**Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”**

**Caracterização do proteoma nuclear de folhas de cana-de-açúcar
(*Saccharum spp*) de 1 e 4 meses de idade**

Danielle Izilda Rodrigues da Silva

Dissertação apresentada para obtenção do título de
mestre em Ciências. Área de concentração: Genética
e Melhoramento de Plantas

**Piracicaba
2012**

Danielle Izilda Rodrigues da Silva
Bióloga

**Caracterização do proteoma nuclear de folhas de cana-de-açúcar (*Saccharum*
spp) de 1 e 4 meses de idade**

Orientador:
Prof. Dr. **Carlos Alberto Labate**

Dissertação apresentada para obtenção do título de
mestre em Ciências. Área de concentração: Genética
e Melhoramento de Plantas

**Piracicaba
2012**

**Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA - ESALQ/USP**

Silva, Danielle Izilda Rodrigues da

Caracterização do proteoma nuclear de folhas de cana-de-açúcar (*Saccharum spp*)
de 1 e 4 meses de idade / Danielle Izilda Rodrigues da Silva.- - Piracicaba, 2012.
170 p: il.

Dissertação (Mestrado) - - Escola Superior de Agricultura "Luiz de Queiroz", 2012.

1. Cana-de-açúcar 2. Folhas - Plantas 3. Programas de predição 4. Proteínas
nucleares I. Título

CDD 633.61
S586c

"Permitida a cópia total ou parcial deste documento, desde que citada a fonte – O autor"

Ao meu pai Wilson que sempre apoiou meus sonhos e nunca mediu esforços para ajudar a realizá-los.

À minha mãe Rosana que esteve ao meu lado em todos os momentos e me ensinou a ser a pessoa que sou hoje.

Aos meus avós Mário, Janet, Noelza e João (*In memoriam*) pelo amor incondicional.

À minha irmã Jéssica por me ajudar a lembrar que os problemas não são tão grandes quanto parecem.

Aos meus animais de estimação, Ponyta, Loba, Tico e Teco que me proporcionaram momentos de alegria, mesmo quando tudo parecia perdido.

Aos meus tios Luiz Carlos, Márcia, Sidney e Nelson pelos momentos de descontração.

DEDICO

AGRADECIMENTOS

À Deus que iluminou meu caminho durante toda a minha vida e em especial durante a execução deste trabalho.

Ao Prof. Dr. Carlos Alberto Labate por acreditar em minha capacidade de realizar esse trabalho, me dar a oportunidade de executá-lo e pelos sábios conselhos.

À FAPESP pelo apoio financeiro representado pelo fornecimento da bolsa de mestrado.

Aos professores e funcionários do Programa de Pós-graduação em Genética e Melhoramento de Plantas.

À minha família que esteve ao meu lado em todos os momentos, seja me mantendo calma nos momentos de aflição, me ajudando a terminar as minhas análises ou simplesmente me proporcionando momentos de alegria e companheirismo.

Às minhas amigas Ilara Budzinski, Simone Tanaka, Hana Silva e Júlia Morosini que me ajudaram não somente na bancada e na elaboração das minhas análises, mas também ao me proporcionar momentos inesquecíveis, recheados de risadas, desabafos e confidências. Jamais esquecerei tudo que passamos juntas, nem a ajuda e apoio que recebi de vocês, que vai muito além do convívio profissional.

Ao Dr. David Moon cuja contribuição para o meu trabalho foi essencial. Seu conhecimento e experiência fizeram toda a diferença, e com certeza contribuíram de forma definitiva com minha formação como pesquisadora.

Ao meu amigo Leonardo Ferreira pela sua paciência e ajuda, pelas correções, piadas, companhia no almoço e por me ensinar que no final as coisas sempre acabam dando certo, por mais tortuoso que seja o caminho.

Ao amigo Luiz Felipe Boaretto por ceder as plantas para que eu começasse meu trabalho, pelos almoços e conversas.

À Simone Guidetti e à Juliana Calderán pela amizade, cumplicidade e pela ajuda na resolução de muitos dos problemas que encontrei.

Aos amigos Ivan Mozol e Sônia Andrade pelas caronas e pelas conversas e ao amigo Fabrício de Moraes pela ajuda com o computador.

À Livia Franceschini que sempre me ajudou nos problemas “bioinformáticos”, que não foram poucos e facilitou em muito o meu trabalho.

À Fernanda Salvato que me deu lições valiosíssimas e guiou o meu trabalho no começo, quando eu ainda não tinha experiência e não sabia qual rumo tomar para chegar a um resultado coerente e confiável.

Ao Fábio C. A. Dossi pela ajuda nas análises em microscopia de fluorescência.

Ao Núcleo de Apoio à Pesquisa em Microscopia Eletrônica Aplicada à Agricultura (NAP/MEPA) da ESALQ pela ajuda nas análises de microscopia de luz.

Aos meus amigos do Laboratório Max Feffer de Genética de Plantas pelo convívio harmonioso em todos os momentos, ajuda nos momentos tensos e risadas nos momentos de descontração.

Ao finalizar, gostaria de agradecer a todos que contribuíram com meu trabalho, mesmo que de forma pequena, pois vocês fizeram a diferença.

MUITO OBRIGADO!

“Que os vossos esforços desafiem as
impossibilidades, lembrai-vos de que as
grandes coisas do homem foram
conquistadas do que parecia
impossível”.

Charles Chaplin

SUMÁRIO

LISTA DE FIGURAS	11
LISTA DE TABELAS	13
LISTA DE SIGLAS.....	15
RESUMO.....	19
ABSTRACT	21
1 INTRODUÇÃO	23
2 OBJETIVOS	31
3 REVISÃO BIBLIOGRÁFICA.....	33
3.1 Cana-de-açúcar e importância econômica.....	33
3.2 Proteômica e Subproteômica	37
3.2.1 Proteômica Vegetal	37
3.2.2 A subproteômica e sua importância	40
3.3 O núcleo e suas proteínas.....	49
3.3.1 Histonas	57
3.3.2 Helicases.....	58
3.3.3 Proteínas de ligação MAR.....	59
3.4 Programas de predição subcelular.....	59
3.4.1 PSORT	61
3.4.2 Plant-mPloc.....	61
3.4.3 WoLF PSORT	61
3.4.4 BaCelLo.....	62
3.4.5 SherLoc2	63
4 MATERIAL E MÉTODOS	65
4.1 Material Vegetal	65
4.2 Isolamento de núcleos.....	65

4.2.1 Tratamento e homogeneização do tecido.....	65
4.2.2 Extração com gradiente de <i>percoll</i>	66
4.3 Verificação da integridade e pureza da fração nuclear isolada.....	68
4.3.1 Microscopia de luz	68
4.3.2 Microscopia de fluorescência.....	68
4.4 Extração, solubilização e quantificação de proteínas	69
4.5 Análise de <i>Western blot</i>	69
4.6 SDS-PAGE	71
4.7 Digestão <i>In gel</i> de proteínas	72
4.8 Identificação de proteínas por espectrometria de massas.....	73
4.9 Análise e processamento dos dados	74
5 RESULTADOS E DISCUSSÃO	77
5.1 Integridade dos núcleos enriquecidos.....	77
5.2 Avaliação da pureza das frações nucleares	80
5.3 Análise das proteínas identificadas na extração teste	82
5.4 Análise das proteínas identificadas nos núcleos isolados de folhas de 1 e 4 meses de idade	87
5.5 Proteínas nucleares características não encontradas nas três repetições	95
5.6 Diferenças entre as proteínas encontradas em folhas e 1 e 4 meses	102
5.7 Aplicação dos programas de predição.....	104
6 CONSIDERAÇÕES FINAIS.....	109
REFERÊNCIAS	111
ANEXO.....	127

LISTA DE FIGURAS

Figura 1 – Áreas de concentração da produção da cana-de-açúcar.....	35
Figura 2 – Fases do desenvolvimento da cana-de-açúcar.....	37
Figura 3 – (a) O núcleo e seus diversos compartimentos; (b) Diferenças entre o nucléolo de vegetais e de mamíferos.....	56
Figura 4 – Fluxograma do processo de enriquecimento e identificação de proteínas nucleares.....	74
Figura 5 – Núcleos de folhas de cana-de-açúcar de 1 mês de idade corados com Orceína Acetoláctica 1%.....	78
Figura 6 – Núcleos de folhas de cana-de-açúcar de 4 meses de idade corados com Orceína Acetoláctica 1%.....	79
Figura 7 – Núcleos de folhas de cana-de-açúcar de 1 mês de idade fotografados com, à esquerda, e sem, à direita, utilização do filtro UV.....	79
Figura 8 – Núcleos de folhas de cana-de-açúcar de 4 meses de idade fotografados com, à esquerda, e sem, à direita, utilização do filtro UV	80
Figura 9 – (a) Análise de <i>Western blot</i> usando os anticorpos da Histona H1.4, PCNA, OEE1, Rubisco, PEPCase. (b) SDS-PAGE contendo o marcador, o extrato protéico total e o extrato protéico nuclear.	81
Figura 10 – Classificação funcional das proteínas obtidas por meio da extração teste.....	86
Figura 11 – 1D SDS – PAGE das proteínas obtidas em folhas de plantas de 1 mês de idade.....	95
Figura 12 – 1D SDS – PAGE das proteínas obtidas em folhas de plantas de 4 meses de idade.....	95

Figura 13 – Classificação funcional das proteínas preferencialmente expressas em núcleos de folhas de cana-de-açúcar de 1 mês de idade.....	103
Figura 14 – Classificação funcional das proteínas preferencialmente expressas em núcleos de folhas de cana-de-açúcar de 1 mês de idade.....	104
Figura 15 – Porcentagem de proteínas encontradas em folhas de plantas de 1 mês que foram preditas nucleares por cada programa de predição.....	106
Figura 16 – Número de proteínas preditas como nucleares por 1, 2, 3, 4, ou 5 programas e seus respectivos programas. Resultados obtidos em folhas de plantas de 1 mês de idade.....	106
Figura 17 – Porcentagem de proteínas encontradas em folhas de plantas de 4 meses que foram preditas nucleares por cada programa de predição.....	107
Figura 18 – Número de proteínas preditas como nucleares por 1, 2, 3, 4, ou 5 programas e seus respectivos programas, em folhas de plantas de 4 meses de idade.....	107
Figura 19 – Classificação funcional das proteínas encontradas em núcleos de folhas de cana-de-açúcar de 1 meses de idade	129
Figura 20 – Classificação funcional das proteínas encontradas em núcleos de folhas de cana-de-açúcar de 4 meses de idade.....	130

LISTA DE TABELAS

Tabela 1 – Distribuição das plantas em relação às suas respectivas repetições.....	65
Tabela 2 – Concentração de anticorpos.....	71
Tabela 3 – Média, desvio padrão e variância do diâmetro nuclear de folhas de plantas de cana-de-açúcar de 1 e 4 meses de idade.....	78
Tabela 4 – Lista de proteínas nucleares características não encontradas nas três repetições, sua localização subcelular de acordo com os programa de predição e sua classificação funcional.....	99
Tabela 5 – Proteínas identificadas na extração teste.....	131
Tabela 6 – Proteínas nucleares identificadas em folhas de 1 mês de idade.....	138
Tabela 7 – Proteínas nucleares identificadas em folhas de cana-de-açúcar de 4 meses de idade.....	151
Tabela 8 – Proteínas nucleares encontradas apenas em folhas de cana-de-açúcar de 1 mês.....	160
Tabela 9 – Proteínas nucleares encontradas apenas em folhas de cana-de-açúcar de 4 meses.....	168

LISTA DE SIGLAS

ACN – Acetonitrila

ADP – Adenosina Difosfato

AGO – Proteínas Argonautas

AKIP-1 – A-kinase-interacting protein 1

AMBIC – Bicarbonato de Amônio

APS – Persulfato de Amônio

ATP – Adenosina Trifosfato

B – BaCelLo

bHLH – basic helix-loop-helix

Btf3 – basic transcription factor

CB – Corpos de Cajal

Cdc2 – Proteína quinase

cDNA – DNA complementar

CF – Centros Fibrilares

cm – centímetro

CRY2 – Criptocromo 2

CTC – Centro de Tecnologia Canavieira

DAPI – 4', 6-diamidino-2-phenylindole

DFCs – Componente Fibrilar Denso

DIC – Delineamento Inteiramente Casualizado

DNA – Ácido Desoxirribonucléico

DTT – DL-Dithiothreitol

EST – Expressed Sequence Tag

FD – fator de transcrição

FT – locus florigeno T

FTSH – metaloprotease

GAPDH – gliceraldeído-3-fosfato desidrogenase

genes rDNA – genes ribossomais

GFP – human (Glu1)-Fibrino-Peptide

GST – glutationa-S-transferase

HCL – Ácido Clorídrico

Hd3a – Hormônio Florigeno

HSP – Heat shock protein

HYL1 – Domínio de ligação ao RNA

IAA – Iodoacetamida

IBGE – Instituto Brasileiro de Geografia e Estatística

ICAT – Isotope-coded affinity tag

KOH – Hidróxido de Potássio

m – metro

MAPKinase 6 – proteína quinase ativada por mitógeno 6

MAPKinase 7 – proteína quinase ativada por mitógeno 7

MAR – Região de ligação à matriz

MFP1 – Proteína filamentosas de ligação MAR

MgCL₂ – Cloreto de Magnésio

mRNAs – RNAs mensageiros

MSN2 – Ativador transcricional

MSN4 – Ativador transcricional

MudPIT – Multidimensional Protein Identification Technology

NaCl – Cloreto de Sódio

NTP – Nucleosídeo Trifosfato

Nups – Nucleoporinas

OEE1 – Proteína 1 do complexo de evolução do oxigênio

OEE2 – Proteína 2 do complexo de evolução do oxigênio

P – PSORT

PBS – Tampão Fosfato-Salino

PCNA – Antígeno nuclear de proliferação celular

PEPCase – Fosfoenol piruvato carboxilase

pH – Potencial hidrogênio iônico

PIPES – 1,4-Piperazinediethanesulfonic acid

PI – Plant-mPloc

PLGS – ProteinLynxGlobalServer

Proteína AS2 – Proteína ASYMMETRIC-LEAVES2

Pré-mRNA – pré RNA mensageiro

Pré-rRNA – pré RNA ribossomal

Proteínas SR – Proteínas ricas em Serina e Arginina

PsbA – Proteína D1 do centro de reação do fotossistema II

PVDF membrane – membrane de fluoreto de polivinilideno

Py – Polipirimidinas

REP 1 – Regulador transcricional de sinalização reduzida do fitocromo

RER – Retículo Endoplasmático Rugoso

RNA – Ácido Ribonucléico

RNAi – RNA de interferência

RoIB – Proteína envolvida em indução de raízes

RONs – Regiões organizadoras de nucléolo

RRM – Sequência de reconhecimento do RNA

RuBisCO – Ribulose-1,5-bifosfato carboxilase oxigenase

S – SherLoc2

SDS – Dodecil sulfato de sódio

SDS – PAGE – Sodium dodecyl sulfate polyacrylamide gel electrophoresis

siRNAs – RNAs de interferência pequenos

snoRNPs – ribonucleoproteínas pequenas nucleolares

SUCEST - Sugarcane Expressed Sequence Tag

SUMO – Pequeno modificador relacionado à ubiquitina

SVM – Support Vector Machines

TAF – Fatores associados à TBP

TBP – Proteína de ligação TATA

TBS-T – Tampão Salino de Tris com Tween 20

TFIID – Fator de transcrição IID

Tris – tris(hidroximetil)aminometano

UNICAMP – Universidade de Campinas

UPLC – Cromatografia Líquida de Alta Pressão

UV – Ultravioleta

W – WoLF PSORT

RESUMO

Caracterização do proteoma nuclear de folhas de cana-de-açúcar (*Saccharum* spp) de 1 e 4 meses de idade

A cana-de-açúcar é uma cultura economicamente importante, cultivada especialmente pelo seu colmo, que constitui a matéria-prima para produtos como o açúcar e o bioetanol. Ademais, a compreensão do proteoma nuclear é essencial para decifrar os mecanismos que governam a regulação gênica. No presente estudo, é demonstrado o isolamento e a identificação através de 1D SDS-PAGE de proteínas nucleares originadas de folhas jovens de plantas de cana-de-açúcar. Os núcleos foram isolados de folhas F+1 frescas de cana-de-açúcar de 1 e 4 meses, usando o protocolo modificado de Folta e Kaufman (2000). O experimento consistiu em um delineamento inteiramente casualizado, com três repetições de 18 plantas cada. Após a purificação usando o gradiente de *percoll*, a integridade do núcleo foi avaliada por meio da coloração com orceína acetolática 1% e com DAPI. Os resultados obtidos revelam os núcleos como esferas uniformes com o diâmetro médio de 5 μ m. As proteínas nucleares foram isoladas usando o reagente TRI Reagent (Sigma) e quantificadas por meio do método de Bradford. As análises de Western blot foram usadas para demonstrar o enriquecimento de proteínas nucleares. As membranas foram incubadas com a RUBISCO, PEPCase, OEE1, Histona e PCNA. A presença da PCNA e da Histona foram detectadas apenas no extrato de proteínas nucleares, já a RUBISCO, a PEPCase e a OEE1 foram detectadas de forma abundante no extrato de proteínas total e reduzida na fração nuclear. Para a caracterização do proteoma nuclear, 60 μ g de proteínas foram separadas por SDS-PAGE e cada canaleta dividida em 20 bandas. As proteínas de cada banda foram digeridas e purificadas. A identificação foi realizada por meio de espectrometria de massas (Synapt G2 HDMS) e analisadas usando o ProteinLynx e o banco de dados SUCEST. Programas como BaCellLo, WoLF PSORT, Plant-mPloc, SherLoc e PSORT foram usados para a predição da localização subcelular das proteínas identificadas. A classe de proteínas identificadas mais abundante se relaciona à montagem de nucleossomos, e é representada principalmente pelas histonas, como H2A.2, H2A.8, H3.3, H2B.1, H2B.2, dentre outras. Ademais, ainda foram encontradas classes menos abundantes relacionadas ao metabolismo do DNA, do RNA, regulação da transcrição, dentre outras. Alguns fatores de transcrição e outras proteínas nucleares típicas também foram identificadas, porém, possivelmente em decorrência de sua baixa abundância, não foram observados em todas as repetições. Os resultados encontrados mostram a aplicabilidade da metodologia para criar um perfil preciso do proteoma nuclear de cana-de-açúcar.

Palavras-chave: Cana-de-açúcar; Núcleo; Subproteoma; Programas de predição

ABSTRACT

Nuclear proteome characterization of one and four-month-old sugarcane (*Saccharum* spp) leaves

Sugarcane is a cash crop, cultivated for its stalks which accumulate sucrose, the raw material for products like sugar and bioethanol. Nuclear proteome comprehension is essential for deciphering the mechanisms that governs genome regulation and function. In the present study, we report the isolation and identification by 1D SDS-PAGE of nuclear proteins from young sugarcane leaves. The nuclei were isolated from fresh tissue of one and four-month-old sugarcane F+1 leaves, using the modified protocol of Folta and Kaufman (2000). The experiment consisted on a completely randomized design, three biological repetitions each with 18 plants. After purification using a *percoll* gradient, nucleus integrity was evaluated by staining with 1% acetolactic orcein and with DAPI. The results obtained reveal nuclei as uniform spheres with an average diameter of 5 μ m. The nuclear proteins were isolated using TRI Reagent (Sigma) and quantified by Bradford. Western blot analysis were used to prove enrichment for nuclear proteins. Membranes were incubated with RUBISCO, PEPCase, OEE1, Histone and PCNA. The presence of PCNA and Histone were detected only in the nuclear fraction. RUBISCO, PEPCase and OEE1 were very abundant in the total protein fraction and reduced in the nuclear fraction. For the characterization of nuclear proteome, 60 μ g of proteins were separated by SDS-PAGE and each lane divided into 20 sections, the proteins from each section were digested and purified. Protein identification was carried out by mass spectrometry (Synapt G2 HDMS) and analyzed using ProteinLynx and SUCEST database. Softwares, such as BaCellLo, WoLF PSORT, Plant-mPloc, SherLoc and PSORT were also used to predict the subcellular localization of the identified proteins. The most abundant protein class is related to the nucleosome assembly. It is represented specially by histones like H2A.2, H2A.8, H3.3, H2B.1, H2B.2, among others. Besides, less abundant classes like the ones related to DNA and RNA metabolism, regulation of transcription and others were also found. Some transcription factors and other typical nuclear proteins were identified as well, but, possibly due to their low abundance, they were not observed in all three repetitions. These results show the applicability of this method to create an accurate sugarcane nuclear proteome profile.

Key words: Sugarcane; Nucleus; Subproteome; Prediction programs

1 INTRODUÇÃO

A cana-de-açúcar (*Saccharum spp*) é uma cultura economicamente importante, cultivada pela produção da sacarose no colmo. Essa cultura contribui para aproximadamente 60% do açúcar branco produzido no mundo inteiro, sendo que sua biomassa também é usada para a produção de bioetanol e geração de energia (ALMARAJ et al., 2010). A alta produtividade e conteúdo de açúcar da cultura transformou-a em uma fonte altamente competitiva de sacarose para alimentação e produção de etanol, tornando a cana uma das culturas mais importantes econômica- e socialmente em países tropicais e subtropicais (CALSA JR., FIGUEIRA, 2007), como o Brasil.

De acordo com o primeiro levantamento da safra 2012/2013, a área cultivada de cana-de-açúcar continua em expansão no Brasil, sendo que as áreas em produção continuam com aumento progressivo nos Estados de Mato Grosso do Sul (12,5%), Goiás (7,9%), Espírito Santo (7,35%), Bahia (5,3%), Mato Grosso (5,5%) e Minas Gerais (3,5%). Até o momento do levantamento, não houve registro por parte dos informantes de nenhuma unidade que iniciará suas atividades na safra 2012/13 (Companhia Nacional de Abastecimento – Conab, 2012). Além disso, a estimativa da produção agrícola de maio de 2012, em nível nacional, apresentou variação positiva de 4,2% em relação à produção obtida em 2011 (Instituto Brasileiro de Geografia e Estatística – IBGE, 2012).

A área cultivada de cana-de-açúcar que será colhida e destinada à atividade sucroalcooleira na safra 2012/13 está estimada em 8.567,2 mil hectares, distribuídos em todos os estados produtores, conforme suas características. São Paulo é o maior produtor, com 51,66% (4.426, 45 mil hectares), seguido por Minas Gerais, com 8,97% (768,64 mil hectares), Goiás, com 8,54% (732,02 mil hectares), Paraná, com 7,17% (614,01 mil hectares), Mato Grosso do Sul, com 6,31% (540,97 mil hectares), Alagoas, com 5,35% (458,09 mil hectares) e Pernambuco, com 3,48% (298,39 mil hectares). Nos demais estados produtores, as áreas são menores, com representações abaixo de 3% (Companhia Nacional de Abastecimento – Conab, 2012).

A produtividade média brasileira está estimada em 70.289 kg/ha, 2,9% maior que na safra 2011/12, que foi de 68.289 kg/ha (Companhia Nacional de

Abastecimento – Conab, 2012). Esse aspecto só reforça o fato de o setor sucroalcooleiro nacional ser referência para os demais países produtores, uma vez que a cana-de-açúcar é produzida em quase todo o país, sendo aproximadamente 60% da produção pertencente a São Paulo. As demais zonas produtoras são Paraná, Triângulo Mineiro e Zona da Mata Nordestina. Além disso, o Brasil domina o ciclo completo da produção de etanol, desde a lavoura de alta produtividade até a instalação dos equipamentos para as destilarias, que estão gerando esse biocombustível a partir da fermentação do caldo extraído da cana-de-açúcar (Ministério da Agricultura Pecuária e Abastecimento – MAPA , 2012).

A importância econômica da cana-de-açúcar no Brasil fez com que essa cultura se tornasse atrativa para as pesquisas, e em setembro de 1998 foi iniciado o projeto SUCEST (Sugarcane Expressed Sequence Tag), que objetivou o sequenciamento aleatório de ESTs e identificou aproximadamente 50.000 genes (VETORE et al., 2001). A diferença da análise do genoma para a do proteoma não é apenas uma questão da complexidade numérica em termos de variantes dos produtos gênicos que podem surgir a partir de um único gene, mas também pelo fato de que um nível de complexidade suplementar é introduzido pela organização supramolecular dos produtos gênicos, decorrentes de interações proteína-proteína ou do direcionamento de proteínas para estruturas subcelulares específicas (DREGER, 2003).

Qualquer discussão a respeito da proteômica em plantas deve ser iniciada considerando-se dois fatos. Primeiramente, os genomas vegetais apresentam dezenas de milhares de genes capazes de expressar um grande número de proteínas, decorrentes, por exemplo, do *splicing* alternativo. Em segundo lugar, as plantas são sistemas vivos extremamente complexos, formados por uma série de órgãos interdependentes, como folha, raiz, caule, flor e semente, além de uma série de tecidos, células e organelas altamente especializadas (HAYNES; ROBERTS, 2007).

A proteômica é o estudo sistemático de muitas e diversas propriedades das proteínas, fornecendo, paralelamente, descrições detalhadas da estrutura, função e controle de sistemas biológicos. Os avanços nos métodos e nas tecnologias têm permitido a expansão da visão dos estudos biológicos, que passaram da análise

bioquímica reducionista de proteínas individuais para estudos proteômicos de larga escala. Dessa forma, a combinação da proteômica e de outros métodos de análise complementares, como a transcriptômica e metabolômica, são componentes essenciais para a emergente abordagem da “Biologia de Sistemas”. Essa área da ciência procura descrever os sistemas biológicos de forma compreensível por meio da integração de diversos tipos de dados para, no futuro, permitir simulações computacionais de sistemas biológicos complexos (PATTERSON; AEBERSOLD, 2003).

A elucidação dos mecanismos moleculares nos quais se baseiam as funções celulares exige o conhecimento global das proteínas expressas, incluindo os produtos de *splicing* alternativo, suas modificações pós-traducionais, localizações subcelulares e sua organização em máquinas moleculares em decorrência das interações proteína-proteína, em um determinado momento durante o ciclo-de-vida de uma célula ou sob determinadas condições celulares. O avanço nas técnicas de espectrometria de massas e bioinformática podem ajudar na realização desses objetivos (BRUNET et al., 2003).

Apesar desse aspecto, a abordagem proteômica clássica pode falhar na descoberta de novos produtos gênicos, que podem ser proteínas de grande abundância em um determinado compartimento subcelular, mas que, no entanto, constituem proteínas de baixa abundância no homogenato celular total (DREGER, 2003b).

O direcionamento de uma proteína para um local subcelular particular é um princípio importante da organização funcional da célula no nível molecular (DREGER, 2003b). Essa localização proteica está, portanto, ligada à função celular e representa um desafio analítico para a pesquisa proteômica, exigindo uma análise proteômica com resolução subcelular (DREGER, 2003). Dessa forma, as organelas subcelulares representam alvos atrativos para a análise proteômica, uma vez que constituem unidades funcionais discretas, apresentando uma composição protéica com reduzida complexidade quando comparadas a células inteiras, de forma que quando proteínas abundantes, como as de citoesqueleto, são removidas, as proteínas de menor abundância, porém específicas para cada organela são reveladas (TAYLOR; FAHY; GHOSH, 2003).

Apesar do interesse pela proteômica subcelular para o melhor entendimento de muitos aspectos da biologia funcional em células vegetais, ainda existem obstáculos principalmente relacionados aos métodos de purificação ou isolamento dessas organelas e à obtenção de quantidades de proteínas suficientes para a análise (MILLAR, 2004).

A caracterização do conteúdo proteico das organelas começou bem antes da introdução do neologismo “proteômica” e do uso da espectrometria de massas. No entanto, o avanço dos métodos de análise de alta capacidade (*“high throughput”*) mudou as escalas de análise nas quais os esforços passaram a ser empregados (BRUNET et al., 2003). Tal caracterização é de grande importância para o entendimento completo das funções da planta, bem como suas rotas biossintéticas e de sinalização (AGRAWAL et al., 2011). Existem três pré-requisitos para o isolamento das organelas: primeiro, a organela deve ser facilmente reconhecível; segundo, ela pode ser purificada; terceiro, o grau de enriquecimento deve ser eficientemente avaliado (AGRAWAL et al., 2011).

A proteômica subcelular é uma das áreas de mais rápido crescimento dentro da pesquisa em proteômica e visa identificar diretamente os conteúdos protéicos dos compartimentos purificados através do fracionamento subcelular (HAYNES; ROBERTS, 2007), constituindo uma interface entre a biologia celular e a proteômica (DREGER, 2003b). A análise proteômica das estruturas subcelulares enriquecidas por meio de fracionamento subcelular, portanto, representa uma análise estratégica que combina os métodos clássicos de fracionamento bioquímico com as ferramentas de identificação de proteínas, considerada, desta forma, um pré-requisito para a detecção de eventos regulatórios importantes, como a translocação de proteínas (DREGER, 2003). Além disso, essa abordagem experimental permite o mapeamento das proteínas de uma estrutura subcelular particular (DREGER, 2003b).

O núcleo é normalmente a organela mais proeminente da célula eucariótica, delimitado por duas membranas concêntricas, que formam o envelope nuclear. Em seu interior encontram-se as moléculas de DNA, polímeros extremamente longos que contêm a informação genética de um organismo (ALBERTS et al., 2010). No núcleo se encontra basicamente toda a informação genética necessária para regular a expressão de proteínas celulares. Ele ajuda no transporte de fatores regulatórios e

produtos gênicos através do poro nuclear, atua na produção de mRNAs e ribossomos e organiza a descompactação do DNA para a replicação e expressão gênica (ABDALLA; RAFUDEEN, 2012).

A estrutura e a organização funcional do núcleo constituem um assunto causador de energéticos debates. De um lado foi proposto que o núcleo apresenta seu próprio núcleo-esqueleto e organelas distintas. Já por outro lado, ele é visto como uma bolsa de membrana, altamente desorganizada, que envolve o DNA e outras moléculas, na qual as estruturas não são consideradas mais que complexos transitórios que se formam e dispersam como resultado das atividades de transcrição, replicação e processamento do RNA em várias regiões do genoma (LAMOND; EARNSHAW, 1998). Nesse aspecto, a compreensão molecular dos princípios de organização do núcleo constitui uma das maiores metas da biologia celular. Deve ser ressaltado que infelizmente ainda se conhece muito pouco acerca da organização do núcleo vegetal, em especial, bem como a estrutura de seus compartimentos (ERHARDT; ADAMSKA; FRANCO, 2010).

O núcleo é composto por diversas estruturas subnucleares, exemplificadas pelos nucléolos, corpos de Cajal e manchas nucleares (*"speckles"*) que constituem organelas subnucleares não delimitadas por membranas e se encontram em estado estacionário. No entanto, respondem de forma dinâmica aos processos fisiológicos básicos, bem como às várias formas de estresse, condições metabólicas alteradas e alterações na sinalização celular. Em coordenação com RNAs não codificantes, modificadores de cromatina e outras "maquinarias", vários corpos nucleares estão envolvidos no seqüestro e modificação de proteínas, processamento de RNAs, montagem de complexos de ribonucleoproteínas, e ainda na regulação epigenética da expressão gênica (MAO; ZHANG; SPECTOR, 2011).

O proteoma nuclear é composto por proteínas de diversas funções, responsáveis pela formação de estruturas essenciais, como a matriz nuclear e os nucléolos, e pelo controle do fluxo e da manutenção da informação genética. Esses fatores-chave são representados especialmente por fatores regulatórios, proteínas envolvidas na manutenção do genoma, na programação e transcrição gênica e ainda no processamento e exportação do RNA (BLAZEK; MEISTERERNST, 2006).

Ademais, as proteínas nucleares apresentam papéis-chave na regulação dos processos nucleares (replicação, transcrição), nas fases do desenvolvimento dos órgãos vegetais e na capacidade de resposta fisiológica por meio da expressão gênica. Dessa forma, assim como o proteoma global da célula, o proteoma nuclear também é dinâmico, mudando sua composição em resposta a estímulos ambientais e intracelulares (PANDEY et al., 2008).

A localização celular das proteínas é normalmente determinada por microscopia, ou por técnicas de fracionamento celular combinadas a técnicas de *blotting* proteico. No entanto, estas técnicas são classificadas como de baixo alcance (*low throughput*), não somente pelo número de amostras processadas, mas também pelo número de proteínas caracterizadas por corrida, o que as tornam limitadas à análise de componentes conhecidos. Nesse aspecto, as técnicas de proteômica combinadas à espectrometria de massas fornecem uma abordagem não-tendenciosa, de alto alcance (*high throughput*), e que pode ser aplicada na determinação da distribuição subcelular do proteoma (BOISVERT *et al.*, 2010). Desta forma, ambas as técnicas apresentam a capacidade de identificar proteínas. Porém o alto alcance podem gerar uma quantidade de dados maior do que o baixo alcance, mas não necessariamente com a mesma precisão.

Um dos problemas enfrentados durante a análise proteômica em nível subcelular é a validação dos dados de localização das proteínas identificadas. Isto se deve ao fato de que as frações subcelulares purificadas na maioria das vezes apresentam proteínas contaminantes derivadas de organelas diferentes da organela-alvo do estudo. Desta forma, há uma incerteza na atribuição da localização celular de um produto gênico desconhecido ou de produtos gênicos antes não descritos como presentes na estrutura subcelular em estudo. Assim, torna-se necessário avaliar os dados por meio de métodos independentes, não relacionados às técnicas de fracionamento (DREGER, 2003; DREGER, 2003b).

O desafio da proteômica organelar é fornecer um contexto funcional para as proteínas identificadas por meio de sua associação a um grupo distinto de proteínas em um meio intracelular definido (TAYLOR; FAHY; GHOSH, 2003).

Além do aspecto da identificação de inúmeras proteínas, o que gera informações a respeito de localização subcelular e de suas isoformas, os estudos de

proteômica de organelas têm levado à grandes descobertas na biologia celular, contribuindo também para remodelar nosso entendimento do funcionamento celular.

