiva, por meio do bloqueio da atividade proteolítica (ABRAHAMSON, 1993). A maioria das cistatinas possui em sua composição uma folha β dobrada, contendo cinco fitas β antiparalelas, organizada em volta de uma α hélice de cinco voltas. A atividade das cistatinas é explicada devido à presença de três pontos de interação envolvidos no reconhecimento da cisteíno peptidase alvo, sendo o primeiro ponto um seguimento N-terminal contendo um resíduo de glicina (ABRAHAMSON *et al.*, 1987). Os outros pontos consistem em um motivo Gln-X-Val-X-Gly, altamente conservado na região central formando um primeiro *loop* (BODE *et al.*, 1988), e uma região contendo um resíduo de triptofano em um segundo *loop*, próximo à extremidade C-terminal (BARRETT, 1987) (Figura 1A). Esses três segmentos juntos, formam uma estrutura em forma de cunha, com boa complementaridade pelo sítio ativo das cisteíno-peptidases do tipo papaína. Ambos os *loops*, das regiões central e C-terminal da cistatina, apresentam tamanhos e formas que se encaixam nas regiões correspondentes aos subsítios S1' e S2' da fenda do sítio ativo da enzima alvo. O segmento N-terminal contendo o resíduo de glicina, interage com o subsítio S2 do sítio ativo da enzima, com uma conformação inapropriada para a proteólise enzimática (BODE *et al.*, 1988; TURK e BODE, 1991; ABRAHAMSON, 1993) (Figura 1B).

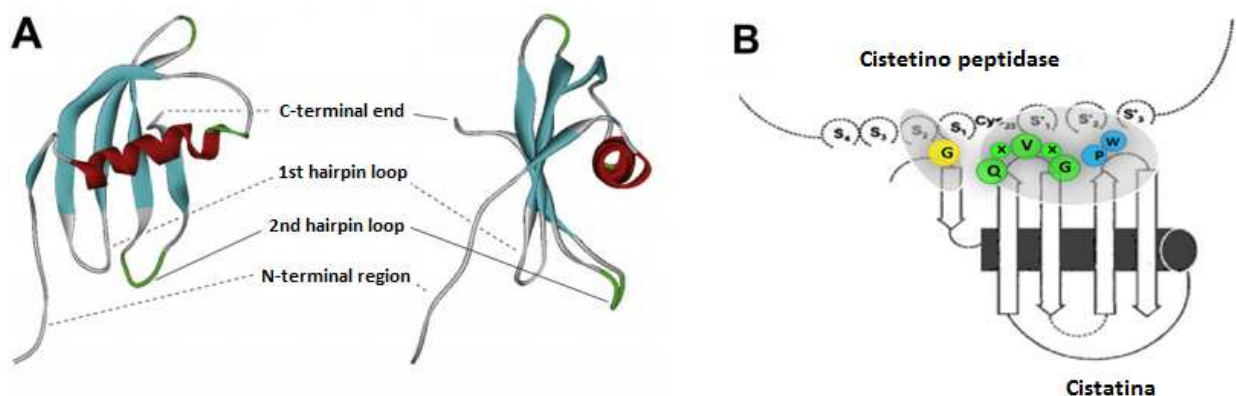


Figura 1: (A) Visão frontal e lateral do modelo estrutural da Oyzacistatina, mostrando em azul as 5 fitas β antiparalelas, e a α -hélice de 5 voltas em vermelho. **(B) Interação entre cistatina e cisteíno peptidase alvo**, em destaque as três regiões da cistatina envolvidas na inibição: em amarelo o segmento N-terminal, contendo o resíduo de glicina (G) conservado, que interage com a peptidase no subsítio S2; em verde a alça central, contendo o motivo Q-X-V-X-G, e em azul a região C-terminal, contendo o resíduo conservado de triptofano (W), que constituem as outras regiões de contato com a peptidase em S1' e S2', respectivamente. Fonte: Modificado de BENCHABANE *et al.*, 2010 e ALVAREZ-FERNANDEZ e ABRAHAMSON, 2006, respectivamente.

As cistatinas de animais foram primeiramente identificadas e caracterizadas. Essas cistatinas foram categorizadas em três famílias, de acordo com a homologia em suas sequências, número de ligações dissulfeto, massas moleculares e localização subcelular (BARRETT, 1987; TURK e BODE, 1991).

As cistatinas do tipo 1, ou estefinas, apresentam massa molecular de aproximadamente 11 kDa, possuem apenas um domínio inibitório e ausência de ligações dissulfeto, não sofrem glicosilação. Cistatinas do tipo 2, ou cistatinas, assim como as estefinas, apresentam um único domínio inibitório, massa molecular de aproximadamente 15 kDa, possuem quatro resíduos de cisteína conservados, portanto fazem duas ligações dissulfeto. Cistatinas do tipo 3, ou cininogênios, possuem múltiplos domínios inibitórios, que possivelmente resultaram de eventos de duplicação, apresentam massas moleculares mais altas, 60 a 120 kDa, são moléculas glicosiladas e também secretadas (BARRETT, 1987; TURK e BODE, 1991; BENCHABANE *et al.*, 2010).

A inclusão de uma quarta família, a das fitocistatinas, as quais pertencem os inibidores de cisteíno-peptidases de plantas, foi primeiramente sugerida por Kondo e colaboradores (1991). Fitocistatinas são inibidores competitivos de cisteíno-peptidases da família C1A, tipo papaína (TURK e BODE, 1991), desempenham importantes papéis nas plantas, desde o controle de diversos processos fisiológicos e celulares, à inibição de cisteíno-peptidases exógenas secretadas por artrópodes herbívoros e patógenos, que digerem ou colonizam os tecidos das plantas (BENCHABANE *et al.*, 2010). Além disso, alguns estudos sugerem que essas proteínas estejam envolvidas na resposta a estresses ambientais bióticos e abióticos (PERNAS *et al.*, 2000; ZHANG; LIU e TAKANO, 2008; BENCHABANE *et al.*, 2010). As fitocistatinas apresentam características que são comuns às famílias 1 e 2 de cistatinas animais, já que o tamanho e a sequência primária das proteínas se assemelham aos da família cistatina, enquanto que se assemelham às estefinas na ausência de pontes dissulfeto e resíduos de cisteína. Ademais, essas proteínas possuem uma sequência consenso L-A-R-F-A-V-X(3)-N, exclusiva de cistatinas de plantas, formando uma α -hélice na região N-terminal. A análise filogenética de cistatinas animais e de plantas mostrou o agrupamento das fitocistatinas em um único ramo, distinto das famílias 1 e 2, justificando sua classificação numa família própria. Outras características comuns atribuídas às fitocistatinas são: ausência de

sítios putativos de glicosilação, e conservação da distância de 17 aminoácidos entre os domínios LARFAV e QxVxG (MARGIS; REIS; VILLERET, 1998).

De um ponto de vista evolutivo, o enovelamento específico de fitocistatinas, a ocorrência de cistatinas com resíduos de cisteína e ausência do motivo L-A-R-F-A-V-X(3)-N em algumas algas, e a identificação de membros da superfamília cistatina em bactérias, suportam fortemente a hipótese de um caminho evolutivo paralelo, e independente para cistatinas e fitocistatinas, envolvendo um ancestral comum que aparece bem antes da separação evolutiva entre cistatinas de plantas e animais (BENCHABANE *et al.*, 2010). As cistatinas de plantas, procedentes desse ancestral comum, teriam então sido submetidas a uma série de mudanças dirigidas por processos evolutivos complementares, incluindo duplicação gênica, *splicing* alternativo e evolução adaptativa, para gerar as diferentes cistatinas codificadas pelo genoma das plantas (CHRISTELLER *et al.*, 2005; KIGGUNDU *et al.*, 2006; MARGIS-PINHEIRO *et al.*, 2008; MARTINEZ e DIAZ, 2008).

Semelhante as cistatinas animais, as cistatinas de plantas podem ser categorizadas em grupos, baseados nas características estruturais ou funcionais, forjadas ao longo do tempo, através de mutações dirigidas por divergência funcional e duplicação gênica de cistatinas de domínio único (Figura 2) (BENCHABANE *et al.*, 2010). Sabe-se, por exemplo, que as "multicistatinas" de batata, tomate, girassol e feijão de corda, que possuem ao longo da cadeia polipeptídica vários domínios inibitórios de cistatina, foram resultantes de eventos de duplicação gênica, seguidos da evolução adaptativa dos domínios inibitórios para gerar hipervariabilidade intramolecular e diversidade funcional contra peptidases do tipo papaína (KIGGUNDU *et al.*, 2006; GOULET *et al.*, 2008). Sabe-se também, que cistatinas com a região C-terminal estendida e peso molecular de aproximadamente 23 kDa, encontradas em algas verdes, e em táxons de plantas de grau superior, são resultantes de eventos de duplicação gênica precoce (MARTINEZ *et al.*, 2005a; MARGIS-PINHEIRO *et al.*, 2008). O domínio C-terminal dessas proteínas teria então sido submetido a uma significativa remodelagem funcional e estrutural, levando gradualmente à perda de atividade contra peptidases do tipo papaína, e adquirindo novas funções, incluindo atividade inibitória contra cisteíno-peptidases do tipo legumaína (Família C13) (MARTINEZ *et al.*, 2007).

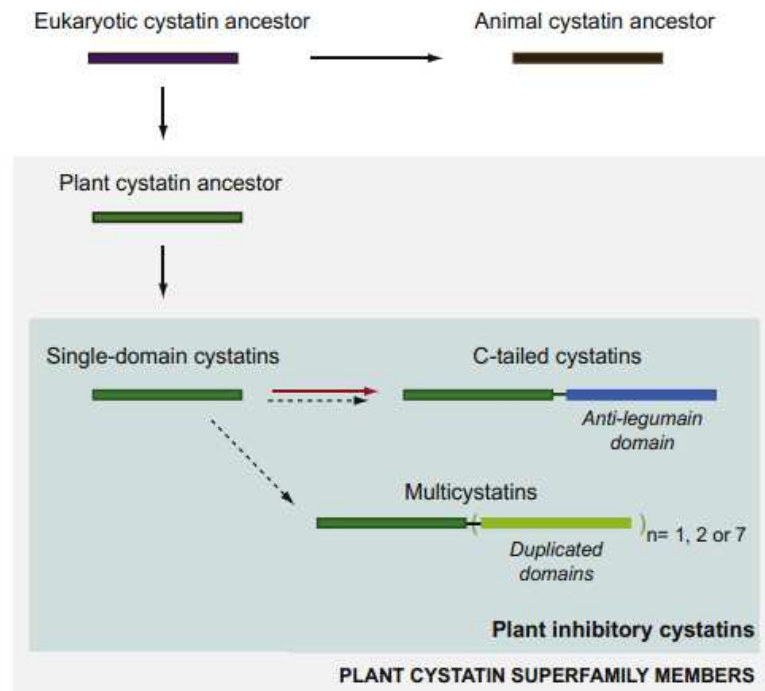


Figura 2: Possível esquema evolutivo para cistatinas de plantas. Setas com traços contínuos indicam mutações baseadas em processos evolutivos, incluindo evolução adaptativa. Setas tracejadas indicam eventos de duplicação gênica levando às cistatinas com multidomínios. A seta vermelha destaca diversificação funcional simultânea à perda de atividade inibitória contra cisteínopeptidases do tipo papaína. Fonte: Modificado de BENCHABANE *et al.*, 2010.

1.3. CISTATINAS DE CANA-DE-AÇÚCAR

As cistatinas de cana-de-açúcar foram primeiramente descritas por Reis e Margis (2001), que identificaram no SUCEST (*Sugarcane Expressed Tags Project*) banco de dados de ESTs da cana-de-açúcar, vinte e cinco prováveis cistatinas, as quais foram classificadas em quatro grupos por meio de análises filogenéticas: O grupo I compreende as cistatinas que apresentam organização estrutural típica das fitocistatinas, ou seja, além das três regiões conservadas da superfamília cistatina, possuem uma sequência consenso na região N-terminal característica de fitocistatinas. O grupo II inclui cistatinas com as mesmas características do grupo I, porém com extensão na região C-terminal. Os membros do grupo III apresentam as regiões conservadas da superfamília cistatina, entretanto não possuem a sequência consenso que caracterizam as fitocistatinas. As cistatinas do grupo IV não possuem o consenso N-terminal LARFAV-X(3)-N, nem o motivo QxVxG conservado entre os membros da superfamília (REIS e MARGIS, 2001).

A primeira cistatina de cana-de-açúcar produzida de forma heteróloga foi denominada Canacistatina, CaneCPI-1, composta por 126 aminoácidos e com massa molecular de aproximadamente 13 kDa (SOARES-COSTA *et al.*, 2002). Esta proteína apresentou atividade inibitória contra cisteíno-peptidases, inclusive catepsinas humanas (OLIVA *et al.*, 2004), e atividade contra o fungo filamentoso *Trichoderma reesei* (SOARES-COSTA *et al.*, 2002). Além da Canacistatina, outras três cistatinas foram isoladas de cana-de-açúcar, produzidas de forma heteróloga e purificadas, as quais foram denominadas CaneCPI-2, CaneCPI-3 e CaneCPI-4. Em ensaios de inibição, estas proteínas mostraram atividade inibitória sobre papaína (GIANOTTI *et al.*, 2005; GIANOTTI *et al.*, 2008); sendo que a proteína CaneCPI-4 apresentou também expressiva atividade inibitória contra as catepsinas humanas B e L (GIANOTTI *et al.*, 2008). Estudos feitos com objetivo de avaliar o potencial terapêutico da CaneCPI-4 como uma droga contra o câncer, mostrou que esta proteína foi capaz de inibir, *in vivo*, o desenvolvimento de células de melanoma (OLIVEIRA *et al.*, 2011). As proteínas recombinantes, CaneCPI-1, CaneCPI-2, CaneCPI-3 e CaneCPI-4, apresentaram também atividade inibitória contra a enzima catepsina L recombinante, denominada r*Sl-CathL*, do inseto praga de cana-de-açúcar, *Sphenophorus levis*, conhecido como bicudo da cana (FONSECA *et al.*, 2012).

Ensaio de inibição realizados com CaneCPI-1, purificada de plantas transgênicas de cana-de-açúcar, que superexpressavam essa proteína, revelaram sua capacidade em inibir a atividade catalítica de cisteíno-peptidases, parcialmente purificadas do intestino médio do bicudo da cana-de-açúcar, e catepsina L humana na ordem de nanomolar (RIBEIRO *et al.*, 2008).

A cana-de-açúcar tem sido alvo de estudos devido à sua importância no cenário econômico nacional e internacional. Pesquisas são realizadas buscando a produção de variedades mais produtivas e menos suscetíveis à perdas, tanto por fatores bióticos, como abióticos. Com base nesses resultados, a busca e caracterização de novas cistatinas de cana-de-açúcar, tornam-se importantes e justificáveis para propósitos industriais, como agentes antifúngicos, na área da saúde e na agricultura. Do ponto de vista agrônomo, é possível melhorar a produção da cultura por meio de ferramentas biotecnológicas, o que auxilia no desenvolvimento de variedades melhor adaptadas ao ataque de pragas e

patógenos. Além disso, no âmbito medicinal, podem apresentar potencial como uma droga para inibir o desenvolvimento de células de melanoma.

1.4. PROJETO GENOMA DA CANA-DE-AÇÚCAR (SUCEST)

O projeto genoma da cana-de-açúcar, denominado SUCEST (*Sugarcane Expressed Tags Project*) (VETTORE *et al.*, 2001), foi proposto com o objetivo de sequenciar ESTs (*Expressed Sequence Tags*) do genoma de cana-de-açúcar, e identificar genes expressos em uma grande coleção de bibliotecas de cDNA. Com o intuito de acessar informações de genes expressos em muitos sistemas biológicos, vinte e seis bibliotecas de cDNA foram construídas a partir de diferentes tecidos e condições, obtidos de trinta diferentes cultivares comerciais de cana-de-açúcar. Provenientes dessas bibliotecas, foram sequenciados 259.325 clones a partir da região 5', sendo que 32.364 desses clones foram também sequenciados a partir da região 3', portanto, o SUCEST gerou, no total, cerca de 291.689 ESTs. Após a trimagem das sequências de baixa qualidade e remoção das regiões do vetor e RNA ribossomal, permaneceram 237.954 ESTs que foram analisados, resultando em 43.141 *clusters* (VETTORE *et al.*, 2001). Das sequências agrupadas em *clusters*, 35,6% não correspondiam às sequências existentes nos bancos de dados públicos até 2001. Uma análise global dos dados do SUCEST indicou que 14.409 *clusters*, 33% do total, contêm pelo menos um clone de cDNA com um inserto completo. A anotação desses *clusters* associou quase 50% dos genes de cana-de-açúcar identificados com proteínas relacionadas ao metabolismo, sinalização celular, transdução de sinal, bioenergia e respostas a estresses. A busca, nesses *clusters*, por domínios proteicos conservados revelou 40.821 sequências de aminoácidos com domínios do banco de dados Pfam. A remontagem das sequências consenso dos 43.141 transcritos revelou 22% de redundância na primeira montagem. Esse resultado mostra que, possivelmente, 33.620 genes foram identificados, e indicou que mais de 90% dos genes de cana-de-açúcar expressos foram etiquetados (VETTORE *et al.*, 2003). Como citado anteriormente, foram identificadas 25 prováveis cistatinas no banco de dados SUCEST, as quais foram reunidas em 4 grupos. Dessa maneira, considerando as novas informações de banco de dados e necessidade de aprofundamento nos resultados encontrados propusemos reanalisar

os dados referentes às cistatinas, assim como caracterizar dois novos clones que as codificam.

2. OBJETIVOS

O objetivo deste trabalho foi identificar e analisar, *in silico*, as sequências das prováveis cistatinas de cana-de-açúcar encontradas no banco de dados SUCEST, assim como clonar e expressar duas novas cistatinas de cana-de-açúcar em sistema de expressão heteróloga e caracterizá-las. Neste sentido, os seguintes objetivos específicos foram propostos:

- Identificar ORFs (open read frames) de prováveis cistatinas no banco de dados do SUCEST;
- Alinhamento das sequências encontradas e clusterização;
- Análise *in silico* das cistatinas de cana-de-açúcar;
- Isolamento das *ORFs* das cistatinas e clonagem em um vetor de expressão;
- Expressão heteróloga em *Escherichia coli* e purificação das proteínas;
- Ensaio de atividade inibitória das cistatinas recombinantes contra cisteíno-peptidases.

3. METODOLOGIA

3.1. ANÁLISES *IN SILICO* DAS SEQUÊNCIAS DAS CISTATINAS DE CANA-DE-AÇÚCAR

As sequências das prováveis cistatinas de cana-de-açúcar foram selecionadas no banco de dados SUCEST, realizando uma busca por similaridade utilizando o programa TBLASTN, que compara uma sequência de proteína contra um banco de dados de nucleotídeos dinamicamente traduzidos nos seis possíveis *frames* (GERTZ *et al.*, 2006). Foram utilizadas como referência as sequências de proteínas das cistatinas de cana-de-açúcar, CaneCPI-1 (SOARES-COSTA *et al.*, 2002), CaneCPI-2, CaneCPI-3 (GIANOTTI *et al.*, 2005) e CaneCPI-4 (GIANOTTI *et al.*, 2008). As sequências selecionadas no banco de dados SUCEST, foram as que apresentavam maiores similaridades com cistatinas, considerando *E-value* menor que $1e^{-3}$, o mínimo comumente aceito para sequências de aminoácidos. As sequências encontradas foram, também, analisadas no programa Blast2GO, uma ferramenta de anotação, que realiza buscas em bancos de dados biológicos como o *Gene Ontology* (GO), *Kegg* e NR (NCBI) (CONESA *et al.*, 2005). Após selecionar as sequências anotadas por esse programa, as mesmas foram alinhadas com o programa MUSCLE no programa MEGA 5.0 (TAMURA *et al.*, 2011), para posterior análise filogenética.

O critério utilizado para a seleção dos clones contendo as prováveis cistatinas, foi a presença das regiões características da superfamília: um resíduo de glicina conservado na região N-terminal, um motivo Gln-X-Val-X-Gly, e um resíduo de triptofano conservado na região C-terminal, sendo que um deles deveria possuir o motivo LARFAV-X(3)-N, típico da subfamília fitocistatina, (SCCCLR1076E04.g) e o outro não (SCVPAM1059C07.g). Os clones foram comprados no Centro Brasileiro de Estocagem de Genes – BCC Center (Jaboticabal).

A predição de possível peptídeo sinal e localização subcelular foram checadas utilizando os programas SignalP 4.1 (PETERSEN *et al.*, 2011) e WoLF PSORT (HORTON *et al.*, 2007), respectivamente. As massas moleculares teóricas, com e sem o possível peptídeo sinal foram preditas utilizando o programa ProtParam.

3.2. ANÁLISE FILOGENÉTICA

As análises evolutivas foram realizadas no programa MEGA 6.0 (TAMURA *et al.*, 2013). A árvore filogenética, para as cistatinas de cana-de-açúcar, foram obtidas por máxima verossimilhança, utilizando o alinhamento resultante do servidor PhyML 3.0 (MILNE *et al.*, 2009), pelo programa MUSCLE, com as configurações *default* para curação de G Blocos. O programa TOPALi V2.5 (MILNE *et al.*, 2004; MILNE *et al.*, 2009) foi utilizado para encontrar o modelo mais apropriado de substituição de aminoácidos e pontuação de probabilidade. O melhor modelo foi determinado como sendo o WAG (DAYHOFF; EYCK; PARK, 1972), com distribuição gama de cinco parâmetros. O suporte estatístico para os nós foi avaliado por *bootstrap* com 1000 pseudorepetições.

3.3. ISOLAMENTO DA REGIÃO CODANTE DAS ORFS QUE CODIFICAM AS DUAS NOVAS CISTATINAS DE CANA-DE-AÇÚCAR

As ORFs correspondentes às cistatinas denominadas CaneCPI-5 e CaneCPI-6 foram amplificadas por PCR (*polymerase chain reaction*) a partir dos clones de cDNA SCVPAM1059C07.g e SCCCLR1076E04.g, respectivamente, utilizando *primers* específicos (Tabela 1). Nestes *primers* foram inseridos sítios para as enzimas de restrição *Nhe*-I e *Bam*H-I, para posterior clonagem direcional em vetor de expressão.

Tabela 1: Sequências dos *primers* utilizados para amplificar a região gênica das ORFs de CaneCPI-5 e CaneCPI-6. Os sítios para as enzimas de restrição *Nhe*-I (GCTAGC) e *Bam*H-I (GGATCC) estão sublinhados.

CaneCPI-5 Forward	5' <u>AGCTAGC</u> CCTCGCCCCGTTCCCGG 3'
CaneCPI-5 Reverse	5' AGGAT <u>CCTCAGT</u> GGGCGCGGGGCG 3'
CaneCPI-6 Forward	5' <u>AGCTAGC</u> GTCTCCGGGTTCCACCTC 3'
CaneCPI-6 Reverse	5' <u>AGGATCC</u> CTAAGAGGATGATGCGAAATCA 3'

Para obtenção do DNA molde, os dois clones bacterianos, SCVPAM1059C07.g e SCCCLR1076E04.g, foram crescidos, em meio LB (USB) contendo o antibiótico ampicilina na concentração final de 100 µg/mL, o DNA plasmidial foi purificado pelo método de lise alcalina como descrito por Sambrook e

Russel, 1989. Foi realizada uma PCR, para cada clone selecionado, em um volume final de 25 μ L, foi utilizado 30 ng de molde, 200 μ M de cada dNTP (Promega), tampão de reação 1x (Tris-HCl 100 mM pH 8,8, MgCl₂ 16 mM, KCl 500 mM), 10 pmoles de cada *primer* e 1,25 U da enzima Taq DNA polimerase (SINAPSE Biotecnologia Ltda.). O protocolo de amplificação foi utilizado em um termociclador T100TM (Bio-Rad) iniciando com uma temperatura de 94°C por 3 minutos para desnaturação do DNA, seguida por 35 ciclos de amplificação de 30 segundos a 94°C, 30 segundos a 55°C, e 60 segundos a 72°C, e com uma extensão final de 10 minutos a 72°C. Os produtos resultantes das duas PCRs foram analisados em gel de agarose 1%, contendo brometo de etídeo 1 μ g/mL e visualizados em luz ultra violeta (UV).

3.4. SUBCLONAGEM E CONSTRUÇÃO DOS VETORES DE EXPRESSÃO

Os produtos da PCR descrita no item 3.3, referentes aos fragmentos correspondentes à região codante para as cistatinas, foram purificados com o kit *Wizard SV Gel and PCR Clean Up System* (Promega), e clonados no vetor de propagação pTZ57R/T (Fermentas) (Figura 3).

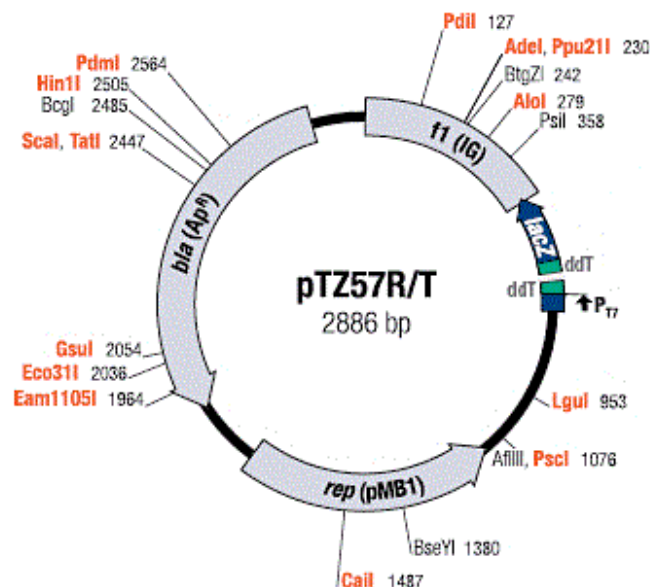


Figura 3: Representação esquemática do plasmídeo pTZ57R/T utilizado como vetor de propagação.

Os produtos da PCR purificados, foram ligados ao vetor utilizando o kit *InsTAclone PCR Cloning Kit* (Thermo Scientific), em um volume final de 30 μ L, e contendo 65 ng do vetor, 78 ng de inserto de CaneCPI-6, e 58 ng de inserto de CaneCPI-5, 5 U da enzima DNA ligase, e tampão de ligação 5x [Tris-HCl 400 mM, $MgCl_2$ 100 mM, DTT 100 mM, ATP 5 mM (pH 7.8 at 25°C)]. As duas ligações foram mantidas a 4°C por 16 h e utilizadas para transformar, pelo método de choque térmico (90 segundos no banho a 42°C, 5 minutos no gelo, adicionando 800 μ L de meio LB aos microtubos, e mantendo por 1 h no banho a 37°C), células competentes de *E. coli* DH5 α (SupE44 Δ lac U169 ϕ 80 lacZ Δ M15 hsd R17 rec A1 end A1 gyrA96 thi-1 relA1), preparadas por tratamento com cloreto de cálcio (Sambrook e Russel, 1989). Os clones bacterianos contendo os vetores recombinantes (pTZ57R/T + ORF) foram selecionados, por *screening* de coloração, em meio LB ágar (USB), contendo 100 μ g/mL do antibiótico ampicilina, 500 mM de IPTG (Isopropil- β -D-tiogalactosídeo) e 40 μ g/mL de X-Gal (5Bromo-4Cloro-3Indolyl-BDGalactoside), substrato da enzima β -galactosidase.

Uma vez selecionados os clones recombinantes, os plasmídeos foram purificados por meio de lise alcalina (Sambrook e Russel, 1989). Estes foram posteriormente digeridos com as enzimas de restrição *Nhe*-I e *Bam*H-I para recuperação dos fragmentos correspondentes às cistatinas CaneCPI-5 e CaneCPI-6 contendo extremidades coesivas, as quais permitiram a posterior subclonagem direcional no vetor de expressão pET28a (Figura 4), previamente digerido com as mesmas enzimas.

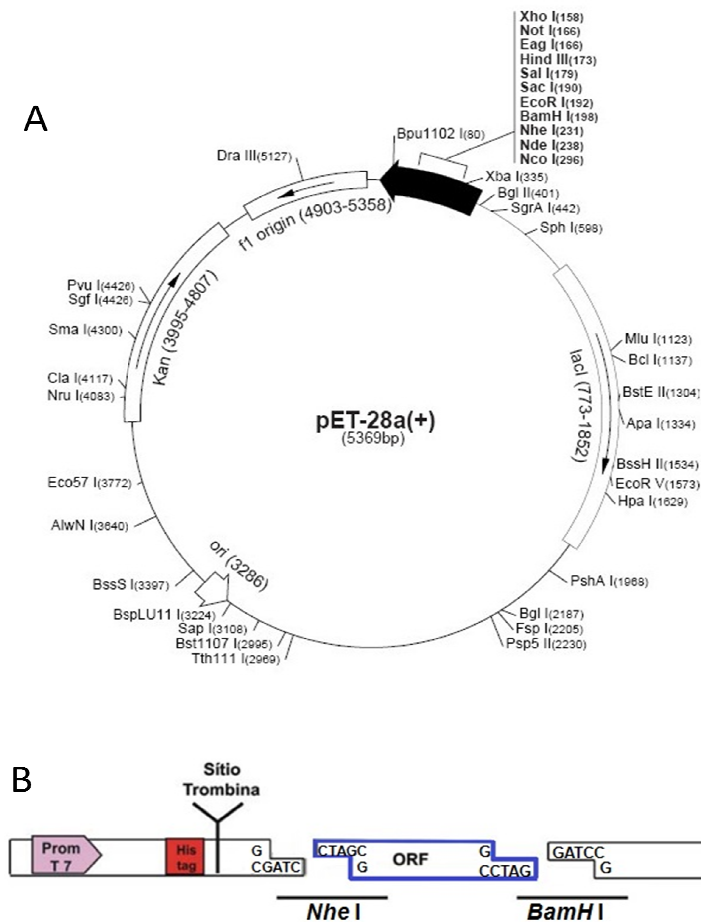


Figura 4: A. Representação esquemática do plasmídeo pET-28a utilizado como vetor de expressão. B. Esquema da construção dos plasmídeos pET-28a contendo os genes de CaneCPI-5 e CaneCPI-6.

Nas reações de digestão para liberação dos fragmentos, e do vetor pET-28a com extremidades coesivas, foram utilizados 2 µg do DNA plasmidial, 1,5 U de cada uma das enzimas *Nhe*-I e *Bam*H-I (Fermentas) e tampão Tango® [Tris-acetate 33 mM (pH 7.9 at 37°C)], acetato de magnésio 10 mM, acetato de potássio 66 mM, BSA 0,1 mg/mL], em um volume final de 50 µL, incubadas a 37°C durante 2 h, e posteriormente analisadas por meio de eletroforese em gel de agarose 1%. Verificada a digestão, os fragmentos correspondentes às cistatinas, e o vetor pET-28a foram recuperados do gel de agarose, purificados com o *kit Wizard SV Gel and PCR Clean Up System* (Promega), e utilizados nas reações de ligação. Para a ligação do produto digerido com o vetor pET-28a, foram utilizados 50 ng do vetor, 16 ng do inserto para reação com CaneCPI-5, e 22 ng do inserto para reação de CaneCPI-6, 1 U da enzima T4 DNA ligase, e tampão da enzima 1x, para um volume final de 10 µL, para as duas ligações, incubando a 4°C durante 16 h. Células de *E.*

coli DH5 α competentes foram transformadas, por choque térmico, com os produtos ligados ao vetor.

Os clones foram selecionados em meio LB ágar contendo 25 $\mu\text{g/mL}$ do antibiótico canamicina. A técnica “PCR de colônia” foi usada para confirmar se o plasmídeo recebido continha o gene inserido. Cinco colônias, de cada construção, foram escolhidas e repicadas com palitos de dente estéreis em 25 μL da reação de PCR. O protocolo de PCR de colônia possuía um ciclo adicional de 95°C por 10 minutos, para uma lise rápida das células bacterianas e liberação do DNA plasmidial. Para as reações foram utilizados os respectivos *primers* específicos, o protocolo de PCR e o programa de amplificação foram os mesmos que os utilizados para a obtenção dos fragmentos (Item 3.3). Os produtos de amplificação foram analisados em gel de agarose 1%.

3.5. EXPRESSÃO RECOMBINANTE DAS CISTATINAS

Para a expressão heteróloga das cistatinas de cana-de-açúcar, CaneCPI-5 e CaneCPI-6, foi utilizada a cepa bacteriana *E. coli* Rosetta (DE3). Os plasmídeos (pET-28a-CaneCPI5 e pET-28a-CaneCPI6), construídos para a expressão heteróloga recombinante (Item 3.4) foram utilizados na transformação, pelo método de choque térmico, de células competentes de *E. coli* Rosetta (DE3). Essas foram plaqueadas em meio LB ágar contendo 25 $\mu\text{g/mL}$ dos antibióticos canamicina e cloranfenicol. Após 16h, apenas uma colônia foi repicada em um pré-inóculo de 5 mL de meio LB líquido, contendo 25 $\mu\text{g/mL}$ dos antibióticos canamicina e cloranfenicol, o qual foi mantido sob agitação de 200 rpm a 37°C durante 16 h. A partir desse período, 5 mL da cultura (pré-inóculo), foi adicionado a 500 mL de meio LB líquido contendo 25 $\mu\text{g/mL}$ dos antibióticos canamicina e cloranfenicol, mantendo sob agitação constante de 200 rpm a 37°C, até atingir densidade ótica de aproximadamente 0,5 (600 nm). A expressão das proteínas foi induzida pela adição de IPTG à uma concentração final de 0,4 mM. Foram coletadas alíquotas da cultura antes da indução por IPTG, e em intervalos de uma hora após a indução para posterior análise da expressão das proteínas em SDS-PAGE 15% (LAEMMLI, 1970). Após 4 h de indução, as células foram coletadas por centrifugação a 5500xg, a 4°C por 10 minutos em centrífuga Sorvall RC5C plus. O material precipitado, contendo as bactérias, foi ressuspenso em tampão de lise, pH 8 (NaCl 250 mM, NaH₂PO₄ 50

mM, Tris 10 mM), e submetidos à lise celular por meio de sonicação, com cinco pulsos de um minuto à potência de 20%, com intervalos de 30 segundos, utilizando o sonicador Sonic Dismembrator 500 (Fisher Scientific). As células lisadas foram submetidas à centrifugação a 11.900xg, a 4°C por 15 minutos, para separar as frações solúvel e insolúvel, coletando alíquotas para posterior análise em SDS-PAGE 15%.

3.6. PURIFICAÇÃO DAS CISTATINAS RECOMBINANTES

As cistatinas foram purificadas da fração solúvel por cromatografia de afinidade, utilizando colunas contendo 3 mL de resina de níquel Ni-NTA Superflow (Qiagen), conforme descrito por Soares-Costa e colaboradores (2002). Para isso, as frações solúveis contendo as cistatinas recombinantes foram purificadas pelas respectivas colunas previamente equilibradas com 5 volumes (15 mL) de tampão de lise. A resina foi então lavada com 3 volumes (9 mL) de tampão de lise, e as proteínas recombinantes foram eluídas, passando pela resina com 2 volumes de tampão de lise, contendo concentrações crescentes de imidazol (10, 25, 50, 75, 100 mM), e com 3 volumes da fração mais concentrada de imidazol (250mM). As frações eluídas foram coletadas e analisadas em SDS-PAGE 15%.

As frações que continham as proteínas purificadas foram dialisadas, para eliminar o imidazol, em membranas (Pierce) 3500 MW contendo 2 L de tampão fosfato salino, PBS (NaCl 1,37 mM, KCl 2,7 mM, Na₂HPO₄ 10 mM, KH₂PO₄ 2 mM), pH 8, a 4°C, realizando duas trocas de 1 h, e uma troca de 16 h. Em seguida as proteínas foram passadas em filtro Millex® GV de 22 µm / 25 mm (Millipore). As concentrações das cistatinas foram determinadas pelo método de Bradford, 1976.

3.7. ENSAIOS DE INIBIÇÃO DAS CISTATINAS CONTRA CISTEÍNO-PEPTIDASES

A atividade inibitória das novas cistatinas de cana-de-açúcar foi testada contra cisteíno-peptidases em espectrofluorímetro Hitachi F-2000. A cistatina CaneCPI-5 teve atividade inibitória testada contra as enzimas catepsina B humana (Calbiochem) e catepsina L do inseto *S. levis*, denominada *rSl-CathL* (FONSECA *et al.*, 2012), produzida em nosso laboratório. A cistatina CaneCPI-6 foi testada apenas contra a enzima catepsina B humana.

Os ensaios foram realizados em cubeta de quartzo de 1 mL, a 37°C e em tampão acetato de sódio 100 mM pH 5,5. Para os ensaios de inibição, as catepsinas B (1,8 nM) e L (11 nM) foram pré-ativadas com 2,5 mM de DTT (dithiothreitol) por 5 min. Para os ensaios com catepsina B foi utilizado 0,01 mM do substrato fluorogênico Z-Phe-Arg-MCA, e para ensaios com a catepsina L, foi utilizado 2 mM do substrato fluorogênico Z-Leu-Arg-MCA. Em seguida concentrações crescentes dos inibidores foram adicionadas a cada 100 segundos e foram feitas leituras das atividades residuais.