Em decorrência desse rápido avanço na pesquisa genômica e proteômica, grandes quantidades de seqüências de DNA e proteínas se acumularam nas bases de dados, o que torna cada vez mais importante o desenvolvimento, por parte dos biólogos computacionais, de ferramentas práticas para extrair de forma eficiente as informações biológicas relevantes a partir das seqüências usadas para a anotação funcional. Uma vez que a função da proteína normalmente está relacionada à sua localização subcelular, a habilidade de predizer sua localização diretamente das seqüências protéicas torna-se muito útil para inferir as funções das proteínas. Assim, nos últimos anos, o interesse pelo desenvolvimento de novas ferramentas computacionais que atinjam esse objetivo tem sido cada vez maior, levando ao surgimento de ferramentas que se baseiam numa grande variedade de algoritmos e alcançando diferentes graus de sucesso para diversos organismos e localizações (YU et al., 2006).

Tendo em vista a importância do estudo da proteômica nuclear e a importância da cana-de-açúcar no mercado brasileiro, o presente trabalho visou permitir o conhecimento acerca do subproteoma de uma das espécies economicamente mais importantes do Brasil, juntamente com a aplicação de diversas ferramentas de bioinformática especializadas na classificação subcelular das proteínas.

2 OBJETIVOS

O projeto de dissertação teve por objetivo caracterizar o proteoma nuclear da cana-de-açúcar, sendo inédito na área, uma vez que se conhece muito pouco a respeito do proteoma de cana-de-açúcar (*Saccharum* spp) e muito menos acerca do subproteoma dessa espécie.

Em decorrência disso, o objetivo principal foi estabelecer um protocolo para o isolamento de núcleos de cana-de-açúcar, bem como verificar a eficiência desse protocolo no enriquecimento de proteínas nucleares por meio de técnicas de microscopia e biologia molecular. A análise bioinformática foi utilizada para verificar a localização celular de todas as proteínas identificadas nos experimentos. Além disso, foram também utilizados cinco programas de predição da localização nuclear das proteínas.

A caracterização do proteoma nuclear foi realizada utilizando-se folhas de cana-de-açúcar em duas idades diferentes, como objetivo de verificar a aplicação do protocolo de isolamento em diferentes estádios fisiológicas e demonstrar sua adequação em pesquisas futuras.

3 REVISÃO BIBLIOGRÁFICA

3.1 Cana-de-açúcar e importância econômica

A cana-de-açúcar (*Saccharum spp.*) pertence à família Poaceae e tem sido cultivada e melhorada ao longo de milhares de anos, desde os tempos pré-históricos, com a domesticação e seleção dos genótipos mais produtivos e continuando com as técnicas atuais de hibridização e engenharia genética (MING et al., 2006). Os relatos do uso da cana-de-açúcar na Índia e China são considerados os mais antigos, existindo pouca dúvida que o homem primitivo já a cultivava. O melhoramento das variedades de cana-de-açúcar também tem uma história tão longa quanto o seu cultivo (JAMES, 2004).

As variedades de cana-de-açúcar cultivadas atualmente são principalmente híbridos interespecíficos complexos entre *Saccharum officinarum*, conhecida como cana nobre, e *Saccharum spontaneum*, com contribuições de *S. robustum*, *S. sinense*, *S. barberi*, e gêneros de gramíneas relacionadas, como *Miscanthus*, *Narenga* e *Erianthus*. Acredita-se que *S. officinarum* originou-se de uma forma domesticada de *S. robustum*, natural da Nova Guiné, que apresenta caules grossos, alto teor de açúcar e baixo teor de fibras (DANIELS; ROACH, 1987). A partir da Nova Guiné, *S. officinarum* se espalhou para a Indonésia, Malásia, China, Índia e Estados Federais da Micronésia e Polinésia, sendo que o primeiro registro da cana-de-açúcar na Índia foi em 325 a.C, realizado por um oficial de Alexandre, o Grande. A distribuição da Polinésia para o Havaí ocorreu juntamente com migrações nativas entre 500 – 1000 d. C. Já da Indonésia para o sul da Arábia e leste da África ocorreu provavelmente em 500 d.C. (MING et al., 2006).

A cultura da cana-de-açúcar foi introduzida no Brasil no período colonial e se transformou em uma das principais da economia do país. Atualmente, o Brasil não é apenas o maior produtor de cana, mas também o primeiro do mundo na produção de açúcar e etanol, e conquista cada vez mais o mercado externo com relação ao uso do biocombustível como alternativa energética (Ministério da Agricultura Pecuária e Abastecimento – MAPA , 2012).

No Brasil, ela é cultivada no sistema convencional em solos bem preparados, com revolvimento de 20 a 30 cm de profundidade. As mudas, propágulos, devem vir de canaviais de 12 a 18 meses, livres de pragas e doenças. A densidade usada é de

12 a 18 gemas por metro linear e a distância entre os sulcos é de 1,20 m, o que facilita a operação de colheita (Companhia Nacional de Abastecimento – Conab, 2012).

O Brasil é responsável por mais da metade do açúcar comercializado no mundo e deve alcançar a taxa média de aumento da produção de 3,25% até 2018/19, e colher 47,34 milhões de toneladas do produto, o que corresponde a um acréscimo de 14,6 milhões de toneladas em relação ao período 2007/2008. Para exportações, o volume previsto para 2019 é de 32,6 milhões de toneladas. Já com relação ao etanol, a produção projetada para 2019 é de 58,8 bilhões de litros, mais do que o dobro da registrada em 2008, sendo 50 bilhões de litros projetados para consumo interno e 8,8 bilhões para exportações (Ministério da Agricultura Pecuária e Abastecimento – MAPA , 2012).

Atualmente, a política nacional para a produção de cana-de-açúcar se orienta na expansão sustentável da cultura, tendo como base os critérios econômicos, ambientais e sociais. O programa de Zoneamento Agroecológico da Cana-de-açúcar (ZAEcana) regula o plantio da cana, levando em consideração o meio ambiente e a aptidão econômica de cada região (Ministério da Agricultura Pecuária e Abastecimento – MAPA , 2012). Esse programa surgiu da necessidade estratégica de se avaliar, indicar e especializar o potencial das terras para a expansão da produção de cana-de-açúcar em regime de sequeiro (sem irrigação plena) para a produção de etanol e açúcar como base para o planejamento do uso sustentável das terras, em harmonia com a biodiversidade (MANZATTO et al., 2009).

A produção da cana-de-açúcar se concentra nas regiões Centro-Sul e Nordeste do Brasil. O mapa abaixo (Figura 1) representa em vermelho as áreas onde se concentram as plantações e usinas produtoras de açúcar, etanol e bioeletricidade, segundo dados oficiais do IBGE, UNICAMP e CTC (Centro de Tecnologia Canavieira) (UNIÃO DA INDÚSTRIA DE CANA-DE-AÇÚCAR – ÚNICA, 2012). Além disso, o maior percentual de expansão da área plantada com cana-de-açúcar está na região Sudeste, responsável por 96,38% do total da área nova agregada, que constitui 357.183 hectares acrescentados à área existente (Companhia Nacional de Abastecimento – Conab, 2012).

Figura 1 – Áreas de concentração da produção da cana-de-açúcar
 Fonte: UNIÃO DA INDÚSTRIA DE CANA-DE-AÇÚCAR – ÚNICA, 2012

A cana-de-açúcar é uma gramínea tropical perene e alta, cuja base perfilha para produzir caules não ramificados de 2 a 4 m ou mais altos e de aproximadamente 5 cm de diâmetro. Suas principais estruturas são o caule ou colmo, as folhas e as raízes. O caule sólido e não ramificado é rústicamente circular na secção transversal, sendo claramente diferenciado em articulações que compreendem o nó e o internó. O nó consiste de uma gema lateral, situada na axila da folha, uma faixa contendo o primórdio da raiz e um anel de crescimento. Normalmente existe apenas uma gema por nó e elas se localizam em lados alternados do caule. As folhas são ligadas ao caule, na base dos nós, alternadamente em duas linhas presentes em lados opostos do colmo. Cada folha consiste de duas partes, a bainha e a lâmina. A bainha é tubular, mais larga na base e circunda fortemente o colmo, sendo separada da lâmina foliar através de uma lígula. A lâmina foliar é uma estrutura fina que se sobrepõe ao colmo na base, mas

tende a ficar menos próxima do mesmo. Após ser plantada, a cana-de-açúcar desenvolve dois tipos de raiz. As primeiras se desenvolvem na faixa da raiz, são finas e ramificadas, sendo chamadas raízes superficiais (*sett roots*). As segundas se desenvolvem a partir dos primórdios da raiz dos perfilhos em desenvolvimento, são grossas, carnosas e menos ramificadas, sendo denominadas raízes de sustentação (*shoot roots*) (JAMES, 2004).

As plantas de cana-de-açúcar apresentam quatro estádios fenológicos: brotação e emergência, perfilhamento, crescimento dos colmos e maturação dos colmos. Na brotação e emergência, o broto, um caule em miniatura ou colmo primário, rompe as folhas da gema e se desenvolve em direção à superfície do solo, surgem as primeiras raízes e aparecem as primeiras folhas. Essa fase ocorre de 20 a 30 dias após o plantio. A segunda fase é denominada perfilhamento, e é definida como a emissão de colmos por uma mesma planta. O processo de perfilhamento é regulado por hormônios e resulta no crescimento de brotos que se desenvolvem em direção à superfície do solo. Os brotos aparecem de 20 a 30 dias após a emergência do colmo primário. É por meio deste processo que ocorre a formação das touceiras de cana-de-açúcar (MARIN, 2012).

Após o auge do perfilhamento, começa o estágio de crescimento dos colmos, no qual aqueles sobreviventes continuam seu crescimento e desenvolvimento, passam a ganhar altura e iniciam o acúmulo de açúcar na base. Nesse estágio o crescimento radicular se torna mais intenso e vigoroso tanto nas laterais quanto em profundidade. O último estágio, maturação dos colmos é dividido em maturação inicial, quando há o crescimento intenso dos colmos, maturação do terço médio, momento em que as touceiras atingem altura igual ou superior a 2 m e as folhas passam a ficar amarelas e secas, e maturação final, que consiste no período de maior atividade de maturação e menor atividade de crescimento, havendo o armazenamento intenso de açúcar (MARIN, 2012).

Figura 2 – Fases do desenvolvimento da cana-de-açúcar

Fonte: MARIN, 2012 apud GASCHO; SHIH, 1983

3.2 Proteômica e Subproteômica

3.2.1 Proteômica Vegetal

A proteômica é o estudo da composição protéica total de um organismo, célula ou organela e surge da necessidade presente em nosso repertório de “ômicas” em identificar e compreender as proteínas e suas funções, em uma célula, tecido, órgão ou organismo (AGRAWAL et al., 2011).

Nesse aspecto, o estudo de complexas questões biológicas por meio da proteômica comparativa está se tornando cada vez mais atrativo, especialmente em decorrência do rápido desenvolvimento das técnicas de identificação e quantificação proteica. Tais técnicas têm permitido o estudo simultâneo de quantidade enorme de proteínas expressas em uma célula, em um determinado momento (ROSE et al., 2004; HAYNES; ROBERTS, 2007).

No entanto, essa rápida geração de dados tem trazido um novo desafio relacionado à atribuição de funções às milhares de proteínas codificadas pelos genomas vegetais, porém deve ser ressaltado ainda que prever o estado funcional das proteínas baseado apenas nos dados de sequenciamento é muito difícil, em

decorrência do fato de que a atividade protéica pode ser regulada pós-traducionalmente, por diversos fatores, como fosforilação, pH, compartimentalização e interações proteína-proteína (KOŁODZIEJEK; VAN DER HOORN, 2010).

Em termos históricos, a proteômica vegetal originou-se há mais de 30 anos com o desenvolvimento da separação multidimensional das proteínas, fundamentada no uso inicial de ferramentas proteômicas básicas, como a eletroforese bi-dimensional, bem como no progresso do conhecimento de seu braço complementar, a espectrometria de massas, que juntos revolucionaram a forma como as proteínas são encontradas e identificadas. Em resumo, a proteômica tem sua origem em diversos campos cuja integração resultou em uma importante tecnologia para a pesquisa biológica na era pós-genômica (AGRAWAL et al., 2011). Dentre esses diversos campos, o rápido avanço da proteômica deve-se principalmente à aplicação combinada de avançadas técnicas de separação, espectrometria de massas e ferramentas de bioinformática na caracterização de proteínas nas complexas misturas biológicas (JUNG et al., 2000; DREGER, 2003b).

As técnicas de proteômica, apesar de relativamente estabelecidas, ainda apresentam diversas dificuldades, exemplificadas especialmente pelo fato de as proteínas serem altamente heterogêneas físico-quimicamente e apresentarem uma estrutura complexa, o que dificulta sua extração, solubilização, manuseio, separação e identificação. Ademais, a inexistência de tecnologias equivalentes ao PCR para amplificar proteínas de baixa abundância faz com que a proteômica apresente ainda mais desafios (BRUNET et al, 2003; ROSE et al., 2004; HAYNES; ROBERTS, 2007).

Outro aspecto a ser mencionado é o fato de o dogma tradicional da biologia celular e molecular, definido por um gene → uma proteína → uma localização e, portanto, uma função, constituir uma definição ultrapassada na era pós-genômica, uma vez que o seqüenciamento de diversos genomas tornou evidente que a complexidade do proteoma excede à do genoma em termos de número de unidades funcionais, isto é, existem mais proteínas do que genes (CARRIE; GIRAUD; WHELAN, 2009).

Esse fato pode ser constatado por estudos realizados em diversas espécies vegetais segundo os quais as mudanças nos níveis dos transcritos não são

completamente acompanhadas pelas mesmas mudanças nos níveis de proteínas (GYGI et al., 1999). Dessa mesma forma, a regulação de eventos celulares pode ocorrer no nível protéico sem mudanças aparentes na abundância do mRNA, uma vez que as modificações pós-traducionais de proteínas podem resultar em um aumento dramático na complexidade protéica, sem aumento conjunto na expressão gênica (ROSE et al., 2004), além da existência de regulação por diversas outras proteínas, como as kinases e fosfatases.

Dessa maneira, diferentemente do que acontece com o genoma e o transcriptoma, um único organismo pode ter diversos proteomas que mudam espacialmente e temporalmente, durante as diversas fases do desenvolvimento de um organismo. Um dos pontos mais atrativos da proteômica é seu potencial para rapidamente caracterizar as diferenças moleculares entre diversas amostras e identificar a presença de produtos gênicos em um determinado lugar e momento (HEAZLEWOOD; MILLAR, 2003).

O simples problema da complexidade constitui um grande desafio para os estudos que visam revelar o complemento protéico do genoma, de forma que um único genoma pode fornecer múltiplos conjuntos protéicos, resultantes de *splicing* alternativo e/ou modificações pós-traducionais (TAYLOR, FAHY, GHOSHY, 2003).

Nos tecidos vegetais, esses múltiplos conjuntos protéicos permitem o desenvolvimento de diversos tipos celulares dentro de um único órgão, de modo que um organismo é composto por diferentes órgãos e tipos de tecidos. Dessa forma, uma abordagem proteômica simples passa a ser insuficiente para o estudo de muitos problemas biológicos, não refletindo a real complexidade da dinâmica protéica existente. Surge, assim, o desafio de desenvolver novas estratégias que possam avaliar a complexidade e diversidade da composição das proteínas, além de suas interações e dinâmicas. Nesse sentido, o estudo do subproteoma torna-se uma estratégia relevante (DREGER, 2003; AGRAWAL et al., 2011).

Com relação à cana-de-açúcar, a análise proteômica encontra-se em fase inicial, e ainda pouco compreendida. Os desafios enfrentados nessa área dizem respeito especialmente à rigidez dos tecidos, sua natureza fibrosa, níveis extremos de enzimas oxidativas, compostos fenólicos, carboidratos (especialmente sacarose),

além da presença de outros metabólitos interferentes, especialmente nos tecidos do colmo (ALMARAJ et al., 2010).

Desta forma, poucos estudos acerca do proteoma da cana-de-açúcar estão disponíveis, sendo a maioria voltada para o estudo do proteoma foliar (SUGIHARTO et al., 2002; RAMAGOPAL, 1994). No entanto trabalhos recentes sobre o proteoma resultante da interação cana-patógeno (LERY et al., 2010), do proteoma do colmo em diferentes estádios de desenvolvimento (BOARETTO, 2012) e do perfil proteômico de peroxidases classe III presentes no colmo (CESARINO et al., 2012) podem ser encontrados.

3.2.2 A subproteômica e sua importância

O aumento no grau de complexidade dos organismos permitiu que eles adquirissem uma gama muito maior de opções para enfrentar os desafios impostos pelo meio ambiente. Tal aumento de complexidade é manifestado em dois níveis principais. Primeiramente existe a diferenciação e especialização das células, de forma a existirem tipos celulares específicos com distintas propriedades que podem, portanto, apresentar produtos gênicos próprios e formar tecidos diferentes. Em segundo lugar, as próprias células são também compartimentalizadas, o que permite o surgimento de ambientes bioquímicos distintos e adequados para diversos processos incluindo a síntese e degradação de proteínas, o fornecimento de metabólitos ricos em energia, a glicosilação proteica e a replicação do DNA dentre outros. Essa estrutura compartimentalizada é suportada por subconjuntos de produtos gênicos direcionados especificamente para uma estrutura subcelular (DREGER, 2003; DREGER, 2003b).

Dessa forma, pode-se considerar que um dos conceitos mais básicos da biologia celular é a compartimentalização dos processos celulares dentro de estruturas subcelulares denominadas organelas. O termo organela foi definido inicialmente no século XIX como referência a entidades morfológicas distintas e até os dias de hoje se reflete no nome dessas estruturas. As palavras núcleo e retículo, por exemplo, vêm do latim e significam “pequena noz” e “pequena rede”, respectivamente, assim como mitocôndria vem do grego “míto” fio + “chóndrion” pequeno grão (ANDREYEV et al., 2010).

A definição clássica do termo organela se refere a várias estruturas delimitadas por membranas de bicamada lipídica com funções especializadas, suspensas no citoplasma da célula eucariótica (JIMÉNEZ, 2009).

Em decorrência desta organização compartimentalizada das células eucarióticas houve a possibilidade de separação e especialização de funções o que facilitou a evolução da diversidade de organismos multicelulares (CAVALIER-SMITH, 2009). Todavia, a criação desta complexa organização intracelular também levou à necessidade de um novo tipo de maquinaria biológica elaborada (MILLAR, 2004), composta de mecanismos responsáveis pelo direcionamento e importação de proteínas em uma variedade de compartimentos celulares. Estes mecanismos de direcionamento parecem ser altamente específicos, e normalmente alcançados por meio de sinais de direcionamento presentes em proteínas (peptídeo-sinal), que são reconhecidos por receptores de importação presentes na superfície das organelas (DYALL; BROWN; JOHNSON, 2004; HAYNES; ROBERTS, 2007).

Além deste mecanismo, existe ainda o direcionamento ambíguo de seqüências, denominado *dual targeting*, um conceito que não está relacionado ao aumento na quantidade de uma determinada proteína, mas sim à expansão nas funções da mesma, uma vez que uma proteína localizada em mais de um ambiente celular pode apresentar diferentes funções relacionadas aos processos bioquímicos existentes nos diferentes locais (CARRIE; GIRAUD; WHELAN, 2009), ou mesmo, no caso de enzimas, mudar seu substrato de ação. Dentro deste tipo de mecanismo, são encontradas ainda proteínas que apresentam a denominada “distribuição em eclipse”, ou seja, são direcionadas a mais de um compartimento subcelular e conseqüentemente participam de diferentes rotas bioquímicas, podendo ter funções completamente diferentes. No entanto, como se encontram-se presentes em enormes quantidades em uma organela e em quantidades menores em outra, sua detecção bioquímica torna-se quase improvável pelos métodos de análise existentes (NAAMATI et al., 2009; REGEV-RUDZKI; PINES, 2007). Um exemplo deste fenômeno é a aconitase (EC 4.2.1.3), uma enzima mitocondrial que participa do ciclo do ácido tricarboxílico e que recentemente foi encontrada em diminutas quantidades no citosol, com funções relacionadas ao desvio do glioxilato (REGEV-RUDZKI et al., 2005).

Esses aspectos tornam evidente o fato de que quando se trata da definição da localização de uma proteína, nem sempre direcionamento, acúmulo e função constituem termos intercambiáveis, de forma que algumas proteínas apresentam localizações fortemente definidas, enquanto outras apresentam baixa especificidade no direcionamento e complexos padrões de acúmulo. Além disso, para algumas proteínas a localização é essencial para a função, já para outras não, ou seja, o direcionamento e o acúmulo de uma proteína não indicam obrigatoriamente o local da função, nem a importância relativa de determinada função para a proteína sob investigação. Também deve ser salientado que uma proteína que apresenta vários alvos pode ter diferentes funções nos diversos locais em que ela se encontra, podendo o papel da proteína em um local ser dominante com relação às outras funções (MILLAR et al., 2009).

A pesquisa proteômica apresenta como objetivo resolver a complexidade biológica codificada pelo genoma. No entanto, a complexidade existente nas células eucarióticas faz com que a caracterização do proteoma em um único passo seja difícil de ser alcançado, não somente em decorrência do número de proteínas expressos em uma única célula, mas também pela faixa de pH ampla, hidrofobicidade, peso molecular e baixa abundância de algumas proteínas. Contudo, algumas vantagens podem ser obtidas em decorrência da arquitetura macromolecular da célula, ou seja, sua divisão em compartimentos subcelulares, para o estabelecimento dos respectivos proteomas organelares (JUNG et al., 2000).

Aliado a esta constatação, o progresso na bioquímica tornou possível atribuir funções biológicas distintas para as diferentes organelas, uma vez que informações detalhadas sobre a localização das reações bioquímicas é crucial para o entendimento do papel de cada organela para a função ou falta de função celular (ANDREYEV *et al.*, 2010). Dessa forma, o refinamento dos métodos e meios usados para o fracionamento subcelular tem permitido a análise da composição de organelas isoladas e o desenvolvimento de ensaios *cell-free* para estudar e reconstituir os complexos processos celulares, além de fornecer informações para o mapeamento do proteoma de estruturas celulares distintas (BRUNET et al., 2003; DREGER, 2003).

As técnicas de fracionamento subcelular desenvolvidas ao longo do tempo permitem a simplificação do proteoma e envolvem basicamente dois passos: o rompimento da parede celular e da membrana; e o fracionamento do homogenato cru, purificando a organela de interesse. O rompimento deve ser controlado para evitar a ruptura excessiva dos compartimentos subcelulares, e pode ser realizado de várias formas, como, por exemplo, por meio de um liquidificador ou de um homogeneizador de tecidos (AGRAWAL et al., 2011). Um aspecto dessas técnicas é o fato de que, em decorrência do protocolo, normalmente aproveita-se apenas a parte da preparação que contém a organela-alvo isolada, sendo o restante desperdiçado (DREGER, 2003b).

O fracionamento por centrifugação é considerado o método de isolamento de organelas mais eficiente, o qual é facilmente estabelecido e pode ser combinado com técnicas proteômicas analíticas (HUBER; PFALLER; VIETTOR, 2003). A purificação e o isolamento de compartimentos subcelulares de células de plantas como *Arabidopsis*, por exemplo, pode ser realizada através de centrifugação diferencial, ou ainda por centrifugação realizada em gradientes de sacarose (mais comum) ou percoll (MILLAR, 2004; HAYNES; ROBERTS, 2007; PASQUALI; FIALKA; HUBER, 1999; HUBER; PFALLER; VIETTOR, 2003).

O processo de fracionamento normalmente é baseado nas diferenças físicas entre as organelas, onde o primeiro passo frequentemente envolve a filtração, usando uma malha de nylon, por exemplo, de forma a remover os debris celulares maiores. Então uma série de centrifugações é usada para o enriquecimento da organela-alvo, eliminando seletivamente outros compartimentos e contaminantes. A velocidade de centrifugação também depende do tamanho e da densidade da organela a ser purificada. Organelas maiores e mais densas são isoladas em forças centrífugas menores. Desta forma, núcleos e cloroplastos são isolados em velocidades menores, enquanto organelas como as mitocôndrias exigem velocidades muito maiores (AGRAWAL et al., 2011). A posição de organelas delimitadas por membrana nos gradientes de densidade é determinada principalmente pela razão entre seu conteúdo protéico e lipídico. Por exemplo, a membrana interna da mitocôndria é rica em proteínas, apresentando alta densidade, enquanto as membranas endossomais são ricas em lipídios e apresentam baixa

densidade. O conteúdo das vesículas e a presença de elementos ligados à organela, como é o caso dos ribossomos no retículo endoplasmático rugoso (RER), também afetam a densidade das partículas (PASQUALI; FIALKA; HUBER, 1999; HUBER; PFALLER; VIETTOR, 2003).

Uma das maiores dificuldades impostas pela utilização dos métodos de fracionamento consiste no fato de que um determinado protocolo se refere somente a uma linhagem celular específica, servindo apenas como pontapé inicial para adaptar o protocolo aos novos sistemas experimentais sob investigação. Assim, deve-se tomar cuidado ao se comparar o proteoma de uma mesma estrutura subcelular derivada de diferentes tipos celulares, uma vez que até o grau de contaminação pode variar (DREGER, 2003; DREGER, 2003b). Outra limitação surge da dificuldade em se verificar o grau de pureza das frações, de maneira que o máximo que se consegue é avaliar o nível de enriquecimento por meio da medida do nível de proteínas esperadas nessas frações quando comparadas ao material inicial, ou ainda pela diminuição ou falta de marcadores característicos de outras organelas (BRUNET et al., 2003).

Dessa forma, com relação ao estudo das proteínas de um compartimento subcelular, surge a proteômica subcelular, subproteômica ou proteômica de organelas, que descreve o estudo das proteínas presentes em uma organela, em um determinado momento durante o período de seu ciclo de vida dentro da célula (AGRAWAL et al., 2011). O estudo do proteoma de compartimentos subcelulares individuais fornece uma caracterização focada na relação entre localização de uma proteína e sua determinada função, o que reduz a complexidade protéica e aumenta a resolução técnica (ANDERSEN; MANN, 2006; DREGER, 2003; JIANG et al., 2005; JUNG et al., 2000; PANDEY et al., 2008), além de constituir ainda a base para o monitoramento de importantes aspectos das mudanças dinâmicas que ocorrem no proteoma, como é o caso da translocação protéica (DREGER, 2003b). Além disso, famílias de produtos gênicos fortemente relacionados são abundantes nas plantas, nas quais, as proteínas codificadas por esses genes frequentemente parecem ser funcionalmente redundantes. No entanto, a diferença crucial entre elas geralmente é o seu destino celular final (HAYNES; ROBERTS, 2007).

A interpretação dos resultados obtidos através de técnicas de subproteômica é complicada em decorrência não somente da natureza dinâmica dos proteomas organelares, como também pelo fato de que muitas proteínas não são exclusivas de um único local, podendo ser encontradas em diferentes compartimentos celulares (BOISVERT et al., 2010). Nesse aspecto, a co-purificação de proteínas “contaminantes” tem sido um dos maiores problemas enfrentados pelas pesquisas em proteômica subcelular, em decorrência da dificuldade de concluir se estes “contaminantes” representam verdadeiros parceiros endógenos ou associações artificiais induzidas pela ruptura celular ou purificação incompleta, uma vez que muitas vezes esses contaminantes, apesar de presentes, apresentam um nível de expressão menor (JIANG et al., 2005; HAYNES; ROBERTS, 2007). Assim, a combinação da abordagem ICAT com técnicas bioinformáticas de predição já foi usada visando diminuir essa ambiguidade (JIANG et al., 2005).

O grau de confiança dos dados obtidos por meio das técnicas de isolamento depende grandemente do grau de purificação e da extensão na qual a contaminação pode ser reconhecida e reduzida ou evitada. Uma variedade de métodos pode ser usada para avaliar o enriquecimento da organela-alvo em diferentes etapas dos protocolos de subproteoma, desde a integridade da organela-alvo, até a presença de proteínas contaminantes. A pureza e a integridade das organelas podem ser determinadas por meio de avaliações microscópicas, uma vez que são rápidas e informativas. Já a pureza dos extratos proteicos pode ser avaliada por meio de métodos como o *Western blot*, ou ainda por meio da determinação da atividade de enzimas consideradas marcadoras da organela-alvo (DREGER, 2003; DREGER, 2003b; AGRAWAL et al., 2011).

A integridade e a qualidade morfológica de uma organela podem ser rapidamente avaliadas por meio de microscopia eletrônica ou de luz, com algumas organelas mais adequadas para este tipo de avaliação do que outras. O núcleo constitui uma organela adequada para tal avaliação por ser grande e facilmente reconhecível, além de ser, por definição, enriquecido para ácidos nucleicos, sendo corado diretamente com substâncias que apresentam afinidade pelo DNA e que podem ser avaliadas por meio de microscopia de fluorescência e/ou de luz. Um exemplo desses corantes é o DAPI (4', 6-diamidino-2-phenylindole), que se liga

diretamente ao DNA fita dupla e é usado tanto em materiais fixados quanto não-fixados, fornecendo uma avaliação da integridade do núcleo, uma vez que núcleos rompidos também são corados sendo facilmente observados. A avaliação da pureza da organela por meio de fluorescência indireta também é usada quando anticorpos específicos para a organela-alvo estão disponíveis em vez de corantes, como é o caso dos microtúbulos do citoesqueleto (AGRAWAL et al., 2011).

A pureza do extrato proteico é geralmente avaliada por meio de *Western blot* (PASQUALI; FIALKA; HUBER, 1999; AGRAWAL et al., 2011), no qual os extratos obtidos são submetidos à separação por eletroforese *in gel*, imobilizados em membranas e incubados com anticorpos específicos para cada organela. A quantidade de anticorpo hibridizado fornece uma identificação indireta seja do enriquecimento ou da contaminação. No primeiro caso, o anticorpo é específico para a organela-alvo, enquanto no segundo o anticorpo é específico para uma organela não-alvo. Essa técnica permite uma estimativa da pureza e pode ainda ser usada para otimizar os protocolos, maximizando a recuperação da organela ou estrutura de interesse (AGRAWAL et al., 2011). Alguns anticorpos usados são: Histona H1, Histona H3, fibrilarina, antígeno nuclear de proliferação celular (PCNA) e regulador transcricional de sinalização reduzida do fitocromo (REP1), como marcadores nucleares e UGPase, V-ATPase, proteína D1 do centro de reação do fotossistema II (PsbA), porina mitocondrial, proteína do complexo de evolução do oxigênio (OEE2 – oxygen-evolving enhancer protein 2) e glutathione-S-transferase (GST) como marcadores de outras organelas (KHAN; KOMATSU, 2004; KALLURI et al., 2009; BAE et al., 2003; ABDALLA; THOMSON; RAFUDEEN, 2009; REPETTO et al., 2008; PANDEY et al., 2006; CHOUDHARY et al. 2009).

A combinação das técnicas de microscopia com os métodos químicos/bioquímicos é recomendada para se alcançar uma alta sensibilidade e confiança na avaliação da pureza do extrato enriquecido para determinada organela. As primeiras são consideradas mais sensíveis, porém mais trabalhosas e caras, enquanto os últimos são menos sensíveis, fornecendo informação no nível de uma única proteína/enzima e não da organela como um todo. No entanto são mais rápidas e baratas (AGRAWAL et al., 2011).

Com relação à identificação das proteínas, uma abordagem que vem sendo largamente usada durante os últimos anos na proteômica por *shotgun* é o uso do gel SDS-PAGE como um método de fracionamento simples, que fornece muitas das vantagens da análise por MudPIT, sem a necessidade de aquisição de mais recursos. Um dos benefícios mais óbvios deste método é o fato de que, além de simples de usar, o gel SDS-PAGE normalmente se encontra disponível para a maioria dos pesquisadores. Outro aspecto é o fato de que os tampões-padrão das amostras separadas por SDS-PAGE solubilizam as proteínas de misturas complexas de maneira mais eficiente do que os tampões normalmente usados no preparo de amostras para o MudPIT. Dessa forma, as proteínas são separadas e o gel é dividido em linhas (bandas), que são retiradas, descoradas, submetidas à redução, alquilação e digestão com tripsina, com os peptídeos resultantes extraídos por subseqüentes eluições (HAYNES; ROBERTS, 2007).

Os resultados obtidos pela separação das proteínas no gel 1D são semelhantes aos obtidos pela análise por MudPIT, sendo que a comparação de diferentes abordagens da proteômica *shotgun* usadas para a análise de organelas de levedura demonstrou que a abordagem por meio de gel SDS-PAGE foi significativamente melhor do que a análise por MudPIT em termos de número de proteínas identificadas (HAYNES; ROBERTS, 2007; BRECI et al., 2005). Além disso, pesquisadores que trabalham com plantas estão entre os que mais utilizam essa tecnologia, obtendo sucesso com relação à proteômica subcelular (REPETTO et al., 2008; MILLAR; HEAZLEWOOD, 2003; BOISVERT et al., 2010).

A proteômica de organelas tem o objetivo de fornecer informações funcionais acerca da estrutura e dinâmica da organização celular (AGRAWAL et al., 2011), uma vez que a definição da localização intracelular de um produto protéico é um aspecto importante na definição de sua função (MILLAR et al., 2009). Atualmente constitui um dos campos da proteômica funcional que mais cresce, já que o foco da proteômica vegetal tem se direcionado cada vez mais para a caracterização de tecidos ou organelas específicas (AGRAWAL; RAKWAL, 2006). Nesse aspecto, além da descrição do proteoma de uma entidade subcelular, a análise das mudanças dinâmicas do mesmo em um nível subcelular leva à produção de um conhecimento significativo com relação aos mecanismos biológicos em diferentes

condições, bem como o entendimento da base molecular da função celular defeituosa (DREGER, 2003).