A constante de inibição (K_i) foi calculada de acordo com a fórmula $V_0/V_i = 1 + [I]/K_i$, onde V_0 é a velocidade da reação na ausência do inibidor, V_i a velocidade da reação na presença do inibidor e $[I]$ corresponde à concentração de inibidor utilizada. Após análise da velocidade de reação em diferentes concentrações dos inibidores foi gerado um gráfico $(V_0 / V_i) - 1$ versus $[I]_0$, cuja inclinação da reta corresponde a $1/K_i$ (BEYNON e BOND, 1990).

3.7 ANÁLISES DE TERMOESTABILIDADE DAS CISTATINAS RECOMBINANTES

As cistatinas recombinantes foram estudadas quanto à sua resistência térmica a altas temperaturas. Estas foram submetidas a fervura (100°C), por 30 minutos, e posteriormente deixadas sobre a bancada até atingirem a temperatura ambiente (diminuição gradual da temperatura). A seguir, a atividade inibitória das cistatinas contra a enzima papaína foi medida em espectrofluorímetro, como descrito no item 3.6, utilizando o substrato fluorogênico Z-Phe-Arg-MCA. A atividade residual da enzima foi calculada a partir dos *slopes* das retas geradas com o programa FL Solutions 2.0, e expressa como porcentagem de atividade em relação ao controle. Cada análise foi feita em triplicata. A atividade inibitória, expressa como porcentagem, foi obtida da diferença entre a atividade residual da papaína na ausência das cistatinas (considerada 100%), e a atividade residual da enzima na presença das cistatinas, submetidas ou não ao tratamento (GIANOTTI, 2008).

4. RESULTADOS E DISCUSSÃO

4.1. ANÁLISES *IN SILICO* DAS SEQUÊNCIAS DAS CISTATINAS DE CANA-DE-AÇÚCAR

Foram encontradas 115 sequências relacionadas com cistatinas no banco de dados de ESTs do SUCEST, associadas às bibliotecas do projeto. Para melhor entendimento desses resultados, segue a descrição das bibliotecas que compõem esse banco de dados (Tabela 2).

Tabela 2: Descrição das bibliotecas que compõem o SUCEST. Fonte: Modificado de VETTORE *et al.*, 2001. Na quarta coluna, encontra-se o número de sequências de cistatinas de cana-de-açúcar encontradas em cada biblioteca de ESTs que integram esse banco de dados.

Código da biblioteca	Nome da biblioteca	Descrição	N ° de Sequências
AD1	<i>G. diazotrophicans</i>	Tecidos de raiz, colmo e meristema apical infectados com <i>Gluconacetobacter diazotrophicans</i>	4
AM1, AM2	Meristema Apical	Meristema apical de plantas jovens	11
CL6	Calo	Pool de calos tratados por 12 h a 4°C e 37°C, no escuro ou na luz	8
FL1, FL3, FL4, FL5, FL8	Flores 1, 3, 4, 5 e 8	Flores em diferentes estágios de desenvolvimento	14
HR1	<i>H. rubrisubalbicans</i>	Tecidos de raiz, colmo e meristema apical de plantas cultivadas <i>in vitro</i> e infectados com <i>Herbaspirillum diazotrophicans</i>	5
LB1, LB2	Gema lateral 1 e 2	Gemas lateral de plantas maduras.	6
LR1, LR2	Cartucho foliar 1 e 2	Cartucho foliar de plantas imaturas	5
LV1	Folha 1	Folhas estioladas de plântulas crescendo <i>in vitro</i>	3
RT1, RT2, RT3	Raiz 1, 2 e 3	Raízes de 0,3 cm de plantas maduras e ápice de raiz	16
RZ1, RZ2, RZ3	Raiz ao ápice	Zonas da raiz ao ápice de plantas jovens	13
SB1	Casca do caule	Superfície do caule de plantas maduras	2
SD1, SD2	Sementes 1 e 2	Sementes em desenvolvimento	16
ST1, ST3	Haste 1 e 3	Primeiro e quarto entrenós de plantas imaturas	12

Do total de 115 sequências encontradas, todas foram sequenciadas pelo projeto a partir da região 5', e 9 dessas sequências, também foram sequenciadas a partir da região 3'.

De um modo geral, os resultados da tabela 2 mostram que as cistatinas são expressas em todos os tecidos (raiz, calo, cartucho foliar, folhas, sementes, meristema, flores e caule), e fases de desenvolvimento (plântulas *in vitro*, plantas imaturas, plantas jovens e plantas maduras) analisadas pelo projeto, tendo pelo menos um representante dessa proteína em cada biblioteca. Apesar de estudos mostrarem que as cistatinas são diferencialmente expressas nos tecidos durante o desenvolvimento das plantas (MASSONNEAU *et al.*, 2005; HWANG *et al.*, 2010; SANTOS-SILVA *et al.*, 2012), sabe-se que esse inibidor de cisteíno-peptidases é ubíquo (BENCHABANE *et al.*, 2010), ou seja, está presente em toda planta e em todas as fases de desenvolvimento. Estudos realizados por Santos-Silva e colaboradores (2012), no qual foi analisado os níveis de expressão gênica da cistatina de cana-de-açúcar, CaneCPI-3, mostrou que este inibidor foi expresso nas quatro fases de desenvolvimentos coletadas: fase de germinação e estabelecimento, fase de perfilhamento, fase de crescimento dos colmos, e início da fase de maturação dos colmos; e em todos os tecidos verificados: folhas, cartucho foliar, broto, raiz, gema lateral, entrenó e meristema apical (SANTOS-SILVA *et al.*, 2012). Portanto, os resultados encontrados na busca por ESTs, corroboram com o descrito pela literatura.

As 115 sequências de cistatinas selecionadas do SUCEST foram analisadas no programa Blast2GO. Todas as sequências apresentaram similaridade com cistatinas, sendo que a maioria apresentou similaridade com sorgo e milho, espécies classificadas como poáceas, mesma família da cana-de-açúcar (Figura 5).

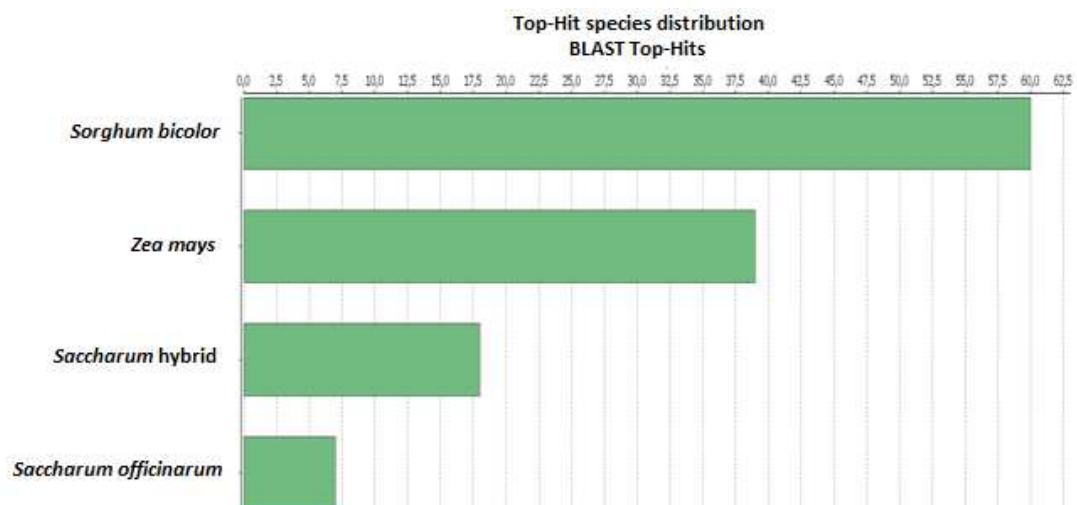


Figura 5: Gráfico extraído do programa Blast2GO, representando as espécies que apresentaram maior similaridade com cistatinas de açúcar cana-de-açúcar.

O Blast2GO faz busca por similaridade em bancos de dados de proteínas, SwissProt e NR (NCBI). O genoma de cana-de-açúcar contém apenas ESTs depositados no NCBI, enquanto que os genomas de sorgo e milho já foram totalmente sequenciados. Isto resulta no fato das sequências provenientes de cana-de-açúcar apresentarem maior número de resultados com sequências de outros gêneros de poáceas do que com seu próprio gênero, *Saccharum*. Neste trabalho, o Blast2GO foi utilizado como uma ferramenta para confirmar se as sequências de cistatinas selecionadas no SUCEST apresentam similaridade com cistatinas em diferentes bancos de dados.

Dois clones, que continham cDNA com fase aberta de leitura para diferentes cistatinas, foram selecionados para estudo. Ambos foram classificados segundo Reis e Margis (2001): o clone **SCCCLR1076E04.g** (CaneCPI-6), que contém o cDNA com a ORF de uma cistatina do grupo II, o qual reúne cistatinas com organização estrutural típica das fitocistatinas, ou seja, além das três regiões conservadas da superfamília, possuem uma sequência consenso na região N-terminal, característica de fitocistatinas, e possui também uma extensão na região C-terminal; e o clone **SCVPAM1059C07.g** (CaneCPI-5), pertence ao grupo III, do qual fazem parte as cistatinas que também apresentam as regiões conservadas da superfamília, entretanto não possuem a sequência consenso típica das fitocistatinas. O clone **SCCCLR1076E04.g** foi identificado em uma biblioteca de cDNAs gerada a partir do cartucho foliar de plantas imaturas (LR1), enquanto que o clone **SCVPAM1059C07.g** foi identificado em uma biblioteca de meristema apical de plantas jovens (AM1) (Tabela 3).

Tabela 3: Resumo das informações de cada clone de cDNA selecionados para estudo. Os códigos das bibliotecas estão destacados em negrito, nos nomes de cada clone. O sufixo .g, na nomenclatura dos clones, significa que o mesmo foi sequenciado na região 5'. Na última coluna da tabela os clones são classificados de acordo com Reis e Margis (2001).

Nome do Clone	Tecido	Nome da Cistatina	Reis e Margis (2001)
SCVPAM1059C07.g	Meristema Apical	CaneCPI-5	Grupo 3
SCCCLR1076E04.g	Cartucho Foliar	CaneCPI-6	Grupo 2

No trabalho publicado por Reis e Margis (2001), o cluster SCCCLR1076E04.g, está classificado como uma cistatina do grupo 4, as quais não possuem o consenso N-terminal LARFAV-X(3)-N, nem o motivo QxVxG, conservado entre os membros da superfamília (REIS e MARGIS, 2001). Entretanto, ao analisar a sequência do clone SCCCLR1076E04.g (*cluster* SCCCLR1076E04.g), figura 7, fica claro que a ORF que codifica essa proteína, presente nesse clone, possui não só o motivo conservado da superfamília, como também, o consenso N-terminal típico das fitocistatinas, mesmo que de forma parcial, com uma substituição do resíduo de leucina por alanina (L → A). Portanto, para a cistatina referente ao clone SCCCLR1076E04.g, é sugerida uma reclassificação, como sendo, na verdade do grupo 2, que possui as características descritas acima.

As sequências do cDNA contendo as ORFs das cistatinas de cana-de-açúcar, CaneCPI-5 e CaneCPI-6, a partir do sequenciamento dos clones de cDNA, pelo projeto SUCEST são mostradas na figura 6 e 7, respectivamente.

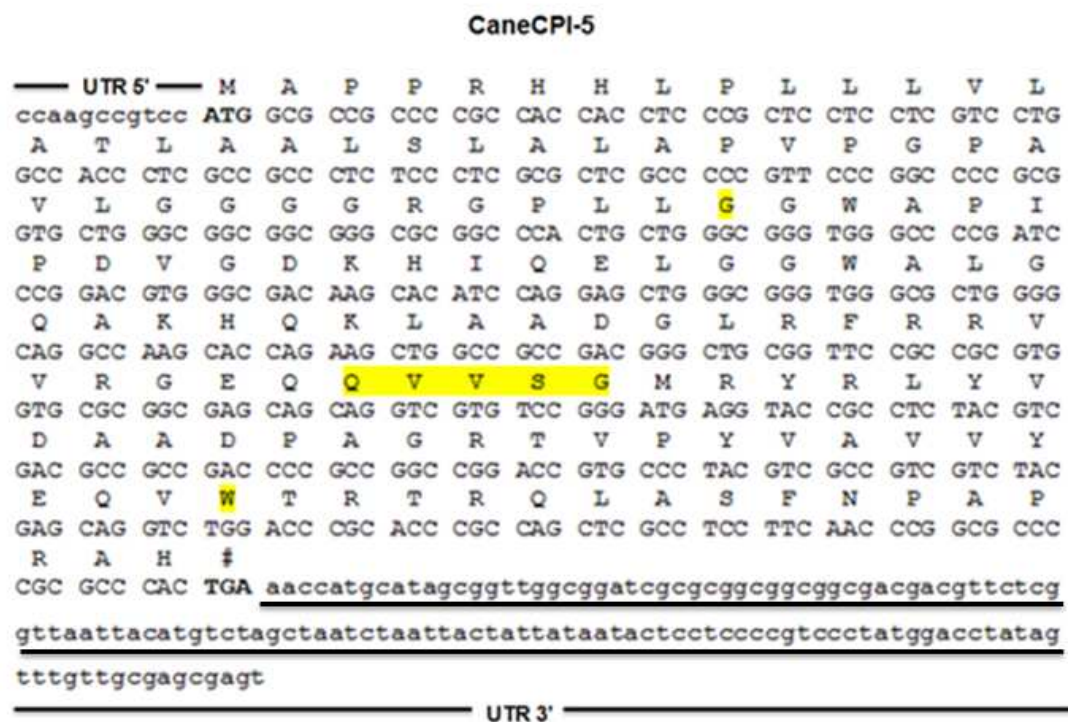


Figura 6: Sequência de nucleotídeos traduzidos do cDNA que codifica a proteína CaneCPI-5. A região codificante da proteína é mostrada a partir do códon de iniciação (ATG), e termina no respectivo códon de terminação (TGA), representado por #. As prováveis regiões conservadas na superfamília cistatina estão destacadas em amarelo na estrutura primária das proteínas. Nota-se que a proteína CaneCPI-5 não possui o motivo conservado LARFAV-3(X)-N, típico de cistatinas de plantas. A região não traduzida (UTR) aparece sublinhada em preto.

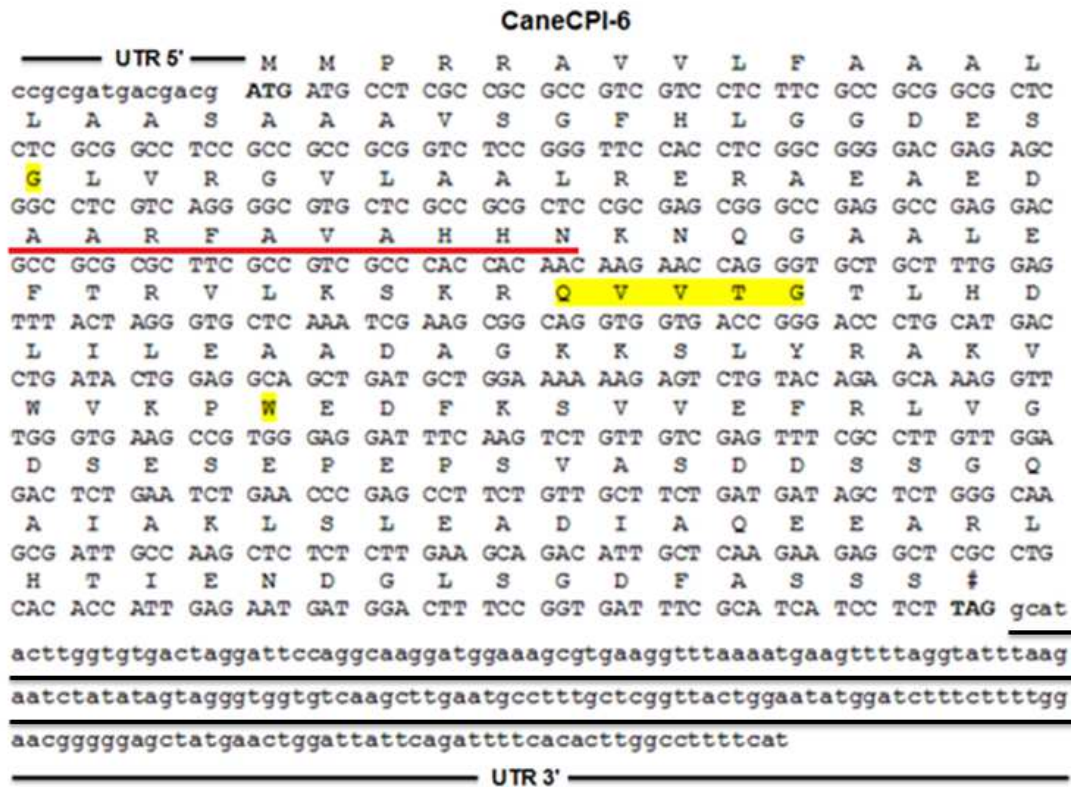


Figura 7: Sequência de nucleotídeos traduzidos do cDNA que codifica a proteína CaneCPI-6. A região codificante da proteína é mostrada a partir do códon de iniciação (ATG), e termina no respectivo códon de terminação (TAG), representado por #. As prováveis regiões conservadas na superfamília cistatina estão destacadas em amarelo na estrutura primária das proteínas. A região conservada típica de fitocistatinas, LARFAV-3(X)-N, aparece sublinhada em vermelho. Nota-se a substituição do resíduo de leucina por alanina (L → A).

A Tabela 4 mostra o tamanho das ORFs e as massas moleculares teóricas das cistatinas por elas codificadas.

Tabela 4: Tamanho das ORFs e massas moleculares teóricas das cistatinas. ⁽¹⁾ Valores referentes às cistatinas contendo a sequência do provável peptídeo sinal; ⁽²⁾ Valores referentes às proteínas de fusão recombinantes (cistatinas sem a sequência do peptídeo sinal + a sequência derivada do vetor pET-28a). A massa molecular teórica foi calculada no programa ProtParam.

Cistatina	Nucleotídeos	Aminoácidos	Massa Molecular (Da)
CaneCPI-5	411	136	14546,9 ⁽¹⁾
			14553,7 ⁽²⁾
CaneCPI-6	525	174	18527,8 ⁽¹⁾
			18896,9 ⁽²⁾

O peptídeo sinal, ou sequência-sinal, é uma sequência sinalizadora que direciona as proteínas aos destinos celulares corretos, e possuem algumas características bem definidas, tais como: é tipicamente constituída de 15 a 60 aminoácidos; é frequentemente encontrada na região N-terminal; cada sequência sinal especifica um destino particular na célula; é composta de cerca de 5 a 10 aminoácidos hidrofóbicos (ALBERTS *et al.*, 2010). Os gráficos gerados pelos programas ProtScale e SignalP revelaram a presença de uma região N-terminal altamente hidrofóbica nas sequências das cistatinas, a qual apresenta as características de um peptídeo sinal (Figura 8 e 9). Esse resultado indica, que as cistatinas CaneCPI-5 e CaneCPI-6 são provavelmente secretadas. O peptídeo sinal presente na maioria desses inibidores sugerem uma localização não citosólica dessas proteínas (ABRAHAM *et al.*, 2006). A localização subcelular das fitocistatinas, ainda era pouco conhecida, com exceção das multicistatinas de batata e tomate encontradas nos vacúolos e no citoplasma; da oryzacistatina do arroz, detectada no citoplasma, nos vacúolos e cloroplastos (MADUREIRA *et al.*, 2006; PRINS *et al.*, 2008; NISSEN *et al.*, 2009); e das cistatinas de cevada, localizadas no retículo endoplasmático e no complexo de Golgi (MARTINEZ *et al.*, 2009). Estudos recentes demonstraram que a cistatina de milho, CC-9, foi detectada no apoplasto, onde inibia cisteíno-peptidases do tipo-papaína (VAN DER LINDE *et al.*, 2012). As análises feitas pelo programa WoLF PSORT sugerem que a localização subcelular mais provável para CaneCPI-5 seja na matriz mitocondrial e nos cloroplastos, e na matriz mitocondrial, nos cloroplastos e na matriz extracelular para proteína CaneCPI-6.

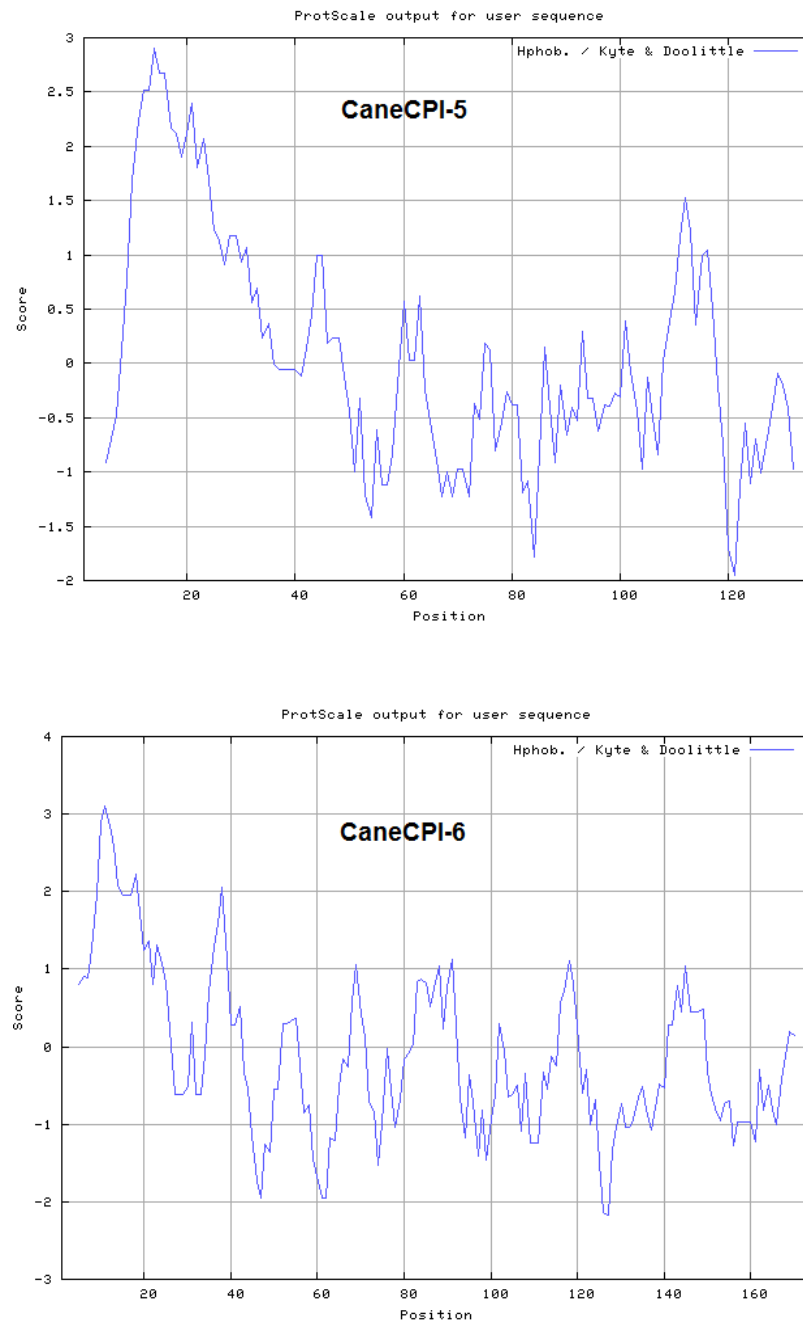


Figura 8: Perfil de hidrofobicidade das cistatinas CaneCPI-5 e CaneCPI-6. Os gráficos gerados com o programa ProtScale mostram a presença de uma região N-terminal altamente hidrofóbica nas sequências das cistatinas, correspondentes aos prováveis peptídeos sinais.

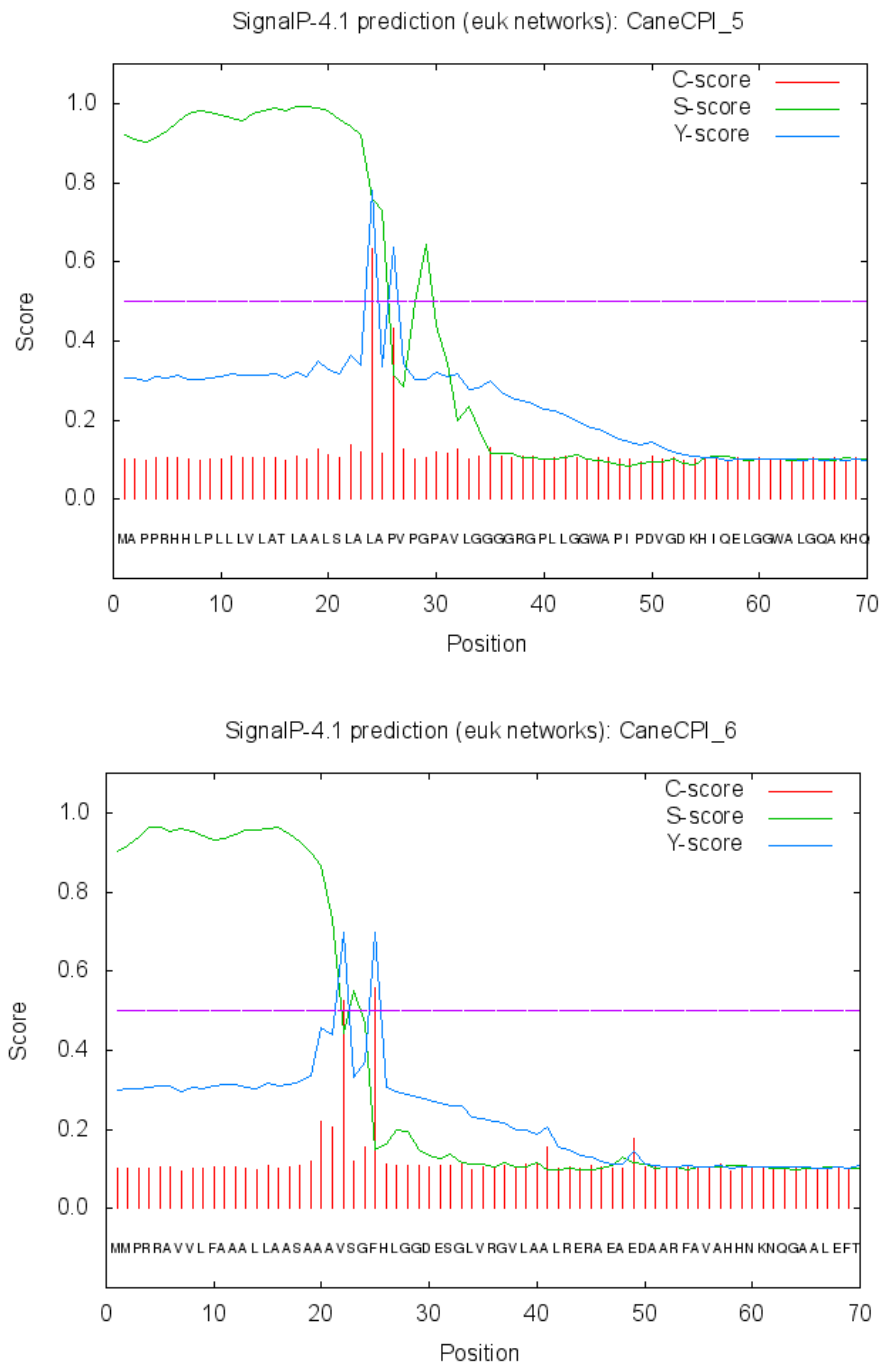


Figura 9: Peptídeos sinais identificados nas seqüências das cistatinas CaneCPI-5 e CaneCPI-6. Os gráficos gerados com o programa SignalP mostram a presença de seqüências com características de peptídeos sinais nas cistatinas. O provável sítio de clivagem do peptídeo sinal é mostrado no pico vermelho. A região do peptídeo sinal, altamente hidrofóbica é mostrada em verde no gráfico.

A Figura 10 mostra o alinhamento das sequências de proteína, de CaneCPI-5 e CaneCPI-6, utilizando o programa MultAlin (CORPET, 1988), destacando as regiões dos prováveis peptídeos sinais.

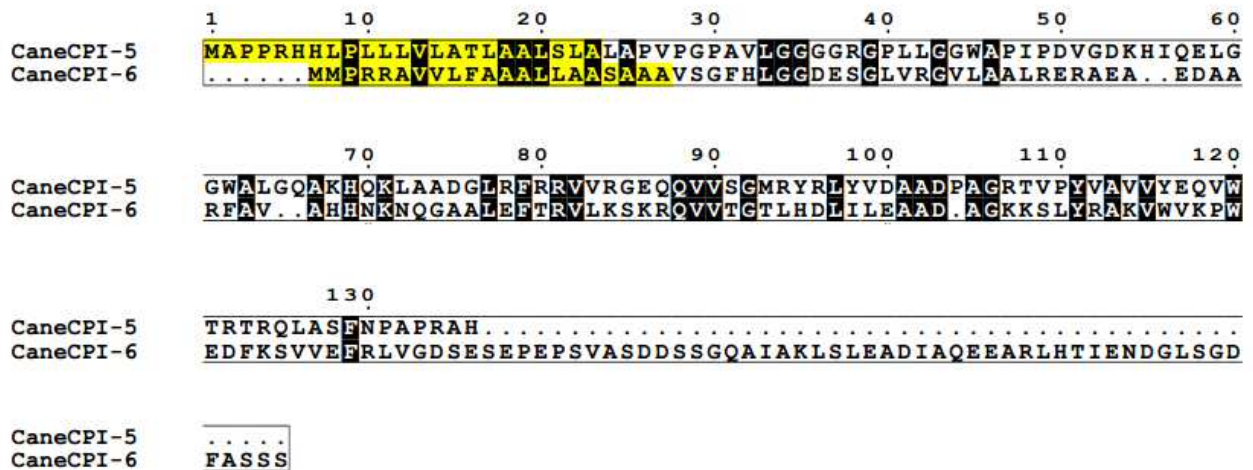


Figura 10: Alinhamento entre as sequências das proteínas CaneCPI-5 e CaneCPI-6. As sequências de aminoácidos foram deduzidas das ORFs contidas nos clones SCVPAM1059C07.g (CaneCPI-5) e SCCCLR1076E04.g (CaneCPI-6). As regiões destacadas em amarelo indicam as prováveis sequências sinalizadoras identificadas pelo programa SignalP.

No banco de dados do NCBI estão depositadas sequências de cistatinas de outras espécies vegetais. Algumas dessas sequências apresentam alta identidade de aminoácidos com as cistatinas de cana-de-açúcar, como pode ser observado na tabela 5 e nas figuras 11 e 12.

Tabela 5: Identidade de sequências de aminoácidos entre as cistatinas de cana-de-açúcar e cistatinas de outras espécies vegetais. A porcentagem de identidade entre as sequências foi obtida com o programa BLAST. Os números de acesso das sequências no banco de dados NCBI estão entre parênteses.

Cistatinas de cana-de-açúcar	Cistatinas de outras espécies vegetais	% de identidade
CaneCPI-5	<i>Zea mays</i> (ACG48521.1)	95
	<i>Sorghum bicolor</i> (XP_002465061.1)	85
	<i>Setaria italica</i> (XP_004983911.1)	75
	<i>Oryza sativa</i> (NP_001050419.1)	70
CaneCPI-6	<i>Zea mays</i> (NP_001151685.1)	94
	<i>Sorghum bicolor</i> (XP_002439792.1)	97
	<i>Setaria italica</i> (XP_004962105.1)	87
	<i>Hordeum vulgare</i> (CAG38124.1)	68

Como esperado, as cistatinas apresentaram maiores identidades de aminoácidos com espécies vegetais da família poácea: *Sorghum bicolor*, sorgo, (PATERSON *et al.*, 2009), *Zea mays*, milho (CHAUBAL *et al.*, 2000), *Setaria italica*, milhete, (DOUST e KELLOGG, 2006), *Oryza sativa*, arroz (ZHAO *et al.*, 1998) e *Hordeum vulgare*, cevada (MARTINEZ *et al.*, 2009)

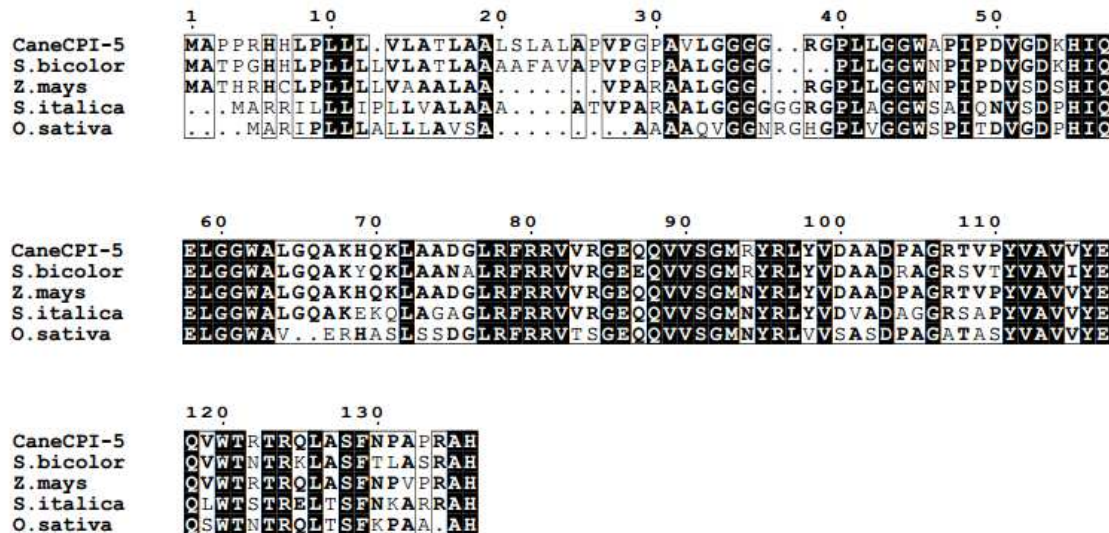


Figura 11: Alinhamento da cistatina CaneCPI-5 com cistatinas de outras espécies vegetais. Alinhamento da sequência de aminoácidos da CaneCPI-5 com outras fitocistatinas de diferentes espécies vegetais. Essas sequências apresentaram maiores porcentagens de identidade na busca realizada com o programa Blast, no banco de dados NR do NCBI.

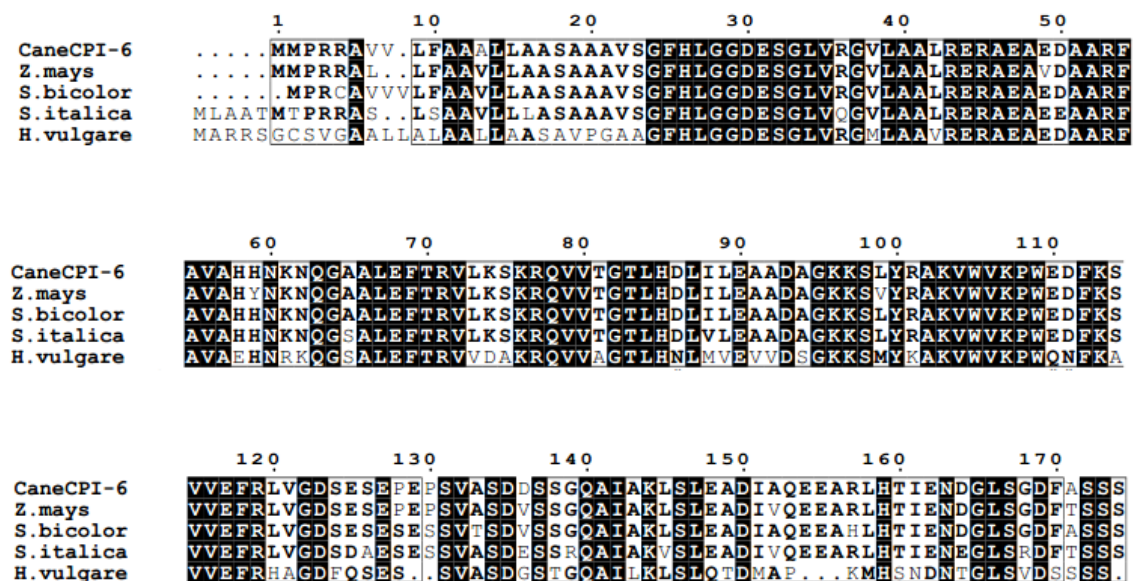


Figura 12: Alinhamento da cistatina CaneCPI-6 com cistatinas de outras espécies vegetais. Alinhamento da sequência de aminoácidos da CaneCPI-6 com outras fitocistatinas de diferentes espécies vegetais. Essas sequências apresentaram maiores porcentagens de identidade na busca realizada com o programa Blast, no banco de dados NR do NCBI.