Esse conhecimento torna-se importante principalmente em decorrência da existência de grupos de proteínas que se encontram associados à determinada estrutura subcelular apenas em certos estados fisiológicos e localizadas em compartimentos diferentes em outros momentos. A variação da localização celular dessas proteínas ocorre devido a vários fatores, como a translocação protéica entre compartimentos, o movimento de proteínas entre a superfície celular e o ambiente intracelular e a troca de proteínas entre nucleoplasma e citoplasma (DREGER, 2003b).

A mudança de foco do estudo de proteoma para subproteoma surgiu principalmente da percepção de que embora promissora, a informação obtida por meio de programas de proteômica de larga escala seja complexa demais para a tecnologia disponível, não se adequando ao estudo de proteínas que apresentam baixo número de cópias (JUNG et al, 2000). Além disso, processos celulares essenciais, como o metabolismo central, as trocas gasosas, a biossíntese de produtos em alta quantidade e qualidade e as rotas de sinalização envolvidas nas respostas de defesa tanto contra o meio ambiente físico quanto contra patógenos invasores são todos compartimentalizados de forma complexa no interior da célula, devendo também ser estudados de acordo com o ambiente celular no qual se encontram (MILLAR, 2004).

As limitações apresentadas pela proteômica clássica vão desde o fato de que, embora o número total de produtos gênicos presentes em uma determinada célula seja estimado em aproximadamente 10.000 (no genoma humano, por exemplo), o número efetivo provavelmente é várias vezes maior em decorrência do *splicing* alternativo e da grande variedade de modificações pós-traducionais existentes (HUBER, 2003), até a existência de grande variação no nível de expressão das diversas proteínas em uma célula ou tecido, de 5 à 6 ordens de magnitude em células de levedura, até mais de 10 ordens de magnitude em determinadas células humanas, o que faz com que proteínas de baixa abundância sejam inevitavelmente mascaradas pelas de alta abundância (PATTERSON; AEBERSOLD, 2003). Um exemplo disso são algumas proteínas regulatórias importantes, como as quinases e

GTPases, além de determinados tipos de receptores de membrana que se encontram presentes em pequenas quantidades, cujo mascaramento faz com que importantes informações sejam perdidas (BRUNET et al., 2003).

Dessa forma, a organização da célula eucariótica em subcompartimentos de funções especializadas fornece uma oportunidade única de ligar os dados de proteômica com as unidades funcionais, o que constitui um passo importante para a compreensão dos mecanismos moleculares que governam o funcionamento celular. A análise do proteoma das organelas pode ajudar na descoberta de novos conceitos da biologia celular, conduzindo à melhor compreensão dos processos celulares complexos que governam o funcionamento de um organismo, bem como sua interação com o ambiente (BRUNET, 2003), além de ajudar na identificação do papel de muitas proteínas, que ainda apresentam, nas bases de dados, a designação de função desconhecida (MILLAR, 2004).

3.3 O núcleo e suas proteínas

O núcleo foi descoberto no século XVII por Antonie van Leeuwenhoek (1632 – 1723) (ERHARDT; ADAMSKA; FRANCO, 2010), e descrito inicialmente por Robert Brown em 1831 (LAMOND; EARNSHAW, 1998). Constitui a estrutura mais proeminente da célula eucariótica. Ele é claramente visível por meio da microscopia de luz e contém a maior parte do DNA presente na célula organizado na forma de cromossomos, constituindo, assim o local da replicação e transcrição do DNA (ERHARDT; ADAMSKA; FRANCO, 2010).

Essa organela é envolvida por uma membrana dupla e constitui a maior diferença entre as células procarióticas e eucarióticas. Além disso, as análises indicam a presença de um constante fluxo de moléculas com funções regulatórias distintas através do envelope, tornando o núcleo um dos órgãos regulatórios mais importantes presentes na célula (ERHARDT; ADAMSKA; FRANCO, 2010).

Desde a primeira observação microscópica do núcleo ficou claro que ele apresenta diversos subcompartimentos e, portanto, está longe de ser uma organela homogênea. O compartimento nuclear mais óbvio é o nucléolo, o local da transcrição ribossomal do DNA e da biogênese dos ribossomos. O restante do núcleo é organizado em regiões ricas em cromatina, que compreende a

heterocromatina condensada, a eucromatina mais dispersa e as regiões de intercromatina (SHAW; BROWN, 2004). O núcleo apresenta ainda outras estruturas ou domínios de diferentes tamanhos, frequências e funções (LAMOND; EARNSHAW, 1998), como os corpos de Cajal ou corpos enovelados, os corpos nucleares, o envelope nuclear e as manchas nucleares (*speckles*). No entanto, a caracterização da organização funcional do núcleo vegetal apresenta um atraso significativo com relação ao conhecimento acerca do núcleo de células de mamíferos (SHAW; BROWN, 2004).

O desenvolvimento das técnicas de microscopia permitiu a descoberta de aproximadamente 30 compartimentos diferentes (DUNDR; MISTELI, 2001). No entanto, o entendimento detalhado da organização nuclear e sua dinâmica nas plantas e da forma como eles impactam a replicação do DNA e a regulação gênica ainda precisam ser elucidados (CHEUNG; REDDY, 2012).

A singularidade do núcleo é demonstrada pelo fato de que todos os seus subcompartimentos não apresentam membranas, sendo entidades auto-organizadas que passam por um estado de desmontagem/remontagem durante a divisão celular. As moléculas nucleares são altamente dinâmicas e estão em constante movimento, sendo sua morfologia totalmente determinada pela interação funcional de seus componentes (MISTELI, 2001). A presença de um alto número de compartimentos intranucleares indica que a relação entre função e localização existe também no interior do núcleo (ERHARDT; ADAMSKA; FRANCO, 2010).

A formação de estruturas no núcleo é influenciada por muitas moléculas e fornece um mecanismo elegante não somente para concentrar fatores quando eles são necessários, mas também para segregar fatores para longe de onde eles são prejudiciais (MISTELI, 2000). Além disso, o movimento das moléculas não é restrito ao interior do núcleo, sendo encontrados relatos que sugerem que proteínas nucleares apresentam funções regulatórias na célula como um todo (BAE et al., 2003; MILLAR et al., 2009).

O envelope nuclear é uma dupla membrana composta por duas bicamadas lipídicas. A membrana nuclear externa é contínua com o retículo endoplasmático, que, por sua vez, apresenta ribossomos aderidos. A membrana interna apresenta um complemento único de proteínas integrais, que interagem com a cromatina e a

lâmina nuclear. Ambas as membranas são perfuradas por grandes complexos multiprotéicos, os poros nucleares, que alcançam o envelope nuclear inteiro e formam canais através do mesmo, abrindo espaço para a troca molecular (ERHARDT; ADAMSKA; FRANCO, 2010).

O envelope é responsável pela separação entre compartimento nuclear e citoplasma e apresenta funções relacionadas à organização da cromatina e regulação gênica (CHEUNG; REDDY, 2012), estando envolvido, dessa forma, na mobilidade nuclear, transdução de sinal, ligação com a cromatina e ativação e repressão transcricional (BORUC; ZHOU; MEIER, 2012).

A lâmina nuclear é uma malha de filamentos protéicos intermediários que mediam a ligação da cromatina com o envelope nuclear durante a intérfase e o desligamento da cromatina durante a mitose. Em plantas, novos estudos ultra-estruturais sugerem a existência de uma estrutura semelhante à lâmina nuclear, sendo os melhores candidatos para proteínas características dessa estrutura uma família de proteínas enroladas em espiral, *coiled-coil proteins*, que são aproximadamente duas vezes maiores que as proteínas da lâmina nuclear de células animais (BORUC; ZHOU; MEIER, 2012).

Os poros nucleares que são complexos multiprotéicos de 40- à 60-MD incorporados ao envelope nuclear (BORUC; ZHOU; MEIER, 2012), por sua vez, regulam o movimento do RNA, de proteínas e de complexos RNA-proteína para dentro e para fora do núcleo. Eles consistem de múltiplas cópias de aproximadamente 30 nucleoporinas (Nups) diferentes organizados em uma estrutura de simetria octagonal (BROHAWN et al., 2009). Mutações nos genes que codificam proteínas constituintes dos complexos poros nucleares e da maquinaria de tráfico nucleocitoplasmático parecem prejudicar as respostas vegetais de defesa, sugerindo a importância da dinâmica de translocação de proteínas para o interior do núcleo na regulação da expressão dos genes vegetais de defesa (CHEUNG; REDDY, 2012).

O espaço nuclear é compartimentalizado pelos corpos nucleares, que são estruturas dinâmicas, porém relativamente estáveis, compostos por proteínas e RNAs e não delimitados por membranas. Sua morfologia é dependente da diferenciação celular e do estado metabólico. Em plantas, os corpos nucleares mais investigados são os nucléolos, os fotocorpos, que contém fitocromos e proteínas

envolvidas na fotomorfogênese, e as manchas (*speckles*) nucleares (REDDY; DAY; GÖHRING; BARTA, 2012).

O nucléolo é a estrutura nuclear mais proeminente. Ele é montado ao redor de *clusters* de genes ribossomais (genes rDNA) repetidos *em tandem* e que são transcritos pela RNA polimerase I (DUNDR; MISTELI, 2001). As regiões contendo as seqüências em tandem de genes rDNA constituem o loci cromossomal denominado de regiões organizadoras de nucléolo (RONs) e são a base da organização estrutural do nucléolo responsáveis pelo direcionamento de todos os componentes de montagem e processamento necessários para a biogênese de ribossomos (CARMO-FONSECA; MENDES-SOARES; CAMPOS, 2000).

Quando observado por microscopia eletrônica, o nucléolo da maioria dos eucariotos superiores apresenta três componentes morfológicos principais: os centros fibrilares, o componente fibrilar denso e o componente granular. Os centros fibrilares abrigam várias seqüências em *tandem* presentes nas RONs. Dentre essas seqüências, apenas uma parte se encontra ativa, sendo aparentemente localizada na região mais periférica, estendendo-se até o interior do componente fibrilar denso (CARMO-FONSECA; MENDES-SOARES; CAMPOS, 2000).

Durante a biogênese das subunidades ribossomais, a RNA polimerase I transcreve os *clusters* repetidos dos genes de rDNA. Os pré-rRNAs, por sua vez, se movem para longe do *template* de DNA e são detectados inicialmente no componente fibrilar denso, onde são clivados e modificados pelas RNPs pequenas nucleolares (snoRNPs) e outras enzimas de processamento. Posteriormente, os rRNAs começam sua rota de montagem juntamente com proteínas ribossomais no DFC e continuam conforme passam pelo componente granular (LAM; TRINKLE-MULCAHY; LAMOND, 2005). Por último as partículas pré-ribossomais são translocadas para o citoplasma através dos poros nucleares (CARMO-FONSECA; MENDES-SOARES; CAMPOS, 2000).

Dessa forma, tal separação morfológica reflete o processo vetorial da biogênese dos ribossomos. Os centros fibrilares são cercados pelo componente fibrilar denso (DFC) e os componentes granulares irradiam a partir dos DFCs (DUNDR; MISTELI, 2001).

Nas células vegetais, o nucléolo parece ser mais esférico, sendo o DFC muito maior do que o das células de mamíferos (até 70% do volume nucleolar) e não muito denso. Diferentemente do que ocorre no DFC dos mamíferos, as unidades de transcrição do rDNA encontram-se dispersas em todo o nucléolo e formam estruturas que lembram árvores de abeto, descritas como “árvores de Natal linearmente compactadas” (GONZÁLEZ-MELENDI et al., 2001).

As manchas nucleares (*speckles*) são estruturas subnucleares enriquecidas em fatores de *splicing* do pré-RNA mensageiro e são localizados nas regiões de intercromatina do nucleoplasma de células de mamíferos. Observações em microscópio de fluorescência, revelam que essas estruturas aparecem de forma pontual e irregular, apresentando diversas formas e tamanhos. No entanto, ao microscópio eletrônico, as mesmas estruturas são observadas como agrupamentos de grânulos de intercromatina. Sua constituição em termos de proteína ou RNA-proteína pode variar, de forma que seus componentes circulam continuamente entre os *speckles* e outras regiões nucleares, incluindo locais de transcrição ativos, o que demonstra a composição dinâmica dessas manchas (LAMOND; SPECTOR, 2003).

O interesse nesses corpos subnucleares tem sido reacendido pelas descobertas recentes de que alguns deles contêm fatores envolvidos na transcrição e processamento do RNA (LAMOND; EARNSHAW, 1998). Além disso, embora aparentemente os *speckles* apresentem apenas alguns genes, a observação de sua freqüente presença próxima à locais de transcrição muito ativos sugere sua provável relação funcional com a expressão gênica (LAMOND; SPECTOR, 2012). Um exemplo é a proteína de *Arabidopsis* ASYMMETRIC-LEAVES2 (AS2), presente especialmente nas manchas nucleares de células epidérmicas foliares e que apresenta um papel-chave na regulação de muitos genes responsáveis pela formação foliar plana simétrica (LUO et al., 2012).

Em plantas, os *speckles* apresentam proteínas SR, ricas em serina e arginina. Essas proteínas constituem uma família de fatores de *splicing* que apresentam um motivo de ligação com o RNA e uma região SR e formam parte do spliceossomo, estando envolvidas na sua montagem e participando do reconhecimento de íntrons e éxons (GRAVELEY, 2000). Assim, de forma similar às células animais, em células vegetais, as proteínas SR também são localizadas em grânulos de intercromatina.

No entanto, diferentemente do que ocorre em animais, os *speckles* de vegetais apresentam quantidades muito pequenas de snRNA (NIEDOJADŁO et al., 2012).

Além das proteínas SR, os *speckles* apresentam ainda snRNPs e proteínas de splicing não-snRNP, fatores de transcrição e a terminação 3' de fatores de processamento e por isso são vistos como áreas de armazenamento e montagem que adquirem fatores de *splicing* para ativar locais de transcrição. Dessa forma, células que estão transcrevendo ativamente possuem *speckles* menores e em maior quantidades, enquanto a inibição da transcrição ou do *splicing* conduzem ao acúmulo de fatores de *splicing* e, portanto, *speckles* maiores (REDDY; DAY; GÖHRING; BARTA, 2012).

Os sinais ambientais e de desenvolvimento podem afetar a dinâmica das manchas nucleares, como é o caso de seu tamanho e do fluxo de algumas moléculas para dentro e para fora dessas estruturas. Alguns aspectos fundamentais relacionados aos *speckles* nucleares, como sua composição, biogênese, e especialização funcional de subpopulações ainda precisam ser esclarecidos (CHEUNG; REDDY, 2012).

Os corpos de Cajal (CBs) são partículas nucleares muito comuns e estão presentes em diversos reinos. São normalmente associados aos nucléolos e parecem estar envolvidos na maturação de snRNP e de ribonucleoproteínas nucleolares pequenas (ERHARDT; ADAMSKA; FRANCO, 2010). Eles foram descobertos há mais de 100 anos por Ramón e Cajal em tecidos neurais de vertebrados e descritos inicialmente como “corpos nucleolares acessórios”. Essa descoberta rendeu a Ramón e Cajal, juntamente com a descoberta do complexo de Golgi por Camillo Golgi, o prêmio Nobel em Fisiologia ou Medicina em 1906 para ambos os pesquisadores. Desde então, os corpos de Cajal têm sido encontrados em uma variedade de núcleos de animais e vegetais, indicando seu envolvimento em processos celulares básicos (GALL, 2000; CIOCE; LAMOND, 2005).

Essas estruturas constituem organelas subnucleares ubíquas em células animais e vegetais e apresentam tamanho que varia de menos de 0,2 μm até 2 μm ou ainda maiores, dependendo do tipo celular ou da espécie (CIOCE; LAMOND, 2005; BASSETT, 2012). Seu número também pode variar de acordo com o tipo celular e com o ciclo celular, normalmente sendo encontrados de 1 a 6 por núcleo

(BASSETT, 2012). Já sua identificação é realizada por meio de sua proteína-assinatura, a proteína coilina e os RNAs CB-associados (scaRNAs) (NIZAMI; DERYUSHEVA; GALL, 2010).

Em decorrência de sua composição de fibrilas espiraladas, revelada por estudos em nível ultraestrutural, os CBs também foram chamados de corpos espiralados (SMOLIŃSKI; KOLOWERZO, 2012).

Esses corpos são dinâmicos, sendo encontrados não somente no nucléolo, mas também no nucleoplasma da maioria das células vivas. Suas funções relacionam-se ao metabolismo de RNA, incluindo a transcrição, o *splicing*, a modificação estável do RNA, biogênese de pequenos RNAs, silenciamento gênico e manutenção de telômeros. Além disso, acredita-se que os CBs formem estações de ancoragem para cada uma das funções, de maneira que os processos individuais são realizados rapidamente e de maneira organizada (BASSETT, 2012).

Em plantas, os CBs são similares aos de núcleos de células animais, sendo que tanto os CBs livres quanto os perinucleolares apresentam elementos do sistema de *splicing* (SMOLIŃSKI; KOLOWERZO, 2012) e aparentemente se encontram presentes em todas as espécies e tipos celulares vegetais (SHAW; BROWN, 2004).

Em células vegetais foram descobertos ainda outros corpos nucleares específicos e que se encontram envolvidos no metabolismo de RNA e em respostas hormonais. Alguns exemplos são os corpos contendo ciclofilina; os corpos contendo o domínio HYL1, que constitui uma proteína de ligação com o RNA de fita dupla e está envolvida no metabolismo de microRNAs; os corpos contendo o domínio de fitocromos, cuja função ainda é desconhecida, mas parece estar envolvida na resposta à luz; e os corpos contendo o domínio AKIP-1 que são encontrados nos núcleos de células-guarda e apresentam função desconhecida, mas podem apresentar função na proteção contra o estresse celular (SHAW; BROWN, 2004).

Dessa forma, tanto na célula eucariótica vegetal quanto na animal, o núcleo é uma estrutura altamente dinâmica e complexa, contendo diversos subcompartimentos, vários tipos de DNA e RNA, e uma grande variedade de proteínas. As interações entre esses componentes dentro do núcleo formam parte de um sistema regulatório complexo que é apenas parcialmente compreendido (ERHARDT; ADAMSKA; FRANCO, 2010).

Figura 3 – (a) O núcleo e seus diversos compartimentos; (b) Diferenças entre o nucléolo de vegetais e de mamíferos

Fonte: Modificado de ERHARDT; ADAMSKA; FRANCO, 2010

O núcleo é uma das mais importantes organelas presentes na célula, uma vez que ele é responsável pela regulação de várias atividades biológicas, incluindo a expressão gênica e o transporte de fatores regulatórios, além da sinalização com diversos objetivos, como a resposta ao estresse (KOMATSU; TANAKA, 2005; REPETTO et al., 2008)

Até recentemente, as técnicas de microscopia e imunoquímica foram as mais usadas no estudo dessa estrutura. No entanto, tais técnicas apresentam uma grande desvantagem relacionada ao fato de que elas não podem identificar as interações de várias moléculas ao mesmo tempo. Assim, para o entendimento global dos mecanismos existentes dentro do núcleo vegetal, é necessária a aplicação de novas técnicas, dentre as quais a proteômica constitui um campo de pesquisa que pode resolver pelo menos em parte o problema do estudo de várias proteínas ao mesmo tempo (ERHARDT; ADAMSKA; FRANCO, 2010).

A subproteômica do núcleo e de seus compartimentos irá facilitar a anotação de proteínas nucleares. Já existem várias bases de dados disponíveis e o seu

desenvolvimento irá contribuir grandemente para melhoras na proteômica celular vegetal (ERHARDT; ADAMSKA; FRANCO, 2010).

3.3.1 Histonas

O DNA genômico é muito frágil. No entanto, ele se encontra estabilizado nas células vivas em decorrência das interações com proteínas e pode, portanto, ser transmitido de forma precisa para as células-filhas. Tal molécula é empacotada em cromossomos que, por sua vez, apresentam basicamente a mesma estrutura tanto em plantas quanto em animais. Os cromossomos encontram-se completamente condensados durante a metáfase e possuem uma variedade de regiões, como a heterocromatina, a eucromatina, as regiões organizadoras de nucléolos (RONs) e os centrômeros (WAKO; FUKUI, 2010).

As histonas são componentes da cromatina eucariótica altamente conservados e são classificadas em quatro subcategorias de histonas centrais: H2A, H2B, H3 e H4 e histona de ligação H1, sendo que as histonas centrais estão entre as proteínas mais abundantes dos organismos eucariotos (YI et al., 2006).

A proteção eficiente e o empacotamento ordenado do DNA genômico são obtidos por meio dos nucleossomos, que são complexos de DNA e proteínas histonas. Os nucleossomos são compostos por histonas centrais, sendo as histonas H2A, H2B, H3 e H4 presentes em pares, estando juntas com a histona de ligação H1. Estas proteínas são evolutivamente conservadas, o que indica seu papel fundamental e sua importância na manutenção do nucleossomo. As histonas de ligação H1 são essenciais na conexão das partículas para formar o cromatossomo, conduzindo à estrutura de cordão de contas, o que ajuda no empacotamento do DNA no interior do núcleo e na proteção do mesmo contra danos físicos e enzimáticos (WAKO; FUKUI, 2010).

Modificações epigenéticas, como acetilação, fosforilação, metilação, ubiquitinação e ribosilação de ADP das histonas centrais altamente conservadas influenciam o potencial genético do DNA. O potencial regulatório enorme das modificações de histonas é ilustrado pela quantidade de marcadores epigenéticos encontrados ao longo do genoma (BÁRTOVÁ et al., 2008). A acetilação e a metilação estáveis, por exemplo, em resíduos específicos de lisina das histonas H3

e H4 apresentam papéis críticos para a regulação epigenética da transcrição e para a definição da estrutura dos cromossomos e de suas regiões eucromáticas ou heterocromáticas (WAKO; FUKUI, 2010).

3.3.2 Helicases

Um grande número de processos genéticos precisa do desespiralamento da dupla fita ou das regiões de bases pareadas de DNA/DNA, RNA/RNA ou RNA/DNA (AUBOURG; KREIS; LECHARNY, 1999).

Assim, apesar dos genomas serem geneticamente estáveis em todos os organismos vivos, incluindo as plantas, o duplex de DNA precisa ser parcialmente desenrolado por um pequeno período de tempo para criar um *template* de DNA, que é necessário para a maioria de suas funções celulares importantes, incluindo a replicação, o reparo, a recombinação e a transcrição (TUTEJA, 2003).

As helicases são proteínas motoras que usam a energia da hidrólise no NTP (nucleosídeo trifosfato) para dissociar as pontes de hidrogênio presentes entre os duplex de ácidos nucleicos e também para realizar a interrupção de outras interações não-covalentes entre pares de base complementares. Todas as helicases compartilham pelo menos três propriedades bioquímicas: (a) ligação a ácidos nucleicos; (b) ligação e hidrólise NTP/dNTP; (c) desespiralamento do duplex de ácidos nucleicos dependente da hidrólise NTP/dNTP na direção 3' → 5' ou 5' → 3' (HALL; MATSON, 1999; TUTEJA; TUTEJA, 2004).

No entanto, apesar da reação bioquímica básica catalisada por todas as helicases ser similar, existem múltiplos mecanismos do desenrolamento dos ácidos nucleicos catalisados pelas helicases que podem ter evoluídos para auxiliar em papéis específicos em vários processos biológicos e para acomodar estados de montagem ativos diferentes (HALL; MATSON, 1999).

As helicases de RNA representam uma grande família de proteínas que estão relacionadas ao metabolismo do RNA e à modulação da sua estrutura, e portanto, estão envolvidas na síntese de RNA, transcrição, *splicing* do pré-mRNA, biogênese de ribossomos, transporte nucleocitoplasmático, tradução e decaimento do RNA (TUTEJA, 2003; CRUZ, KRESSLER; LINDER, 1999).

3.3.3 Proteínas de ligação MAR

A hipótese da matriz nuclear propõe a existência de uma armação estrutural no núcleo eucariótico que seja similar ao citoesqueleto, sendo que processos nucleares como replicação, transcrição e *splicing* são organizados espacialmente e associados à esta armação. Além disso, a cromatina se liga à matriz nuclear por meio de fragmentos de DNA específicos denominados regiões de ligação à matriz nuclear (MARs). As proteínas filamentosas de ligação MAR, como a MFP1, por sua vez, estão associadas a estruturas semelhantes aos *speckles*, que se encontram na periferia nuclear, e apresentam papel na ligação da cromatina a locais específicos da periferia nuclear (GINDULLIS; MEIER, 1999).

Dessa forma, as proteínas de ligação MAR são consideradas proteínas longas e enroladas em espiral que se ligam à dupla fita de DNA e são específicos de plantas. Além disso, tais proteínas apresentam um domínio N-terminal contendo dois domínios hidrofóbicos conservados, um domínio central em haste enrolada em espiral e um domínio C-terminal de ligação ao DNA (SAMANIEGO et al., 2006).

3.4 Programas de predição subcelular

A localização subcelular fornece informação acerca da função proteica e contribui para o entendimento das interações entre proteínas, além da compreensão das rotas de sinalização dentro da célula (BRAMEIER; KRINGS; MACCALLUM). Assim, tal conhecimento é essencial para elucidar a função proteica. No entanto, é difícil determinar a localização subcelular de uma proteína nas células eucarióticas por procedimentos experimentais “*high-throughput*” utilizando somente a identificação da proteína. Nesse aspecto, os procedimentos computacionais tornam-se necessários para a anotação da localização subcelular das proteínas em projetos genômicos de larga escala (PIERLEONI et al., 2006).

As predições são valiosas por serem fáceis de usar, além de poder predizer localizações que são inesperadas e desta forma estabelecerem novas linhas de investigação experimental (MILLAR et al., 2009).

No entanto, a habilidade de direcionamento de uma proteína pode ser avaliada tanto pela predição quanto por análise experimental. Nesse aspecto, os preditores podem fornecer diferentes resultados em decorrência do tipo de algoritmo

de predição usado (redes neurais, máquinas de vetores de apoio, dentre outros), do conjunto de seqüências de entrada usadas no seu desenvolvimento, da gama de localizações preditas em cada programa, da hidrofobicidade predita da seqüência de entrada e da significância do limite (*cut-off*) que é usado na análise dos dados de saída (MILLAR et al., 2009; EMANUELSSON et al., 2007; CHOU; SHEN, 2008).

A forma mais simples de atribuir os dados genômicos obtidos por meio de estudos em plantas a um contexto relevante para a biologia celular é pelo do uso de algoritmos de bioinformática para prever a localização das proteínas. Existe um grande número de programas disponíveis para esse objetivo, baseados primariamente nas similaridades de seqüência, de forma que as proteínas podem ser preditas como localizadas no núcleo, mitocôndria, cloroplasto, dentre outras organelas. No entanto, a comparação de dados obtidos por análises de sequenciamento de larga escala freqüentemente mostra pouco consenso entre os programas, de forma que apenas um pequeno grupo de proteínas são consistentemente preditos por vários programas como localizados em uma determinada organela ou compartimento (HEAZLEWOOD et al., 2004; TANAKA et al., 2004; HAYNES; ROBERTS, 2007). A razão para tal discrepância pode ser, em primeiro aspecto, a falta de informação de alta qualidade para desenvolver os algoritmos, reforçada pela pouca sobreposição entre os dados experimentais e predições consenso (aproximadamente 50% ou menos). Apesar desses aspectos, as ferramentas de bioinformática ainda são utilizadas, podendo frequentemente apontar para a localização correta de uma proteína, ou pelo menos fornecer informações suficientes para direcionar novos experimentos (HAYNES; ROBERTS, 2007).

Apesar desses aspectos, uma vez que produtos gênicos antes desconhecidos, ou mesmo conhecidos porém não antes relacionados à organela em questão, são identificados no nível proteico, uma análise *in silico* da estrutura primária pode revelar dicas importantes para uma potencial função da proteína. Dessa forma, com relação à localização subcelular de um produto gênico, existem diversas ferramentas de predição disponíveis e que podem ser acessadas pela internet (DREGER, 2003). Alguns exemplos são WoLF PSORT (<http://wolfpsort.org/>), Plant-mPloc (<http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/>), BaCellO

(<http://gpcr.biocomp.unibo.it/bacello/pred.htm>), SherLoc2 (<http://abi.inf.uni-tuebingen.de/Services/SherLoc2>), e PSORT (<http://psort.hgc.jp/form.html>), que foram escolhidos por apresentarem o núcleo como uma de suas localizações preditas e por serem de acesso livre.

3.4.1 PSORT

O PSORT prediz a localização subcelular de proteínas explorando o conhecimento acerca da classificação protéica. Assim, ele precisa de uma sequência completa de aminoácidos e a origem da sequência, que pode ser bactéria Gram-positiva, bactéria Gram-negativa, levedura, animal ou planta. De acordo com a categoria selecionada, o programa atribui um conjunto de locais candidatos. O programa então calcula os valores das variáveis em destaque que refletem as várias características da sequência. Então, ele tenta interpretar o conjunto de valores obtidos e estima a probabilidade da proteína ser classificada em cada local candidato e no final ele apresenta alguns dos locais mais prováveis (NAKAI; HORTON, 1999).

3.4.2 Plant-mPloc

O Plant-mPloc foi desenvolvido pela integração da informação obtida por meio do *Gene Ontology*, juntamente com a informação dos domínios funcionais e com a informação evolutiva sequencial por três modos diferentes de composição de pseudo aminoácidos. Ele pode ser usado para identificar proteínas vegetais que se encontram em 12 locais subcelulares: membrana celular, parede celular, cloroplasto, citoplasma, retículo endoplasmático, extracelular, complexo de Golgi, mitocôndria, núcleo, peroxissomo, plastideos e vacúolo. Comparado a outros métodos de predição da localização subcelular de proteínas vegetais, o Plant-mPloc é poderoso e flexível, apresentando a capacidade de lidar com proteínas que apresentam múltiplas localizações subcelulares (CHOU; SHEN, 2010).

3.4.3 WoLF PSORT

O WoLF PSORT é uma extensão do programa de predição de localização subcelular de proteínas PSORT II. Ele converte as sequências de aminoácidos das

proteínas em características de localização numéricas, com base em sinais classificadores, classificação de aminoácidos e motivos funcionais, como os motivos de ligação ao DNA. Depois da conversão, um simples classificador de vizinho k-mais próximo é usado para a predição. A evidência de cada predição é mostrada de duas formas: (a) uma lista de proteínas de localização conhecida com características de localização mais similares à sequência em questão; (b) uma tabela com informações detalhadas acerca de características individuais de localização. A primeira página que o serviço oferece como resposta à sequência em questão apresenta o sumário do resultado em apenas uma linha (HORTON et al., 2007).

Esse programa classifica as proteínas em mais de dez localizações subcelulares, incluindo a localização dupla, como é o caso de proteínas que mudam de localização entre citoplasma e núcleo. De acordo com os estudos de validação cruzada, a precisão e a especificidade do programa é aproximadamente 70% para o núcleo, mitocôndria, citosol, membrana plasmática, compartimento extracelular (em plantas) e cloroplasto. Em outras localizações, como os peroxissomos e o complexo de Golgi, a sensibilidade é muito baixa, mas mesmo assim útil em alguns casos (HORTON et al., 2007).

3.4.4 BaCelLo

O BacelLo prediz cinco classes de localização subcelular: rota secretória, citoplasma, núcleo, mitocôndria e cloroplasto. Ele é baseado em diferentes Máquinas de Vetores de Apoio, ou SVM (*Support Vector Machines*), organizadas em uma árvore de decisão. O sistema explora a informação originada da sequência resíduo e da informação evolutiva presente nos perfis de alinhamento. Além disso, analisa a composição completa da sequência e a composição tanto do C- quanto do N-terminal. O conjunto de desenvolvimento é curado, com o objetivo de diminuir a redundância. Três preditores reino específicos se encontram presentes: animais, plantas e fungos, sendo que a certeza na distribuição das proteínas vegetais chega a atingir 67% (PIERLEONI et al., 2006).

3.4.5 SherLoc2

O SherLoc 2 é um sistema de predição de localização subcelular de proteínas de alta precisão. Ele é aplicável em proteínas de animais, plantas e fungos e cobre as principais localizações subcelulares dos organismos eucariotos. O programa integra várias características baseadas em sequência, bem como características baseadas em texto e incorpora perfis filogenéticos e os termos do Gene Ontology, com o objetivo de melhorar a performance da predição. O SherLoc 2 consegue alcançar uma precisão geral na classificação de até 95% em validação cruzada de 5 vezes (BRIESEMEISTER et al., 2009).

4 MATERIAL E MÉTODOS

4.1 Material Vegetal

Plantas de cana-de-açúcar (*Saccharum spp*), cultivar SP 80 – 3280, foram cultivadas em vasos de 10 litros e mantidas em casa de vegetação em delineamento experimental inteiramente casualizado (DIC), constituído por 54 plantas. Foram utilizadas apenas as folhas F+1 de plantas de 1 e 4 meses de idade, onde cada repetição biológica consistiu-se de 18 plantas e, portanto, 18 folhas, segundo a Tabela 1.

As folhas foram coletadas, armazenadas em gelo e transportadas para o laboratório para utilização imediata, uma vez que o protocolo utilizado exige material fresco para que os núcleos extraídos possam estar intactos.

Tabela 1 – Distribuição das plantas em relação às suas respectivas repetições

Repetição	Plantas
1	27A, 1B, 20B, 11A, 23B, 21B, 19A, 27B, 18A, 8A, 11B, 26B, 2B, 17A, 25A, 26A, 13B, 9A
2	7A, 28B, 18B, 20A, 22A, 24A, 4A, 3B, 12B, 16B, 10A, 5B, 22B, 6B, 16A, 17B, 13A, 7B
3	5A, 24B, 4B, 2A, 12A, 6A, 15B, 14B, 15A, 21A, 10B, 8B, 28B, 9B, 3A, 1A, 14A, 25B

4.2 Isolamento de núcleos

Os núcleos foram isolados de acordo com o protocolo de Folta e Kaufman (2000), com modificações, conforme descrito a seguir.

4.2.1 Tratamento e homogeneização do tecido

Em cada repetição biológica utilizou-se entre 29 e 35g de material vegetal, (folhas +1 de plantas de 1 e 4 meses de idade). O tecido foi mantido em gelo e cortado com uma tesoura de floricultura (Supercort - Tramontina) em pedaços pequenos de aproximadamente 2cm² e colocados simultaneamente em um béquer contendo éter gelado, que constitui um solvente para clorofila, também mantido em gelo. O volume de éter utilizado foi o suficiente para cobrir totalmente o tecido, permanecendo a ± 4 -5 cm acima do limite de tecido.

O tecido foi mantido no éter por 10 min, sob agitação nos primeiros 5 min e picotado de forma irregular em pedaços menores, sem agitação, nos 5 min seguintes.

O éter foi retirado e o tecido lavado duas vezes em tampão de extração (1,0 M hexilenoglicol (2-metil-2,4-pentandiol), 0,5 M PIPES-KOH (pH 7,0), 10 mM MgCl₂, 5mM 2-mercaptoetanol), e em seguida ressuspendido no mesmo tampão de extração, em volume duas vezes superior ao volume do tecido. O material ressuspendido foi homogeneizado por três vezes de 30s com ajuda de um liquidificador em potência mínima.

O homogeneizado foi filtrado usando-se uma camada dupla de *miracloth* com poros de 22-25 µm de diâmetro, com o objetivo de eliminar materiais particulados. Em seguida adicionou-se Triton X-100 (25%), gota a gota, com leve agitação, de forma que sua concentração final atingisse 1% (v/v). O homogeneizado foi submetido à extração com gradiente de percoll.

Todos os passos da extração foram realizados a 4°C para prevenir ao máximo a degradação do tecido.

4.2.2 Extração com gradiente de *Percoll*

Nesta etapa, 6 mL de *Percoll* 80% (v/v) preparado em tampão de gradiente (0,5 M hexilenoglicol (2-metil-2,4-pentandiol), 0,5 M PIPES-KOH (pH 7,0), 10 mM MgCl₂, 5mM 2-mercaptoetanol, 1% Triton X-100) foram inicialmente pipetados na parte mais inferior de um tubo de acrílico de fundo redondo (50 mL). Imediatamente acima desta camada colocou-se, gentilmente, uma camada de 6 mL de percoll 35% também preparado em tampão de gradiente.

O homogeneizado foi colocado vagarosamente acima do gradiente montado (35% percoll), para evitar que as fases se misturassem. Os tubos foram então centrifugados a 2000xg a 4°C por 1 h, em rotor A-4-81 da centrífuga Eppendorf® 5810 R.

Após a centrifugação, a camada de núcleos consistiu-se de uma faixa escura localizada entre as camadas de *Percoll* 80% e 35%. A fração de homogeneizado encontrada acima da camada de *Percoll* 35% foi retirada cuidadosamente com uma pipeta *Pasteur* até a proximidade do *Percoll* 35%. A camada de núcleos foi então

removida tanto da interface quanto das paredes do tubo cuidadosamente com a ajuda de uma pipeta *Pasteur*, sendo que em sua extremidade foi colocada uma ponteira (200 μ L), de forma a tornar o processo mais preciso, evitando ao máximo a remoção conjunta de *Percoll* sem a presença dos mesmos.

A fração nuclear foi diluída em 35 mL de tampão de gradiente e inserida novamente em um tubo de acrílico de 50 mL. No fundo do mesmo (abaixo da fração nuclear diluída em tampão de gradiente) colocou-se ainda uma camada de 6mL *percoll* 35% em tampão de gradiente. Os tubos foram novamente centrifugados a 2000xg a 4°C por 20 min.

O pellet enriquecido de núcleos foi ressuspensionado em aproximadamente 35 mL de tampão de gradiente e submetido novamente à extração com gradiente de *percoll*. Desta forma, o gradiente foi mais uma vez montado, de maneira idêntica à anterior, isto é, 6 mL de *percoll* 80% (v/v) no fundo do tubo e 6 mL de *percoll* 35% (v/v) adicionado imediatamente acima do *percoll* 80%, ambos preparados em tampão de gradiente. O *pellet* ressuspensionado foi colocado de forma cuidadosa acima do *percoll* 35% e os tubos foram centrifugados a 2000xg e 4°C por 1 h.

Após a centrifugação, formou-se novamente a camada mais escura entre as camadas de *percoll* 80% e 35%. A fração encontrada acima do *percoll* 35% foi retirada e a camada de núcleos encontrada na interface e na parede do tubo, entre as duas camadas de *percoll* foi removida cuidadosamente, da mesma forma descrita anteriormente.

Esta fração nuclear também foi ressuspensionada em aproximadamente 35 mL de tampão de gradiente e colocada em tubo de acrílico de 50 mL, onde 6 mL de *percoll* 35% em tampão de gradiente também foi adicionado, como anteriormente, no fundo do tubo, abaixo da fração de núcleos ressuspensionados. Os tubos também foram centrifugados a 2000xg e 4°C por 20 min.

O *pellet* de núcleos obtido foi ressuspensionado em cerca de 3 volumes de tampão de armazenagem (50mM Tris-HCl (pH 7,8), 10mM 2-mercaptoetanol, 20% glicerol, 5mM MgCl₂, 0,44M sacarose) e armazenado a -80°C até a utilização.

Todas as centrifugações foram realizadas utilizando-se o *swing bucket*, com velocidades mínimas tanto de aceleração quanto desaceleração, diminuindo, desta forma, a chance de romper o gradiente de *percoll*.

4.3 Verificação da integridade e pureza da fração nuclear isolada

A integridade dos núcleos isolados foi verificada por meio da coloração da fração nuclear com orceína aceto láctica e *4,6-diamidino-2-phenylindole hydrochloride* (DAPI), seguida de observação em microscópio de luz e de fluorescência, respectivamente.

4.3.1 Microscopia de luz

Para a microscopia de luz, as lâminas foram preparadas utilizando-se 30 µL da suspensão nuclear e 5 µL de orceína aceto láctica a 1%.

A fração nuclear permaneceu no corante por 20 min. Em seguida, utilizaram-se 10 µL da suspensão já corada para a preparação da lâmina.

As observações, captura de imagens e medidas de diâmetro nuclear foram feitas em microscópio Zeiss Axioskop 2, com câmera digital acoplada a um computador no Núcleo de Apoio à Pesquisa em Microscopia Eletrônica Aplicada à Agricultura (NAP/MEPA) da ESALQ.

4.3.2 Microscopia de fluorescência

Nesta etapa, os núcleos isolados foram analisados por coloração com *4,6-diamidino-2-phenylindole hydrochloride* (DAPI) 0,1 µL/mL.

Desta forma, 30 µL de suspensão nuclear foram corados por 15 min em 30 µL de DAPI (Sigma) 0,1 µL/mL preparado em 0,1M PBS (pH 7,4). Em seguida, a suspensão nuclear foi lavada duas vezes com PBS. Então, os núcleos foram recuperados por centrifugação e ressuspensos em 100 µL de PBS.

Aproximadamente 30 µL foram utilizados na preparação da lâmina, sendo esta preparada por meio da colocação da suspensão na lâmina, com a posterior colocação da lamínula, realizando-se leve pressão, para melhor distribuição do material.

As observações e captura de imagens foram realizadas em microscópio de epifluorescência Zeiss Axio Imager M2, equipado com filtro nº49, apropriado à visualização da coloração DAPI, e com câmera digital acoplada a um computador no Laboratório de Interações em Insetos (IInLab) da ESALQ.

4.4 Extração, solubilização e quantificação de proteínas

As proteínas nucleares foram extraídas da suspensão enriquecida de núcleos utilizando-se o reagente TRI[®] (Sigma) de acordo com as instruções do fabricante com algumas modificações.

Desta forma, 100 µL de cada amostra foi utilizado para cada 150 µL de reagente TRI[®]. Quatro lavagens com 0,3M de guanidina hidrocloreídica em 95% de etanol e duas lavagens com etanol 100% foram procedidas.

As proteínas foram solubilizadas em tampão de solubilização (7M Uréia, 2M Tiuréia, 10mM DTT, 0,4% Triton) durante repetidos ciclos de pipetagem alternados com *spin* em centrífuga e 20 min de sonicação.

A quantificação das proteínas foi realizada utilizando-se do Kit BioRad Protein Assay (BioRad)[®], de acordo com o método de Bradford (BRADFORD, 1976).

Inicialmente uma curva-padrão de albumina sérica bovina (BSA) foi estabelecida, em espectrofotômetro (U-3300) Hitachi[®] a 595nm. Em seguida, prepararam-se triplicatas em cubetas de vidro contendo 5 µL de extrato protéico, 800 µL de reagente Bradford diluído em água milliQ (1:4) e 1 mL de água milliQ. Leituras de absorbância foram realizadas a 595 nm. A concentração das proteínas totais foi obtida pela interpolação dos valores obtidos com a curva-padrão.

4.5 Análise de *Western blot*

Para as análises de *Western blot*, as proteínas foram separadas em gel SDS-PAGE 12,5%, sendo inicialmente fervidas por 1 min, juntamente com tampão de desnaturação (457µL água milliQ, 125 µL 0,5M Tris-HCl (pH 6,8), 100 µL Glicerol, 200 µL SDS 10%, 50 µL 2-mercaptoetanol, 50 µL de azul de bromofenol 1%). A corrida de eletroforese foi realizada a 40V por meia hora, para que as proteínas entrem no gel de separação, e à 80V pelo restante do tempo.

As amostras foram eletrotransferidas para uma membrana PVDF Hybond-P (Amershan Pharmacia Biotech[®]), pré-equilibrada com metanol por 5 min, e usando-se tampão de transferência gelado (49 mM Tris, 39 mM glicina, 20% metanol). A transferência foi efetuada em cuba Mini-TransBlot (BioRad[®]), e a corrida realizada por 12 h, usando voltagem constante de 20V, à 4°C.

Após a transferência, a membrana foi lavada duas vezes por 10 min em TBS-T (8g NaCl, 20 mL 1M Tris-HCL (pH 7,6), 980 mL água milliQ, Tween 20 0,1% (v/v)) e então mantida em solução bloqueadora (5g leite desnatado, 100 mL TBS-T) por 1h. A solução bloqueadora foi descartada e a membrana lavada duas vezes em TBS-T.

O anticorpo primário diluído em TBS-T foi adicionado e a membrana incubada sob agitação suave e constante por 1 h. Após a incubação, a solução contendo o anticorpo foi retirada e a membrana lavada quatro vezes com TBS-T (1x por 15 min e 3x por 5 min). O anticorpo secundário foi então adicionado, também diluído em TBS-T e a membrana novamente incubada sob agitação suave e constante por 1 h, sempre à temperatura ambiente.

A solução com anticorpo secundário foi drenada e a membrana novamente lavada quatro vezes com TBS-T (1x por 15 min e 3x por 5 min).

Para a detecção, utilizou-se o kit ECL Plus - Western Blotting Detection Reagents (Amershan Biosciences®). Primeiramente retirou-se o excesso de TBS-T, em seguida pipetou-se sobre a membrana a mistura das soluções A e B do Kit, na proporção 40:1. A membrana, juntamente com as soluções, foi incubada por 5 min à temperatura ambiente. O excesso de reagente de detecção foi drenado. A membrana que se encontrava em plástico foi fechada e colocada em um cassete de revelação juntamente com o filme, que foi revelado nos tempos adequados para cada anticorpo.

Cinco anticorpos primários foram utilizados, dois dos quais marcadores nucleares, Anti-acetyl-Histone H1.4 (Ac-Lys²⁶) (Sigma®) e Anti-PCNA (proliferating cell nuclear antigen) (Sigma®), e três marcadores não nucleares (cloroplasto, citoplasma), RuBisCO, PEPCase (fosfoenol piruvato carboxilase), OEE1 (oxygen evolving enhancer protein1). O anticorpo secundário utilizado foi o anti-IgG de coelho, conjugado de peroxidase (Sigma®).

Nos géis foram utilizados 30µg de proteínas nucleares e 30µg de proteínas totais para o *Western blot* da Rubisco, PEPCase, OEE1 e Histona. No entanto, para o *Western blot* da PCNA utilizaram-se 60µg de proteínas nucleares e 60µg de proteínas totais, pelo fato da PCNA ser menos abundante na célula.

As concentrações dos anticorpos foram otimizadas, sendo estabelecidas como ideais as concentrações apresentadas na tabela 2.

Tabela 2 – Concentração de anticorpos

Anticorpo	Concentração
Anti-acetyl-Histone H1.4 (Ac-Lys ²⁶)	1:5000
Anti-PCNA	1:5000
Anti-RuBisCO	1:20000
Anti-PEPCase	1:10000
Anti-OEE1	1:5000
Anti-IgG	1:10000

4.6 SDS-PAGE

Após a quantificação, as amostras foram fervidas por 1 min juntamente com tampão de desnaturação e separadas em gel SDS-PAGE, composto por uma porção menor de 4% denominada *stacking* ou gel de entrada (750 µL de Acrilamida/Bis-Acrilamida (30%/2,67%), 3,05 mL água milliQ, 1,25 mL 0,5M Tris-HCl (pH 6,8), 25 µL de APS 10%, 10 µL de temed) e uma porção maior de 12,5% chamada *running* ou gel de separação (4 mL de Acrilamida/Bis-Acrilamida (30%/2,67%), 3,35 mL água milliQ, 2,50 mL 1,5M Tris-HCl (pH 8,8), 50 µL de APS 10%, 10 µL de temed). A corrida de eletroforese foi realizada a 40V por meia hora, período em que a amostra se encontrava no gel *stacking*, e à 80V pelo restante do tempo (amostra no *running*).

Ao término da corrida, os géis permaneceram por 1h em solução fixadora (40% (v/v) etanol, 10% (v/v) ácido acético) e em seguida foram mantidos *overnight* em solução de *Coomassie Brilliant Blue* (G250) preparada de acordo com Candiano et al., 2004. O excesso de corante foi retirado por meio de sucessivas lavagens com água milliQ. Os géis foram digitalizados em scanner Image Scanner 3 – GE.

As repetições biológicas de cada tratamento (plantas de 1 e 4 meses) foram aplicadas no mesmo gel, com duas canaletas contendo 30µg de proteína para cada repetição, de maneira que, ao final, cada repetição foi composta por 60µg de proteína, como demonstrado no item 5.4.

4.7 Digestão *In gel* de proteínas

Cada repetição dos géis foram divididos em 20 bandas similares (mesma posição no gel), com exceção do gel da extração-teste, que foi dividido em 30 bandas. Cada banda foi retirada e cortada, com o auxílio de um bisturi, em segmentos de aproximadamente 1mm³. Os fragmentos foram acondicionados em tubos de microcentrífuga, nos quais em seguida adicionou-se solução de descoloração (50% (v/v) acetonitrila (ACN) e 25mM de bicarbonato de amônio (AMBIC) pH 8) por sucessivas vezes, sempre com agitação no *vortex*, até a retirada total do corante.

A solução de descoloração foi removida e os fragmentos de gel foram desidratados com 100% ACN. Em seguida, os fragmentos foram reduzidos por reidratação com solução 20 mM DTT em 50 mM AMBIC a 56 °C por 40 min. Para a etapa de alquilação, o líquido excedente foi retirado e substituído por 55mM IAA em 50mM AMBIC. Nesta etapa, os fragmentos foram mantidos no escuro por 30 min à temperatura ambiente.

O sobrenadante foi removido e os fragmentos lavados com 25 mM AMBIC e novamente desidratados com 100% ACN. O resíduo remanescente no gel foi evaporado à temperatura ambiente.

Após a desidratação, os fragmentos foram reidratados com solução de 50 ng.µL⁻¹ de tripsina (Promega V5111) preparada em 50 mM AMBIC, pH 8,0, por 15 min a 4° C, sendo usados 400ng de tripsina para cada banda. Após a reidratação completa, os fragmentos foram cobertos com 25 mM de AMBIC, e a digestão realizada a 37 °C por 16 h.

A ação da tripsina foi interrompida pela adição de solução bloqueadora (50% (v/v) ACN, 5% (v/v) ácido fórmico, 45% (v/v) água milliQ). Os peptídeos foram então eluídos dos fragmentos por sucessivas lavagens com as soluções de eluição I (60% (v/v) metanol, 1% (v/v) ácido fórmico, 39% (v/v) água milliQ) e eluição II (50% (v/v) ACN, 1% (v/v) ácido fórmico, 49% (v/v) água milliQ), até a lavagem final com 100% ACN. Em cada etapa de eluição, os fragmentos foram sonicados (sonicador Branson 3510) por 15 min a 40°C e o sobrenadante de cada lavagem foi recolhido em um *vial*, sendo o volume reduzido em concentrador a vácuo (Eppendorf® modelo 5301) à temperatura ambiente.

Os peptídeos concentrados foram ressuspensos em 30 μL de formiato de amônio 20 mM, pH 10 e submetidos à análise por espectrometria de massas.

4.8 Identificação de proteínas por espectrometria de massas

A mistura de peptídeos foi submetida ao fracionamento por cromatografia líquida de ultra pressão em sistema UPLC 2D nanoACQUITY (Waters®). Os peptídeos foram separados usando-se uma pré-coluna C18 (XBridge C18 300 μm x 50mm), 5 μm seguida por uma coluna *trap* de fase reversa C18 (Symmetry C18, 180 μm x 20mm, 5 μm) e finalmente uma coluna analítica (BEH130 C18, 150 μm x 100mm, 1.7 μm). A configuração dessas colunas permitiu o fracionamento em duas dimensões utilizando-se um sistema binário de bombas e um gradiente de pH para a eluição dos peptídeos. O gradiente começou com 3% de B1 (100% Acetonitrila) e aumentou para 65% B em 5 min, seguido por equilíbrio com 97% A1 (100% Água, 0,1% ácido fórmico). Os peptídeos foram eluídos em coluna de captura com um gradiente linear que iniciou com 10% B2 (100% Acetonitrila, 0,1% ácido fórmico), aumentou para 40% por 26 min e depois 85% por 10 min para eluição dos peptídeos através da coluna analítica. Adotou-se uma vazão de 400 $\text{nL}\cdot\text{min}^{-1}$ com tempo de execução total de 45 min. As colunas foram re-equilibradas para a condição inicial de 90% A2 (100% água, 0,1% ácido fórmico) e 10% B2. Nano-ESI-MS/MS foi realizado no Synapt G2 HDMS (Waters®), empregando-se ionização positiva. Como referência para o espectrômetro de massa, foi utilizado 100 fmol de GFP (*human (Glu¹)-Fibrino-Peptide*) a um fluxo constante de 0.35 $\mu\text{L}\cdot\text{min}^{-1}$.

Figura 4 – Fluxograma do processo de enriquecimento e identificação de proteínas nucleares

4.9 Análise e processamento dos dados

Os dados foram processados utilizando-se o programa ProteinLynxGlobalServer (PLGS – Waters) versão 2.5v, com os seguintes parâmetros de tolerância: mínimo de ions fragmentos consecutivos (y^+ , b^+) por peptídeo igual a dois; mínimo de fragmentos consecutivos por proteína igual a 5; mínimo de peptídeos (identificados pelos critérios anteriores) por proteína igual a

1; apenas uma clivagem perdida permitida. Foram considerados *scores* acima de 30,000, sendo que o *score* é proporcional à probabilidade logarítmica natural de uma proteína ser aleatória.

A identificação dos clusters foi realizada por meio do projeto de EST de cana-de-açúcar, o SUCEST – *Sugarcane Expressed Sequence Tag Project* – (<http://sucest-fun.org/wsapp>). Em casos em que o SUCEST não apresentou resultados, a sequência da proteína fornecida pelo mesmo foi alinhada manualmente contra o Uniprot (<http://www.uniprot.org/>), uma vez que este foi o banco escolhido para a identificação de homólogos no SUCEST, pelo fato daquele consistir em um banco de proteínas curado. O banco SUCEST foi escolhido por constituir o maior projeto de EST de cana-de-açúcar, que gerou cerca de 240.000 ESTs de 26 bibliotecas de cDNA de cana de diferentes tecidos, como folhas, raízes e entrenós, e estádios (VETTORE et al., 2001) e, representa, portanto, o banco de dados mais completo para análises moleculares em plantas de cana-de-açúcar.

Para a predição subcelular das proteínas nucleares foram escolhidos cinco programas de acesso livre, que apresentam o núcleo como uma de suas organelas preditas, além de permitirem a especificação para células vegetais. Os programas foram WoLF PSORT (W) (<http://wolfpsort.org/>), Plant-mPloc (PI) (<http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/>), BaCelLo (B) (<http://gpcr.biocomp.unibo.it/bacello/pred.htm>), SherLoc2 (S) (<http://abi.inf.uni-tuebingen.de/Services/SherLoc2>) e PSORT (P) (<http://psort.hgc.jp/form.html>). As proteínas consideradas nucleares por pelo menos um destes programas foram apresentadas.

A classificação funcional das proteínas foi realizada com base no *Gene Ontology*, pela utilização do *Interpro* (<http://www.ebi.ac.uk/interpro/>) e *Pfam* (<http://pfam.sanger.ac.uk/>) (BURGE et al., 2012; HUNTER et al., 2011; QUEVILLON et al., 2005).

Foi realizada uma análise estatística descritiva, uma vez que os dados são qualitativos.

5 RESULTADOS E DISCUSSÃO

5.1 Integridade dos núcleos enriquecidos

Na área da subproteômica, um dos critérios mais importantes a ser levado em consideração é a pureza do compartimento analisado. Em decorrência disso, a integridade do proteoma subcelular é altamente dependente da qualidade do protocolo usado para o isolamento (PANDEY et al., 2006). Em plantas, a separação de núcleos altamente purificados é uma tarefa difícil, uma vez que este procedimento pode comprometer a quantidade de proteínas obtidas (PANDEY et al., 2006; CHOUDHARY et al., 2009).

No presente trabalho, os núcleos foram isolados por meio da utilização do gradiente de *percoll* de acordo com Folta e Kaufman (2000) com modificações. Inicialmente, duplicou-se o tempo de centrifugação visando a formação de uma faixa de núcleos isolados mais uniformes. Em uma segunda etapa, adicionou-se mais um ciclo de centrifugações, com o objetivo de reduzir a presença de debris celulares provenientes de outras organelas e que ainda pudessem estar presentes.

Tal procedimento foi executado pois, de acordo com Millar (2004), os gradientes precisam ser repetidos e refinados para reduzir a contaminação de forma mais completa, especialmente com relação à retirada da faixa enriquecida com a organela de interesse, e, que nesse caso, se localiza na interface entre o *percoll* 35 e 80%.

A integridade dos núcleos isolados foi observada através da coloração com orceína acetoláctica 1%, que constitui um corante usado em preparações nucleares por possuir afinidade pelo DNA (TONZETICH, 2004). Além deste corante, os núcleos também foram analisados por meio da coloração com DAPI e da sua observação em microscópio de fluorescência.

Os núcleos corados com orceína acetoláctica 1% foram observados em microscópio de luz (Figuras 4 e 5) e apresentaram o diâmetro médio de 5 μm (Tabela 3). Os baixos valores de desvio-padrão e variância demonstram a uniformidade dos dados (Tabela 3). Choudhary et al. (2009) encontraram o diâmetro nuclear médio de ~3-5 μm em plântulas de arroz, semelhante ao resultado encontrado por este trabalho em plantas de cana-de-açúcar. No entanto, estudos em

outras espécies vegetais observaram uma variedade de diâmetros. Abdalla; Baker e Rafudeen (2010) estudaram proteínas nucleares de plantas de *Xerophyta viscosa* submetidas à desidratação e observaram núcleos intactos e uniformes, com o diâmetro aproximado de 2-3 μm . Já Pandey et al. (2006) ao estudar o proteoma nuclear de plantas de grão-de-bico (*Cicer arietinum* L.) encontraram núcleos intactos e esféricos, com o diâmetro aproximado de 20 μm , resultado semelhante ao encontrado por Khan e Komatsu (2004) para plantas de arroz. Ademais, Folta e Kaufman (2000), observaram que núcleos de ervilha também apresentam diâmetro aproximado de 20 μm , podendo chegar a 30 μm e são, portanto, maiores quando comparados aos núcleos de *Arabidopsis*, que apresentam diâmetro variando entre 5 e 10 μm .

Além disso, as imagens obtidas pela fotografia dos núcleos corados usando DAPI com e sem a utilização do filtro UV (Figuras 6 e 7) permitem a sobreposição das mesmas, o que indica a integridade e a ausência de contaminantes. Imagens semelhantes foram encontradas por Choudhary et al. (2009), Abdalla; Baker e Rafudeen (2010), Pandey et al. (2006) e BAE et al. (2003).

Tabela 3 – Média, desvio padrão e variância do diâmetro nuclear de folhas de plantas de cana-de-açúcar de 1 e 4 meses de idade

Tratamento	Média	Desvio Padrão	Variância
1 mês	4,96	0,49	0,24
4 meses	4,92	0,41	0,17

Figura 5 – Núcleos de folhas de cana-de-açúcar de 1 mês de idade corados com Orceína Acetoláctica 1%

Figura 6 – Núcleos de folhas de cana-de-açúcar de 4 meses de idade corados com Orceína Acetoláctica 1%

Figura 7 – Núcleos de folhas de cana-de-açúcar de 1 mês de idade fotografados com, à esquerda, e sem, à direita, utilização do filtro UV

Figura 8 – Núcleos de folhas de cana-de-açúcar de 4 meses de idade fotografados com, à esquerda, e sem, à direita, utilização do filtro UV

5.2 Avaliação da pureza das frações nucleares

Em decorrência da grande quantidade de DNA e RNA presentes nos núcleos, as proteínas foram separadas dos ácidos nucleicos pela utilização do reagente TRI (Sigma) (BAE et al., 2003).

O enriquecimento para proteínas nucleares foi avaliado por *Western blot* com anticorpos para histona H1.4 e PCNA, que são marcadores nucleares, e com anticorpos para Rubisco, PEPCase, OEE1, que, por serem marcadores de outras organelas, indicaram o grau de contaminação do extrato protéico.

A detecção dos anticorpos para histona H1.4 e PCNA apenas na fração correspondente às proteínas nucleares indica a eficácia do protocolo de isolamento de núcleos. Já a menor detecção de anticorpos para Rubisco, PEPCase, OEE1 detectada no extrato protéico nuclear quando comparado ao extrato de proteína total demonstra a presença de pequena quantidade de contaminantes, levando à conclusão de que houve o enriquecimento para proteínas nucleares.

Os anticorpos para a histona H1 e para a PCNA já haviam sido usados como marcadores nucleares em estudos com *Xerophyta viscosa* (ABDALLA; THOMSON; RAFUDEEN, 2009; ABDALLA; BAKER; RAFUDEEN, 2010). As histonas como marcadores nucleares também já foram usados por Repetto et al. (2008), Bae et al. (2003), Kalluri et al. (2009) e Khan e Komatsu (2004) em *Medicago truncatula*, *Arabidopsis*, *Populus* e *Oriza sativa*, respectivamente. Já com relação aos

marcadores de outras organelas, o anticorpo da OEE2 foi usado por BAE et al. (2003).

O número de anticorpos usados seja como marcadores nucleares ou marcadores de outras organelas também variou nos diversos estudos encontrados, indo de 1 à 5 anticorpos. Choudhary et al. (2009), Khan e Komatsu (2004), Kalluri et al. (2009) usaram apenas um marcador nuclear como anticorpo, já Abdalla; Thomson e Rafudeen (2009) e Abdalla; Baker e Rafudeen (2010) usaram dois marcadores nucleares e um representando a contaminação por outras organelas. Além disso, Bae et al (2003) usaram dois anticorpos como marcadores nucleares e dois como marcadores de outras organelas, enquanto Pandey et al (2006) usaram dois marcadores nucleares e Repetto et al. (2008) usaram apenas um anticorpo como marcador nuclear e quatro anticorpos representando outras organelas.

Dessa forma, os anticorpos usados no presente trabalho forneceram uma indicação consistente do grau de contaminação e de enriquecimento, o que permitiu a conclusão de que o material foi enriquecido para proteínas nucleares.

Figura 9 – (a) Análise de *Western blot* usando os anticorpos da Histona H1.4, PCNA, OEE1, Rubisco, PEPCase. (b) SDS-PAGE contendo o marcador, o extrato protéico total e o extrato protéico nuclear

5.3 Análise das proteínas identificadas na extração-teste

Antes da aplicação do protocolo de Folta e Kaufman (2000) para a extração dos núcleos de folhas de cana-de-açúcar de 1 e 4 meses de idade alguns testes foram realizados (dados não apresentados) com o objetivo de otimização do protocolo, permitindo a obtenção de um extrato nuclear o mais puro possível. Inicialmente, o protocolo foi executado exatamente como descrito pelos autores, em seguida foram realizadas as modificações como descritas no item 5.1., o que permitiu a obtenção de uma faixa contendo núcleos mais nítida na segunda etapa, e uma faixa ainda mais definida na terceira etapa. Essa mudança no protocolo permitiu a diminuição na contaminação, porém levou também à diminuição na quantidade de núcleos recuperados.

Nesta extração-teste, as proteínas usadas (100 µg) foram inicialmente separadas em gel 1D SDS – PAGE, o qual foi dividido em 30 bandas. Os peptídeos obtidos após a digestão e eluição foram seqüenciados no Synapt G2 HDMS (Waters®). O seqüenciamento das proteínas permitiu a obtenção, por meio da busca de similaridades entre os peptídeos seqüenciados e os clusters depositados no SUCEST, de 902 *clusters* distribuídos nas 30 bandas. Os clusters mencionados no presente trabalho referem-se ao agrupamento de ESTs de cana-de-açúcar gerados automaticamente pelo SUCEST, o qual forneceu as proteínas homólogas aos *clusters* encontrados, de acordo com o Uniprot. Nas análises foram usadas as proteínas homólogas uma vez que, a seqüência do cDNA deve estar completa pois, de maneira contrária, regiões importantes para o correto direcionamento subcelular das proteínas como os peptídeos-sinal, podem estar ausentes (DREGER, 2003). Tais proteínas foram utilizadas nos programas de predição de localização subcelular, bem como no *Pfam* e *InterPro*, usados para a classificação das mesmas em categorias funcionais (ANEXO C).

Nessa extração, as categorias funcionais mais representativas foram fotossíntese, montagem de nucleossomos e tradução, que representaram 23,17%; 15,53% e 13,97% das proteínas identificadas, respectivamente. Apesar da categoria mais representativa ter sido fotossíntese, a categoria de montagem de nucleossomos, exemplificada pelas histonas também apresentou alta porcentagem de clusters dentre o total encontrado (Figura 10).

Além disso, metabolismo de DNA (0,34%), metabolismo de RNA (0,45%), biogênese de ribossomos (0,22%), regulação da transcrição DNA-dependente (0,23%) e regulação da transcrição RNA-dependente (0,11%) que constituem classes funcionais contendo representantes nucleares típicos também foram encontradas (Figura 10).

A extração permitiu ainda a obtenção de alguns fatores de transcrição, como o BTF3, que se encontra na categoria regulação da transcrição DNA-dependente. Esta proteína foi isolada inicialmente como um fator de transcrição geral, necessário para a transcrição dependente da RNA polimerase II. Posteriormente descobriu-se sua função como subunidade-B de complexos associados a polipeptídeos nascentes, sendo implicadas na regulação da localização protéica durante a tradução (YANG et al., 2007). Em plantas, algumas subunidades Btf3 podem estar associadas à morte celular mediada pela reação de hipersensibilidade (HR) ou envolvidas na regulação do estresse biótico no interior do núcleo (HUH; KIM; PAEK, 2012).

A proteína quinase ativada por mitógeno 7 (MAPKinase 7), envolvida na fosforilação, também foi encontrada. A cascata de MAPKinases é considerada uma forma de transdução de sinal conservada evolutivamente e está envolvida na transdução de sinais extracelulares para o núcleo para uma resposta celular apropriada. Esta cascata consiste essencialmente de três componentes, a MAPK quinase quinase (MAPKKK), a MAPK quinase (MAPKK) e uma MAPK conectada uma à outra pelo evento da fosforilação. A sinalização pela cascata de MAP quinases pode conduzir a diversas respostas celulares como divisão celular, diferenciação e estresses. Em plantas, elas são representadas por famílias multigênicas e estão envolvidas na transmissão de estímulos e na regulação do sistema de defesa antioxidante em resposta à sinalização de estresse (SINHA et al., 2011). Em arroz, por exemplo, a MAP quinase BWMK1 fosforila o fator de transcrição OsEREBP1 que, por sua vez, liga-se a promotores gênicos relacionados à patogênese. Esta MAP quinase apresenta um domínio C-terminal estendido, que é essencial para sua atividade como quinase e para a translocação para o núcleo (CHEONG et al., 2003).

No presente estudo, a categoria de metabolismo de RNA foi representada pela subunidade maior do fator de *splicing* U2af. Este fator de *splicing* é essencial e apresenta papéis críticos no reconhecimento do local de *splice* 3' (WANG; BRENDDEL, 2006). Ele apresenta importância no *splicing* do pré-mRNA na maioria dos eucariotos e realiza sua função em decorrência da presença de dois motivos de reconhecimento do RNA (RRM) consecutivos que reconhecem uma sequência de polipirimidinas (Py) no local de *splice* 3' (JENKINS; LAIRD; KIELKOPF, 2012). No entanto, em plantas a sequência Py não está presente, fazendo com que provavelmente o reconhecimento do local de *splice* 3' dependa mais do fator de *splicing* U2af. Neste caso, tais fatores apresentam regiões de interação com outros U2af e com proteínas SR (proteínas ricas em serina e arginina), sendo provável que essas proteínas interajam com alguma proteína U2AF que, por sua vez, pode interagir com alguma subunidade maior U2AF específica. O balanço entre diferentes complexos proteicos U2AF-SR nos pré-mRNAs pode ser importante para o *splicing* de íntrons particulares (WANG; BRENDDEL, 2006).

Em *Arabidopsis*, níveis de expressão alterados desses fatores de *splicing* causam fenótipos pleiotrópicos que incluem o tempo de florescimento, morfologia foliar e o formato da flor, o que sugere a importância de suas funções em plantas superiores (WANG; BRENDDEL, 2006).