A Figura 13 mostra o alinhamento das cistatinas CaneCPI-5 e CaneCPI-6, com as outras cistatinas de cana-de-açúcar: CaneCPI-1, CaneCPI-2, CaneCPI-3, e CaneCPI-4, todas expressas e caracterizadas pelo nosso grupo de pesquisa. Nota-se, que existe bastante diferença ao longo das sequências de aminoácidos, porém há alta identidade de aminoácidos nos domínios conservados. O domínio é geralmente conservado por ser uma região importante em termos de mecanismo de inibição. Nas cistatinas, particularmente, uma glicina na região N-terminal, o domínio QxVxG , e o resíduo de triptofano na região C-terminal, são importantes na interação com a protease. A cistatina CaneCPI-6 tem sua região C-terminal mais estendida, quando comparada às cistatinas CaneCPI-1, 2, 4 e 5, porém essa extensão é menor quando comparada à CaneCPI-3. Além disso, diferente de CaneCPI-3, a proteína CaneCPI-6 não possui, nessa região, o motivo conservado SNSL, envolvido com a inibição de cisteíno-peptidases do tipo leguminaína. Dentre essas cistatinas de cana-de-açúcar, as que possuem maior similaridade em sua estrutura primária, são as proteínas CaneCPI-4 e CaneCPI-5. O alinhamento entre essas duas cistatinas revelou alta identidade de aminoácidos, principalmente próximo aos domínios conservados (Figura 14).



Figura 13: Alinhamento das cistatinas de cana-de-açúcar. O motivo SNSL, na proteína CaneCPI-3, importante para inibir cisteíno-peptidases do tipo legumína, está destacado em vermelho.



Figura 14: Alinhamento das cistatinas de cana-de-açúcar, CaneCPI-4 e CaneCPI-5. Essas cistatinas possuem alta identidade de aminoácidos, principalmente próximo aos motivos conservados.

4.2 ANÁLISE FILOGENÉTICA

Com o intuito de estabelecer uma relação filogenética entre as sequências das prováveis cistatinas de cana-de-açúcar, identificadas no SUCEST (item 4.1), 59 delas foram selecionadas para construção da árvore filogenética. Sequências repetidas ou truncadas foram retiradas da análise. Sete sequências de cistatinas de *Arabidopsis* (*AtCYS1* - *AtCYS7*) foram utilizadas como controle do tipo *outgroup*. As proteínas foram reunidas em 4 principais grupos, identificados por diferentes cores na árvore filogenética (Figura 15), as sequências pertencentes a cada grupo foram analisadas *in silico*, e caracterizadas (Figura 16). O Anexo 1 mostra a identificação dos clones do SUCEST e seus respectivos códigos usados na construção da árvore filogenética.

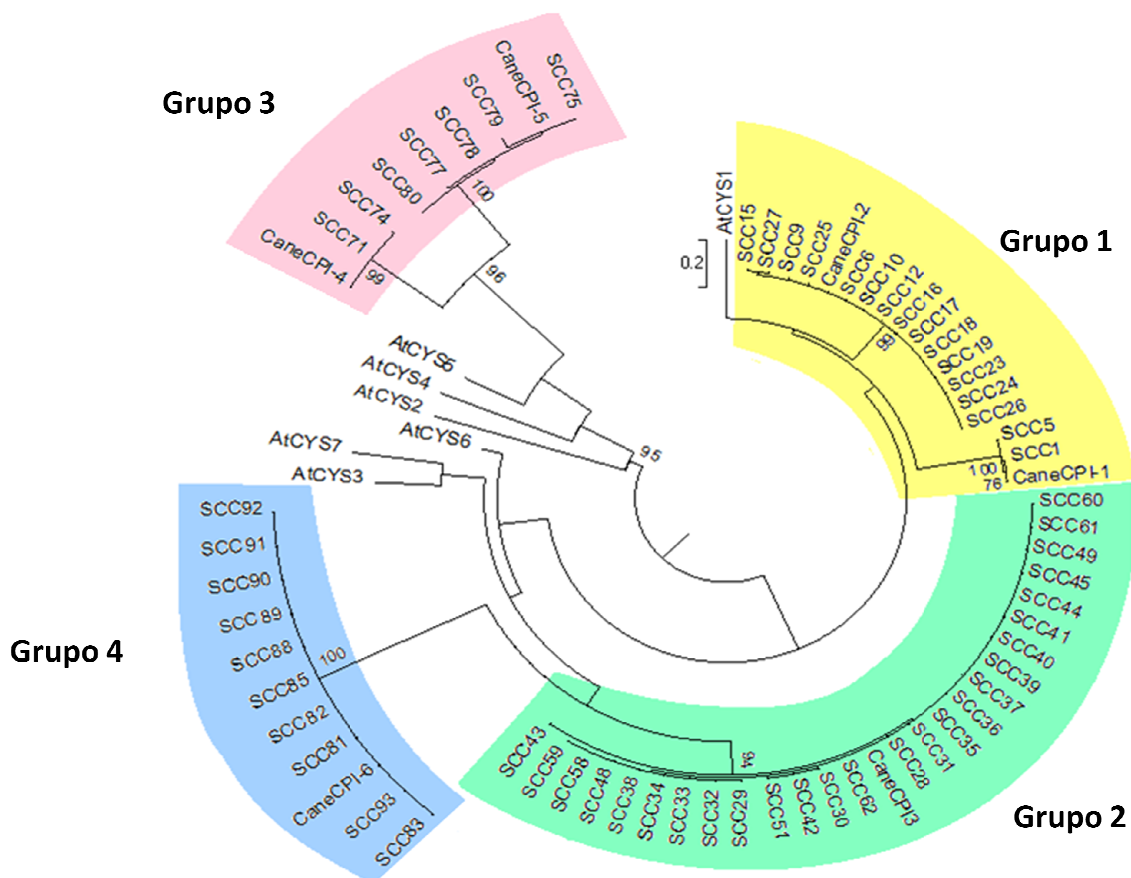


Figura 15: Árvore filogenética correspondentes às prováveis cistatinas de cana-de-açúcar. O método de máxima verossimilhança foi utilizado para construção filogenética a partir do alinhamento obtido pelo algoritmo MUSCLE e posteriormente corrigido pela matriz WAG com distribuição G de 5 parâmetros sem sítios invariantes. Os 4 grupos estão representados por diferentes cores. Amarelo: Grupo 1; Verde: Grupo 2; Rosa: Grupo 3; Azul: Grupo 4. Sete sequências de cistatina de *Arabidopsis* foram usadas como *outgroup*: *AtCYS1* (GI: 15239883), *AtCYS2* (GI: 330253521), *AtCYS3* (GI: 330254798), *AtCYS4* (GI: 332658361), *AtCYS5* (GI: 332008148), *AtCYS6* (GI: 332641680), e *AtCYS7* (GI: 332003446).

De acordo com as análises filogenéticas uma nova classificação das prováveis cistatinas de cana-de-açúcar foi possível:

Grupo 1: Nesse grupo estão reunidas as cistatinas que possuem os três motivos conservados na superfamília cistatina, além do motivo conservado LARFAV-X(3)-N, exclusivo de fitocistatinas. Portanto, as cistatinas do grupo 1 são consideradas fitocistatinas típicas. Fazem parte desse grupo as cistatinas CaneCPI-1 e CaneCPI-2.

Grupo 2: No grupo dois reúnem-se as cistatinas que, assim como no grupo 1, possuem os três motivos conservados da superfamília, além do motivo conservado típico das fitocistatinas. As cistatinas desse grupo, possuem ainda, sua região C-terminal estendida, onde está presente o motivo conservado SNSL, importante para inibição de legumina. Portanto, as cistatinas de cana-de-açúcar do grupo 2, são inibidores com capacidade bifuncional, podendo inibir cisteíno-peptidases do tipo papaína e legumina. Faz parte desse grupo a cistatina CaneCPI-3.

Grupo 3: As cistatinas do grupo 3 possuem os três motivos conservados entre a superfamília cistatina, mas carecem do motivo conservado típico das fitocistatinas. Fazem parte desse grupo as cistatinas CaneCPI-4 e CaneCPI-5.

Grupo 4: Nesse grupo encontram-se as cistatinas de cana-de-açúcar que possuem os três motivos conservados na superfamília. Possuem o motivo típico das fitocistatinas LARFAV-X(3)-N, porém com a substituição do resíduo de leucina por um de alanina (L → A). As cistatinas do grupo 4, também possuem a região C-terminal estendida, porém não foi identificada nenhuma característica particular nessa região. A proteína CaneCPI-6 faz parte desse grupo.

região C-terminal (NAGATA *et al.*, 2000; PAVLOVA e BJIÖRK, 2003; VALADARES *et al.*, 2010). As cistatinas dos grupos 1 e 2 possuem a sequência consenso LARFAV-X(3)-N, típica da família das fitocistatinas; as do grupo 3 possuem esse consenso de forma parcial, com a substituição do resíduo de leucina por um de alanina (L → A); e as do grupo 4 carecem dessa sequência. A função desse consenso ainda não foi bem esclarecida, mas sabe-se que os aminoácidos nessa região podem variar: [LVI]-[AGT]-[RKE]-[FY]-[AS]-[VI]-X-[EDQV]-[HYFQ]-N, formando uma α -hélice na região N-terminal (MARGIS *et al.*, 1998; VAN DER LINDE *et al.*, 2012). Um estudo recente, feito por Valadares e colaboradores (2010), sugere que o papel desse motivo é prover uma complementaridade ideal para os resíduos hidrofóbicos na folha- β das fitocistatinas, essencial para estabilizar a estrutura terciária (VALADARES *et al.*, 2010). Por fim, as cistatinas dos grupos 2 e 4, possuem sua região C-terminal estendida. As cistatinas do grupo 2, possuem ainda, um motivo conservado SNSL, importante para a inibição das cisteíno-peptidases do tipo legumaina. Essas cistatinas, com a região C-terminal estendida, possuem massa molecular de aproximadamente 23 kDa, encontradas em algas verdes e em táxons de plantas de grau superior, são resultantes de eventos de duplicação gênica precoce, já na evolução de *Viridiplantae*, e bem antes da irradiação de plantas terrestres (MARTINEZ *et al.*, 2005a; MARGIS-PINHEIRO *et al.*, 2008). O domínio C-terminal dessas proteínas teria então sido submetido a uma significativa remodelagem funcional e estrutural, levando gradualmente à perda de atividade contra peptidases do tipo papaína, e adquirindo novas funções incluindo atividade inibitória contra cisteíno-peptidases do tipo legumaina (Família C13) (MARTINEZ *et al.*, 2007). Por fim, não foi encontrada nenhuma característica particular na extensão C-terminal das cistatinas pertencentes ao grupo 4, que é menor quando comparada com a extensão C-terminal do grupo 2.

4.3. AMPLIFICAÇÃO DAS ORFS QUE CODIFICAM AS CISTATINAS CANECPI-5 E CANECPI-6, E CONSTRUÇÃO DOS VETORES DE EXPRESSÃO

A expressão e a solubilidade de outras cistatinas de cana-de-açúcar contendo o peptídeo sinal foram muito reduzidas, provavelmente devido à alta hidrofobicidade dessa sequência (GIANOTTI *et al.*, 2008). Portanto, foi decidido expressar as novas cistatinas de cana-de-açúcar sem as sequências correspondentes aos prováveis peptídeos sinais. As ORFs foram obtidas por PCR e a análise dos produtos das duas amplificações em gel de agarose mostrou a presença do fragmento correspondente a cada uma das cistatinas nos respectivos tamanhos esperados, de aproximadamente 345 pb para CaneCPI-5 e 465 pb para CaneCPI-6 (Figura 17).

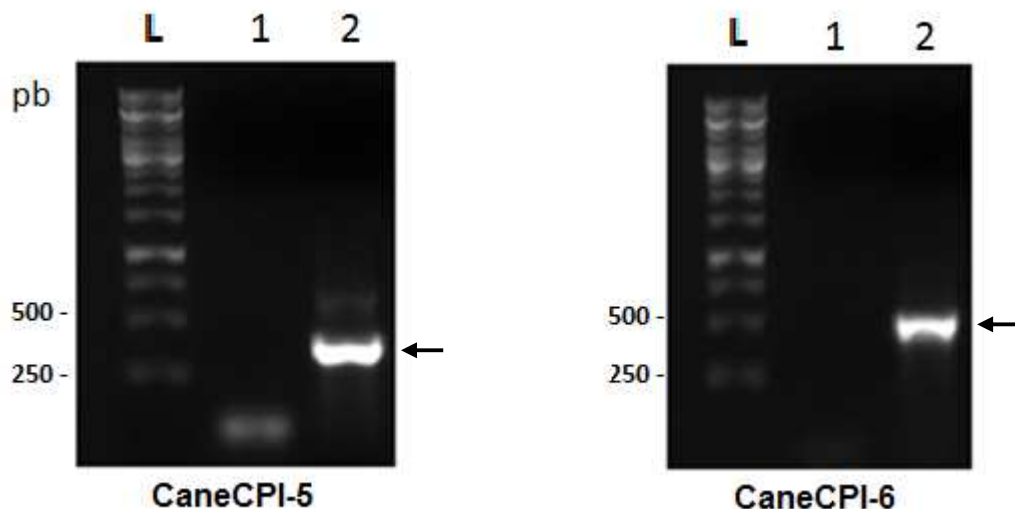


Figura 17: Amplificação das ORFs das cistatinas da cana-de-açúcar. As regiões gênicas das cistatinas CaneCPI-5 e CaneCPI-6 foram amplificadas por PCR a partir dos clones SCVPAM1059C07.g e SCCCLR1076E04.g, respectivamente, utilizando primers específicos. Nas colunas do gel: **(L)** 1 Kb DNA *ladder* (Invitrogen), **(1)** controle negativo da PCR, e **(2)** produto de PCR correspondente às ORFs de cada cistatina. **pb**: pares de bases. As setas indicam as bandas amplificadas de aproximadamente 345 pb para CaneCPI-5 e 465 pb para CaneCPI-6. Gel de agarose 1%.

Os produtos da PCR (Figura 17) foram ligados ao plasmídeo de propagação pTZ57R/T (Figura 3). Devido a enzima Taq DNA polimerase adicionar um nucleotídeo adenina (A) na extremidade 3' dos fragmentos amplificados, a ligação destes ao vetor pTZ57R/T é facilitada pela presença de um nucleotídeo timina (T) em ambas as extremidades 3' do plasmídeo. Neste vetor, o sítio para a clonagem do fragmento de PCR está inserido na região que codifica o peptídeo α da enzima β -galactosidase, permitindo a identificação dos clones recombinantes por uma seleção

branco/azul. Dessa maneira as colônias recombinantes brancas foram selecionadas, e confirmadas por PCR de colônia, utilizando os *primers* do vetor, M13 *forward* e *reverse*, confirmando um produto de aproximadamente 490 pb para CaneCPI-5, e 610 pb para CaneCPI-6. (Figura 18).

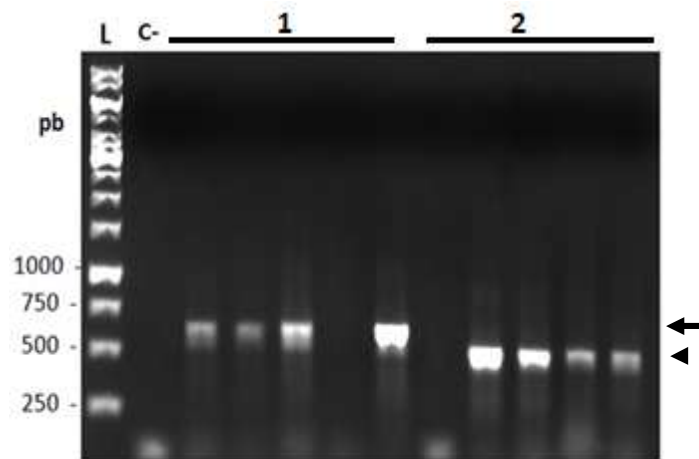


Figura 18: PCR de colônia de *E. coli* DH5- α transformadas com o vetor pTZ57R/T com as ORFs de CaneCPI-5 e CaneCPI-6 clonadas. (1) indica os fragmentos correspondentes à cistatina CaneCPI-6, e (2) indica os fragmentos correspondentes à cistatina CaneCPI-5, ambas com sequência do vetor pTZ57R/T. (C-): controle negativo: (L): 1 kb DNA ladder (Invitrogen) pb: pares de bases. A cabeça de seta indica as bandas amplificadas de aproximadamente 490 pb para CaneCPI-5 e a seta indica as bandas amplificadas de aproximadamente 610 pb para CaneCPI-6. Gel de agarose 1%.

Como outras cistatinas de cana-de-açúcar, CaneCPI-1 (SOARES-COSTA *et al.*, 2002), CaneCPI-2 e CaneCPI-3 (GIANOTTI *et al.*, 2006), já haviam sido expressas e purificadas de forma eficiente usando o vetor pET-28a, neste trabalho foi adotada a mesma estratégia para a obtenção das proteínas CaneCPI-5 e CaneCPI-6. O sistema pET (Novagen) foi desenvolvido para clonagem e expressão de proteínas recombinantes em *E. coli*. Os genes de interesse são clonados nos plasmídeos pET sob controle transcricional do promotor do bacteriófago T7. A expressão é induzida pela adição de IPTG ou lactose à cultura bacteriana, devido as células hospedeiras proverem uma fonte de RNA polimerase T7 sob o controle do promotor lacUV5. Desta forma a presença do IPTG induz a expressão da RNA polimerase T7, que por sua vez, liga-se ao promotor T7 e ativa a expressão do gene clonado no vetor pET.

A figura 19 mostra os resultados das reações de digestão dos vetores de propagação pTZ57R/T contendo os fragmentos correspondentes às regiões gênicas que codificam as novas cistatinas recombinantes.

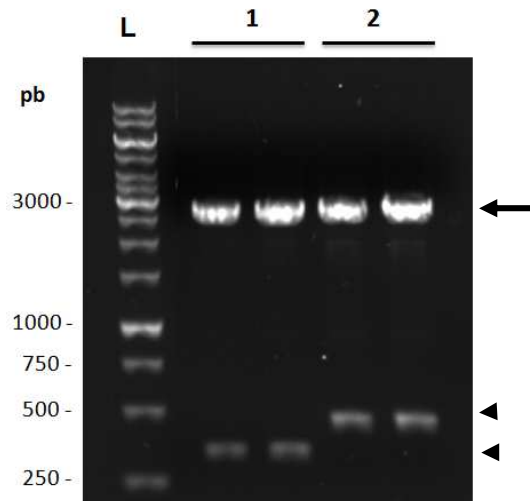


Figura 19: Reações de digestão dos vetores pTZ57R/T, contendo as ORFs de CaneCPI-5 e CaneCPI-6. Nas colunas do gel: (L): 1 Kb DNA *ladder* (Invitrogen); (1): Plasmídeo de propagação pTZ57R/T clivado com as enzimas de restrição *Nhe* I e *Bam*HI e o fragmento correspondente à ORF da proteína CaneCPI-5 (~345 pb); (2): Plasmídeo pTZ57R/T clivado com as enzimas de restrição *Nhe* I e *Bam*HI e o fragmento correspondente à ORF da proteína CaneCPI-6 (~465 pb). pb: pares de bases. A seta indica a banda de 2886 pb do vetor pTZ57R/T, e as cabeças das setas indicam os insertos clonados. Gel de agarose 1%.