A subunidade 3 do fator de especificidade da polidadenilação e quebra também se encontrou presente na classe de metabolismo de RNA. A quebra e a poliadenilação do mRNA precursor é um processo essencial para a maturação do mRNA e consiste da quebra e poliadenilação na extremidade 3' recém gerada (ZHAO; XING; LI, 2009).

O fator 11 de resposta à auxina foi o representante da classe de regulação da transcrição DNA-dependente. A sinalização relacionada às auxinas apresenta papel-chave em muitos processos de crescimento e desenvolvimento, a maioria iniciados ou mediados pela expressão gênica regulada por esse hormônio. Duas famílias de fatores de transcrição são necessárias para regular a expressão destes genes, sendo que a família dos fatores de resposta à auxina funciona por meio de sua ligação a elementos de resposta à auxina nos promotores de genes de resposta à auxina, ativando ou reprimindo tais genes e recrutando uma segunda família de

fatores de transcrição, os repressores Aux/IAA, que conferem uma resposta gênica da auxina (GUILFOYLE; HAGEN, 2007).

A proteína de reparo de DNA recA foi uma das representantes da categoria do metabolismo do DNA. As proteínas de reparo serão melhor discutidas no item 5.5.

Tais resultados, juntamente com os obtidos pelas técnicas de microscopia e *Western blot*, permitiram a conclusão de que o protocolo otimizado pode ser usado para o eficiente enriquecimento de núcleos de cana-de-açúcar, sendo utilizado como metodologia padrão para o desenvolvimento dos experimentos.

Figura 10 – Classificação funcional das proteínas obtidas por meio da extração teste

5.4 Análise das proteínas identificadas nos núcleos isolados de folhas de 1 e 4 meses de idade

Para analisar a composição do proteoma nuclear de folhas de cana-de-açúcar de 1 mês foram utilizados 60 µg de proteínas em cada repetição. As proteínas foram separadas em gel 1D SDS – PAGE, que foi dividido em 20 bandas. Os peptídeos obtidos, após a digestão e eluição das proteínas presentes em cada banda, foram seqüenciados no Synapt G2 HDMS (Waters®). O mesmo procedimento foi realizado para as triplicatas biológicas.

Em seguida, três listagens de clusters foram obtidas, como descrito no item 5.3, cada uma referente a uma repetição. Em decorrência disso, apenas os clusters que apareceram em pelo menos duas repetições foram considerados nas análises.

Assim, neste experimento foram encontrados 1759 clusters presentes nas 20 bandas, sendo que todos apresentaram scores acima de 30,000. Esses clusters também foram identificados da maneira já descrita.

Os clusters encontrados foram classificados funcionalmente com o auxílio das bases de dados *InterPro* e *Pfam*, com a classificação funcional baseada, portanto, no *Gene Ontology* (ANEXOS A e D). No entanto, essa classificação indica apenas uma função provável, já vez a função de muitas proteínas ainda não foi estabelecida experimentalmente (PANDEY et al., 2006).

No experimento realizado com plantas de 1 mês de idade, foram encontradas 38 classes funcionais, onde a montagem de nucleossomos foi a classe mais representativa, responsável por 21,96% dos clusters encontrados. Essa classe é composta basicamente pelas histonas. A segunda classe mais encontrada foi a relacionada à tradução, que apresentou 18% dos clusters encontrados, sendo as proteínas ribossomais as principais proteínas desse grupo. A terceira classe mais abundante foi a de proteínas fotossintéticas, que representaram 13,64% dos clusters identificados (Figura 18).

As classes de metabolismo de DNA, metabolismo de RNA, regulação da transcrição DNA-dependente, regulação da transcrição RNA-dependente e biogênese de ribossomos também se encontraram presentes e representaram 0,06%, 0,11%, 0,06%, 0,23% e 0,4% dos clusters encontrados, respectivamente. A classe de morfogênese foliar representou 0,06% (Figura 18).

Apesar das organelas serem compartimentos distintos que realizam funções celulares específicas, existem mecanismos complexos de comunicação intracelular, além de locais de contato entre elas, o que torna difícil avaliar o significado biológico de proteínas normalmente associados a uma organela, mas detectadas no proteoma de outra. Apesar destas proteínas poderem ser contaminantes, resultantes de problemas nos métodos de fracionamento subcelular, elas também podem ser biologicamente significantes (TAYLOR, FAHY, GHOSHY, 2003).

A análise da composição do proteoma nuclear de folhas de cana-de-açúcar de 4 meses de idade foi realizada da mesma forma em que no tratamento anterior, sendo encontrados 992 clusters, todos com *scores* acima de 30,000.

Da mesma forma que no tratamento anterior, foram geradas três listas de proteínas referentes às três repetições, onde apenas os clusters que apareceram em duas ou três repetições foram considerados nas análises (ANEXOS B e E).

Nesta avaliação as classes mais representativas também foram montagem de nucleossomos, tradução e fotossíntese, que apresentaram 24,9%, 17,24% e 11,69%, respectivamente. A classe de morfogênese foliar também se encontrou presente, sendo responsável por 0,11% dos clusters identificados. As classes de metabolismo de RNA, regulação da transcrição DNA-dependente, regulação da transcrição RNA-dependente e biogênese de ribossomos representaram 0,21%, 0,02%, 0,11% e 0,71% dos clusters identificados, respectivamente (Figura 19).

É interessante notar que tanto em 1 quanto em 4 meses as classes de proteínas mais abundantes foram as mesmas, variando somente na porcentagem de proteínas encontradas. Deste modo, a classe mais abundante foi montagem de nucleossomos, representada principalmente pelas histonas.

As histonas, como já mencionado no item 3.3.1, estão entre as proteínas mais abundantes dos organismos eucariotos (YI et al., 2006) e apresentam função relacionada à proteção eficiente e ao empacotamento ordenado do DNA genômico (WAKO; FUKUI, 2010). A fosforilação da Ser139 da histona variante H2AX, por exemplo, apresenta importância no reparo do DNA em plantas (DRURY et al., 2012).

Dentre as histonas encontradas em plantas de 1 mês estão: H4, H2B.5, H2A.6, H2B.2, H2AXb, H2B.1, H2A.4, H2A variante 3, H2A, H2B.11, H2A variante 1, H2A.8, H3.3, H3.2.

Já aos 4 meses as histonas mais encontradas foram: H4, H2B.5, H2B.2, H2B.1, H2B.11, H2A.6, H2A variante 1, H2AXb, H2A.4, H2A variante 3, H2A, H2A.8, H2A.2, H3.2, H3, e H3.3.

A segunda classe mais encontrada, refere-se à tradução e foi representada pelas proteínas ribossomais. Tal resultado também era esperado, uma vez que como também mencionado no item 3.3, a biogênese dos ribossomos ocorre no núcleo, pois os rRNAs começam sua rota de montagem juntamente com as proteínas ribossomais no DFC, que faz parte da região organizadora de nucléolo e apenas no final essas partículas pré-ribossomais são translocadas para o citoplasma através dos poros nucleares (CARMO-FONSECA; MENDES-SOARES; CAMPOS, 2000; LAM; TRINKLE-MULCAHY; LAMOND, 2005). Essa classe de proteínas também foi encontrada como abundante na análise do proteoma nuclear de arroz por Khan e Komatsu (2004)

Dentre as proteínas nucleares encontradas nas plantas de 1 mês estão as proteínas MAR que, como mencionado no item 3.3.3, são proteínas longas e enroladas em espiral que se ligam à dupla fita de DNA e são específicas de plantas. Apresentam um domínio N-terminal com dois domínios hidrofóbicos conservados, um domínio central em haste enrolada em espiral e um domínio C-terminal de ligação ao DNA (SAMANIEGO et al., 2006).

A terceira classe de proteínas mais encontrada em ambos os tratamentos foi a fotossíntese, representada por proteínas dos fotossistemas, proteínas de ligação à clorofila a/b e que, juntamente com proteínas como a Rubisco foram consideradas contaminantes, apesar da possibilidade de estarem realizando alguma forma de regulação no interior do núcleo. Proteínas consideradas contaminantes semelhantes às encontradas no presente trabalho também já foram observadas em outros estudos. Dentre elas se encontram a própria Rubisco, tubulinas, malato desidrogenase, fosfoglicerato quinase (BAE et al., 2003), frutose bi-fosfato hidrolase, metaloprotease FTSH ATP-dependente, ferredoxina, anidrase carbônica, proteínas de ligação à clorofila a/b, OEE1, ATP sintase dentre outras (PANDEY et al., 2006)

As fontes mais comuns de contaminação cruzada são o retículo endoplasmático, a membrana plasmática, os elementos do citoesqueleto e os

grandes complexos protéicos citoplasmáticos, como os proteassomos (HUBER; PFALLER; VIETTOR, 2003).

Neste aspecto, duas classes representativas foram aquelas relacionadas à proteólise (5,8% em 1 mês e 4,95% em 4 meses), que pode ser exemplificada pelos proteassomos, e a relacionada a estrutura (3% em 1 mês e 2,32% em 4 meses), representada principalmente pela actina. Apesar de os proteassomos, como já mencionado, constituírem fontes comuns de contaminação cruzada, as moléculas de actina podem ser fontes de contaminação mas podem também representar proteínas nucleares não-típicas e que, portanto, não apresentam nenhum sinal comum que indica sua localização nuclear.

A actina já foi implicada em vários processos nucleares, como o remodelamento da cromatina (é um dos principais componentes do complexo de remodelação da cromatina), transcrição, transporte nuclear e estrutura nuclear geral. Apesar de não existirem muitas informações a respeito da actina em núcleos de células vegetais, já foi demonstrada sua presença no nucleoplasma, nos *speckles* e no nucléolo de núcleos de células de *Arabidopsis*, provando sua presença também nos núcleos vegetais (KANDASAMY; MCKINNEY; MEAGHER, 2010).

Ademais, a presença da actina no núcleo já foi proposta há mais de 30 anos atrás. No entanto, os processos nucleares que necessitam da actina foram identificados apenas recentemente, portanto, além das funções já mencionadas, a actina ainda é incorporada em ribonucleoproteínas recém-sintetizadas, influencia a organização da cromatina e age em conjunção com diferentes tipos de proteínas de ligação à actina, que regulam sua função e constituem uma ponte para as interações entre a actina e os outros componentes nucleares (VISA; PERCIPALLE, 2010).

As ubiquitinas constituem as principais representantes da classe funcional proteólise. No entanto podem estar envolvidas em outras funções que justificam sua presença no interior dos núcleos. Um exemplo disso é a monoubiquitinação¹ da histona (HUB1), que constitui uma ubiquitina E3 ligase não convencional, que não está envolvida na degradação protéica, mas sim na modificação da histona H2B implicada na ativação transcricional em plantas (HIMANEN et al., 2012).

A ubiquitinação dessa histona está envolvida ainda na estabilidade dos nucleossomos (CHANDRASEKHARAN, HUANG, SUN, 2009) e parece possuir papel

na regulação da metilação da histona H3 e na expressão gênica, possivelmente em decorrência de um mecanismo de *crosstalk* trans-histona evolutivamente conservado (CHANDRASEKHARAN, HUANG, SUN, 2010). Em leveduras, a monoubiquitinação de histonas H2B facilita o *crosstalk* entre a cromatina e o *splicing* do pré-mRNA por meio da modulação da distribuição de íntrons e éxons relacionadas às modificações de histonas (SHIEH et al., 2011).

Além disso, mesmo as proteínas envolvidas na via da ubiquitina-proteassoma podem estar localizadas no núcleo em determinado momento do ciclo celular. De acordo com Ciechanover (1998), este processo apresenta importantes papéis em uma grande variedade de processos celulares básicos, sendo que seus substratos variam desde reguladores de crescimento e do ciclo celular, componentes das rotas de transdução de sinais até proteínas que sofreram mutações ou foram danificadas após a tradução. Um exemplo é o receptor de luz azul CRY2, que é ubiquitinado em resposta à luz azul e então é degradado pelo proteassomo 26S no interior do núcleo, onde ele age como receptor de luz azul para promover o florescimento sob condições de dia longo. O mesmo estudo demonstra que um fotoreceptor pode completar seu ciclo de vida pós-traducional (da modificação protéica à função e posterior degradação) no interior do núcleo (YU et al., 2007).

A ubiquitinação e a SUMOilação da PCNA (antígeno nuclear de proliferação celular), uma proteína-chave na replicação e no reparo do DNA, controla precisamente a forma como as células eucarióticas respondem a diferentes tipos de dano do DNA e como as rotas celulares de tolerância ou de reparo do dano do DNA são selecionadas para lidar com os prejuízos causados ao mesmo (CHEN; BOZZA; ZHUANG, 2011).

A SUMO, pequeno modificador relacionado à ubiquitina, também está envolvida em diversos processos biológicos que ocorrem no núcleo, como a transcrição, o reparo do DNA, a estabilidade do genoma, a organização da cromatina, importação pelo núcleo (MEIER, 2012; MILLER, VIERSTRA, 2011). A sumoilação parece representar ainda um mecanismo rápido e global para manipular reversivelmente as funções cromossômicas em vegetais, especialmente durante estresses ambientais (MILLER et al., 2010).

Dentre as proteínas mais encontradas, como possíveis contaminantes, está a gliceraldeído-3-fosfato desidrogenase (GAPDH), normalmente conhecida pela sua função na glicólise. No entanto, além desta função, ela também participa do metabolismo do ferro, do transporte pela membrana, da biossíntese das histonas, manutenção da integridade do DNA e sinalização celular mediada por receptores. Assim, a GAPDH não é apenas uma proteína citosólica, pois também pode estar localizada na membrana, no núcleo, polissomos, retículo endoplasmático e complexo de Golgi (SIROVER, 2012), o que justifica sua presença no extrato protéico obtido no presente estudo. Dentre suas funções nucleares, a GAPDH participa ainda do reparo do DNA, regulação da transcrição gênica e *turnover* do DNA telomérico (TRISTAN et al., 2010).

Em células humanas, a translocação da GAPDH para o núcleo é mediada pela acetilação de três resíduos específicos de lisina (117, 227 e 251) (VENTURA et al., 2010). Em vegetais, especificamente, a expressão temporária da GAPDH fusionada com a proteína verde fluorescente demonstrou a localização da enzima no núcleo de protoplastos de *A. thaliana*. A localização nuclear e sua ligação com o DNA parece ocorrer, em numerosos sistemas, sob condições de estresse, indicando que esse mecanismo pode estar envolvido em rotas de sinalização, além da possível participação em sistemas protetores relacionados à formação de espécies reativas de oxigênio e nitrogênio, bem como na diminuição e proteção do metabolismo pela modificação de resíduos essenciais de cisteína (HOLTGREFE et al., 2008).

As proteínas 14-3-3 foram encontradas no presente estudo de maneira representativa tanto em plantas de 1 mês quanto de 4 meses de idade. Elas constituem proteínas de ligação à fosfoserina que regulam a atividade de uma grande variedade de alvos através de interações proteína-proteína diretas. Em plantas, elas regulam as H⁺-ATPases da membrana e as enzimas do metabolismo do carbono e do nitrogênio. No entanto, além dessas funções, as proteínas 14-3-3 em vegetais se ligam a diversos fatores de transcrição e à outras proteínas de sinalização e apresentam papéis na regulação do desenvolvimento vegetal e nas respostas aos estresses. A presença das proteínas 14-3-3 em diversas organelas

como citoplasma, núcleo e cloroplasto sugere ainda que isoformas específicas provavelmente interajam com diferentes alvos (ROBERTS, 2003).

Nos núcleos em especial, os fatores de transcrição, proteínas quinases e proteínas fosfatases são regulados por essas proteínas, de forma que as interações podem afetar tanto a importação quanto a exportação de moléculas pelo núcleo e apresentam importância para várias proteínas envolvidas no desenvolvimento e nas rotas de sinalização em resposta ao estresse (ROBERTS, 2003). Alguns exemplos são os ativadores transcricionais MSN2 e MSN4 (BECK; HALL, 1999), a proteína quinase Cdc2 (CHAN et al., 1999), a proteína RoIB envolvida na indução de raízes adventícias (MORIUCH et al., 2004).

Dentre os processos em que tais proteínas estão envolvidas, pode-se citar ainda sua participação na inibição eficiente de fatores de transcrição fosforilados por brassinoesteróides (GAMPALA et al., 2007), uma vez que a interação de proteínas 14-3-3 com os reguladores transcricionais BZR1 e BZR2 faz com que eles sejam redistribuídos para o citoplasma e a inibição seja alcançada (de VRIES, 2007). Além disso, podem atuar como receptores intracelulares para o hormônio florigeno Hd3a de arroz, onde sua interação com o homólogo FT (*flowering locus T*) Hd3a produz um complexo que transloca para o núcleo e liga-se ao fator de transcrição FD1 e induz a transcrição do OsMADS15, homólogo ao APETALA1 de *A. thaliana*, o que conduz ao florescimento (TAOKA et al., 2011).

Esses exemplos demonstram claramente a presença destas proteínas no núcleo, o que indica que são esperadas em tal compartimento e não constituem fontes de contaminação.

A presença da família das proteínas de choque térmico (HSPs – heat shock proteins) no extrato nuclear também não foi surpresa, uma vez que sua existência no núcleo havia sido descrita. Esta família de proteínas é altamente conservada e constitui-se de proteínas nucleares atípicas, uma vez que se localizam no núcleo seletivamente e não invariavelmente. Sua presença no núcleo normalmente está relacionada à resposta a uma variedade de estresses (KNOWLTON; SALFITY, 1996). No entanto, apesar da propriedade de acúmulo nuclear das HSPs em resposta ao estresse ter sido documentada há aproximadamente 30 anos, as

funções nucleares de diversas HSPs, como as HSP70s ainda não foi muito bem estudada (KOSE; FURUTA; IMAMOTO, 2012).

Desta forma, as chaperonas moleculares nucleares e/ou citosólicas HSP90 e HSC70, da família das proteínas heat shock, por exemplo, são proteínas conservadas e essenciais nos eucariotos, que apresentam funções relacionadas à imunidade inata e à tolerância ao estresse abiótico em plantas superiores e constituem uma maquinaria importante para o fechamento estomático, sendo essenciais para integrar sinais provenientes do meio ambiente biótico e abiótico (CLÉMENT *et al.*, 2011).

A descoberta de carreadores nucleares de importação de proteínas HSP (KOSE; FURUTA; IMAMOTO, 2012) confirma a presença dessas proteínas no interior do núcleo, juntamente com o fato dessas moléculas interagirem com moléculas de sinalização como receptores nucleares de hormônios (WANG *et al.*, 2004).

Ademais, mesmo quando o núcleo é adicionalmente purificado, ele continua a carregar uma quantidade considerável de membranas contaminantes, como as originadas do retículo endoplasmático, ao longo de todos os passos dos procedimentos de purificação (HUBER; PFALLER; VIETTOR, 2003), o que indica que a presença de alguns contaminantes também é esperada, uma vez que objetiva-se o enriquecimento de determinada organela e não seu isolamento total.

Além disso, é importante não somente identificar a presença de uma proteína em determinada organela ou estrutura, mas também medir sua abundância relativa nos diversos locais, além de avaliar a forma como esta localização subcelular pode variar entre diferentes compartimentos, sob diferentes condições tanto fisiológicas quanto de crescimento (BOISVERT *et al.*, 2010), e apesar desta não ter sido a abordagem escolhida no presente trabalho, ela constitui um próximo passo para o estudo do proteoma nuclear de cana-de-açúcar.

Figura 11 – 1D SDS – PAGE das proteínas obtidas em folhas de plantas de 1meses de idade

Figura 12 – 1D SDS – PAGE das proteínas obtidas em folhas de plantas de 4 meses de idade

5.5 Proteínas nucleares características não encontradas nas três repetições

Algumas proteínas nucleares típicas foram encontradas apenas em uma ou outra repetição e por isso não foram consideradas nas análises estatísticas (Tabela 4). No entanto, em decorrência de sua importância, serão discutidas a seguir.

Em 1 mês de idade, na categoria metabolismo de RNA, encontraram-se algumas helicases, como as DEAD-box ATP-dependente RNA helicases 3, 12 e 15 e a RNA helicase ATP-dependente uap56. As helicases de RNA representam uma grande família de proteínas que estão envolvidas no metabolismo do RNA e na modulação da sua estrutura como mencionado no item 3.3.2 (TUTEJA, 2003; CRUZ, KRESSLER; LINDER, 1999).

A família DEAD-box de RNA helicases é composta por proteínas ubíquas em quase todos os organismos e envolvidas em praticamente todos os processos relacionados ao RNA. Elas são caracterizadas por dois domínios RecA-semelhantes ligados *em tandem*, que apresentam 11 motivos conservados e por seqüências

flanqueadoras nas extremidades amino e carboxi-terminais altamente variáveis. Por esse motivo, são consideradas proteínas modulares de múltiplos domínios (BANROQUES et al., 2011). Em *Arabidopsis* e milho, por exemplo, a helicase de RNA RH3 atua no splicing de íntrons específicos e contribui ainda para a montagem da partícula ribossomal 50S (ASAKURA et al., 2012).

Os fatores de transcrição atuam por meio da ligação dos elementos cis em regiões promotoras de genes-alvo e, dessa forma, ativam ou reprimem sua expressão (OSORIO et al., 2012). Os fatores de transcrição bHLH128 e ILR3, também conhecido como bHLH105, foram encontrados na idade de 1 mês.

A família de fatores de transcrição bHLH é encontrada nos três reinos eucarióticos e constitui uma das maiores famílias de fatores de transcrição. Estão envolvidos em diversos processos regulatórios, especialmente na geração da diversidade fenotípica vegetal, o que indica sua possível função chave em processos de crescimento e desenvolvimento (CARRETERO-PAULET et al., 2010). Em plantas de soja pelo menos 14 fatores de transcrição bHLH estão envolvidos em respostas a estresses ambientais (OSORIO et al., 2012).

No presente estudo, RNA polimerases foram encontradas em 1 mês de idade. Este grupo constitui enzimas relacionadas à produção do RNA, como RNA polimerase II, RNA polimerase direcionada pelo DNA, RNA polimerase cid Poli (A). Os eucariotos normalmente possuem três RNA polimerases dependentes de DNA, Pol I-III, que são especializadas para a síntese de diferentes tipos de RNA e dessa forma, são essenciais para sua viabilidade (MARCUSSEN et al., 2010). A RNA polimerase II (Pol II), por exemplo, é responsável pela síntese de mRNA em eucariotos e sua transcrição constitui o primeiro passo na expressão gênica e um ponto central da regulação celular (CRAMER; BUSHNELL; KORNBERG, 2001). A fosforilação de determinadas repetições conservadas no domínio C-terminal da subunidade maior dessa enzima apresenta papel central na regulação da transcrição e no processamento co-transcricional do RNA (HAJHEIDARI et al., 2012).

As ribonucleoproteínas e outras proteínas relacionadas ao splicing, como o fator de splicing de pré-mRNA SF2 e a proteína LSm3 U6 associada ao snRNA também foram encontradas, algumas em apenas uma idade e outras nas duas. O *splicing* alternativo do pré-mRNA é um fenômeno largamente espalhado nos

eucariotos multicelulares e representa uma das principais formas para a expansão funcional do transcriptoma. Vários membros das ribonucleoproteínas são conhecidos como reguladores de *splicing* em animais e alguns de seus homólogos parecem apresentar funções similares em vegetais (WACHTER; RÜHL; STAUFFER, 2012). Ademais, a montagem dos pré-spliceossomos é responsável pela seleção de íntrons para o *splicing*. As snRNPs U1 e U2, por exemplo, reconhecem locais de *splicing* 5' (SHAO et al., 2012). O complexo snRNA U2/U6 é conservado e um componente essencial do spliceossomo ativo, que interage com o pré-mRNA e com outros fatores de *splicing* essenciais, promovendo a catálise do *splicing* (BURKE et al., 2012).

Algumas proteínas relacionadas ao metabolismo do DNA, como a helicase de reparo do DNA foram encontradas no presente estudo tanto em 1 quanto em 4 meses. As proteínas de reparo e recombinação do DNA (DRR) protegem os organismos contra os danos genéticos causados tanto por agentes ambientais, quanto agentes genotóxicos, por meio da remoção das lesões do DNA ou ajudando na sua tolerância. Em plantas, os genes DRR são muito similares aos encontrados em outros eucariotos e mesmo procariotos. Além disso, estudos de localização subcelular sugerem que essas proteínas residam em sua maioria no núcleo (SINGH et al., 2010). A helicase de reparo do DNA UVH6, em especial, apresenta papel no reparo de fotoprodutos da radiação ultravioleta, sendo necessária, desta forma, para a resistência a este tipo de radiação (LIU et al., 2003).

A proteína quinase ativada por mitógenos 6 foi encontrada na idade de 4 meses e representou a classe das MAPKinases, que como já tratado no item 5.3, faz parte da cascata de MAPKinases, considerada uma forma de transdução de sinal conservada evolutivamente e envolvida na transdução de sinais extracelulares para o núcleo para uma resposta celular apropriada (SINHA et al., 2011).

Nas plantas dessa mesma idade também foi observada a presença do fator de iniciação da transcrição TFIID. Esta proteína é um fator de reconhecimento do promotor necessário para que a RNA polimerase II inicie a transcrição. O TFII é composto por fatores altamente conservados que incluem a proteína de ligação TATA (TBP) e aproximadamente 14 fatores associados à TBP (TAFs) (LAGO et al., 2005).

A proteína fitocromo também foi encontrada em plantas de 1 mês e está envolvida no controle de muitas respostas fisiológicas modulando diretamente a expressão gênica, onde o evento regulatório chave nessa rota de transdução de sinal é a translocação deste fotoreceptor do citoplasma para o núcleo controlada pela luz (FRANKHAUSER; CHEN, 2008).

Ademais, a proteína argonauta e a proteína de ligação ao RNA rica em glicina foram encontradas no presente trabalho em plantas desta mesma idade. As proteínas argonautas (AGO) recrutam pequenos RNAs para formar o centro dos complexos efetores de RNA de interferência (RNAi) (MI et al., 2008). Em *Arabidopsis*, a proteína AGO4 e a RNA polimerase IV (Pol IV) são necessárias para a metilação guiada por 24 siRNAs (RNAs de interferência pequenos) no interior do núcleo (LI et al., 2006). As proteínas de ligação ao RNA ricas em glicina apresentam uma sequência de reconhecimento de RNA (RRM) no N-terminal e uma região rica em glicina no C-terminal. Em *Arabidopsis* e arroz, elas são funcionalmente conservadas e apresentam função como chaperonas de RNA durante o processo de adaptação ao frio (KIM et al., 2010).

Além dessas proteínas, também foram encontradas proteínas relacionadas à manutenção estrutural de cromossomos, recombinação meiótica, ligação ao telômero, *RING finger*, proteína de cinetócoro, proteínas homeobox, de retrotransposons, dentre outras, todas relacionadas às funções de metabolismo do RNA, regulação da transcrição, sinalização, metabolismo de DNA e algumas com funções ainda desconhecidas.

A presença dessas proteínas indica a eficiência do protocolo no isolamento dos núcleos, uma vez que estas constituem proteínas de baixa abundância na célula, não sendo detectadas com facilidade em extratos de proteínas totais. A presença delas em apenas uma repetição pode ser decorrente de sua pequena abundância ou limitações técnicas, uma vez que o enriquecimento de proteínas nucleares foi comprovado de diversas maneiras diferentes.

Tabela 4 – Lista de proteínas nucleares características não encontradas nas três repetições, sua localização subcelular de acordo com os programa de predição e sua classificação funcional (continua)

Função	Identificação	Idade - Repetição	Cluster	e-value	B	S	W	P	PI	E.C.
Fosforilação	sp C0LGP9 IMK3_ARATH. Probable leucine-rich repeat receptor-like protein kinase IMK3	1 mês - RA	SCQSRT2032B03	3,00E-55	X	nd	nd	nd	X	<u>EC=2.7.11.1</u>
função desconhecida	sp Q336X9 MPK6_ORYSJ. Mitogen-activated protein kinase 6	4 mês - RA	SCVPCL6042B11	0	X	nd	nd	nd	X	<u>EC=2.7.11.24</u>
	sp Q9CAJ9 Y1385_ARATH. BTB/POZ domain-containing protein At1g63850	4 mês - RB	SCCCLR1076H05	1,00E-170	X	nd	X	nd	X	
	sp Q8RWU5 FBL3_ARATH. F-box/LRR-repeat protein 3 O S=Arabidopsis thaliana GN=FBL3 PE=2 SV=1	1 mês - RA	SCAGLR1043F06	5,00E-41	nd	nd	X	nd	X	
	sp Q93W93 FBK22_ARATH. F-box/kelch-repeat protein At1g55270	4 mês - RA	SCBGLR1047G04	5,00E-35	nd	nd	X	X	X	
	sp Q9SKK0 EBF1_ARATH. EIN3-binding F-box protein 1	4 mês - RB	SCJLLR1104C02	1,00E-132	X	nd	X	nd	X	
iniciação da transcrição	sp Q9SMY8 FBL15_ARATH. F-box/LRR-repeat protein 15	1 mês - RC	SCCCCL4007A08	0	X	nd	X	nd	X	
	sp Q9D3R9 TAF7L_MOUSE. Transcription initiation factor TFIID subunit 7-like	4 mês - RA	SCQSHR1021A12	2,00E-13	X	X	X	nd	X	
	sp O93309 SMC3_XENLA. Structural maintenance of chromosomes protein 3	1 mês - RC	SCVPLR1049A01	0	X	nd	nd	X		
metabolismo do DNA	sp P42698 DR111_ARATH. DNA-damage-repair/toleration protein DRT111, chloroplastic	1 mês - RB	SCCCST1001B07	2,00E-21	X	X	X	X	nd	
	sp Q8W4M7 ERCC2_ARATH. DNA repair helicase UVH6	4 mês - RA	SCBGRT3073E07	8,00E-45	X	X	nd	nd	X	<u>EC=3.6.4.12</u>
	sp Q9LL45 TBP1_ORYSJ. Telomere-binding protein 1	4 mês - RC	SCVPCL6042A04	1,00E-152	X	X	nd	nd	X	
	sp Q9M4A1 SPO12_ARATH. Meiotic recombination protein SPO11-2	1 mês - RB	SCJFAM1066H07.b	2,00E-42	X	X	nd	nd	nd	
	tr Q2QXP9 Q2QXP9_ORYSJ Retrotransposon protein, putative	1 mês - RB	SCSFAD1122E08	4.0×10 ⁻¹¹	X	X	X	nd	X	
	sp Q0DM51 RH3_ORYSJ. DEAD-box ATP-dependent RNA helicase 3, chloroplastic	1 mês - RB	SCUTLR1058C12	6,00E-47	X	X	nd	nd	X	<u>EC=3.6.4.13</u>
metabolismo do RNA	sp O13792 SUB2_SCHPO. ATP-dependent RNA helicase uap56	1 mês - RC	SCCCAM1C01F11	2,00E-68	X	X	X	nd	X	<u>EC=3.6.4.13</u>

Tabela 4 – Lista de proteínas nucleares características não encontradas nas três repetições, sua localização subcelular de acordo com os programa de predição e sua classificação funcional (continuação)

Função	Identificação	Idade - Repetição	Cluster	e-value	B	S	W	P	PI	E.C.
	sp O22315 SRSF1_ARATH. Pre-mRNA-splicing factor SF2	4 mês - RC	SCRLRZ3039E01	6,00E-50	X	X	X	X	X	
	sp O74326 CID11_SCHPO. Poly(A) RNA polymerase cid11	1 mês - RC	SCQSRT2032H08	1,00E-19	X	nd	nd	nd	X	<u>EC=2.7.7.19</u>
	sp P62311 LSM3_MOUSE. U6 snRNA-associated Sm-like protein LSM3	1 mês - RC	SCCCLR2002D01	1,00E-23	X	nd	nd	nd	X	
	sp Q00416 SEN1_YEAST. Helicase SEN1	1 mês - RC	SCCCCL3080D02	5,00E-61	X	X	X	X		<u>EC=3.6.4.-</u>
	sp Q109G2 RH12_ORYSJ. DEAD-box ATP-dependent RNA helicase 12	1 mês - RC	SCCCCL3080F04.b	0	X	X	nd	nd	X	<u>EC=3.6.4.13</u>
	sp Q54PB2 MRD1_DICDI. Multiple RNA-binding domain-containing protein 1	1 mês - RA	SCCCLR1C02H04	1,00E-64	nd	X	X	X	nd	
	sp Q5JK84 RH15_ORYSJ. DEAD-box ATP-dependent RNA helicase 15	1 mês - RC	SCVPRT2080F12	0	X	X	nd	nd	X	<u>EC=3.6.4.13</u>
	sp Q5JK84 RH15_ORYSJ. DEAD-box ATP-dependent RNA helicase 15	1 mês - RC	SCRFLR1034B02	0	X	X	nd	nd	X	<u>EC=3.6.4.13</u>
	sp Q6K972 AGO1C_ORYSJ. Protein argonaute 1C	1 mês - RB	SCEZLR1009H11	7,00E-59	X	X	nd	nd	X	
processos celulares	sp Q10198 NDC80_SCHPO. Kinetochore protein ndc80	4 mês - RC	SCJFLR1013D10	6,00E-08	X	nd	nd	nd		
Proteólise	sp P0CH30 RING1_GOSHI. E3 ubiquitin-protein ligase RING1	1 mês - RB	SCCCCL3080D06	1,00E-36	X	X	X	nd	X	<u>EC=6.3.2.-</u>
regulação da transcrição DNA-dependente	sp Q00IB6 CPL4_ARATH. RNA polymerase II C-terminal domain phosphatase-like 4	1 mês - RB	SCUTSB1033D03	2,00E-45	X	X	X	nd	nd	<u>EC=3.1.3.16</u>
	sp Q6ENT2 RPOA_SACOF. DNA-directed RNA polymerase subunit alpha	1 mês - RC	SCAGLR2033C03	1,00E-170	X	nd	nd	nd	nd	<u>EC=2.7.7.6</u>
	sp Q8H102 BH128_ARATH. Transcription factor bHLH128	1 mês - RB	SCQSRT2034E01	9,00E-31	X	X	X	nd	X	
	sp Q9SY59 NFXL1_ARATH. NF-X1-type zinc finger protein NFXL1	4 mês - RB	SCCCLR1022A06	0	X	X	X	X		<u>EC=6.3.2.-</u>
	sp Q9FH37 ILR3_ARATH. Transcription factor ILR3	1 mês - RA	SCCCRT1003E06	3,00E-45	X	X	X	nd	X	
regulação da transcrição RNA-dependente	sp Q9SSQ4 FRS6_ARATH. Protein FAR1-RELATED SEQUENCE 6	1 mês - RB	SCBFLR1026D05	8,00E-53	X	nd	X	X	X	
	sp Q9SZL8 FRS5_ARATH. Protein FAR1-RELATED SEQUENCE 5	1 mês - RC	SCCCST2003G01	5,00E-32	X	nd	X	X	X	

Tabela 4 – Lista de proteínas nucleares características não encontradas nas três repetições, sua localização subcelular de acordo com os programas de predição e sua classificação funcional (conclusão)

Função	Identificação	Idade - Repetição	Cluster	e-value	B	S	W	P	PI	E.C.
resposta ao estresse	sp P10979 GRPA_MAIZE. Glycine-rich RNA-binding, abscisic acid-inducible protein	1 mês – RB	SCMCFL5004A10	3,00E-35	X	X	X	nd	X	
Sinalização	sp P93526 PHYA_SORBI. Phytochrome a	1 mês - RB	SCCCCL3080H06	0	nd	nd	nd	nd	X	
	sp Q0E3C3 KNOS2_ORYSJ. Homeobox protein knotted-1- like 2	4 mês - RC	SCEPSB1133D04	1,00E-136	X	X	X	X	X	
	sp Q0E3C3 KNOS2_ORYSJ. Homeobox protein knotted-1-like 2	4 mês - RC	SCEPRT2048D04	1,00E-136	X	X	X	X	X	
	sp Q9ZNU6 DET1_SOLLC. Light-mediated development protein DET1	1 mês - RB	SCCCLR1065C07	7,00E-97	X	X	nd	nd	X	
ubiquitinação de proteínas	sp O81902 PUB8_ARATH. U-box domain-containing protein 8	1 mês - RC	SCSFLR2009D04	2,00E-06	X	nd	X	nd	X	<u>EC=6.3.2.-</u>
	sp Q7M3S9 RNGB_DICDI. RING finger protein B	1 mês - RC	SCCCRZ1C01H12	2,00E-07	X	X	X	nd	X	

*nd = localização não nucleares (mitocôndria, cloroplasto, citoplasma, peroxissomo, membrana plasmática, retículo endoplasmático, complexo de Golgi); X = localização nuclear; B = BaCellLo; S = SherLoc2; W = WoLF PSORT; P = PSORT; PI = Plant-mPloc

5.6 Diferenças entre as proteínas encontradas em folhas de 1 e 4 meses

Para as folhas de plantas com 1 mês, foram encontrados 970 clusters, enquanto em plantas de 4 meses foram encontrados 206 clusters (ANEXOS F e G).