A figura 20 mostra o resultado da PCR de colônia, feita a partir de células competentes de *E. coli* DH5- α , transformadas com o vetor pET-28a, contendo as ORFs de CaneCPI-5 e CaneCPI-6. Para a reação de PCR foram utilizados os *primers* do vetor T7 *promoter* e T7 *terminator*, confirmando os produtos de aproximadamente 610 pb para CaneCPI-5, e 730 pb para CaneCPI-6.

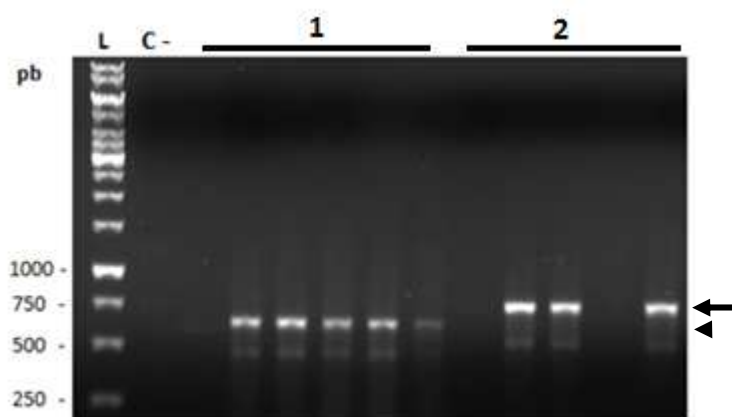


Figura 20: PCR de colônia de *E. coli* DH5- α transformadas com o vetor pET-28a com ORFs de CaneCPI-5 e CaneCPI-6 subclonadas. (1) indica os fragmentos correspondentes à cistatina CaneCPI-5, e (2) indica os fragmentos correspondentes à cistatina CaneCPI-6, ambas com sequência do vetor pET-28a. L: 1 kb (C-): controle negativo: (L): 1 kb DNA ladder (Invitrogen) pb: pares de bases. A cabeça de seta indica as bandas amplificadas de aproximadamente 610 pb para CaneCPI-5 e a seta indica as bandas amplificadas de aproximadamente 730 pb para CaneCPI-6. Gel de agarose 1%.

4.4. EXPRESSÃO HETERÓLOGA E PURIFICAÇÃO DAS CISTATINAS RECOMBINANTES

As ORFs correspondentes às cistatinas foram clonadas no vetor de expressão pET-28a (figura 4). Este vetor permite a clonagem direcional do gene de interesse em fase com uma sequência, que codifica uma “cauda” de seis histidinas na extremidade C ou N-terminal. Esta “His-Tag” facilita a purificação da proteína de fusão por cromatografia de afinidade e posteriormente pode ser removida, se necessário, uma vez que entre a sequência da proteína de interesse e as histidinas existe um sítio para a protease trombina. Dessa maneira, as proteínas, CaneCPI-5 e CaneCPI-6 foram expressas em fusão com a cauda de histidinas (His-Tag), na região N-terminal, derivada do vetor pET-28a. Considerando que os resíduos de histidina possuem afinidade pelo níquel, foi possível purificá-las em etapa única por esse procedimento.

A análise em SDS-PAGE 15%, das amostras coletadas periodicamente a cada 1 hora, após o início da indução, revelou que a partir da primeira hora, as proteínas recombinantes já eram expressas, e que não há muita diferença nos níveis de expressão das amostras coletadas em 2, 3 e 4 horas (Figura 21).

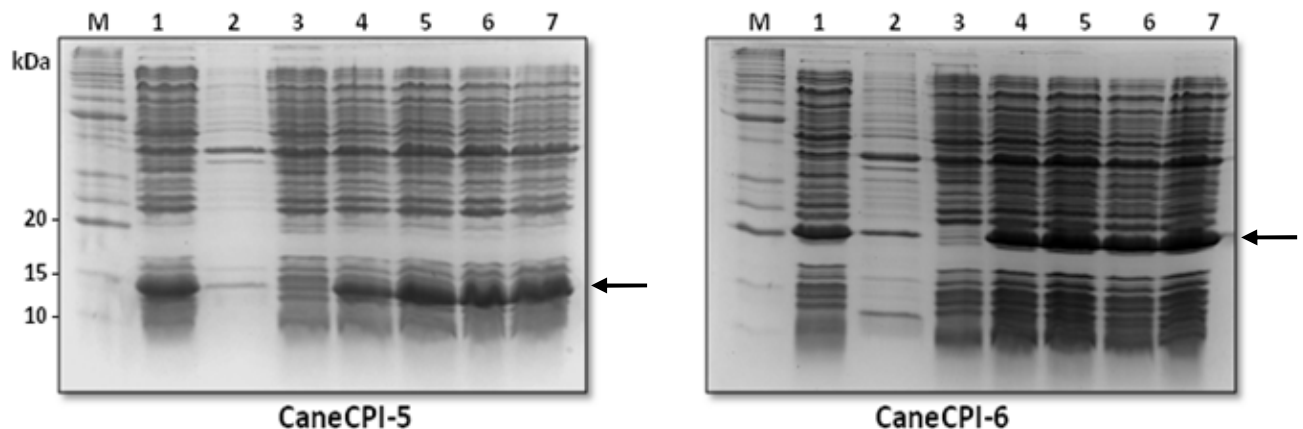


Figura 21: Expressão heteróloga e coleta periódica a cada 1 hora após a indução das cistatinas CaneCPI-5 e CaneCPI-6. Análise em SDS-PAGE 15%, das amostras coletadas das frações solúveis (1), e insolúveis (2), não induzidas (3), com 1 hora de indução (4), com 2 horas de indução (5), com 3 horas de indução (6), e com 4 horas de indução (7). M: BenchMark™ Protein Ladder (Invitrogen). As setas indicam as bandas referentes às cistatinas CaneCPI-5 (~14,5 kDa) e CaneCPI-6 (~19 kDa).

A comparação dos padrões proteicos dos clones bacterianos, após 4 h de indução, e não induzidos, analisado em SDS-PAGE 15%, revelou presença de bandas de tamanhos esperados (His-Tag + cistatina) nos clones induzidos com IPTG, referentes à superexpressão das proteínas (Figura 22, poços 1 e 2).

As frações solúveis e insolúveis resultante da lise celular das bactérias, mostraram que, ambas as cistatinas, CaneCPI-5 e CaneCPI-6, encontram-se em sua maior parte na fração solúvel (Figura 22, poços 3 e 4). As frações solúveis foram utilizadas para a purificação das cistatinas recombinantes por cromatografia de afinidade em colunas contendo resina de níquel. Essas frações foram, então, adicionadas às colunas e as proteínas foram eluídas em tampão de lise contendo imidazol. O imidazol, que também apresenta forte afinidade pelo níquel, compete com as histidinas pela ligação a esse metal, liberando a proteína de fusão.

A análise em SDS-PAGE revelou que a maior parte das proteínas CaneCPI-5 e CaneCPI-6 foram eluídas nas frações que continham 250 mM de imidazol (Figura 22, poço 5).

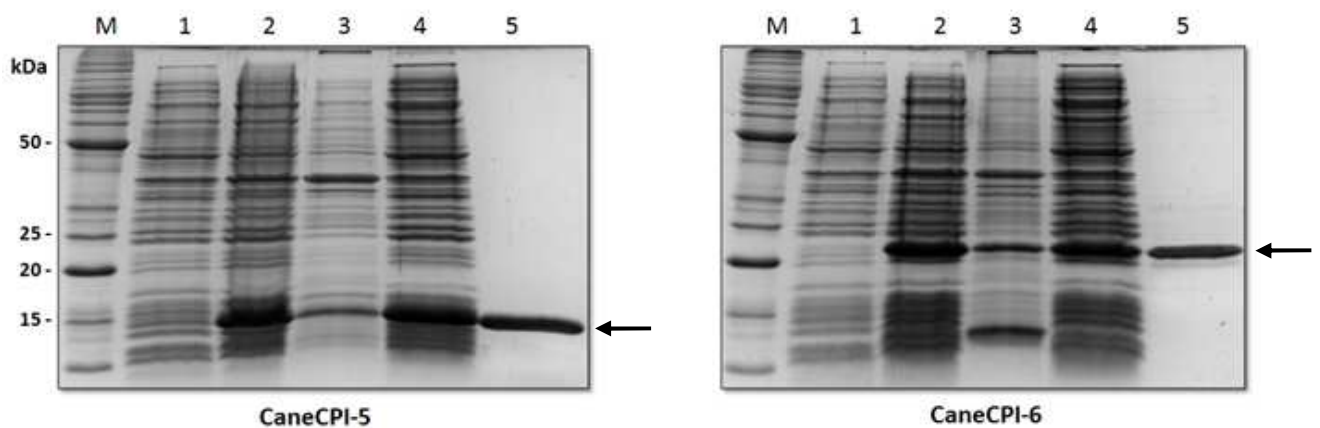


Figura 22: Expressão heteróloga e purificação das cistatinas CaneCPI-5 e CaneCPI-6. Análise em SDS-PAGE 15% das amostras coletadas das culturas não induzidas (1), e induzidas (2), das frações insolúveis (3), e solúveis (4), resultante da lise celular, e das purificações em coluna de níquel (5). M: BenchMark™ Protein Ladder (Invitrogen). As setas indicam as bandas referentes às cistatinas CaneCPI-5 (~15,5 kDa) e CaneCPI-6 (~19 kDa).

O rendimento das cistatinas purificadas, foi de aproximadamente 16 mg por litro de cultura celular, para CaneCPI-5, e de aproximadamente 8,0 mg por litro de cultura celular, para CaneCPI-6. Os resultados obtidos são satisfatórios, uma vez que foi possível purificar as proteínas de interesse em concentrações e pureza suficientes para a realização dos ensaios experimentais. A tabela 6 mostra uma comparação

entre os rendimentos das seis cistatinas de cana-de-açúcar expressas, em *E. coli*, e purificadas no Laboratório de Biologia Molecular (LBM-DGE/UFSCar). Mesmo com alta identidade de aminoácidos na estrutura primária das proteínas CaneCPI-4 e CaneCPI-5, nota-se que o rendimento da cistatina CaneCPI-5 (~16 mg / L de cultura) foi maior do que em CaneCPI-4 (~2 mg / L de cultura). Sabe-se que a proteína CaneCPI-4 tende a se agregar, o que a torna pouco solúvel quando expressa de maneira heteróloga em células de bactéria (GIANOTTI *et al.*, 2008). Esse resultado não foi observado para CaneCPI-5, sugerindo que apesar de serem similares, as diferenças dos aminoácidos ao longo de sua estrutura primária refletem em uma maior solubilidade de CaneCPI-5. O fato da proteína CaneCPI-5 ser mais solúvel em relação a CaneCPI-4 é vantajoso do ponto de vista industrial. Os rendimentos de CaneCPI-5 e CaneCPI-6 são mais similares aos de CaneCPI-2 (~20 mg / L de cultura), e CaneCPI-1 (~10 mg / L de cultura), respectivamente.

Tabela 6: Rendimento das cistatinas de cana-de-açúcar. Rendimento expresso em miligramas de proteína por litro de cultura induzido.

Cistatina	Rendimento por L de cultura	Referência
CaneCPI-1	10,00 mg	SOARES-COSTA <i>et al.</i> , 2002
CaneCPI-2	20,00 mg	GIANOTTI <i>et al.</i> , 2006
CaneCPI-3	22,00 mg	
CaneCPI-4	2,00 mg	GIANOTTI <i>et al.</i> , 2008
CaneCPI-5	16,00 mg	Este estudo
CaneCPI-6	8,00 mg	

4.5 INIBIÇÃO DAS ATIVIDADES DAS CISTEÍNO-PEPTIDASES CATEPSINA B HUMANA E *RSL-CATHL* PELAS NOVAS CISTATINAS RECOMBINANTES.

A atividade inibitória das cistatinas recombinantes contra a enzima catepsina B, foi medida em espectrofluotímetro Hitachi F2000, utilizando o substrato fluorogênico Z-Phe-Arg-MCA. Os resultados, sumarizados na figura 23, mostram que as cistatinas recombinantes, CaneCPI-5 e CaneCPI-6, apresentaram diferenças nas atividades inibitória contra catepsina B, sendo que somente a cistatina CaneCPI-5 foi capaz de inibir eficientemente a atividade dessa cisteíno peptidase.

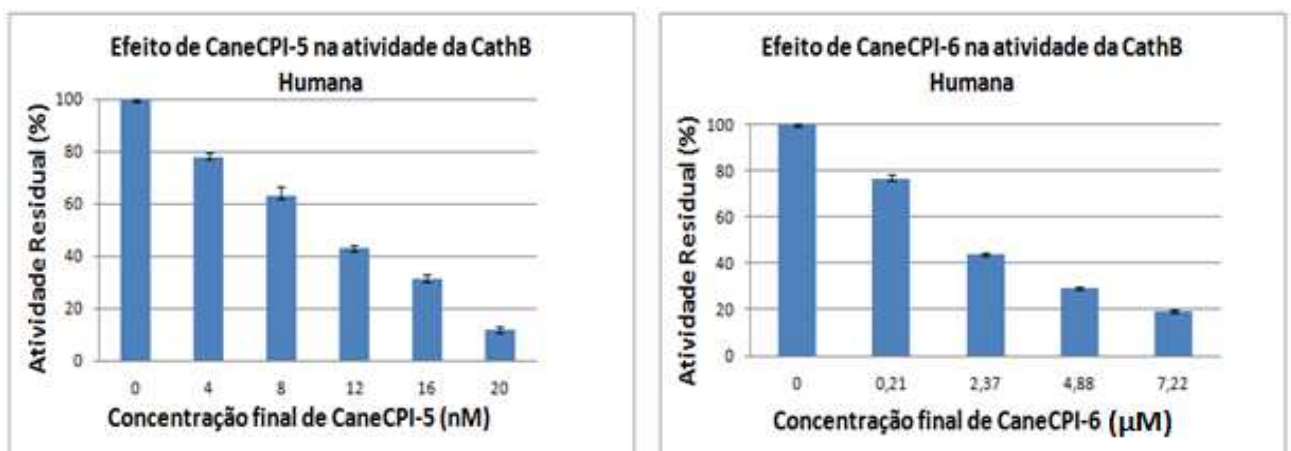


Figura 23: Inibição da catepsina B humana pelas cistatinas recombinantes. A atividade inibitória das cistatinas CaneCPI-5 e CaneCPI-6 contra a enzima catepsina B humana foi medida em ensaio fluorométrico utilizando o substrato Z-Phe-Arg-MCA. Os gráficos representam as curvas de inibição de cada inibidor, obtida pelo cálculo da atividade residual da catepsina B humana à medida que o inibidor foi adicionado. Os valores estão expressos como a média \pm o desvio padrão de três experimentos independentes.

O valor da constante de inibição (K_i), também denominado constante de dissociação do complexo enzima-inibidor, expressa a afinidade do inibidor pela enzima (CHENG e PRUSOFF, 1973). Para um inibidor ser considerado potente, sua afinidade pela enzima alvo deve ser tão alta, que o seu valor de K_i é menor ou igual a concentração total de enzima utilizada no ensaio de inibição (COPELAND, 2005). Considerando os valores de K_i encontrados para a inibição de catepsina B (Tabela 7), assim como a concentração de enzima usada nos ensaios (1,8 nM), podemos verificar que, dentre os dois inibidores testados, a cistatina CaneCPI-6 não foi capaz de inibir catepsina B de forma eficiente, porém CaneCPI-5 apresentou uma considerável atividade inibitória contra essa enzima.

Tabela 7: Constantes de Inibição (Ki) das cistatinas CaneCPI-5 e CaneCPI-6 contra a catepsina B humana (1,8 nM).

Inibidor	Ki
CaneCPI-5	6,87 nM
CaneCPI-6	1,83 μ M

Ao comparar as constantes de inibição (Ki), obtidas nesse estudo, com as de outras cistatinas de cana-de-açúcar (Tabela 8) contra a catepsina B humana, é verificado, que o poder inibitório de CaneCPI-5 é similar ao de CaneCPI-4, apresentando, a mesma ordem de grandeza, em nanomolar. A comparação entre as estruturas primárias de CaneCPI-4 e CaneCPI-5 (Figura 14), revelou que essas cistatinas tem alta identidade de aminoácidos, principalmente nas regiões envolvidas com a ligação do inibidor à enzima alvo, sugerindo que essas proteínas inibem catepsina B por um mecanismo de ação semelhante, o que reflete em valores comparáveis do Ki. Além disso, essas proteínas pertencem ao mesmo grupo. Em relação à proteína CaneCPI-6, é verificado uma atividade inibitória mais próxima de CaneCPI-1, 2 e 3, na ordem de micromolar. Os resultados das análises *in silico* (Figura 13), revelaram que a proteína CaneCPI-6 tem sua região C-terminal mais estendida quando comparadas às demais, com exceção de CaneCPI-3, que além de ser mais estendida na mesma região, possui um motivo SNSL importante para inibição de legumaína. Esse motivo conservado não foi identificado na estrutura primária de CaneCPI-6, portanto, apesar de ambas apresentarem Ki similares, contra a enzima catepsina B, e possuírem extensão C-terminal, não era esperado que CaneCPI-6 fosse capaz de inibir cisteíno-peptidases do tipo legumaína. Essa hipótese foi confirmada por análises experimentais, onde uma concentração de 0,7 mM de CaneCPI-6, não foi capaz de inibir a atividade de uma legumaína de cana-de-açúcar.

Tabela 8: Comparação entre as constantes de inibição (Ki) obtidas neste estudo com outras cistatinas de cana-de-açúcar contra catepsina B humana. As atividades de CaneCPI-4 e CaneCPI-5 estão destacadas por serem as mais eficientes.

Inibidor	Ki	[Catepsina B] nM	Referência
CaneCPI-1	0,125 μ M	0,5	OLIVA <i>et al.</i> , 2004
CaneCPI-2	0,47 μ M		
CaneCPI-3	2,44 μ M	0,5	GIANOTTI <i>et al.</i> , 2008
CaneCPI-4	0,83 nM		
CaneCPI-5	6,87 nM		
CaneCPI-6	1,83 μ M	1,8	Este estudo

A capacidade de inibir a atividade enzimática das catepsinas, além de outras cisteíno-peptidases, sugere que as cistatinas recombinantes estão em suas conformações corretas. Os resultados obtidos nesse trabalho revelam que apenas a proteína CaneCPI-5 foi capaz de inibir eficientemente a atividade da catepsina B humana, com um K_i similar ao de CaneCPI-4 (GIANOTTI *et al.*, 2008).