A classe funcional de montagem do nucleossomo foi a mais representativa nos dois tratamentos, constituindo 22,92% dos clusters de folhas de 1 mês e 41,1% dos clusters encontrados em folhas de 4 meses. A segunda classe funcional mais abundante foi tradução, apresentando 17,85% e 13,86% dos clusters exclusivos de plantas de 1 mês e 4 meses de idade, respectivamente (Figuras 13 e 14).

As classes de metabolismo de RNA e biogênese de ribossomos constituíram 0,5% e 0,99% dos clusters exclusivos de plantas de 4 meses (Figura 14). Já as classes metabolismo de DNA, metabolismo de RNA, regulação da transcrição RNA-dependente e biogênese de ribossomos apareceram nos clusters exclusivos de plantas de 1 mês, representando 0,1%, 0,1%, 0,31% e 0,21% dos resultados encontrados, respectivamente (Figura 13).

Com relação ao total de classes funcionais encontradas, nas plantas de 4 meses foram observadas apenas 33 classes, estando ausentes 6 classes: ciclo do ácido tricarboxílico, homeostase celular redox, isomerização peptidil-prolil de proteínas, macropinocitose, metabolismo de nucleotídeos, e metabolismo de DNA. Já a classe ciclo celular está presente em plantas de 4 meses, mas ausente em plantas de 1 mês.

Apesar destas diferenças não há a possibilidade de determinar se essas proteínas são preferencialmente expressas em uma idade ou outra, uma vez que o número de proteínas identificadas foi diferente, provavelmente em decorrência de outros aspectos não relacionados à diferenças fisiológicas das plantas.

Figura 13 – Classificação funcional das proteínas preferencialmente expressas em núcleos de folhas de cana-de-açúcar de 1 mês de idade

Figura 14 – Classificação funcional das proteínas preferencialmente expressas em núcleos de folhas de cana-de-açúcar de 4 mês de idade

5.7 Aplicação dos programas de predição

Nos programas de predição de localização subcelular, foi utilizada a proteína homóloga ao cluster identificado, pois, como mencionado no item 5.3, segundo Dreger (2003), a sequência do cDNA deve estar completa, uma vez que de maneira contrária, regiões importantes para o correto direcionamento subcelular das proteínas, como os peptídeos-sinal, podem estar faltando.

Foram utilizados cinco programas de predição desenvolvidos de formas diferentes.

Em plantas de 1 mês, 49,75% dos clusters encontrados foram classificadas como nucleares por pelo menos um programa de predição, sendo que destes,

36,68% foram classificados nucleares pelos cinco programas (Figura 16). Além disso, os programas que mais apresentaram semelhanças em seus resultados são o BaCcelo e o Plant-mPloc.

Em plantas de 1 mês, o programa que mais apresentou resultado nuclear positivo foi o BaCcelo (36,73%), enquanto o programa mais rigoroso na classificação nuclear positiva foi o SherLoc, seguido pelo WolfPsort, correspondendo a 22,68% e 25,24% respectivamente (Figura 15).

Em plantas de 4 meses de idade, 52,62% dos clusters encontrados foram classificados como nucleares por um ou mais programas de predição, sendo que dentre esses, 34,1% apresentaram resultado nuclear positivo nos cinco programas de predição. Neste caso, os programas que apresentaram resultados mais similares foram o PSORT e o Plant-mPloc (Figura 18).

Nos clusters encontrados em plantas de 4 meses de idade, o programa responsável pelo maior número de classificações nucleares foi Plant-mPloc (39,71%). Já o SherLoc e o WolfPsort apresentaram novamente as menores quantidades de clusters classificados como nucleares (25,5% e 26,41% respectivamente) e constituem os programas mais rigorosos na classificação (figura 17).

O tráfego núcleo-citoplasmático de proteínas é um processo regulatório central que sustenta a integridade do espaço nuclear das células eucarióticas por meio da interação entre numerosos fatores. Desta forma, apesar do progresso na caracterização experimental de alguns sinais de localização nuclear, a sua presença por si só continua um indicador não confiável da real translocação protéica (MEHDI et al., 2011), o que demonstra o fato de que algumas proteínas consideradas não nucleares pelos programas de predição podem estar presentes no núcleo em algum momento do ciclo celular, mesmo sem apresentar os sinais óbvios de localização nuclear. A translocação destas proteínas para o núcleo pode ocorrer, por exemplo, por meio de transportadores como as chaperonas, que direcionam as proteínas, após sua síntese para o compartimento celular correto.

Figura 15 – Porcentagem de proteínas encontradas em folhas de plantas de 1 mês que foram preditas nucleares por cada programa de predição

Figura 16 – Número de proteínas preditas como nucleares por 1, 2, 3, 4, ou 5 programas e seus respectivos programas. Resultados obtidos em folhas de plantas de 1 mês de idade; B = BaCCello; S = SherLoc2; W = WolfPsort; P = PSORT; PL = Plant-mPloc

Figura 17 – Porcentagem de proteínas encontradas em folhas de plantas de 4 meses que foram preditas nucleares por cada programa de predição

Figura 18 – Número de proteínas preditas como nucleares por 1, 2, 3, 4, ou 5 programas e seus respectivos programas, em folhas de plantas de 4 meses de idade; B = BaCCello; S = SherLoc2; W = WoLF PSORT; P = PSORT; PL = Plant-mPloc

6 CONSIDERAÇÕES FINAIS

O presente trabalho atingiu seu objetivo, que foi o estabelecimento de um protocolo para o enriquecimento de proteínas nucleares de cana-de-açúcar e a consequente caracterização, até o momento inédita, do proteoma nuclear de folhas dessa espécie que se encontravam em dois estádios fenológicos distintos. A eficiência desse protocolo foi avaliada de algumas maneiras independentes e todas demonstraram a aplicabilidade do mesmo em plantas de cana-de-açúcar, em diversas condições, de desenvolvimento ou mesmo submetidas à algum estresse, uma vez que a análise das proteínas mostra que apesar da possibilidade de detecção de proteínas diferenciais, a porcentagem de proteínas nucleares e de proteínas consideradas contaminantes permaneceu semelhante nos dois tratamentos, indicando a adequação do protocolo para estudos futuros.

Ademais, semelhante ao observado em outros trabalhos, foram encontradas muitas proteínas de sinalização o que reflete o papel do núcleo na sinalização e regulação gênica e, que juntamente com a presença de proteínas nucleares típicas, como as histonas, além de alguns fatores de transcrição, não encontrados em todas as repetições possivelmente em decorrência de sua baixa abundância, corroboram mais uma vez com a eficiência do protocolo otimizado durante o presente estudo. Além disso, através da identificação de proteínas antes não consideradas como nucleares, o trabalho possibilitou a criação de questionamentos, acerca da localização dessas proteínas, que possam ser esclarecidos em trabalhos futuros, uma vez que como discutido, já existem relatos do envolvimento de proteínas, antes consideradas contaminantes, como a GAPDH, em processos nucleares.

Em outras espécies o estudo do proteoma nuclear já foi usado para comparar diferentes situações, porém normalmente relacionadas a estresses hídrico e por baixa temperatura. No entanto, apesar destes estudos não se referirem a diferenças no proteoma nuclear em diversos estágios do desenvolvimento, os resultados obtidos no presente trabalho demonstram que o estudo do proteoma dessa organela também pode ser aplicado para esse tipo de análise.

Ademais, apesar das idades escolhidas serem referentes à estádios fenológicos distintos, as diferenças quanto às proteínas identificadas foram muito pequenas, uma vez que ambas as idades se encontram inseridas no período germinativo do desenvolvimento das plantas de cana-de-açúcar. No entanto, a

observação de que existem algumas diferenças no núcleo de plantas de cana-de-açúcar em diferentes estádios fenológicos fornece a base para estudos futuros que possam explorar de maneira mais detalhada essas diferenças em especial em períodos do desenvolvimento da planta completamente distintos, como por exemplo nas idades de 3 meses e 1 ano, referentes ao período germinativo e de maturação respectivamente.

REFERÊNCIAS

- ABDALLA, K. O.; THOMSON, J. A.; RAFUDEEN, M. S. Protocols for nuclei isolation and nuclear protein extraction from the resurrection plant *Xerophyta viscose* for proteomic studies. **Analytical Biochemistry**, New York, v. 384, p. 365 – 367, 2009.
- ABDALLA, K. O.; BAKER, B.; RAFUDEEN, M. S. Proteomics analysis of nuclear proteins during dehydration of the resurrection plant *Xerophyta viscose*. **Plant Growth Regulation**, Dordrecht, v. 62, p. 279 – 292, 2010.
- ABDALLA, K. O.; RAFUDEEN, M. S. Analysis of the nuclear proteome of the resurrection plant *Xerophyta viscose* in response to dehydration stress using iTRAQ with 2DLC and tandem mass spectrometry. **Journal of Proteomics**, Amsterdam, v. 75, p. 2361 – 2374. 2012
- AGRAWAL, G. K.; BOURGUIGNON, J.; ROLLAND, N.; EPHRITIKHINE, G.; FERRO, M.; JAQUINOD, M.; ALEXIOU, K. G.; CHARDOT, T.; CHAKRABORTY, N.; JOLIVET, P.; DOONAN, J. H.; RAKWAL, R. Plant Organelle Proteomics: Collaborating for optimal cell function. **Mass Spectrometry Reviews**, New York, v. 30, p. 772 – 853, 2011. ***Completar
- AGRAWAL, G. K.; RAKWAL, R. Rice proteomics: a cornerstone for cereal food crop proteomes. **Mass Spectrometry Reviews**, New York v. 25, p. 1 – 53. 2006.
- ALBERTS, B.; BRAY, D.; HOPKIN, K.; JOHNSON, A.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WALTER, P. **Essential Cell Biology**. 3. ed. New York: Garland Science, Taylor & Fancis Group, 2010. 841 p.
- AMALRAJ, R. S.; SELVARAJ, N.; VELUSWAMY, G. K.; RAMANUJAN, R. P.; MUTHURAJAN, R.; PALANIYANDI, M.; AGRAWAL, G. K.; RAKWAL, R.; VISWANATHAN, R. Sugarcane proteomics: establishment of a protein extraction method for 2-DE in stalk tissues and initiation of sugarcane proteome reference map. **Electrophoresis**, Kyoto, v. 31, p. 1959-1974, 2010.
- ANDERSEN, J. S.; MANN, M. Organellar proteomics: turning inventories into insights. **EMBO reports**, Oxford, v. 7, n. 9, p. 874 – 879, 2006.
- ANDREYEV, A. Y.; SHEN, Z.; GUAN, Z.; RYAN, A.; FAHY, E.; SUBRAMANIAM, S.; RAETZ, C. R. H.; BRIGGS, S.; DENNIS, E. A. Application of Proteomic Marker Ensembles to Subcellular Organelle Identification. **Molecular & Cellular Proteomics**, Bethesda, v. 9, p. 388 – 402. 2010
- ASAKURA, Y.; GALARNEAU, E.; WATKINS, K. P.; BARKAN, A.; WIJK, K. J. V. Chloroplast RH3 DEAD Box RNA Helicases in Maize and Arabidopsis Function in Splicing of Specific Group II Introns and Affect Chloroplast Ribosome Biogenesis. **Plant Physiology**, Lancaster, v. 159, p. 961 – 974, 2012.
- AUBOURG, S.; KREIS, M.; LECHARNY, A. The DEAD box RNA helicase family in *Arabidopsis thaliana*. **Nucleic Acids Research**, London, v. 27, n. 2, p. 628 -636, 1999.

- BAE, M. S.; CHO, E. J.; CHOI, E. Y.; PARK, O. K. Analysis of the Arabidopsis nuclear proteome and its response to cold stress. **The Plant Journal**, Hoboken, v. 36, p. 652 – 663, 2003.
- BANROQUES, J.; CORDIN, O.; DOÈRE, M.; LINDER, P.; TANNER, N. K. Analyses of the Functional regions of DEAD-Box RNA “Helicases” with Deletion and Chimera Constructs Tested *In Vivo* and *In Vitro*. **Journal of Molecular Biology**, London, v. 413, n. 2, p. 451 – 472, 2011.
- BÁRTOVÁ, E.; KREJČÍ, J.; HARNIČAROVÁ, A.; GALIOVÁ, G.; KOZUBEK, S. Histone Modifications and Nuclear Architecture: A Review. **Journal of Histochemistry & Cytochemistry**, Palo Alto, v. 56, n. 8, p. 711 – 721, 2008.
- BASSET, C. L. Cajal Bodies and Plant RNA Metabolism. **Critical Reviews in Plant Sciences**, London, v.31, n.3, p. 258 – 270, 2012.
- BECK, T.; HALL, M. N. The TOR signalling pathway controls nuclear localization of nutrient-regulated transcription factors. **Nature**, London, v. 402, p. 689 – 692, 1999.
- BLAZEK, E.; MEISTERERNST, M. A functional proteomics approach for the detection of nuclear proteins based on derepressed improtin alpha. **Proteomics**, Wheinheim, v. 6, p. 2070 – 2078, 2006.
- BOARETTO, L. F. **Análise do transcriptoma e proteoma do colmo de cana-de-açúcar relacionada ao metabolismo de sacarose**. 2012. 178 p. Tese (Doutorado em Fisiologia e Bioquímica de Plantas) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2012.
- BOISVERT, F. M.; LAM, Y. W.; LAMONT, D.; LAMOND, A. I. A Quantitative Proteomics Analysis of Subcellular Proteome Localization and Changes Induced by DNA damage. **Molecular & Cellular Proteomics**, Bethesda, v. 9, p. 457 – 470. 2010.
- BORUC, J.; ZHOU, X.; MEIER, I. Dynamics of the Plant Nuclear Envelope and Nuclear Pore. **Plant Physiology**, Lancaster, v. 158, p. 78 – 86, 2012.
- BRADFORD, M. M. A rapid and sensitive method for the quantitation of microgram quantities pf proteins utilizing the principle of protein-dye binding. **Analytical Biochemistry**, New York, v.72, p. 248 – 254, 1976.
- BRAMEIER, M.; KRINGS, A.; MACCALLUM, R. M. NucPred – Predicting nuclear localization of proteins. **Bioinformatics**, Oxford, v. 23, n. 9, p. 1159 – 1160, 2007.
- BRECI, L.; HATTRUP, E.; KEELER, M.; LETARTE, J.; JOHNSON, R.; HAYNES, P. A. Comprehensive proteomics in yeast using chromatographic fractionation, gas phase fractionation, protein gel electrophoresis, and isoeletric focusing. **Proteomics**, Wheinheim, v. 5, p. 2018 – 2028, 2005.
- BRIESEMEISTER, S.; BLUM, T.; BRADY, S.; LAM, Y.; KOHLBACHER, O.; SHATKAY, H. SherLoc2: A High-Accuracy Hybrid Method for Predicting Subcellular Localization of Proteins. **Journal of Proteome Research**, Washington, v. 8, p. 5363 – 5366, 2009.

BROHAWN, S. G.; PARTRIDGE, J. R.; WHITTLE, J. R. R.; SCHWARTZ, T. U. The Nuclear Pore Complex Has Entered the Atomic Age. **Structure**, London, v. 17, p. 1156 – 1168, 2009.

BRUNET, S.; THIBAUT, P.; GAGNON, E.; KEARNEY, P.; BERGERON, J. J. M.; DESJARDINS, M. Organelle proteomics: looking at less to see more. **Trends in Cell Biology**, Cambridge, v. 13, p. 629 – 638, 2003.

BURGE, S., KELLY, E., LONSDALE, D., MUTOWO-MUELLENET, P., MCANULLA, C., MITCHELL, A., SANGRADOR-VEGAS, A., YONG, S., MULDER, N., HUNTER, S. Manual GO annotation of predictive protein signatures: the InterPro approach to GO curation. **Database**, Oxford, 2012. Disponível em: < Database 2012; doi: 10.1093/database/bar06>. Acesso em: 20 mar. 2012.

BURKE, J. E.; SASHITAL, D. G.; ZUO, X.; WANG, Y. X.; BUTCHER, S. E. Structure of the yeast U2/U6 snRNA complex. **RNA**, New York, v. 18, p. 673 – 683, 2012.

CALSA JR., T.; FIGUEIRA, A. Serial analysis of gene expression in sugarcane (*Saccharum* spp) leaves revealed alternative C4 metabolism and putative antisense transcripts. **Plant Molecular Biology**, Dordrecht, v. 63, p. 745 – 762, 2007.

CANDIANO, G.; BRUSCHI, M.; MUSANTE, L.; SANTUCCI, L.; GHIGGERI, G. M.; CARNEMOLLA, B.; ORECCHIA, P.; ZARDI, L.; RIGHETTI, P. G. Blue silver: A very sensitive colloidal Coomassie G-250 staining for proteome analysis. **Electrophoresis**, Kyoto, v. 25. p.1327-1333, 2004.

CARRETERO-PAULET, L.; GALSTYAN, A.; ROIG-VILLANOVA, I.; MARTÍNEZ-GARCÍA, J. F.; BILBAO-CASTRO, J. R.; ROBERTSON, D. L. Genome-Wide Classification and Evolutionary Analysis of the bHLH Family of Transcription Factors in Arabidopsis, Poplar, Rice, Moss and Algae. **Plant Physiology**, Lancaster, v. 153, p. 1398 – 1412, 2010.

CARMO-FONSECA, M.; MENDES-SOARES, L.; CAMPOS, I. To be or not to be in the nucleolus. **Nature Cell Biology**, London, v. 2, p. E107 – E112, 2000.

CARRIE, C.; GIRAUD, E.; WHELAN, J. Protein transport in organelles: Dual targeting of proteins to mitochondria and chloroplasts. **The FEBS Journal**, Hoboken, v. 276, p. 1187 – 1195, 2009.

CAVALIER-SMITH, T. Predation and eukaryote cell origins: A coevolutionary perspective. **The International Journal of Biochemistry & Cell Biology**, Amsterdam, v. 41, p. 307 – 322, 2009.

CESARINO, I.; ARAÚJO, P.; MAYER, J. L. S.; LEME, A. F. P.; MAZZAFERA, P. Enzymatic activity and proteomic profile of class III peroxidases during sugarcane stem development. **Plant Physiology and Biochemistry**, Amsterdam, v.55, p. 66 – 76, 2012.

CHAN, T.A.; HERMEKING, H.; LENGAUER, C.; KINZLER, K. W.; VOGELSTEIN, B. 14-3-3Sigma is required to prevent mitotic catastrophe after DNA damage. **Nature**, London, v. 401, p. 616 – 620, 1999.

CHANDRASEKHARAN, M. B.; HUANG, F.; SUN, Z. W. Ubiquitination of histone H2B regulates chromatin dynamics by enhancing nucleosome stability. **Proceedings of the national academy of sciences of the United States of America**, Washington, v. 106, n. 39, p. 16686 – 16691, 2009.

CHANDRASEKHARAN, M. B.; HUANG, F.; SUN, Z. W. Histone H2B ubiquitination and beyond – Regulation of nucleosome stability, chromatin dynamics and the trans-histone H3 methylation. **Epigenetics**, Dordrecht, v. 5, n. 6, p. 460 – 468, 2010.

CHEN, J.; BOZZA, W.; ZHUANG, Z. Ubiquitination of PCNA and Its Essential Role in Eukaryotic Translesion Synthesis. **Cell Biochemistry and Biophysics**, Totowa, v. 60, p. 47 – 60, 2011.

CHEONG, Y. H.; MOON, B. C.; KIM, J. K.; KIM, C. Y.; KIM, M. C.; KIM, I. H.; PARK, C. Y.; KIM, J. C.; PARK, B. O.; KOO, S. C.; YOON, H. W.; CHUNG, W. S.; LIM, C. O.; LEE, S. Y.; CHO, M. J. BWMK1, a Rice Mitogen-Activated Protein Kinase, Locates in the Nucleus and Mediates Pathogenesis-Related Gene Expression by Activation of a Transcription Factor. **Plant Physiology**, Lancaster, v. 132, p. 1961 – 1972, 2003.

CHEUNG, A. Y.; REDDY, A. S. N. Editorial – Nuclear Architecture and Dynamics: Territories, Nuclear Bodies, and Nucleocytoplasmic Trafficking. **Plant Physiology**, Lancaster, v. 258, p. 23 – 25, 2012

CHOU, K. C.; SHEN, H. B. Cell-PLoc: a package of Web servers for predicting subcellular localization of proteins in various organisms. **Nature Protocols**, London, v. 3, n. 2, p. 153 – 162, 2008.

CHOU, K. C.; SHEN, H. B. Plant-mPLoc: A Top-Down Strategy to Augment the Power for Predicting Plant Protein Subcellular Localization. **PLoS ONE**, San Francisco, v. 5, n. 6, 2010. Disponível em: < <http://www.plosone.org/article/info:doi/10.1371/journal.pone.0011335> >. Acesso em: 2 jul. 2012.

CHOUDHARY, M. K.; BASU, D.; DATTA, A.; CHAKRABORTY, N.; CHAKRABORTY, S. Dehydration-responsive Nuclear Proteome of Rice (*Oryza sativa* L.) Illustrates Protein Network, Novel Regulators of Cellular Adaptation, and Evolutionary Perspective. **Molecular & Cellular Proteomics**, Bethesda, v. 8, p. 1579 – 1598, 2009.

CIECHANOVER, A. The ubiquitin-proteasome pathway: on protein death and cell life. **The EMBO Journal**, London, v.17, n. 24, p. 7151 – 7160, 1998.

CIOCE, M.; LAMOND, A. I. Cajal Bodies: A Long History of Discovery. **Annual Review of Cell and Developmental Biology**, Palo Alto, v. 21, p. 105 – 131, 2005.

CLÉMENT, M.; LEONHARDT, N.; DROILLARD, M. J.; REITER, I.; MONTILLET, J. L.; GENTY, B.; LAURIÈRE, C.; NUSSAUME, L.; NOËL, L. D. The Cytosolic/Nuclear HSC70 e HSP90 Molecular Chaperones Are Important for Stomatal Closure and Modulate Absciscic Acid-Dependent Physiological Responses in Arabidopsis. **Plant Physiology**, Lancaster, v. 156, p. 1481 – 1492, 2011.

COMPANHIA NACIONAL DE ABASTECIMENTO. **Acompanhamento da Safra Brasileira: Cana-de-Açúcar. Primeiro levantamento, abril/2012.** Brasília: Conab, 2012. 19 p.

CRAMER, P.; BUSHNELL, D. A.; KORNBERG, R. D. Structural Basis of Transcription: RNA Polymerase II at 2.8 Ångstrom Resolution. **Science**, Washington, v. 292, p. 1863 – 1876, 2001.

DANIELS, J.; ROACH, B. T. Taxonomy and evolution. In: HEINZ, D. J. T. (Ed.). **Sugarcane Improvement through Breeding.** Amsterdam: Elsevier, 1987. chap. 2, p. 7 – 87.

de la CRUZ, J.; KRESSLER, D.; LINDER, P. Unwinding RNA in *Saccharomyces cerevisiae*: DEAD-box proteins and related families. **Trends in Biochemical Sciences**, Cambridge, v. 24, p. 192 – 198, 1999.

de VRIES, S. C. 14-3-3 Proteins in Plant Brassinosteroid Signalling. **Developmental cell**, Cambridge, v. 13, p. 162 - 164, 2007.

DREGER, M. Subcellular Proteomics. **Mass Spectrometry Reviews**, New York. v. 22, p. 27 – 56, 2003.

DREGER, M. Proteome analysis at the level of subcellular structures. **European Journal of Biochemistry**, Berlin, v. 270, p. 589 – 599, 2003b.

DRURY, G. E.; DOWLE, A. A.; ASHFORD, D. A.; WATERWORTH, W. M.; THOMAS, J.; WEST, C. E. Dynamics of Plant Histone Modifications in Response to the DNA Damage. **Biochemical Journal**, London, v. 445, n. 3, p. 393 – 401, 2012.

DUNDR, M.; MISTELI, T. Functional architecture in the cell nucleus. **Biochemical Journal**, London, v. 356, p. 297 – 310, 2001.

DYALL, S. D.; BROWN, M. T.; JOHNSON, P. J. Ancient Invasions: From Endosymbionts to Organelles. **Science**, Washington, v. 304, p. 253 – 257, 2004.

EMANUELSSON, O.; BRUNAK, S.; VON HEIJNE, G.; NIELSEN, H. Locating proteins in the cell using TargetP, SignalP and related tools. **Nature Protocols**, London, v. 2, n. 4, p. 953 – 971, 2007.

ERHARDT, M.; ADAMSKA, I.; FRANCO, O. L. Plant nuclear proteomics – inside the cell maestro. **The FEBS Journal**, Hoboken, v. 277, p. 3295 – 3307, 2010.

FRANKHAUSER, C.; CHEN, M. Transposing phytochrome into the nucleus. **Trends in Plant Science**, Oxford, v. 13, n. 11, p. 596 – 601, 2008.

FOLTA, K. M.; KAUFMAN, L. S. Preparation of transcriptionally active nuclei from etiolated *Arabidopsis thaliana*. **Plant Cell Reports**, Berlin, v. 19, p. 504 – 510, 2000.

GALL, J. G. Cajal Bodies: The First 100 Years. **Annual Review of Cell and Developmental Biology**, Palo Alto, v. 16, p. 273 – 300, 2000.

GAMPALA, S. S.; KIM, T. W.; HE, J. X.; TANG, W.; DENG, Z.; BAI, M. Y.; GUAN, S.; LALONDE, S.; SUN, Y.; GENDRON, J. M.; CHEN, H.; SHIBAGAKI, N.; FERL, R. J.; EHRHARDT, D.; CHONG, K.; BURLINGAME, A. L.; WANG, Z. Y. An Essential Role for 14-3-3 Proteins in Brassinosteroid Signal Transduction in Arabidopsis. **Developmental cell**, Cambridge, v. 13, p. 177 – 189, 2007.

GRAVELEY, B. R. Sorting out the complexity of SR protein functions. **RNA**, New York, v. 6, p. 1197 – 1211, 2000.

GINDULLIS, F.; MEIER, I. Matrix Attachment Region Binding Protein MFP1 Is Localized in Discrete Domains at the Nuclear Envelope. **Plant Cell**, Rockville, v. 11, p. 1117 – 1128, 1999.

GONZÁLEZ-MELENDI, P.; WELLS, B.; BEVEN, A. F.; SHAW, P. J. Single ribosomal transcription units are linear, compacted Christmas trees in plant nucleoli. **The Plant Journal**. Hoboken, v. 27, n. 3, p. 223 – 233, 2001.

GUILFOYLE, T. J.; HAGEN, G. Auxin response factors. **Current Opinion in Plant Biology**, London, v.10, p. 453 – 460, 2007.

GYGI, S. P.; ROCHON, Y.; FRANZA, B. R.; AEBERSOLD, R. Correlation between protein and mRNA abundance in yeast. **Molecular and Cellular Biology**, Washington, v. 19, p. 1720 – 1730. 1999

HAJHEIDARI, M.; FARRONA, S.; HUETTEL, B.; KONCZ, Z.; KONCZ, C. CDKF;1 and CDKD Protein Kinases Regulate Phosphorylation of Serine Residues in the C-Terminal Domain of *Arabidopsis* RNA Polymerase II. **Plant Cell**, Rockville, v. 24, p. 1626 – 1642, 2012.

HALL, M. C.; MATSON, S.W. Helicase motifs: the engine that powers DNA unwinding. **Molecular Microbiology**, Oxford, v. 34, n.5, p. 867 – 877, 1999.

HEAZLEWOOD, J. L.; MILLAR, A.H. Integrated plant proteomics – putting the green genomes to work. **Functional Plant Biology**, Melbourne, v.30 p. 471– 482, 2003.

HEAZLEWOOD, J. A.; TONTI-FILIPPINI, J. S.; GOUT, A. M.; DAY, D. A.; WHELAN, J. Experimental Analysis of the Arabidopsis Mitochondrial Proteome Highlights Signaling and Regulatory Components, Provides Assessment of Targeting Prediction Programs, and Indicates Plant-Specific Mitochondrial Proteins. **Plant Cell**, Rockville, v. 16, p. 241 – 256, 2004.

HAYNES, P. A.; ROBERTS, T. H. Subcellular shotgun proteomics in plants: Looking beyond the usual suspects. **Proteomics**, Weinheim, v. 7, p. 2963 – 2975, 2007.

- HIMANEN, K.; WOLOSZYNSKA, M.; BOCCARDI, T. M.; DE GROEVE, S.; BRUNO, L.; VUYLSTEKE, M.; LIJSEBETTENS, M. V. Histone H2B monoubiquitination is required to reach maximal transcript levels of circadian clock genes in *Arabidopsis*. **The Plant Journal**, Hoboken, 2012 Disponível em: <<http://onlinelibrary.wiley.com/doi/10.1111/j.1365-313X.2012.05071.x/abstract?systemMessage=Wiley+Online+Library+will+be+disrupted+on+4+August+from+10%3A00-12%3A00+BST+%2805%3A00-07%3A00+EDT%29+for+essential+maintenance>>. Acesso em: 26 jun. 2012.
- HOLTGREFE, S.; GOHLKE, J.; STARMANN, J.; DRUCE, S.; KLOCKE, S.; ALTMANN, B.; WOJTERA, J.; LINDERMAYR, C.; SCHEIBE, R. Regulation of plant cytosolic glyceraldehyde-3-phosphate dehydrogenase isoforms by thiol modifications. **Physiologia Plantarum**, Sweden, v. 133, p. 211 – 228, 2008.
- HORTON, P.; PARK, K. J.; OBAYASHI, T.; FUJITA, N.; HARADA, H.; ADAMS-COLLIER, C. J.; NAKAI, K. WoLF PSORT: protein localization predictor. **Nucleic Acids Research**, London, v. 35, p. w585 – w587, 2007.
- HUBER, L. A. Is proteomics heading in the wrong direction? **Nature Reviews Molecular Cell Biology**, London, v. 4, p. 74 – 80, 2003
- HUBER, L. A.; PFALLER, K.; VIETTOR, I. Organelle Proteomics: Implications for Subcellular Fractionation in Proteomics. **Circulation Research – Journal of the America Heart Association**, Los Angeles, v. 92, p. 962 – 968, 2003.
- HUH, S. U.; KIM, K. J.; PAEK, K. H. *Capsicum annum* basic transcription factor 3 (CaBtf3) regulates transcription of pathogenesis-related genes during hypersensitive response upon *Tobacco mosaic virus* infection. **Biochemical and Biophysical research Communications**, New York, v. 417, p. 910 – 917, 2012.
- HUNTER, S.; JONES, P.; MITCHELL, A.; APWEILER, R.; ATTWOOD, T. K.; BATEMAN, A.; BERNARD, T.; BINNS, D.; BORK, P.; BURGE, S.; DE CASTRO, E.; COGGILL, P.; CORBETT, M.; DAS, U.; DAUGHERTY, L.; DUQUENNE, L.; FINN, R. D.; FRASER, M.; GOUGH, J.; HAFT, D.; HULO, N.; KAHN, D.; KELLY, E.; LETUNIC, I.; LONSDALE, D.; LOPEZ, R.; MADERA, M.; MASLEN, J.; MCANULLA, C.; MCDOWALL, J.; MCMENAMIN, C.; MI, H.; MUTOWO-MUELLENET, P.; MULDER, N.; NATALE, D.; ORENGO, C.; PESSEAT, S.; PUNTA, M.; QUINN, A. F.; RIVOIRE, C.; SANGRADOR-VEGAS, A.; SELENGUT, J. D.; SIGRIST, C. J. A.; SCHEREMETJEV, M.; TATE, J.; THIMMAJANARTHANAN, M.; THOMAS, P. D.; WU, C. H.; YEATS, C.; YONG, S. Y. InterPro in 2011: new developments in the family and domain prediction database. **Nucleic Acids Research**, London, v. 40, p. D306 – D312, 2012. Disponível em: <<http://nar.oup.com/gkr948>>. Acesso em: 20 mar. 2012.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE. **Levantamento Sistemático da Produção Agrícola**: pesquisa mensal de previsão e acompanhamento das safras agrícolas no ano civil. Rio de Janeiro, 2012. 130 p.
- JAMES, G. L. An Introduction to Sugarcane. In: JAMES, G. L. (Ed.). **Sugarcane**. Oxford: Blackwell Science Ltd, 2004. chap. 1, p. 1 – 19.

JENKINS, J. L.; LAIRD, K. M.; KIELKOPF, C. L. A Broad Range of Conformations Contribute to the Solution Ensemble of the Essential Splicing Factor U2AF.

Biochemistry, Washington, v. 51, p. 5223 – 5225, 2012.

JIANG, X. S.; DAI, J.; SHENG, Q. H.; ZHANG, L.; XIA, Q. C.; WU, J. R.; ZENG, R. A Comparative Proteomic Strategy for Subcellular Proteome Research: ICAT approach coupled with bioinformatics prediction to ascertain rat liver mitochondrial proteins and indication of mitochondrial localization for catalase. **Molecular & Cellular**

Proteomics, Bethesda, v. 4, p. 12 – 34, 2005.

JIMÉNEZ, C. R. Editorial - Sorting and zooming: Subcellular proteomics is booming!

Journal of Proteomics, Amsterdam, v. 72, p. 1 – 3, 2009.

JUNG, E.; HELLER, M.; SANCHEZ, J. C.; HOCHSTRASSER, D. F. Proteomics meets cell biology: The establishment of subcellular proteomes. **Electrophoresis**, Kyoto, v. 21, p. 3369 – 3377, 2000.

KANDASAMY, M. K.; MCKINNEY, E. C.; MEAGHER, R. B. Differential Sublocalization of Actin Variants Within the Nucleus. **Cytoskeleton**, Hoboken, v. 67, p. 729 – 743, 2010.

KALLURI, U. C.; HURST, G. B.; LANKFORD, P. K.; RANJAN, P.; PELLETIER, D. A. Shotgun proteome profile of *Populus* developing xylem. **Proteomics**, Weinheim, v. 9, p. 4871 – 4880, 2009.

KHAN, M. M. K.; KOMATSU, S. Rice proteomics: recent developments and analysis of nuclear proteins. **Phytochemistry**, New York, v. 65, p. 1671 – 1681, 2004.

KIM, J. Y.; KIM, W. Y.; KWAK, K. J.; OH, S. H.; HAN, Y. S.; KANG, H. Glycine-rich RNA-binding proteins are functionally conserved in *Arabidopsis thaliana* and *Oryza sativa* during cold adaptation process. **Journal of Experimental Botany**, Oxford, v. 61, n.9, p. 2317 – 2325, 2010.

KNOWLTON, A. A.; SALFITY, M. Nuclear localization and the heat shock proteins. **Journal of Bioscience**, Bangalore, v. 21, n. 2, p. 123 – 132, 1996.

KOŁODZIEJEK, I.; VAN DER HOORN, R. A. Mining the active proteome in plant science and biotechnology. **Current opinion in Biotechnology**, London, v. 21, p. 225 – 233, 2010

KOSE, S.; FURUTA, M.; IMAMOTO, N. Hikeshi, a Nuclear Import Carrier for Hsp70s, Protects Cells from Heat Shock-Induced Nuclear Damage. **Cell**, Cambridge, v. 149, p. 578 – 589, 2012.

LAGO, C.; CLERICI, E.; DRENI, L.; HORLOW, C.; CAPORALI, E.; COLOMBO, L.; KATER, M. M. The Arabidopsis TFIID factor AtTAF6 controls pollen tube growth. **Developmental Biology**, Orlando, v. 285, p. 91 – 100, 2005.

LAM, Y. W.; TRINKLE-MULCAHY, L.; LAMOND, A. I. The nucleolus. **Journal of Cell Science**, London, v.118, p. 1335 - 1337, 2005.

LAMOND, A. I.; EARNSHAW, W. C. Structure and Function in the Nucleus. **Science**, Washington, v. 280, p. 547 – 553, 1998.

LAMOND, A. I.; SPECTOR, D. L. Nuclear Speckles: A model for nuclear organelles. **Nature Reviews: Molecular Cell Biology**, London, v. 4, p. 605 – 612, 2003.

LEREY, L. M. S.; HEMERLY, A. S.; NOGUEIRA, E. M.; VON KRÜGER, W. M. A.; BISCH, P. M. Quantitative Proteomic Analysis of the Interaction Between the Endophytic Plant-Growth-Promoting Bacterium *Gluconacetobacter diazotrophicus* and Sugarcane. **Molecular Plant-Microbe Interactions**, New York, v. 24, p. 562 – 576, 2011.

LI, C. F.; PONTES, O.; EL-SHAMI, M.; HENDERSON, I. R.; BARNATAVICHUTE, Y. V.; CHAN, S. W. L.; LAGRANGE, T.; PIKAARD, C. S.; JACOBSEN, S. E. An ARGONAUTE4-Containing Nuclear Processing Center Colocalized with Cajal Bodies in *Arabidopsis thaliana*. **Cell**, Cambridge, v. 126, p. 93 – 106, 2006.

LIU, Z.; HONG, S. W.; ESCOBAR, M.; VIERLING, E.; MITCHELL, D. L.; MOUNT, D. W.; HALL, J. D. Arabidopsis UVH6, a Homolog of Human XPD and Yeast RAD3 DNA Repair Genes, Functions in DNA Repair and Is Essential for Plant Growth, **Plant Physiology**, Lancaster, v. 132, p. 1405 – 1414, 2003.

LUO, L.; ANDO, S.; SASABE, M.; MACHIDA, C.; KURIHARA, D.; HIGASHIYAMA, T.; MACHIDA, Y. Arabidopsis ASYMMETRIC LEAVES2 protein required for leaf morphogenesis consistently forms speckles during mitosis of tobacco BY-2 cells via signals in its specific sequence. **Journal of Plant Research**, Dordrecht. Disponível em: <<http://www.springerlink.com/content/h970336615951170/?MUD=MP>>, 2012. Acesso em: 20 jul. 2012.

MARCUSSEN, T.; OXELMAN, B.; SKOG, A.; JAKOBSEN, K. S. Evolution of plant RNA polymerase IV/V genes: evidence of subneofunctionalization of duplicated NRPD2/NRPE2-like paralogs in *Viola* (Violaceae). **BMC Evolutionary Biology**, London, v. 10, n. 45, 2010. Disponível em: <<http://www.biomedcentral.com/1471-2148/10/45/abstract/>>. Acesso em: 26 jun. 2012.

MARIN, F. R. **Fenologia** – Cana-de-açúcar. Disponível em:<http://www.agencia.cnptia.embrapa.br/gestor/cana-de-acucar/arvore/CONTAG01_68_22122006154840.html>. Acesso em: 03 jul. 2012

MAO, Y. S.; ZHANG, B.; SPECTOR, D. L. Biogenesis and function of nuclear bodies. **Trends in genetics**, Cambridge, v. 27, n. 8, p. 295 – 306, 2011.

MANZATTO, C. V.; ASSAD, E. D.; BACCA, J. F. M.; ZARONI, M. J.; PEREIRA, S. E. M. **Zoneamento Agroecológico da Cana-de-Açúcar**: expandir a produção, preservar a vida, garantir o futuro. Rio de Janeiro: Embrapa Solos, 2009. 55 p.

MEHDI, A. M.; SEHGAL, M. S. B.; KOBE, B.; BAILEY, T. L.; BODÉN, M. A probabilistic model of nuclear import proteins. **Bioinformatics**, Oxford, v. 27, n. 9, p. 1239 – 1246, 2011.

MEIER, I. mRNA export and sumoylation – Lessons from plants. **Biochimica et Biophysica Acta**, Amsterdam, v. 1819, p. 531 – 537, 2012.

MI, S.; CAI, T.; HU, Y.; CHEN, Y.; HODGES, E.; NI, F.; WU, L.; LI, S.; ZHOU, H.; LONG, C.; CHEN, S.; HANNON, G. J.; QI, Y. Sorting of Small RNAs into *Arabidopsis* Argonaute Complexes is Directed by the 5' Terminal Nucleotide. **Cell**, Cambridge, v. 133, p. 116 – 127, 2008.

MILLAR, A. H.; HEAZLEWOOD, J. L. Genomic and Proteomic Analysis of Mitochondrial Carrier Proteins in *Arabidopsis*. **Plant Physiology**, Lancaster, v. 131, p. 443 – 453, 2003.

MILLAR, A. H. Location, location, location: surveying the intracellular real state through proteomics in plants. **Functional Plant Biology**, Melbourn, v. 31, p. 563 – 571, 2004.

MILLAR, A. H.; CARRIE, C.; POGSON, B.; WHELAN, J. Exploring the Function-Location Nexus: Using Multiple Lines of Evidence in Defining the Subcellular Location of Plant Proteins. **Plant Cell**, Rockville, v. 21, p. 1625 – 1631, 2009.

MILLER, M. J.; BARRETT-WILT, G. A.; HUA, Z.; VIERSTRA, R. D. Proteomic analyses identify a diverse array of nuclear processes affected by small ubiquitin-like modifier conjugation in *Arabidopsis*. **Proceedings of the national academy of sciences of the United States of America**, Washington, v. 107, n. 38, p. 16512 – 16517, 2010.

MILLER, M. J.; VIERSTRA, R. D. Mass spectrometric identification of SUMO substrates provides insights into heat stress-induced SUMOylation in plants. **Plant Signaling & Behavior**, Austin, v. 6, n. 1, p. 130 – 133, 2011.

MING, R.; MOORE, P. H.; WU, K. K.; D'HONT, A.; GLASZMANN, J. C.; TEW, T. L.; MIRKOV, T. E.; SILVA, J.; JIFON, J.; RAI, M.; SCHNELL, R. J.; BRUMBLEY, S. M.; LAKSHMANAN, P.; COMSTOCK, J. C.; PATERSON, A. H. Sugarcane Improvement through Breeding and Biotechnology. In JANICK, J. (Ed.). **Plant Breeding Reviews**. Hoboken: John Wiley & Sons, 2006. chap. 2, p. 15 – 118.

Ministério da Agricultura Pecuária e Abastecimento – MAPA. Disponível em: <<http://www.agricultura.gov.br/vegetal/culturas/cana-de-acucar>>. Acesso em: 03 jul.2012.

Ministério da Agricultura Pecuária e Abastecimento – MAPA. Disponível em: <<http://www.agricultura.gov.br/vegetal/culturas/cana-de-acucar/saiba-mais>>. Acesso em: 03 jul. 2012. MISTELI, T. Cell biology of transcription and pre-mRNA splicing: nuclear architecture meets nuclear function. **Journal of Cell Science**, London, v. 113, p. 1841– 184 , 2000.

MISTELI, T. The concept of self organization in cellular architecture. **The Journal of Cell Biology**, New York, v. 155, p. 181 – 185, 2001.

MORIUCHI, H.; OKAMOTO, C.; NISHIHAMA, R.; YAMASHITA, I.; MACHIDA, Y.; TANAKA, N. Nuclear localization and interaction of RoIB with plant 14-3-3 proteins correlates with induction of adventitious roots by the oncogene rolB. **The Plant Journal**, Hoboken, v. 38, p. 260 – 275, 2004.

NAAMATI, A.; REGEV-RUDZKI, N.; GALPERIN, S.; LILL, R.; PINES, O. Dual Targeting of Nfs1 and Discovery of Its Novel Processing Enzyme, Icp55. **The Journal of Biological Chemistry**, Rockville, v. 284, p. 30200 – 30208, 2009.

NAKAI, K. HORTON, P. PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. **Trends in Biochemical Sciences**, Cambridge, v. 24, n. 1, p. 34 – 35, 1999.

NIEDOJADŁO, J., MIKULSKI, Z.; DEŁEŃKO, K.; SZMIDT-JAWORSKA, A.; SMOLIŃSKI, D. J.; EPSTEIN, A. L. The perichromatin region of the plant cell nucleus is the area with the strongest co-localization of snRNA and SR proteins. **Planta**, Berlin, v. 236, n. 2, p. 715 – 726, 2012. Disponível em: < <http://www.springerlink.com/content/p3395u368146657u/>>. Acesso em: 26 jun. 2012.

NIZAMI, Z.; DERYUSHEVA, S.; GALL, J. G. The Cajal Body and Histone Locus Body. **Cold Spring Harbor Perspectives in Biology**, Cary, v. 75, p. 313 – 320, 2010. Disponível em: < <http://cshperspectives.cshlp.org/content/early/2010/05/24/cshperspect.a000653.full.pdf+html> >. Acesso em: 01 jul. 2012.

OSORIO, M. B.; BÜCKER-NETO, L.; CASTILHOS, G.; TURCHETTO-ZOLET, A. C.; WIEBKE-STROHM, B.; BODANESE-ZANETTINI, M. H.; MARGIS-PINHEIRO, M. Identification and *in silico* characterization of soybean trihelix-GT and bHLH transcription factors involved in stress responses. **Genetics and Molecular Biology**, Ribeirão Preto, v. 35, p. 233 – 246, 2012.

PASQUALI, C.; FIALKA, I.; HUBER, L. A. Subcellular Fractionation, electromigration analysis and mapping of organelles. **Journal of Chromatography B**, Amsterdam, v. 722, p. 89 – 102, 1999

PANDEY, A.; CHOUDHARY, M. K.; BHUSHAN, D.; CHATTOPADHYAY, A.; CHAKRABORTY, S.; DATTA, A.; CHAKRABORTY, N. The Nuclear Proteome of Chickpea (*Cicer arietinum* L.) Reveals Predicted and Unexpected Proteins. **Journal of Proteome Research**, Washington, v. 5, p. 3301 – 3311, 2006.

PANDEY, A.; CHAKRABORTY, S.; DATTA, A.; CHAKRABORTY, N. Proteomics Approach to Identify Dehydration Responsive Nuclear Proteins from Chickpea (*Cicer arietinum* L.). **Molecular & Cellular Proteomics**, Bethesda, v. 7, p. 88 – 107, 2008.

PATTERSON, S. D.; AEBERSOLD, R. H. Proteomics: the first decade and beyond. **Nature Genetics**, New York, v.33, p. 311 – 323, 2003.

PIERLEONI, A.; MARTELLI, P. L.; FARISELLI, P.; CASADIO, R. BaCellLo: a balanced subcellular localization predictor. **Bioinformatics**, Oxford, v. 22, n. 14, p. e408 – e416, 2006.

QUEVILLON, E.; SILVENTOINEN V.; PILLAI S.; HARTE N.; MULDER N.; APWEILER R.; LOPEZ R.

InterProScan: protein domains identifier. **Nucleic Acids Research**, London, v. 33, p. W116 – W120, 2005. Disponível em: < http://nar.oxfordjournals.org/content/33/suppl_2/W116.long >. Acesso em: 25 mar. 2012.

RAMAGOPAL, S. Protein variation accompanies leaf dedifferentiation in sugarcane (*Saccharum officinarum*) and is influenced by genotype. **Plant Cell Reports**, Berlin, v. 13, p. 692 – 697, 1994.

REDDY, A. S. N.; DAY, I. S.; GÖHRING, J.; BARTA, A. Localization and Dynamics of Nuclear Speckles in Plants. **Plant Physiology**, Lancaster, v. 158, p. 67 – 77, 2012.

REGEV-RUDZKI, N.; PINES, O. Eclipsed distribution: A phenomenon of dual-targeting of protein and its significance. **BioEssays**, Cambridge, v. 29, p. 772 – 782, 2007.

REGEV-RUDZKI, N.; KARNIELY, S.; BEN-HAIM, N. N.; PINES, O. Yeast Aconitase in Two Locations and Two Metabolic Pathways: Seeing Small Amounts is Believing. **Molecular Biology of the Cell**, Bethesda, v. 16, p. 4163 – 4171, 2005.

REPETTO, O.; ROGNIAUX, H.; FIRNHABER, C.; ZUBER, H.; KÜSTER, H.; LARRÉ, C.; THOMPSON, R.; GALLARDO, K. Exploring the nuclear proteome of *Medicago truncatula* at the switch towards seed filling. **The Plant Journal**, Hoboken, v. 56, p. 398 – 410, 2008.

ROBERTS, M. R. 14-3-3 Proteins find new partners in plant cell signalling. **Trends in Plant Science**, Oxford, v. 8, n. 5, p. 218 – 223, 2003.

ROSE, J. K. C.; BASHIR, S.; GIOVANNONI, J. J.; JAHN, M. M.; SARAVANAN, R. S. Tackling the plant proteome: practical approaches, hurdles and experimental tools. **The Plant Journal**, Hoboken, v. 39, p. 715 – 733, 2004.

SAMANIEGO, R.; JEONG, S. Y.; MEIER, I.; de la ESPINA, S. M. D. Dual location of MAR-binding, filament-like protein 1 in *Arabidopsis*, tobacco and tomato. **Planta**, Berlin, v. 223, p. 1201 – 1206, 2006.

SHAO, W.; KIM, H. S.; CAO, Y.; XU, Y. Z.; QUERY, C. C. A U1-U2 snRNP Interaction Network during Intron Definition. **Molecular and Cellular Biology**, Washington, v. 32, n. 2, p. 470 – 478, 2012.

SHAW, P. J.; BROWN, J. W.S. Plant nuclear bodies. **Current Opinion in Plant Biology**, London, v. 7, p. 614 – 620, 2004.

SHIEH, G. S.; PAN, C. H.; WU, J. H.; SUN, Y. J.; WANG, C. C.; HSIAO, W. C.; LIN, C. Y.; TUNG, L.; CHANG, T. H.; FLEMING, A. B.; HILLYER, C.; LO, Y. C.; BERGER, S. L.; OSLEY, M. A.; KAO, C. F. H2B ubiquitylation is part of chromatin architecture that marks exon-intron structure in budding yeast. **BMC Genomics**, London, v. 12, 2011, Disponível em: < <http://www.biomedcentral.com/1471-2164/12/627> >. Acesso em: 02 jul. 2012.

SINGH, S. K.; ROY, S.; CHOUDHURY, S. R.; SENGUPTA, D. N. DNA repair and recombination in higher plants: insights from comparative genomics of Arabidopsis and rice. **BMC Genomics**, London, v. 11, 2010. Disponível em: < <http://www.biomedcentral.com/1471-2164/11/443/>>. Acesso em: 28 ago. 2012.

SINHA, A. K.; JAGGI, M.; RAGHURAM, B.; TUTEJA, N. Mitogen-activated protein kinase signaling in plants under abiotic stress. **Plant Signaling & Behavior**, Austin, v. 6, n. 2, p. 196 – 206, 2011.

SIROVER, M. A. Subcellular Dynamics of Multifunctional Protein Regulation: Mechanisms of GAPDH Intracellular Translocation. **Journal of Cellular Biochemistry**, New York, v. 113, p. 2193 – 2200, 2012.

SMOLIŃSKI, D. J.; KOŁOWERZO, A. mRNA accumulation in the Cajal Bodies of the diplotene larch microsporocyte. **Chromosoma**, Dordrecht, v.121, p. 37 – 48, 2012.

SPECTOR, D. L.; LAMOND, A. I. Nuclear Speckles. **Cold Spring Harbor Perspectives in Biology**, Cary, v. 3, 2011. Disponível em: < <http://cshperspectives.cshlp.org/content/3/2/a000646.full.pdf+html> >. Acesso em: 26 jun. 2012.

SUGIHARTO, B.; ERMAWATI, N.; MORI, H.; AOKI, K.; YONEKURA-SAKAKIBARA, K.; YAMAYA, T.; SUGIYAMA, T.; SAKAKIBARA, H. Identification and Characterization of a Gene Encoding Drought-Inducible Protein Localizing in the Bundle Sheath Cell of Sugarcane. **Plant Cell Physiology**, Kyoto, v. 43, p. 350 – 354, 2002.

TAOKA, K. I.; OHKI, I.; TSUJI, H.; FURUITA, K.; HAYASHI, K.; YANASE, T.; YAMAGUCHI, M.; NAKASHIMA, C.; PURWESTRI, Y. A.; TAMAKI, S.; OGAKI, Y.; SHIMADA, C.; NAKAGAWA, A.; KOJIMA, C.; SHIMAMOTO, K. 14-3-3 proteins act as intracellular receptors for rice Hd3a florigen. **Nature**, London, v. 476, p. 332 – 338, 2011.

TANAKA, N.; FUJITA, M.; HANDA, H.; MURAYAMA, S.; UEMURA, M.; KAWAMURA, Y.; MITSUI, T.; MIKAMI, S.; TOZAWA, Y.; YOSHINAGA, T.; KOMATSU, S. Proteomics of the rice cell: systematic identification of the protein populations in subcellular compartments. **Molecular Genetics & Genomics**, Berlin, v. 271, p. 566 – 576, 2004.

TAYLOR, S. W.; FAHY, E.; GHOSH, S. S. Global organellar proteomics. **TRENDS in Biotechnology**, Amsterdam, v. 21, p. 82 – 88, 2003.

TONZETICH, J. Orcein Staining and the Identification of Polytene Chromosomes. In HENDERSON, D. S. (Ed.) **Drosophila Cytogenetics Protocols**. Totowa: Humana Press, 2004. chap. 11, p. 249 – 256.

TRISTAN, C.; SHAHANI, N.; SEDLAK, T. W.; SAWA, A. The diverse functions of GAPDH: Views from different subcellular compartments. **Cellular Signalling**, Oxford, v. 23, p. 317 – 323, 2011.

TUTEJA, N. Plant DNA helicases: the long unwinding road. **Journal of Experimental Botany**, Oxford, v. 54, n. 391, p. 2201 – 2214, 2003.

TUTEJA, N.; TUTEJA, R. Unraveling DNA helicases: motif, structure, mechanism and function. **European Journal of Biochemistry**, Berlin, v. 271, p. 1849 – 1863, 2004.

UNIÃO DA INDÚSTRIA DE CANA-DE-AÇÚCAR – ÚNICA. Disponível em: <<http://www.unica.com.br/content/show.asp?cntCode={D6C39D36-69BA-458D-A95C-815C87E4404D}>>. Acesso em: 03 jul. 2012.

UNIÃO DA INDÚSTRIA DE CANA-DE-AÇÚCAR – ÚNICA. Disponível em: <<http://www.unica.com.br/linhadotempo/>>. Acesso em: 03 jul. 2012.

VENTURA, M.; MATEO, F.; SERRATOSA, J.; SALAET, I.; CARUJO, S.; BACHS, O.; PUJOL, M. J. Nuclear translocation of glyceraldehyde-3-phosphate dehydrogenase is regulated by acetylation. **The International Journal of Biochemistry & Cell Biology**, Amsterdam, v. 42, p. 1672 – 1680, 2010.

VETTORE, A. L.; SILVA, F. R.; KEMPER, E. L.; ARRUDA, P. **Genetics and Molecular Biology**, Ribeirão Preto, v. 24, p. 1 – 7, 2001

VISA, N.; PERCIPALLE, P. Nuclear Functions of Actin. **Cold Spring Harbor Perspectives in Biology**, Cary, v. 2, n. 4, a000620, 2010.

WACHTER, A.; RÜHL, C.; STAUFFER, E. The role of polypyrimidine tract-binding proteins and other hnRNP proteins in plant splicing regulation. **Frontiers in Plant Science**, Lausanne, v. 3, 2012. Disponível em: <http://www.frontiersin.org/Plant_Genetics_and_Genomics/10.3389/fpls.2012.00081/abstract>. Acesso em: 28 jul. 2012.

WAKO, T.; FUKUI, K. Higher Organization and Histone Modification of the Plant Nucleus and Chromosome. **Cytogenetic and Genome Research**, New York, v. 129, p. 55 – 63, 2010.

WANG, W.; VINOCUR, B.; SHOSEYOV, O.; ALTMAN, A. Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. **Trends in Plant Science**, Oxford, v. 9, n. 5, p. 244 – 252, 2004.

WANG, B. B.; BRENDDEL, V. Molecular Characterization and Phylogeny of U2AF³⁵ Homologs in Plants. **Plant Physiology**, Lancaster, v. 140, p. 624 – 636, 2006.

YANG, K. S.; KIM, H. S.; JIN, U. H.; LEE, S. S.; PARK, J. A.; LIM, Y. P.; PAI, H. S. Silencing of NbBTF3 results in developmental defects and disturbed gene expression in chloroplasts and mitochondria of higher plants. **Planta**, Berlin, v. 225, p. 1459 – 1469, 2007.

YU, C. S.; CHEN, Y. C.; LU, C. H.; HWANG, J. K. Prediction of Protein Subcellular Localization. **Proteins: Structure, Function, and Bioinformatics**. **Proteins**, New York, v. 64, n. 3, p. 643 – 651, 2006.

YU, X.; KLEJNOT, J.; ZHAO, X.; SHALITIN, D.; MAYMON, M.; YANG, H.; LEE, J.; LIU, X.; LOPEZ, J.; LIN, C. *Arabidopsis* Cryptochrome 2 Completes Its Posttranslational Life Cycle in the Nucleus. **Plant Cell**, Rockville, v. 19, p. 3146 – 3156, 2007.

YI, H. C.; SARDESAI, N.; FUJINUMA, T.; CHAN, C. W.; VEENA; GELVIN, S. B. Constitutive Expression Exposes Functional Redundancy between Arabidopsis Histone H2A Gene HTA1 and Other H2A Gene Family Members. **Plant Cell**, Rockville, v. 18, p. 1575 – 1589, 2006.

ZHAO, H.; XING, D.; LI, Q. Q. Unique Features of Plant Cleavage and Polyadenylation Specificity Factor Revealed by Proteomic Studies. **Plant Physiology**, Lancaster, v. 151, p. 1546 – 1556, 2009.

ANEXO

ANEXO A

Figura 19 – Classificação funcional das proteínas encontradas em núcleos de folhas de cana-de-açúcar de 1 mês de idade

Figura 20 – Classificação funcional das proteínas encontradas em núcleos de folhas de cana-de-açúcar de 4 meses de idade.

ANEXO C

Tabela 5 – Proteínas identificadas na extração teste

(continua)

Função	Identificação	Banda	Cluster	Vezes	B	S	W	P	PI	E.C.
desenvolvimento	sp Q337C0 U496D_ORYSJ. UPF0496 protein 4	17	SCSFAD1108E01 (3E-80)	1	X	nd	nd	nd	X	
divisão de peroxissomos	sp Q5VRJ8 PX115_ORYSJ. Peroxisomal membrane protein 11-5	9	SCCCRZ2001B03 (1E-107)	1	nd	nd	X	nd	nd	
estrutura	sp P02582 ACT1_MAIZE. Actin-1	12	SCEZLR1009B09 (0)	1	nd	nd	nd	X	nd	
fixação do carbono	sp P15804 CAPP3_SORBI. Phosphoenolpyruvate carboxylase 3	1; 2; 3; 4; 5; 6; 7; 8; 9; 11; 12; 19	SCRULR1020D09 (0)	12	nd	nd	nd	X	nd	<u>EC=4.1.1.31</u>
fosforilação	sp Q67C40 MPK7_ORYSJ. Mitogen-activated protein kinase 7	13	SCEZAM2031C06 (1E-49)	1	X	nd	nd	nd	X	<u>EC=2.7.11.24</u>
fotossíntese	sp O65101 PSAH_MAIZE. Photosystem I reaction center subunit VI, chloroplastic	25; 26	SCJFLR2035B05 (3E-34)	2	X	nd	nd	X	nd	
	sp O65107 PSAN_MAIZE. Photosystem I reaction center subunit N, chloroplastic (Fragment)	26; 27; 28	SCACLR1129A11 (5E-43)	3	X	nd	nd	nd	nd	
	sp P12329 CB21_MAIZE. Chlorophyll a-b binding protein 1, chloroplastic	1; 2; 3; 4; 5; 6; 7; 8; 9; 11; 12; 13; 18; 18; 20; 21; 22; 23	SCJFLR1074A11 (1E-141)	19	X	nd	nd	nd	nd	
	sp P27524 CB4A_SOLLC. Chlorophyll a-b binding protein CP24 10A, chloroplastic	1; 8; 12; 20; 21; 22; 23	SCJFSB1014F03.b (4E-45)	7	X	nd	nd	nd	nd	
	sp P27525 CB4B_SOLLC. Chlorophyll a-b binding prot ein CP24 10B, chloroplastic	1; 8; 12; 20; 21; 22; 23	SCACLR2007B05 (1E-108) / SCRLLV1051H04 (7E-25)	8	X	nd	nd	nd	nd	
	sp Q41048 PSBQ1_MAIZE. Oxygen-evolving enhancer protein 3-1, chloroplastic	8; 12	SCBGLR1114G09 (1E-67)	2	X	nd	nd	nd	nd	
função desconhecida	sp Q8H6A5 TCTP_MAIZE. translationally-controlled tumor protein homolog	19	SCCCLR2001G09 (2E-80) / SCEZLR1031E10 (2E-80)	2	nd	nd	X	nd	nd	

Tabela 5 – Proteínas identificadas na extração teste

(continuação)

Função	Identificação	Banda	Cluster	Vezez	B	S	W	P	PI	E.C.
metabolismo de carboidratos metabolismo de proteínas metabolismo de proteínas metabolismo do DNA	sp Q9LS39 CCR4C_ARATH. Carbon catabolite repressor protein 4 homolog 3	5	SCEPLB1043H09 (9E-66)	1	nd	nd	X	nd	X	<u>EC=3.1.13.4</u>
	tr A2E1P4 A2E1P4_TRIVA Beige/BEACH domain containing protein	11	SCRFAM2128F04 (5.0×10-1)	1	X	nd	nd	nd	nd	
	tr C5WRY9 C5WRY9_SORBI Putative uncharacterized protein Sb01g000680	19	SCJFRT1012A02 (3.0×10-75)	1	X	nd	nd	nd	nd	
	tr C5XF99 C5XF99_SORBI Putative uncharacterized protein Sb03g029860	11	SCCCAM2C05B03 (1.0×10-119)	1	X	X	X	X		
	tr C5XHI8 C5XHI8_SORBI Putative uncharacterized protein Sb03g032060	11	SCRUAD1064H05.b (4.0×10-12)	1	X	X	nd	nd	X	
	tr C5XIC1 C5XIC1_SORBI Putative uncharacterized protein Sb03g033020	5	SCBFST3135C09.b (1.0×10-122)	1	X	nd	nd	X	X	
	tr C5XMJ3 C5XMJ3_SORBI Putative uncharacterized protein Sb03g037470	5	SCCCT1004A06 (2.0×10-71)	1	X	nd	nd	X		
	tr C5Y9P5 C5Y9P5_SORBI Putative uncharacterized protein Sb06g018500	26	SCVPFL3048C02 (3.0×10-91)	1	nd	nd	nd	nd	X	
	tr C5YBP8 C5YBP8_SORBI Putative uncharacterized protein Sb06g000920	11	SCCCST2002G06 (7E-76)	1	X	nd	nd	nd	nd	
	tr C5YVN9 C5YVN9_SORBI Putative uncharacterized protein Sb09g010000	8; 9; 11; 12	SCACSB1122B11 (1.0×10-103)	4	X	nd	nd	X	X	
	tr C9YSZ6 C9YSZ6_STRSW Putative integral membrane protein	9	SCCCSD1001D11 (3.0)	1	X	nd	nd	nd	nd	
	tr E6Q3E5 E6Q3E5_9ZZZZ Putative Beta-galactosidase	16	SCEQSD1077B02 (6.7×10-1)	1	X	nd	nd	nd	nd	
	sp Q6H795 CLPD1_ORYSJ. Chaperone protein ClpD1, chloroplastic	1; 3; 4; 5; 6	SCQGRT1041D10 (9E-73)	5	X	nd	nd	nd	nd	
	tr C5YF84 C5YF84_SORBI Putative uncharacterized protein Sb06g014590	1; 3; 4; 5; 6	SCQGLR1062F08 (0)	5	nd	nd	nd	X	nd	
	sp Q9LEF5 SSRP1_MAIZE. FACT complex subunit SSRP1	6	SCCCCL3120G05 (0)	1	X	X	X	X	X	
	sp Q9ZUP2 RECA3_ARATH. metabolismo do DNA protein recA homolog 3, mitochondrial	9	SCSBLB1034B03 (5E-47)	1	nd	nd	nd	nd	X	

Tabela 5 – Proteínas identificadas na extração teste

(continuação)