Quando comparada a outras cisteíno-peptidases, a catepsina B contém um elemento estrutural, denominado *loop* de oclusão (“*occluding loop*”), o qual bloqueia a fenda do sítio ativo (MUSIL *et al.*, 1991). Como as cistatinas inibem as catepsinas através de ligação competitiva ao sítio ativo, a conformação deste *loop* na catepsina B, é incompatível com a ligação das cistatinas. Essa influência do *loop* na acessibilidade ao sítio ativo é refletida em valores altos de K_i para a maioria das cistatinas, quando comparados àqueles obtidos contra cisteíno-peptidases que carecem de tal *loop* (BARRETT *et al.*, 1986; KONDO *et al.*, 1990; ABRAHAMSON, 1994; OHTSUBO *et al.*, 2005; MARTINEZ *et al.*, 2005b).

Estudos realizados por Abraham e colaboradores (2006), mostraram que as variações de aminoácidos das cistatinas de cevada, nos três motivos conservados envolvidos com a inibição da peptidase, afetam a interação enzima-inibidor (ABRAHAM *et al.*, 2006). A Figura 10 mostra que as estruturas primárias das duas cistatinas são bastante diferentes, inclusive nas regiões onde a proteína se liga ao sítio ativo da enzima para que haja inibição. Enquanto CaneCPI-5 possui um resíduo de serina (S) no motivo QVVS_G, a cistatina CaneCPI-6 possui uma treonina (T), QVVT_G, diferente não apenas de CaneCPI-5, mas também de todas as outras cistatinas de cana-de-açúcar já expressas e caracterizadas. Ademais, possui sua região C-terminal estendida. A diferença entre a capacidade das duas cistatinas em inibir catepsina B humana, reflete possíveis diferenças conformacionais, decorrentes das particularidades estruturais entre as duas proteínas.

A cistatina encontrada no ovo de galinha também é capaz de inibir eficientemente a catepsina B, com valor de K_i igual a 1,7 nM (BARRETT *et al.*, 1986). A inibição ocorre por um mecanismo de dois passos, o qual requer o deslocamento do *loop* de oclusão. No primeiro passo ocorre uma interação inicial da região N-terminal da cistatina com os subsítios S2 e S3 do sítio ativo da enzima. No passo seguinte, há o deslocamento do *loop* de oclusão conforme a cistatina ancorada introduz os seus dois *loops*, o central e o da região C-terminal, nos subsítios S' da enzima (NYCANDER *et al.*, 1998; PAVLOVA *et al.*, 2000). Como foi dito anteriormente, o

motivo conservado na superfamília cistatina, Q-X-V-X-G, é Q-V-V-S-G em CaneCPI-5, assim como na proteína CaneCPI-4 (GIANOTTI *et al.*, 2008), e na cistatina do ovo de galinha (BARRETT *et al.*, 1986). A substituição do resíduo de serina por alanina (S → A) no motivo Q-X-V-X-G da cistatina do ovo de galinha reduziu a inibição da catepsina B (AUERSWALD *et al.*, 1992), e uma variante similar da cistatina da cana-de-açúcar CaneCPI-4 (CaneCPI-4 S58A), construída por meio de mutação sítio-dirigida, apresentou uma redução de ~30% na atividade inibitória contra a catepsina B humana (GIANOTTI, 2008), indicando que o resíduo de serina, presente no motivo conservado, QVVSG, dessas cistatinas, seja importante no segundo passo da inibição dessa enzima.

Inibidores da enzima catepsina B são de grande interesse biotecnológico, uma vez que, estas enzimas estão envolvidas com o desenvolvimento de tumores e doenças neurodegenerativas (VALADARES *et al.*, 2013).

A atividade inibitória da proteína recombinante CaneCPI-5 foi também testada contra a cisteíno peptidase *rSL-CathL*, utilizando o substrato fluorogênico Z-Leu-Arg-MCA. O perfil de inibição para *rSL-CathL* é mostrado na figura 24. O poder inibitório de CaneCPI-5 foi semelhante aos de outras cistatinas de cana-de-açúcar testadas contra a mesma enzima, sob as mesmas condições (FONSECA *et al.*, 2012) (Tabela 9).

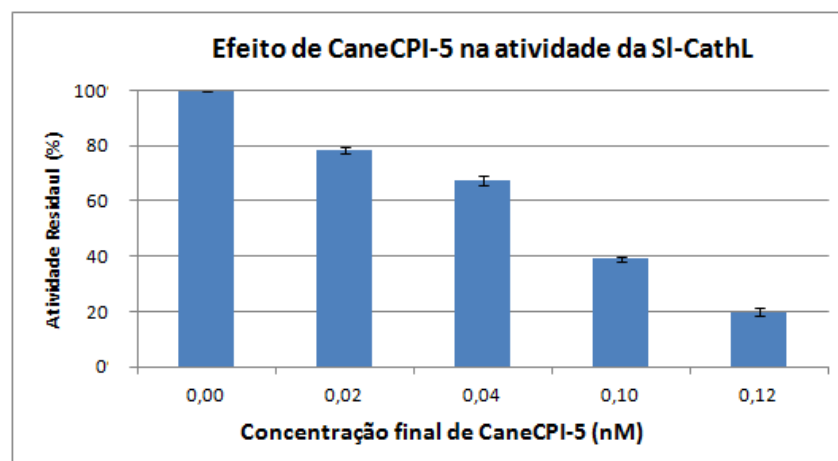


Figura 24: Inibição da catepsina L de *Sphenophorus levis* (*rSI-CathL*) pela cistatina recombinante CaneCPI-5. A atividade inibitória da cistatina CaneCPI-5 contra a enzima catepsina L de *S. levis* foi medida em ensaio fluorométrico utilizando o substrato Z-Leu-Arg-MCA. O gráfico representa a curva de inibição do inibidor, obtida pelo cálculo da atividade residual da catepsina L à medida que o inibidor foi adicionado. Os valores estão expressos como a média ± o desvio padrão de três experimentos independentes.

O valor da constante de inibição (Ki), calculado para a proteína CaneCPI-5 (0,059 nM), contra a catepsina L recombinante do inseto, foi semelhante ao encontrado para a inibição de catepsina L humana, pela cistatina CaneCPI-4 (0,021 nM) (GIANOTTI *et al.*, 2008).

Tabela 9: Comparação entre as constantes de inibição (Ki), obtidas neste estudo, com outras cistatinas de cana-de-açúcar, contra a enzima *rSL-CathL* (11 nM). A atividade de CaneCPI-5 está destacada por ser a mais eficiente.

Inibidor	Ki	Referência
CaneCPI-1	2,3 nM	FONSECA <i>et al.</i> , 2012
CaneCPI-2	0,9 nM	
CaneCPI-3	0,3 nM	
CaneCPI-4	0,2 nM	
CaneCPI-5	0,059 nM	Este estudo

A tabela 9 mostra que a proteína CaneCPI-5 é o inibidor mais potente contra a atividade da catepsina L recombinante, do inseto *Sphenophorus levis*. Esse inseto, conhecido como bicudo da cana, foi identificado em 1978 (VANIN, 1990), e desde então, se tornou uma das principais pragas da cana-de-açúcar no Brasil, com sua ocorrência aumentando rapidamente devido à expansão das plantações (FONSECA *et al.*, 2012). Em nosso laboratório (LBM/DGE - UFSCar) foram desenvolvidas canas-de-açúcar, geneticamente modificadas, super produzindo o inibidor CaneCPI-1 (RIBEIRO *et al.*, 2008). Experimentos recentes mostraram que as plantas desenvolvidas são mais resistentes ao ataque desse inseto.

Alguns fitopatógenos e insetos secretam cisteíno-peptidases para facilitar o processo de patogênese e degradação de proteínas da planta (HABIB e FAZILI, 2007). Estudos descrevem o potencial de inibidores de peptidases recombinantes como um potente componente antidigestivo para proteger plantações contra o ataque de patógenos e insetos (HAQ, *et al.*, 2004; SCHLÜTER *et al.*, 2010). As cistatinas têm sido mencionadas como agentes protetores contra vários artrópodes herbívoros e nematóides que parasitam as raízes (ARAI *et al.*, 2002; BENCHABANE *et al.*, 2010). Assim como a maioria das proteínas inibidoras de peptidases em plantas, as fitocistatinas atuam como um pseudo substrato para acessar o sítio ativo da peptidase. Após a inibição, a peptidase alvo não é mais capaz de clivar ligações peptídicas, o que resulta em uma inibição prejudicial das funções digestivas da proteína nos insetos herbívoros pragas, sendo assim, responsável pela resistência de

várias plantas transgênicas que superexpressão os inibidores recombinantes (MUNGER *et al*, 2012).

O besouro *S. levis* produz essencialmente catepsina L como sua principal enzima digestiva para degradação de proteínas no intestino (FONSECA *et al.*, 2012). Quando desafiado com o inibidor de cisteíno-peptidases, cistatina N da soja, as larvas do besouro *Callosobruchus maculatus* são capazes de reajustar a expressão de suas enzimas digestivas, e passam a produzir a catepsina B1 em taxas bem mais elevadas. A catepsina B é insensível ao inibidor, provavelmente devido à presença do *loop* de oclusão que impede o acesso do inibidor ao sítio ativo da enzima, portanto essa catepsina constitui uma enzima do sistema de contra-defesa do inseto a fatores anti-nutricionais (KOO *et al.*, 2008).

Neste trabalho, foi demonstrado que a proteína CaneCPI-5 é o inibidor mais eficiente, *in vitro*, contra a catepsina L recombinante do besouro *S. levis*, ademais é capaz de inibir eficientemente a catepsina B humana. Considerando esses resultados sugere-se que a produção de uma linhagem de cana-de-açúcar, geneticamente modificada, super produzindo este inibidor endógeno, poderia reduzir eficientemente os danos causados por este inseto.

4.6 ANÁLISES DE TERMOESTABILIDADE DAS CISTATINAS RECOMBINANTES

As atividades inibitórias das cistatinas recombinantes foram testadas quanto à resistência a altas temperaturas (Figura 25). Após serem submetidas à temperatura de 100°C durante 30 minutos, a cistatina CaneCPI-5 mostrou uma redução de aproximadamente 3% da sua atividade inibitória e a cistatina CaneCPI-6 de aproximadamente 6%. A característica das cistatinas de serem estáveis a altas temperaturas foi confirmada para as cistatinas da cana-de-açúcar recombinantes. O mesmo já tinha sido observado para a Orizacistatina, uma cistatina de arroz, cuja atividade inibitória permanece inalterada após 30 minutos a 100°C (ARAI *et al.*, 2002). Recentemente, estudos de termoestabilidade realizados com uma cistatina de seringueira, *HbCPI*, mostraram que sua atividade inibitória era inalterada até 90°C, e diminuía muito pouco quando a temperatura atingia 100°C (BANGRAK e CHOTIGEAT, 2011).

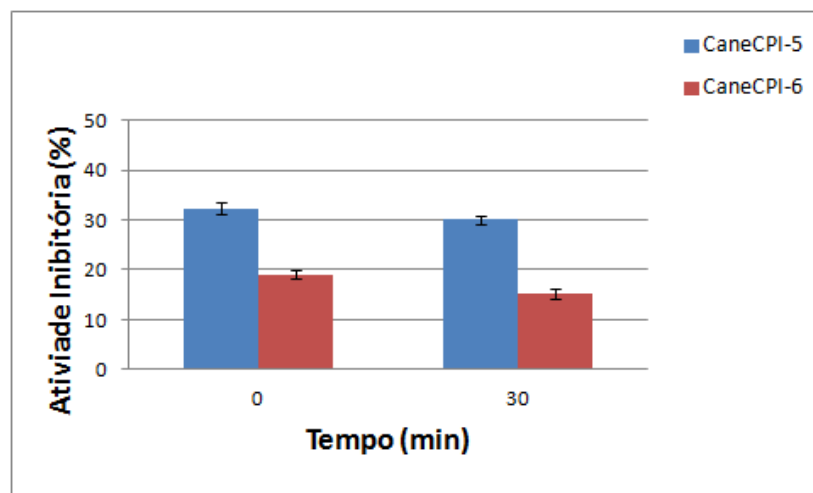


Figura 25: Estabilidade térmica das cistatinas recombinantes de cana-de-açúcar. Depois de serem submetidas à temperatura de 100°C durante 30 minutos, as cistatinas (4 nM) foram testadas em ensaios de inibição contra a enzima papaína (50 nM) utilizando um espectrofluorímetro. A atividade inibitória, expressa como porcentagem, foi obtida da diferença, entre a atividade residual da papaína na ausência das cistatinas (considerado 100%), e a atividade residual da enzima na presença da cistatina não submetidas à fervura (0 minuto) ou submetidas à fervura por 30 minutos. Cada medida foi realizada em triplicata. Os valores estão expressos como a média \pm o desvio padrão de três experimentos independentes.

5. CONCLUSÕES

De acordo com os resultados obtidos neste trabalho, conclui-se:

- As análises filogenéticas realizadas com as prováveis cistatinas de cana-de-açúcar, encontradas no banco de dados SUCEST, possibilitou uma nova classificação para estas proteínas;
- A expressão das proteínas recombinantes usando o sistema heterólogo *E. coli* Rosetta mostrou-se eficiente, possibilitando a obtenção das proteínas puras e em quantidades suficientes para os ensaios experimentais;
- As proteínas CaneCPI-5 e CaneCPI-6 pertencem aos grupos 3 e 4 respectivamente, portanto possuem consideráveis diferenças em sua estrutura primária. Esse fato pode ter refletido em diferentes capacidades de inibir catepsina B humana;
- A proteína CaneCPI-5 é a cistatina de cana-de-açúcar com melhor poder inibitório, *in vitro*, contra a catepsina L recombinante do inseto *S. levi*, dentre todas já testadas;
- A proteína CaneCPI-6 foi a primeira cistatina de cana-de-açúcar do grupo 4 produzida em sistema de expressão heteróloga;
- As duas novas cistatinas recombinantes de cana-de-açúcar, CaneCPI-5 e CaneCPI-6, mostram-se termoestáveis.

6. CONSIDERAÇÕES FINAIS E PERSPECTIVAS

As catepsinas participam em importantes processos fisiológicos, tais como, processamento e apresentação de antígenos, remodelamento ósseo, cicatrização, entre outros (MCGRATH, 1999; TURK *et al.*, 2000; DICKINSON, 2002). Elas também estão envolvidas em uma variedade de processos patológicos, incluindo osteoporose, artrite reumatóide, osteoartrite, doença de Alzheimer e câncer (TURK *et al.*, 2000; VASILJEVA *et al.*, 2007; GIANOTTI *et al.*, 2008; VALADARES *et al.*, 2010; OLIVEIRA *et al.*; 2011). Neste sentido, torna-se interessante testar a atividade inibitória das cistatinas da cana-de-açúcar contra outras catepsinas.

Além disso, alguns estudos têm mencionado importantes efeitos pleiotrópicos para cistatinas recombinantes expressas em plantas transgênicas, incluindo principalmente, fenótipos com tolerância ao ataque de artrópodes herbívoros e patógenos, e contra estresses bióticos e abióticos (MUNGER *et al.*, 2012; VAN DER LINDE, *et al.*, 2012; EASON *et al.*, 2014). Portanto, o uso de fitocistatinas para produção de plantas transgênicas mais resistentes a patógenos, insetos e estresses abióticos vem se tornando cada vez mais crescente (URWIN *et al.*, 2003; HAQ *et al.*, 2004; HABIB e FAZILI, 2007; ZHANG *et al.*, 2008). Por este ponto de vista, o estudo das fitocistatinas torna-se importante como uma ferramenta biotecnológica para o desenvolvimento de variedades melhor adaptadas contra o ataque de insetos e agentes patogênicos, e alguns tipos de estresses, como alta salinidade, seca, oxidação e frio.

7. REFERÊNCIAS BIBLIOGRÁFICAS

- ABRAHAM, Z.; *et al.* Structural and functional diversity within the cystatin gene family of *Hordeum vulgare*. **Journal of Experimental Botany**, v. 57, n. 15, p. 4245–4255, 2006.
- ABRAHAMSON, M. Cystatins—Protein inhibitors of papain-like cysteine proteinases. **Journal of the Brazilian Association of the Advancement of Science**, v. 45, n. 5, p. 299-304, 1993.
- ABRAHAMSON, M. Cystatins. **Methods Enzymology**, v. 244, p. 685-700, 1994.
- ABRAHAMSON, M.; *et al.* Identification of the Probable Inhibitory Reactive Sites of the Cysteine Proteinase Inhibitors Human Cystatin C and Chicken Cystatin. **The Journal of Biology Chemistry**, v. 262 n. 20, p. 9688-9694, 1987.
- ALBERTS, B.; *et al.* **Biologia Molecular da célula**. 5. ed. Porto Alegre: Artmed, 2010.
- ALVAREZ-FERNANDEZ; ABRAHAMSON, M. Structural and protease inhibitory features of cystatins. In: ZEROVNIK, E.; JERALA, N. K. (eds) **Human Stefins and Cystatins**. Nova York: Nova Science Publishers, Inc., 2006. p. 23-42.
- ARAI, S.; *et al.* Plant seed cystatins and their target enzymes of endogenous and exogenous origin. **Journal of Agricultural and Food Chemistry**, v. 50, n. 22, p. 6612 – 6617, 2002
- AUERSWALD, E. A.; *et al.* Recombinant chicken egg white cystatin variants of the QLVSG region. **European Journal of Biochemistry**, v. 209, n. 3, p. 837-845, 1992.
- BANGRAK, P.; CHOTIGEAT, W. Molecular cloning and biochemical characterization of a novel cystatin from *Hevea* rubber latex. **Plant Physiology and Biochemistry**, v. 49, n. 3, p. 244–250, 2011.
- BARRETT, A. J. The Cystatins, a new class of protease inhibitors. **Trends in Biochemical Sciences**, v. 12, p. 193-196, 1987.
- BARRETT, A. J.; RAWLINGS, N. D.; WOESSNER, J. F. **Handbook of Proteolytic Enzymes**. 2. ed. Elsevier Academic Press, 2004. 2368 p.
- BARRETT, A. J.; *et al.* Nomenclature and classification of the proteins homologous with the cysteine-proteinase inhibitor chicken cystatin. **Biochemical Journal**, v. 236, p. 312, 1986.
- BENCHABANE, M.; *et al.* Plant cystatins. **Biochimie**, Québec, v. 92, p.1657-1666, 2010.
- BEYNON, R. J.; BOND, J. S. **Proteolytic enzymes: a practical approach**. New York: Oxford University Press, 1989. 259 p.
- BODE, W.; *et al.* The 2.0 Å X ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. **The EMBO Journal**, v. 7, n. 8, p. 2593-2599, 1988
- BRADFORD, M. M. A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. **Analytical Biochemistry**, v. 72, p. 248-254, 1976.
- BUGG, T. **Introduction to Enzyme and Coenzyme Chemistry**. 2. ed. Oxford: Blackwell Publishing, 2004.
- CHAUBAL, R.; *et al.* Two male-sterile mutants of *Zea Mays* (Poaceae) with an extra cell division in the anther wall. **American Journal of Botany**, v. 87, n. 8, p. 1198 - 1201, 2000.
- CHENG, Y.; PRUSOFF, W. H. Relationship between the inhibition constant (K₁) and the concentration of inhibitor which causes 50 per cent inhibition (I₅₀) of an enzymatic reaction. **Biochemical Pharmacology**, v. 22, p. 3099-3108, 1973.

- CHRISTELLER, J.T.; *et al.* The squash aspartic proteinase inhibitor SQAPI is widely present in the Cucurbitales, comprises a small multigene family, and is a member of the phytocystatin family. **Journal of Molecular Evolution**, v. 63, n. 6, p. 747-757, 2006.
- CONESA, A.; *et al.* Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. **BIOINFORMATICS**, v. 21, n. 18, p. 3674–3676, 2005.
- COPELAND, R. A. **Evaluation of enzyme inhibitors in drug discovery: a guide for medicinal chemists and pharmacologists**. New Jersey: John Wiley & Sons, Inc., 2005.
- DANILES, J.; ROACH, B.T. Taxonomy and evolution, In: HEINZ, D.J. (Ed). **Sugarcane improvement through breeding**. Amsterdam: Elsevier. p. 7-84, 1987.
- DAYHOFF, M.O.; EYCK, R.V.; PARK, C.M. A model of evolutionary change in proteins. In: DAYHOFF M.O., **Atlas of protein sequence and structure**. Washington, DC: National Biomedical Research Foundation, 1972. p. 89–99.
- DICKINSON, D. P. Cysteine Peptidases of Mammals: their biological roles and potential effects in the oral cavity and other tissues in healthy and disease. **Critical Reviews in Oral Biology and Medicine**, v. 13, n. 3, p. 238-275, 2002.
- DOUST, A. N.; KELLOGG, E. A. Effect of genotype and environment on branching in weedy green millet (*Setaria viridis*) and domesticated foxtail millet (*Setaria italica*) (Poaceae). **Molecular Ecology**, v. 15, p. 1335–1349, 2006.
- EASON, J. R.; *et al.* Overexpression of the protease inhibitor BoCPI-1 in broccoli delays chlorophyll loss after harvest and causes down-regulation of cysteine protease gene expression. **Postharvest Biology and Technology**, v. 97, p. 23 - 31, 2014.
- FONSECA, F. P. P.; *et al.* Recombinant expression, localization and *in vitro* inhibition of midgut cysteine peptidase (SI-CathL) from sugarcane weevil, *Sphenophorus levis*. **Insect Biochemistry and Molecular Biology**, v. 42, p. 58-69, 2012.
- GERTZ, E. M.; *et al.* Composition-based statistics and translated nucleotide searches: Improving the TBLASTN module of BLAST. **BMC Biology**, v. 4, n. 41, 2006.
- GIANOTTI, A. **Produção recombinante e estudos funcionais de três novas cistatinas da cana-de-açúcar e sua utilização em estudos de inibição da adesão, proliferação, migração e invasão celular**. 2004. 119 f. Tese (Doutorado em Genética e Evolução) - Curso de Genética e Evolução, Departamento de Departamento de Genética e Evolução, Universidade Federal de São Carlos, São Carlos, 2008.
- GIANOTTI, A.; *et al.* Recombinant expression, purification, and functional analysis of two novel cystatins from sugarcane (*Saccharum officinarum*). **Protein Expr Purif**, v. 47, p. 483-489, 2006.
- GIANOTTI, A.; *et al.* Inhibitory effect of the sugarcane cystatin CaneCPI-4 on cathepsins B and L and human breast cancer cell invasion. **Biological Chemistry**. v. 389, p. 447 – 453, 2008.
- GOULET, M.-C.; *et al.* Tailoring the specificity of a plant cystatin toward herbivorous insect digestive cysteine proteases by single mutations at positively selected amino acid sites. **Plant Physiology**, v. 146, p. 1010 - 1019, 2008.
- HABIB, H.; FAZILI, K. M. Plant protease inhibitors: a defense strategy in plants. **Biotech Mol Biol Rev**, v. 2, n. 3, p. 68-85, 2007.
- HAQ, S. K.; ATIF, S. M.; KHAN, R. H. Protein proteinase inhibitor genes in combat against insects, pests, and pathogens: natural and engineered phytoprotection. **Arch Biochem Biophys**, v. 431, p. 145-159, 2004.
- HORTON, P.; *et al.* WoLF PSORT: protein localization predictor. **Nucleic Acids Research**, v. 35, p. 585 - 587, 2007.

- KIGGUNDU, A.; *et al.* Modulating the proteinase inhibitory profile of a plant cystatin by single mutations at positively selected amino acid sites. **The Plant Journal**, v. 48, n. 3, p. 403-413, 2006.
- KONDO, H.; *et al.* Gene organization of oryzacystatin-II, a new cystatin superfamily member of plant origin, is closely related to that of oryzacystatin-I but different from those of animal cystatins. **FEBS Lett**, v. 278, n. 1, p. 87-90, 1991.
- KONDO, H.; *et al.* Two distinct species in rice seeds with different specificities against cysteine proteinases. **Journal of Biological Chemistry**, v. 265, n. 26, p. 15832-15837, 1990.
- KOO, Y. D.; *et al.* Functional expression of an insect cathepsin B-like counter-defence protein. **Insect and Molecular Biology**, v. 17, p. 235-245, 2008.
- LAEMMLI, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. **Nature**, v. 227, n. 5259, p. 680-685, 1970.
- LEUNG-TOUNG, R.; Li, W.; Tam T. F.; Kaarimian, K. Thiol-dependent enzymes and their inhibitors: a review. **Current Medicinal Chemistry**, v. 9, p. 979-1002, 2002.
- MARGIS, R.; REIS, E. M.; VILLERET, V. Structural and Phylogenetic Relationships among Plant and Animal Cystatins. **Archives of Biochemistry and Biophysics**, v. 359, n. 1, p. 24-30, 1998.
- MARGIS-PINHEIRO, M.; *et al.* Molecular evolution and diversification of plant cysteine proteinase inhibitors: New insights after the poplar genome. **Molecular Phylogenetics and Evolution**, v. 49, p. 349-355, 2008.
- MARTINEZ, M.; *et al.* Comparative phylogenetic analysis of cystatin gene families from arabidopsis, rice and barley. **Molecular Genetics and Genomics**, v. 273, p. 423-432, 2005a.
- MARTINEZ, M.; *et al.* The strawberry gene *Cyf1* encodes a phytocystatin with antifungal properties. **Journal of Experimental Botany**, v. 56, n. 417, p. 1821-1829, 2005b.
- MARTINEZ, M.; *et al.* Carboxy terminal extended phytocystatins are bifunctional inhibitors of papain and legumain cysteine proteinases. **FEBS letters**, v. 58, p. 2914-2918, 2007.
- MARTINEZ, M.; DIAZ, I. The origin and evolution of plant cystatins and their target cysteine proteinases indicate a complex functional relationship. **BMC Evolutionary Biology**, v. 8, p. 198, 2008.
- MARTINEZ, M.; *et al.* Characterization of the Entire Cystatin gene family in barley and their target cathepsin L-Like Cysteine-Proteases, partners in the hordein mobilization during seed germination. **Plant Physiology**, v. 151, p. 1531-1545, 2009.
- McGRATH, M. E. The lysosomal cysteine proteases. **Annual Review of Biophysics & Biomolecular Structure**, v. 28, p. 181-204, 1999.
- MILNE, I.; *et al.* TOPALi: software for automatic identification of recombinant sequences within DNA multiple alignments. **Bioinformatics**, v. 20, n. 11, p. 1806-1807, 2004.
- MILNE, I.; *et al.* TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. **Bioinformatics**, v. 25, n. 1, p. 126-127, 2009.
- MUSIL, D.; *et al.* The refined 2.15 Å X-ray crystal structure of human liver cathepsin B: the structural basis for its specificity. **The EMBO Journal**, v. 10, n. 9, p. 2321-2330, 1991.
- MUNGER, A.; *et al.* Beneficial 'unintended effects' of a cereal cystatin in transgenic lines of potato, *Solanum tuberosum*. **BMC plant Biology**, v. 12, n. 198, p. 1 - 12, 2012.
- NAGATA, K.; *et al.* Three-dimensional solution structure of oryzacystatin-I, a cysteine proteinase inhibitor of the rice, *Oryza sativa* L. japonica. **Biochemistry**, v. 39, p. 14753-14760, 2000.

- NYCANDER, M.; *et al.* Two-step mechanism of inhibition of cathepsin B by cystatin C due to displacement of the proteinase occluding loop. **FEBS Letters**, v. 422, n. 1, p. 61-64, 1998.
- OLIVA, M. L. V.; *et al.* Inhibitory selectivity of canecystatin: a recombinant cysteine peptidase inhibitor from sugarcane. **Biochemical and Biophysical Research Communications**, v. 320, n. 4, p. 1082 - 1086, 2004.
- OLIVEIRA J. P.; *et al.* Sugarcane Cystatin CaneCPI-4 inhibits Melanoma Growth by Angiogenesis Disruption. **Journal of Cancer Science & Therapy**, v. 3, n. 7, p. 161 - 167, 2011.
- OTTO, H. H.; SCHIRMEISTER, T. Cysteine Proteases and Their Inhibitors. **Chemical reviews**. v. 97, p. 133-171, 1997.
- OHTSUBO, S.; *et al.* Molecular cloning and characterization of Oryzacystatin-III, a novel member of phytocystatins in rice (*Oryza sativa* L. japonica). **Journal of Agricultural and Food Chemistry**, v. 53, p. 5218-5224, 2005.
- PATERSON, A. H.; *et al.* The *Sorghum bicolor* genome and the diversification of grasses. **Nature**, v. 457, p. 551 - 556, 2009.
- PAULA, F. F. P. **Biologia molecular aplicada a identificação de alvos para o controle de *Sphenophorus levis*, o bicudo da cana-de-açúcar.** 2007. 184 f. Tese (Doutorado em Genética e Evolução) - Curso de Genética Evolutiva e Biologia Molecular, Departamento de Departamento de Genética e Evolução, Universidade Federal de São Carlos, São Carlos, 2012.
- PAVLOVA, A.; *et al.* Cystatin inhibition of cathepsin B requires dislocation of the proteinase occluding loop. Demonstration by release of loop anchoring through mutation of His110. **FEBS Letters**, v. 487, n. 2, p. 156-160, 2000.
- PAVLOVA, A.; BJORK, I. Grafting of Features of Cystatins C or B into the N-Terminal Region or Second Binding Loop of Cystatin A (Stefin A) Substantially Enhances Inhibition of Cysteine Proteinases. **Biochemistry**, v. 42, p. 11326-11333, 2003.
- PETERSEN, T. N.; *et al.* SignalP 4.0: discriminating signal peptides from transmembrane regions. **Nature Methods**, v. 8, p. 785-786, 2011.
- PERNAS, M.; *et al.* A chestnut seed cystatin differentially effects against cysteine proteinases from closely related pests. **Plant Molecular Biology**, v. 38, n. 6, p. 1235-1242, 1998.
- PERNAS, M.; *et al.* Der p 1 and Der f 1, the highly related and major allergens from house dust mites, are differentially affected by a plant cystatin. **Clinical & Experimental Allergy**, v. 30, p. 972-978, 2000.
- RAWLINGS, N. D.; BARRETT, A. J. Families of cysteine proteases. **Methods in Enzymology**, v. 244, p. 461-486, 1994.
- RAWLINGS, N. D.; MORTON, F. R.; BARRETT, A. J. An Introduction to Peptidases and the Merops Database. In: POLAINA, J.; MACCABE, A.P. (eds). **Industrial Enzymes**, p. 161-179, 2007.
- RAWLINGS, N.D.; WALLER, M.; BARRET, A.J.; BATEMAN, A. MEROPS: the peptidase database. **Nucleic Acids Research**, v. 42, p. D503-D509, 2014.
- REIS, E. M.; MARGIS, R. Sugarcane phytocystatins: Identification, classification, and expression pattern analysis. **Genetics and Molecular Biology**, v. 24, n. 1-4, p. 291-296, 2001.
- RIBEIRO, C. W.; *et al.* Production of a His-tagged Canecystatin in transgenic sugarcane and subsequent purification. **Biotechnology Progress**, v. 24, p. 1060 - 1066, 2008.
- SAJID, M.; MCKERROW, J. H. Cysteine proteases of parasitic organisms. **Molecular and Biochemical Parasitology**, v. 120, n. 1, p. 1-21, 2002.

- SAMBROOK, J.; RUSSEL, D. W. **Molecular Cloning: A laboratory Manual**. New York: Cold Spring Harbor Laboratory Press, 1989.
- SCHLÜTER, U.; *et al.* Recombinant protease inhibitors for herbivore pest control: a multitrophic perspective. **Journal Experimental Botany**, v. 61, p. 4169 – 4183, 2010.
- SOARES-COSTA, A.; *et al.* A sugarcane cystatin: recombinant expression, purification, and antifungal activity. **Biochemical and Biophysical Research Communications**, v. 296, n. 5, p. 1194-1199, 2002.
- TAMURA, K.; *et al.* MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. **Molecular Biology and Evolution**, v. 28, p. 2731–2739, 2011.
- TURK, V.; BODE, W. The cystatins: protein inhibitors of cysteine proteinases. **FEBS**, v. 285, n. 2, p. 213-219, 1991.
- TURK, B.; TURK, D.; TURK, V. Lysosomal cysteine proteases: more than scavengers. **Biochimica et Biophysica Acta**, v. 1477, n. 1-2, p. 98-111, 2000.
- URWIN, P. E.; GREEN, J.; ATKINSON, H. J.; Expression of a plant cystatin confers partial resistance to Globodera, full resistance is achieved by pyramiding with natural resistance. **Mol Breeding**, v. 12, p. 263-269, 2003.
- VALADARES, N. F.; *et al.* Molecular determinants of improved cathepsin B inhibition by new cystatins obtained by DNA shuffling. **BMC Structural Biology**, v. 10 p. 1472-6807, 2010.
- VALADARES, N. F.; *et al.* X-ray crystallography and NMR studies of domain-swapped canecystatin-1. **The FEBS Journal**, v. 4, p. 1028-1038, 2013.
- VALUEVA, T. A.; MOSOLOV, V. V. Role of inhibitors of proteolytic enzymes in plant defense against phytopat hogenic microorganisms. **Biochemistry**, v. 69, n. 11, p.1305-1309, 2004.
- VAN DER LINDE, K.; *et al.* The maize cystatin CC9 interacts with apoplastic cysteine proteases. **Plant Signaling & Behavior**, v. 7, n. 11, p. 1397–1401, 2012.
- VANIN, S. A. A new species of Sphenophorus schoenherr from Brazil (Coleoptera, Curculionidae, Rhynchophorinae). **Revista Brasileira de Entomologia**, v. 34, n. 4, p. 697-701, 1990.
- VASILJEVA, O.; *et al.* Emerging roles of cysteine cathepsins in disease and their potential as drug targets. **Current Pharmaceutical Design**, v. 13, p. 387- 403, 2007.
- VETTORE, A. L.; *et al.* The libraries that made SUCEST. **Genetics and Molecular Biology**, v. 24, p. 1 – 7, 2001.
- VETTORE, A. L.; *et al.* Analysis and functional annotation of an expressed sequence tag collection for tropical crop sugarcane. **Genome Research**, v. 13, n. 12, p. 2725-2735, 2003.
- ZHANG, X.; LIU, S.; TAKANO, T. Two cysteine proteinase inhibitors from Arabidopsis thaliana, AtCYSa and AtCYSb, increasing the salt, drought, oxidation and cold tolerance. **Plant Mol Biol.**, Tokyo, v. 68, p.131-143, 2008.
- ZHAO, Z.; *et al.* Distinguishing rice (*Oryza sativa* poaceae) from wild *Oryza* species through phytolith analysis, II Finalized method. **Economic Botany**, v. 54, n. 2, p. 134 - 145, 1998.

ANEXO I

TABELA DE IDENTIFICAÇÃO DOS CLONES DO SUCEST COM OS RESPECTIVOS
CÓDIGOS USADOS NA CONSTRUÇÃO DA ÁRVORE FILOGENÉTICA

SUCEST	Código	SUCEST	Código
SCJFST1015E11.b	SCC1	SCRUFL3064A04.g	SCC42
SCACCL6009E04.g	SCC5	SCVPHR1092F06.g	SCC43
SCEQRT3019F11.g	SCC6	SCMCST1054D03.g	SCC44
SCSBHR1053A07.g	SCC9	SCMCLR1122B04.g	SCC45
SCEQRT1025E03.g	SCC10	SCBGAM1093G07.g	SCC48
SCJLLB2077F03.g	SCC12	SCUTSD2086H12.g	SCC49
SCJLLR1108B02.g	SCC15	SCCCST1003G05.g	SCC51
SCCCRT3006B03.g	SCC16	SCBFRZ2049B11.g	SCC58
SCACHR1039F05.g	SCC17	SCVPFL4C11C01.g	SCC59
SCSFSD1064E05.g	SCC18	SCQGRZ3010H08.g	SCC60
SCRFSD2022E11.g	SCC19	SCJFRZ3C01H10.g	SCC61
SCBFSD2038H11.g	SCC23	SCCCFL4001H08.g	SCC62
SCRURT2007H08.g	SCC24	SCMCRT2087B02.g	SCC71
SCRLLV1026B04.g	SCC25	SCCCCL7002G06.g	SCC74
SCSBSD2032H05.g	SCC26	SCVPAM1059C07.g	SCC75
SCJLLB2077F03.b	SCC27	SCCCRZ3002G10.g	SCC77
SCSGFL1081B07.g	SCC28	SCAGCL6015G10.g	SCC78
SCQGFL1096E02.g	SCC29	SCCCLB2010E04.g	SCC79
SCPILB2021A09.g	SCC30	SCRUFL1111H10.g	SCC80
SCSBST3096B04.g	SCC31	SCCCLR1076E04.g	SCC81
SCEZST3150F10.g	SCC32	SCEPRZ1008H05.g	SCC82
SCMCRT2086B07.g	SCC33	SCAGSD1040G10.g	SCC83
SCRFLB2058G01.g	SCC34	SCEZRT2019D04.g	SCC85
SCMCRZ3067B01.g	SCC35	SCQGRT1041E01.g	SCC88
SCCCFL5065E04.g	SCC36	SCACRZ3036D05.g	SCC89
SCSGFL4036B11.g	SCC37	SCACRZ3036D08.g	SCC90
SCBGHR1061D02.g	SCC38	SCEZRZ3094E05.g	SCC91
SCAGFL8040C03.g	SCC39	SCSGST3121C03.g	SCC92
SCUTRZ3070E04.g	SCC40	SCEZRT3067C11.g	SCC93
SCRFST3145F09.g	SCC41		