Função	Identificação	Banda	Cluster	Vezez	B	S	W	P	PI	E.C.
metabolismo do RNA	tr C5XAM3 C5XAM3_SORBI Putative uncharacterized protein Sb02g022780	1	SCVPAM1056G07 (1.0×10 ⁻¹⁶²)	1	X	nd	nd	nd	nd	
	sp Q2QZL4 U2A2B_ORYSJ. Splicing factor U2af large subunit B	3	SCUTFL3079A09 (5E-9)	1	X	X	X	X	nd	
	sp Q8BG79 C19L2_MOUSE. CWF19-like protein 2	7	SCEPAM1022H01 (3E-54)	1	X	X	X	X	X	
	sp Q9C952 CPSF3_ARATH. Cleavage and polyadenylation specificity factor subunit 3-l	12	SCMCFL8030F07 (1E-76)	1	X	nd	nd	nd	nd	<u>EC=3.1.27.-</u>
microtubule-based movement montagem de nucleossomos	tr C5WPG0 C5WPG0_SORBI Putative uncharacterized protein Sb01g039540	9	SCEQSB1C11G05 (3.0×10 ⁻⁵⁵)	1	X	X	X	X	X	
	sp O81635 ATK4_ARATH. Kinesin-4	26	SCJFRZ1005G10 (1E-125)	1	X	nd	X	nd	X	
	sp P0CG89 H4_SOYBN. Histone H4	22; 23; 24; 25; 26	SCCCLR2001D01 (2E-35) / SCCCRZ2001G01 (1E-36) / SCCCLR2002D04 (2E-39) / SCAGLB1069E02 (4E-39) / SCAGLB1070C07 (1E-39) / SCCCRZ1004C05 (2E-39) / SCSBFL1043F08 (2E-39) / SCCCLR2001D12 (4E-38) / SCSGHR1071E06 (2E-18) / SCCCLR2002G09 (2E-39) / SCJLLR2020B06 (1E-39) / SCJFLR1013A12 (9E-37) / SCEPLR1051H12 (3E-39) / SCCCLR1078H02 (2E-39) / SCJFRZ2009F02 (2E-39) / SCJLLR1011E11 (9E-37) / SCRFLR2038C05 (4E-42) / SCUTLR2015B06 (3E-37) / SCVPRZ2035C06 (2E-35) / SCSBLB1033H09 (1E-39) / SCCCLR2003E01 (1E-39) / SCSGRT2062A10 (1E-38) / SCJLHR1028C12 (1E-29)	115	X	X	X	X	X	
	sp P27807 H2B1_WHEAT. Histone H2B.1	22	SCCCLR2C02B06 (1E-46)	1	X	X	X	X	X	
	sp P54348 H2B5_MAIZE. Histone H2B.5	22	SCCCLR1C04B01 (1E-36) / SCCCLR2002G11 (5E-37) / SCCCLR1048C02 (2E-36) / SCEQLR1050A02 (2E-38) / SCBGFL3095E08 (1E-36) / SCUTLR2015D03 (8E-37) / SCEPLB1042A02 (7E-37) / SCCCLR1024A09 (1E-36) / SCCCRZ1001H04 (1E-36) / SCJLLR2020F01 (7E-37)	10	X	X	X	X	X	
	sp Q28DR4 H4_XENTR. Histone H4	25	SCAGLR2026F05 (7E-27)	1	X	X	X	X	X	
	sp Q43216 H2B5_WHEAT. Histone H2B.5	22	SCMCLR1032A12 (7E-16)	1	X	X	X	X	X	

Tabela 5 – Proteínas identificadas na extração teste

(continuação)

Função	Identificação	Banda	Cluster	Vezes	B	S	W	P	PI	E.C.
processo de oxidação-redução	sp Q6ZBP3 H2B2_ORYSJ. Histone H2B.2	22	SCBGLR1002E03 (1E-43) / SCMCRT2108A05 (4E-43) / SCQGLR1041C10 (9E-44) / SCQGLR1062G09 (1E-43) / SCSFFL3090D03 (7E-44) / SCJLLR1101H08 (4E-44) / SCCCZR1004F04 (5E-42) / SCEZLB1014E02 (4E-43) / SCCCLR1068G11 (5E-44) / SCSGAM1096G11 (8E-44) / SCSGFL5C05H11 (2E-33)	11	X	X	X	X	X	
	sp Q943L2 H2B11_ORYSJ Histone H2B.11	22	SCVPLR1049C05 (1E-37)	1	X	X	X	X	X	
	sp O04130 SERA_ARATH. D-3-phosphoglycerate dehydrogenase, chloroplastic	9	SCCCLR2002G02 (0)	1	nd	nd	nd	X	nd	<u>EC=1.1.1.95</u>
	sp Q69TY4 PR2E1_ORYSJ. Peroxiredoxin-2E-1, chloroplastic	23	SCEZRZ1014F10 (7E-84)	1	nd	nd	nd	nd	X	<u>EC=1.11.1.15</u>
	sp Q7FB12 NDHM_ORYSJ. NAD(P)H-quinone oxidoreductase subunit M, chloroplastic	20	SCEPSD2005D11 (1E-72)	1	nd	nd	nd	nd	X	<u>EC=1.6.5.-</u>
processos biossintéticos	sp P34106 ALA2_PANMI. Alanine aminotransferase 2	12	SCCCZR1002E01 (1E-137)	1	X	nd	nd	nd	nd	<u>EC=2.6.1.2</u>
processos celulares	sp P04907 GSTF3_MAIZE. Glutathione S-transferase 3	21	SCJFLR1073H08 (1E-91) / SCCCLR1001A07 (2E-92) / SCEZRT3070A09 (7E-71)	3	X	nd	nd	nd	nd	<u>EC=2.5.1.18</u>
processos metabólicos	sp Q9M158 Y4105_ARATH. Uncharacterized protein At4 g01050, chloroplastic	11	SCEZHR1047F05 (2E-68)	1	X	nd	nd	nd		
	sp P43309 PPO_MALDO. Polyphenol oxidase, chloroplastic	9	SCEZLB1007D07 (7E-45) / SCEPAM1051D07 (5E-66)	2	nd	nd	nd	X	nd	<u>EC=1.10.3.1</u>
	tr A5EGJ0 A5EGJ0_BRASB Putative Short-chain dehydrogenase/reductase	3	SCUTAD1030C02 (2.7×10 ⁻¹)	1	X	nd	nd	nd	nd	<u>EC=1.1.1.62</u>
	tr A6DSB2 A6DSB2_9BACT Monooxygenase family protein	12	SCRLAM1012D03 (2.7×10 ⁻¹)	1	X	X	nd	X	nd	
proteólise	sp Q8LQJ8 FTSH5_ORYSJ. ATP-dependent zinc metalloprotease FTSH 5, mitochondrial	8; 9	SCCCLR1078F01 (0)	2	X	nd	nd	nd	X	<u>EC=3.4.24.-</u>
	sp Q9LSU3 PSA6_ORYSJ. Proteasome subunit alpha type-6	19	SCCCLR1001B04 (1E-133)	1	X	nd	nd	nd	X	<u>EC=3.4.25.1</u>

Tabela 5 – Proteínas identificadas na extração teste

(continuação)

Função	Identificação	Banda	Cluster	Vezes	B	S	W	P	PI	E.C.
regulação da transcrição DNA-dependente	sp Q6PC91 BT3L4_DANRE. Transcription factor BTF3 homolog 4	3	SCUTSD2025B03 (1E-11)	1	X	X	nd	X	X	
	sp Q8S983 ARFK_ORYSJ. Auxin response factor 11	5	SCCCFL5061G06 (1E-107)	1	X	X	X	nd	X	
regulação da transcrição RNA-dependente	sp Q8S0J7 IM30_ORYSJ. Probable membrane-associated 30 kDa protein, chloroplastic	16	SCUTLR2015H12 (1E-151)	1	nd	nd	nd	X	nd	
	sp P09189 HSP7C_PETHY. Heat shock cognate 70 kDa protein	6	SCCCCL3120G07 (0) / SCCCLR1C01D05 (0) / SCAGLR1043F02 (0)	3	nd	nd	nd	X	nd	
	sp P27322 HSP72_SOLLC. Heat shock cognate 70 kDa protein 2	7; 8	SCCCRZ1002G01 (0)	2	nd	nd	nd	X	nd	
	sp Q0J4P2 HSP81_ORYSJ. Heat shock protein 81-1	1	SCEZLB1006B07 (1E-100) / SCCCLR1075E09 (0)	2	X	nd	nd	X	nd	
sinalização	sp P49106 14331_MAIZE. 14-3-3-like protein GF14-6	16; 17	SCCCLR1022D05 (1E-142)	3	X	nd	nd	nd	X	
	sp Q6ZKC0 14333_ORYSJ. 14-3-3-like protein GF14-C	16; 17	SCEQRT1031D02 (1E-113) / SCCCRZ1001D02 (1E-134) / SCBFLR1026E02 (1E-131) --> - 3	9	X	nd	nd	nd	X	
	sp Q92562 FIG4_HUMAN. Polyphosphoinositide phosphatase	7	SCJFRT1059F07 (2E-44)	1	X	nd	nd	X	nd	<u>EC=3.1.3.-</u>
	sp Q9SP07 1433_LILLO. 14-3-3-like protein	16; 17	SCAGLB1069G08 (2E-76)	3	X	nd	nd	nd	X	
tradução	sp A1E9W9 RR7_SORBI. 30S ribosomal protein S7, chloroplastic	22	SCCCLR2C01E06 (8E-81)	1	nd	nd	nd	X	nd	
	sp O48557 RL17_MAIZE. 60S ribosomal protein L17	21	SCJFLR1017D02 (8E-82) / SCJFRZ2005H06 (6E-88)	2	nd	nd	X	nd	nd	
	sp O64937 EF1A_ORYSJ. Elongation factor 1-alpha	1; 3; 5; 8; 11; 12	SCVPLB1020E07 (0) / SCCCLR1022B09 (0) / SCJLRT1006G01 (0) / SCEZRZ1013H01 (0)	24	nd	nd	nd	X	X	
	sp O65731 RS5_CICAR. 40S ribosomal protein S5 (Fragment)	20; 21; 22	SCACLR1127E08 (3E-94) / SCRFLR2038B04 (9E-72) / SCUTSD2088D06 (8E-60)	5	X	nd	nd	nd	nd	
	sp O80360 RK3_TOBAC. 50S ribosomal protein L3, chloroplastic (Fragment)	19	SCEQLB1066D08 (2E-97)	1	nd	nd	X	X	nd	
	sp P19950 RS141_MAIZE. 40S ribosomal protein S14	24	SCCCLR1048A02 (5E-62) / SCCCLR1C03A02 (1E-61)	2	X	nd	nd	nd	nd	

Tabela 5 – Proteínas identificadas na extração teste

(continuação)

Função	Identificação	Banda	Cluster	Vezez	B	S	W	P	PI	E.C.
	sp P19951 RS142_MAIZE. 40S ribosomal protein S14	24	SCRUFL3066B06.b (9E-58) / SCCCRT2002B03 (6E-63)	2	X	nd	nd	nd	nd	
	sp P34823 EF1A2_DAUCA. Elongation factor 1-alpha	1; 3; 5; 8; 11; 12	SCSBRT3039F07 (8E-45)	6	nd	nd	nd	nd	X	
	sp P41378 IF4A_WHEAT. Eukaryotic initiation factor 4A	12	SCRFR3057C02 (2E-74)	1	X	nd	nd	nd	X	<u>EC=3.6.4.13</u>
	sp P42798 R15A1_ARATH. 40S ribosomal protein S15a-1	25	SCCCRZ2C01G03 (7E-68) / SCCCLR2004D07 (9E-68) / SCMCST1058E10 (1E-67)	3	nd	nd	nd	nd	X	
	sp P49199 RS8_ORYSJ. 40S ribosomal protein S8	19	SCCCLR1048F11 (5E-94) / SCQSST3114E08 (3E-41) / SCCCCR22002D10 (5E-78)	3	nd	nd	nd	X	nd	
	sp P49210 RL9_ORYSJ. 60S ribosomal protein L9	21	SCCCLR1069A12 (7E-90) / SCBFRZ2045D04 (7E-94)	2	X	nd	nd	nd	nd	
	sp P50883 RL121_ARATH. 60S ribosomal protein L12-1	22; 23	SCAGLR1021E04 (2E-80) / SCCCLR1C01C09 (4E-77)	4	nd	nd	nd	nd	X	
	sp Q03033 EF1A_WHEAT. Elongation factor 1-alpha	12	SCCCLB1021F05 (3E-89) / SCJFAD1012C08 (8E-90)	2	nd	nd	nd	X	X	
	sp Q08699 RS14_PODCA. 40S ribosomal protein S14	24	SCSBS2054A09 (6E-38)	1	X	nd	nd	nd	nd	
	sp Q0DK10 RL11_ORYSJ. 60S ribosomal protein L11	22	SCCCRZ2001F06 (7E-88) / SCVPLR2019F01 (5E-88) / SCSGLR1045F12 (3E-87)	3	nd	nd	nd	X	nd	
	sp Q41741 IF4A_MAIZE. Eukaryotic initiation factor 4A	12	SCCCLR1C07G11 (0)	1	X	X	nd	nd	X	<u>EC=3.6.4.13</u>
	sp Q41803 EF1A_MAIZE. Elongation factor 1-alpha	1; 3; 5; 8; 11; 12	SCEQLR1007G03 (0) / SCCCL3004F01.b (0)	12	nd	nd	nd	X	X	
	sp Q5N1T5 EFP_SYNP6. Elongation factor P	20	SCQSAM2099H05 (5E-44)	1	X	nd	nd	nd	nd	
	sp Q6Z2Z4 IF4A3_ORYSJ. Eukaryotic initiation factor 4A-3	12	SCCCLR1072G07 (0)	1	X	X	nd	nd	X	<u>EC=3.6.4.13</u>
	sp Q8LJU5 RS7_ORYSJ. 40S ribosomal protein S7	20	SCVPRZ2038F11 (5E-95)	1	X	nd	nd	X	nd	
	sp Q9AYP4 RS10_ORYSJ. 40S ribosomal protein S10	22	SCCCLR2C01F11 (3E-50) / SCMCST1051E11 (2E-42) / SCACLR1126F04 (5E-60)	3	X	nd	X	nd	X	
	sp Q9C285 RL12_NEUCR. 60S ribosomal protein L12	22	SCEQRT1024D04 (7E-77)	1	nd	nd	nd	X	X	
	sp Q9LZ17 RS174_ARATH. 40S ribosomal protein S17-4	23	SCBGLR1002D11 (8E-54) / SCRLFL1011G08 (2E-48) / SCCCLR1C05F01 (6E-54)	3	X	nd	X	X	X	

Tabela 5 – Proteínas identificadas na extração teste

(conclusão)

Função	Identificação	Banda	Cluster	Vezes	B	S	W	P	PI	E.C.
transporte	sp Q9XET4 RS7_SECCE. 40S ribosomal protein S7	20	SCVPLR2005E09 (2E-96) / SCCCCL4010B12 (8E-98)	2	X	nd	nd	X	nd	
	sp Q9ZUT9 RS51_ARATH. 40S ribosomal protein S5-1	20; 21; 22	SCCCCL3001E09.b (7E-93)	3	X	nd	nd	nd	nd	
	tr B6UH09 B6UH09_MAIZE 40S ribosomal protein S14	24	SCJLRZ1022H11 (3.0×10 ⁻²³)	1	X	nd	nd	nd	X	
	tr C5WYH8 C5WYH8_SORBI Putative uncharacterized protein Sb01g019490	6	SCRLLR1016B06 (1.0×10 ⁻¹⁵³)	1	nd	nd	nd	X		
	tr C5Y0E9 C5Y0E9_SORBI Putative uncharacterized protein Sb04g030890	22; 23	SCACLR2014B04 (2E-84)	2	X	nd	nd	nd	X	
	sp P42056 VDAC2_SOLTU. Mitochondrial outer membrane protein porin of 36 kDa	3	SCCCSB1002A08 (2E-50)	1	X	nd	nd	nd	nd	
	sp Q10D38 MRS2I_ORYSJ. Magnesium transporter MRS2- I	1	SCJFRZ2028F12 (1E-163)	1	X	nd	nd	nd	nd	

Tabela 6 – Proteínas nucleares identificadas em folhas de 1 mês de idade

(continua)

Função	Identificação	Banda	Cluster	vezes	B	S	W	P	PI	E.C.
biogênese de ribossomos	sp P35685 RL7A_ORYSJ. 60S ribosomal protein L7a	11	SCVPLR1028C12 (1E-101)	1	X	nd	nd	X	nd	
	sp P37107 SR54C_ARATH. Signal recognition particle 54 kDa protein, chloroplastic	8	SCQGRT1042A05 (0)	1	nd	nd	nd	X	nd	
fixação do carbono	sp P15804 CAPP3_SORBI. Phosphoenolpyruvate carboxylase 3	1; 2; 3; 4; 5; 6; 7; 8; 9; 10; 11; 12; 13; 14; 15 19; 20	SCRULR1020D09 (0)	15	nd	nd	nd	X	nd	<u>EC=4.1.1.31</u>
fotossíntese	sp O65107 PSAN_MAIZE. Photosystem I reaction center subunit N, chloroplastic (Fragment)		SCACLR1129A11 (5E-43)	2	X	nd	nd	nd	nd	
	sp O80634 PPL2_ARATH. PsbP-like protein 2, chloroplastic	16	SCJLLR1102C07 (1E-55)	1	nd	nd	X	nd	nd	
	sp P12329 CB21_MAIZE. Chlorophyll a-b binding protein 1, chloroplastic	1; 2; 3; 4; 5; 6; 7; 8; 9; 10; 11; 12; 13; 14; 15; 16; 19	SCJFLR1074A11 (1E-141)	17	X	nd	nd	nd	nd	
	sp P27524 CB4A_SOLLC. Chlorophyll a-b binding protein CP24 10A, chloroplastic	1; 9; 10; 13; 14; 15	SCJFSB1014F03.b (4E-45)	6	X	nd	nd	nd	nd	
	sp P27525 CB4B_SOLLC. Chlorophyll a-b binding protein CP24 10B, chloroplastic	1; 8; 9; 10; 13; 14; 15	SCACLR2007B05 (1E-108)	7	X	nd	nd	nd	nd	
	sp Q41048 PSBQ1_MAIZE. Oxygen-evolving enhancer protein 3-1, chloroplastic	9; 10	SCBGLR1114G09 (1E-67)	2	X	nd	nd	nd	nd	
	função desconhecida	sp P94063 HAL3B_ARATH. Probable phosphopantothenoyl cysteine decarboxylase	7	SCQSRT2031G03 (2E-32)	1	X	nd	nd	X	nd
sp Q8H6A5 TCTP_MAIZE. translationally-controlled tumor protein homolog		13; 14	SCCCLR2001G09 (2E-80) / SCEZLR1031E10 (2E-80)	4	nd	nd	X	nd	nd	

Tabela 6 – Proteínas nucleares identificadas em folhas de 1 mês de idade

(continuação)

Função	Identificação	Banda	Cluster	vezes	B	S	W	P	PI	E.C.
	tr A2E1P4 A2E1P4_TRIVA Beige/BEACH domain containing protein	8	SCRFAM2128F04 (5.0×10 ⁻¹)	1	X	nd	nd	nd	nd	
	tr C0PAK4 C0PAK4_MAIZE Putative uncharacterized protein	13; 14; 15; 19	SCRUAD1133C09.b (3.0×10 ⁻⁷⁵)	4	nd	nd	nd	X	nd	
	tr C5WRY9 C5WRY9_SORBI Putative uncharacterized protein Sb01g000680	12; 13	SCJFRT1012A02 (3.0×10 ⁻⁷⁵)	2	X	nd	nd	nd	nd	
	tr C5WVY3 C5WVY3_SORBI Putative uncharacterized protein Sb01g045750	16	SCEPSD1004E09 (6E-54)	1	nd	nd	nd	nd	X	
	tr C5XIC1 C5XIC1_SORBI Putative uncharacterized protein Sb03g033020	1	SCBFST3135C09.b (1.0×10 ⁻¹²²)	1	X	nd	nd	X	X	
	tr C5YBP8 C5YBP8_SORBI Putative uncharacterized protein Sb06g000920	8	SCCST2002G06 (7E-76)	1	X	nd	nd	nd	nd	
	tr C5YJN5 C5YJN5_SORBI Putative uncharacterized protein Sb07g028880	12	SCCCT1C06B08 (8E-79)	1	nd	nd	nd	X	X	
	tr C5YTS7 C5YTS7_SORBI Putative uncharacterized protein Sb08g007230	10	SCAGFL3023D11 (0)	1	X	nd	nd	nd	nd	
	tr C5YVN9 C5YVN9_SORBI Putative uncharacterized protein Sb09g010000	7; 8; 9	SCACSB1122B11 (1.0×10 ⁻¹⁰³)	3	X	nd	nd	X	X	
glicólise	sp Q40677 ALFC_ORYSJ. Fructose- bisphosphate aldolase, chloroplastic	10	SCMCLV1030C12 (2E-60) / SCRULR1020D11 (0)	2	X	nd	nd	nd	nd	
homeostase celular redox	sp O64628 TXND9_ARATH. Thioredoxin domain-containing protein 9 homolog	13; 19	SCEZRZ1012C04 (3E-61)	2	X	X	X	X	X	
metabolismo da membrana	sp Q6H7M7 OEP80_ORYSJ. Outer envelope protein of 80 kDa	7	SCEPAM2014E05 (0)	1	X	nd	nd	nd	nd	
metabolismo de ácidos nucleicos	sp P82277 RRP2_SPIOL. 30S ribosomal protein 2, chloroplastic	14	SCRLRZ3039A08 (1E-47)	1	nd	X	nd	nd	nd	
metabolismo de aminoácidos	sp Q9SJ05 AROK_ARATH. Shikimate kinase, chloroplastic	13	SCQSST3114G03 (2E-20)	1	X	nd	nd	nd	nd	<u>EC=2.7.1.71</u>
metabolismo de ATP	sp Q43433 VATB2_GOSHI. V-type proton ATPase subunit B2 (Fragment)	8	SCQSLR1089G05 (1E-24)	1	X	nd	nd	nd	nd	
	sp Q9SWE7 VATE_CITLI. V-type proton ATPase subunit E	10; 11	SCCCLB1025B04 (5E-90) / SCEPLB1042D04 (6E-77)	3	X	nd	nd	nd	X	
	sp Q9XGM1 VATD_ARATH. V-type proton ATPase subunit D	11	SCEPLB1044F03 (1E-103)	1	nd	nd	nd	nd	X	

Tabela 6 – Proteínas nucleares identificadas em folhas de 1 mês de idade

(continuação)

Função	Identificação	Banda	Cluster	vezes	B	S	W	P	PI	E.C.
metabolismo de carboidratos	sp B0UV30 TAL_HAES2. Transaldolase	10	SCACLR2022D05 (4E-41)	1	X	nd	nd	nd	nd	<u>EC=2.2.1.2</u>
	sp Q6H795 CLPD1_ORYSJ. Chaperone protein ClpD1, chloroplastic	1; 5	SCQGRT1041D10 (9E-73)	2	X	nd	nd	nd	nd	
	tr C5YF84 C5YF84_SORBI Putative uncharacterized protein Sb06g014590	1; 5; 6	SCQGLR1062F08 (0)	3	nd	nd	nd	X	nd	
metabolismo do ATP	sp P05494 ATPAM_MAIZE. ATP synthase subunit alpha, mitochondrial	8	SCCCLB1004B02 (7E-91)	1	X	nd	nd	nd	nd	
metabolismo do DNA	sp Q9M7J4 MFP1_TOBAC. MAR-binding filament-like protein 1-1	5	SCCCFL5093B04 (4E-35)	1	X	X	nd	nd	X	
metabolismo do RNA	sp P19683 ROC4_NICSY. 31 kDa ribonucleoprotein, chloroplastic	11	SCAGLR1043B12 (3E-75)	1	nd	X	nd	nd	nd	
	sp Q9ZUU4 ROC1_ARATH. Ribonucleoprotein At2g37220, chloroplastic	13	SCCCST1008B08 (2E-67)	1	nd	X	nd	nd	nd	
montagem de nucleossomos	sp P0CG89 H4_SOYBN. Histone H4	1; 10; 15; 16; 18; 19	SCAGLB1069E02 (1E-39) / SCAGLB1070C07 (1E-39) / SCCCLR1078H02 (2E-39) / SCCCLR2001D01 (2E-35) / SCCCLR2001D12 (4E-38) / SCCCLR2002D04 (2E-39) / SCCCLR2002G09 (2E-39) / SCCCLR2003E01 (1E-39) / SCCCZR1004C05 (2E-39) / SCCCZR2001G01 (1E-36) / SCEPLR1051H12 (3E-39) / SCJFLR1013A12 (9E-37) / SCJFRZ2009F02 (2E-39) / SCJLHR1028C12 (1E-29) / SCJLLR1011E11 (9E-37) / SCJLLR2020B06 (1E-39) / SCRFLR2038C05 (4E-42) / SCSBFL1043F08 (2E-39) / SCSBLB1033H09 (1E-39) / SCSGHR1071E06 (2E-18) / SCSGRT2062A10 (1E-38) / SCUTLR2015B06 (3E-37) / SCVPRZ2035C06 (2E-35)	134	X	X	X	X	X	
	sp P19177 H2A_PETCR. Histone H2A	10; 15; 16; 17; 18	SCEZAM1082G06 (4E-38) / SCCCLR1001C03 (3E-37) / SCUTLR2008G03 (1E-38)	11	X	X	X	X	X	
	sp P27807 H2B1_WHEAT. Histone H2B.1	10; 13; 14; 15; 16; 17	SCCCLR2C02B06 (1E-46)	6	X	X	X	X	X	

Tabela 6 – Proteínas nucleares identificadas em folhas de 1 mês de idade

(continuação)

Função	Identificação	Banda	Cluster	vezes	B	S	W	P	PI	E.C.
	sp P54348 H2B5_MAIZE. Histone H2B.5	9; 10; 13; 14; 15; 16; 17	SCBGFL3095E08 (1E-36) / SCCCLR1024A09 (1E-36) / SCCCLR1048C02 (2E-36) / SCCCLR1C04B01 (1E-36) / SCCCLR2002G11 (5E-37) / SCCCZR1001H04 (1E-36) / SCEPLB1042A02 (7E-37) / SCEQLR1050A02 (2E-38) / SCJLLR2020F01 (7E-37) / SCUTLR2015D03 (8E-37) / SCCCLR1C04B07 (8E-59)	61	X	X	X	X	X	
	sp P59226 H32_ARATH Histone H3.2	17	SCAGFL8013E01 (3E-54) / SCAGLR2026E02 (7E-59) / SCBFLR1026A06 (6E-59) / SCBGLR1002H11 (9E-59) / SCCCLR1001F12 (1E-58) / SCCCLR1068F12 (2E-58) / SCCCLR2001D10 (2E-54) / SCCCZR2001G08 (4E-54) / SCEQLB1063F02 (1E-58) / SCEQLB1066C12 (2E-56) / SCJFRZ2009C12 (3E-54) / SCMCLR1032D12 (2E-54) / SCQGLR1019F06 (8E-57) / SCQGLR1062C03 (4E-54) / SCVPLB1016B08 (8E-59) / SCVPLB1017H05 (5E-57)	1	nd	X	X	X	X	
	sp P68428 H32_WHEAT. Histone H3.2	17	SCAGLR2026F05 (7E-27)	16	nd	X	X	X	X	
	sp Q28DR4 H4_XENTR. Histone H4	18; 19	SCCCLR1001G05 (5E-46) / SCCCLR1068C05 (7E-43) / SCEPAM1020H08 (3E-31)	2	X	X	X	X	X	
	sp Q2QPG9 H2AXB_ORYSJ. Probable histone H2AXB	10; 15; 16	SCMCLR1032A12 (7E-16)	7	X	X	X	nd	X	
	sp Q43216 H2B5_WHEAT. Histone H2B.5	10; 13; 14; 15; 16; 17	SCCCLR2002H03 (2E-39)	6	X	X	X	X	X	
	sp Q6L500 H2A4_ORYSJ. Probable histone H2A.4	10; 15; 16; 17; 18	SCBGLR1002E03 (1E-43) / SCCCLR1068G11 (5E-44) / SCCCZR1004F04 (5E-42) / SCEZLB1014E02 (4E-43) / SCJLLR1101H08 (4E-44) / SCMCR2108A05 (4E-43) / SCQGLR1041C10 (9E-44) / SCQGLR1062G09 (1E-43) / SCSFFL3090D03 (7E-44) / SCSGAM1096G11 (8E-44) / SCSGFL5C05H11 (2E-33)	5	X	X	X	X	X	
	sp Q6ZBP3 H2B2_ORYSJ. Histone H2B.2	10; 13; 14; 15; 16; 17	SCJFRZ2010D04 (9E-35) / SCJFRZ2014A08 (8E-35) / SCVPLR2019A02 (1E-34)	66	X	X	X	X	X	
	sp Q6ZL42 H2A2_ORYSJ. Probable histone H2A.2	15; 16		6	X	X	X	nd	X	

Tabela 6 – Proteínas nucleares identificadas em folhas de 1 mês de idade

(continuação)

Função	Identificação	Banda	Cluster	vezes	B	S	W	P	PI	E.C.
	sp Q71H73 H33_VITVI. Histone H3.3	16	SCBGFL3095F09 (1E-70) / SCJFRZ2006G04 (1E-70) / SCBGLR1002A05 (2E-70) / SCBGLR1023F03 (2E-70) / SCCCCL3120E10 (2E-70) / SCCCCL7C02A12 (2E-65) / SCCCFL5003D06 (1E-70) / SCCCLR2001F08 (2E-70) / SCCCZR2002A05 (2E-70) / SCEQLR1094C02 (2E-70) / SCJFLR1013C11 (9E-71) / SCJFLR1035H03 (1E-70) / SCJFRT1060F02 (2E-70) / SCRFLR1055C06 (2E-70) / SCSGFL4037A02 (6E-57) / SCUTLR2008H04 (1E-70) / SCVPLR2027D12 (8E-40) / SCVPRZ2039D03 (2E-70)	18	nd	X	X	X	X	
	sp Q75L11 H2A6_ORYSJ. Probable histone H2A.6	10; 15; 16; 17; 18	SCAGLR1021A12 (5E-35) / SCCCFL3003C11 (3E-38) / SCCCLB1024F05 (7E-37) / SCCCLR1048F07 (2E-41) / SCCCLR2001C09 (4E-39) / SCCCLR2001G07 (3E-35) / SCCCLR2C02C12 (7E-40) / SCCCZR2C03C04 (2E-41)	21	X	X	X	X	X	
	sp Q84MP7 H2AV3_ORYSJ. Probable histone H2A variant 3	10; 14; 15; 16; 18	SCEPLR1008D05 (1E-57) / SCJLLR1101A02 (4E-47)	10	nd	X	X	nd	X	
	sp Q8H7Y8 H2AV1_ORYSJ. Probable histone H2A variant 1	14; 15; 16; 18	SCCCCL4007C09 (8E-54)	4	X	X	X	nd	X	
	sp Q943L2 H2B11_ORYSJ Histone H2B.11	10; 13; 14; 15; 16; 17	SCVPLR1049C05 (1E-37)	6	X	X	X	X	X	
	tr B6UIQ0 B6UIQ0_MAIZE Histone H2A	15; 16	SCRLFL1005C02 (2.0×10 ⁻⁵⁹)	2	X	X	X	nd	X	
	sp A2ZK26 H2A8_ORYSI. Probable histone H2A.8	15; 15; 16	SCJLLR1101F06 (2E-35) / SCJLLR2013B11 (2E-35)	4	nd	X	nd	nd	X	
morfogênese foliar	sp O24412 PSD7_ARATH. Probable 26S proteasome non-ATPase regulatory subunit 7	10	SCJFRZ2027F12 (1E-160)	1	X	nd	nd	nd	nd	
processos biossintéticos	sp O04974 LEU1B_SOLPN. 2-isopropylmalate synthase B	7	SCQGLB1039D07 (2E-61)	1	nd	nd	nd	X	nd	<u>EC=2.3.3.13</u>
	sp P34106 ALA2_PANMI. Alanine aminotransferase 2	8; 9	SCCCRZ1002E01 (1E-137)	2	X	nd	nd	nd	nd	<u>EC=2.6.1.2</u>
	sp P72580 PREA_SYNY3.Prenyltransferase	10	SCCCAM1002F08 (1E-61)	1	X	nd	nd	nd	nd	EC=2.5.1.-

Tabela 6 – Proteínas nucleares identificadas em folhas de 1 mês de idade

(continuação)

Função	Identificação	Banda	Cluster	vezes	B	S	W	P	PI	E.C.
processos celulares	sp P04907 GSTF3_MAIZE. Glutathione S-transferase 3	13	SCCCLR1001A07 (2E-92) / SCEZRT3070A09 (7E-71) / SCJFLR1073H08 (1E-91)	3	X	nd	nd	nd	nd	<u>EC=2.5.1.18</u>
	sp P24203 YJIA_ECOLI. Uncharacterized GTP-binding protein YjiA	01; 08	SCRULB1060G07 (7E-08)	2	X	nd	nd	nd	nd	
	sp Q9FM19 HIR1_ARATH. Hypersensitive-induced response protein 1	10; 11	SCCCRZ1002F04 (1E-137) / SCJFRT1008C05 (1E-127) / SCRUD1061C12 (3E-49)	4	X	nd	nd	nd	nd	
	sp Q9M158 Y4105_ARATH. Uncharacterized protein At4 g01050, chloroplastic	7; 8; 9	SCEZHR1047F05 (2E-68)	3	X	nd	nd	nd		
processos de oxidação-redução	sp O04130 SERA_ARATH. D-3-phosphoglycerate dehydrogenase, chloroplastic	7	SCCCLR2002G02 (0) / SCJFRT1005D01 (6 e-06)	2	nd	nd	nd	X	nd	<u>EC=1.1.1.95</u>
	sp P18123 CATA3_MAIZE. Catalase isozyme 3	07; 08	SCCCLR1048H09 (0) / SCCCAD1004H02 (3E-91)	3	X	nd	nd	nd	nd	<u>EC=1.11.1.6</u>
	sp Q69TY4 PR2E1_ORYSJ. Peroxiredoxin-2E-1, chloroplastic	15; 16	SCEZRZ1014F10 (7E-84)	2	nd	nd	nd	nd	X	<u>EC=1.11.1.15</u>
	sp Q7FB12 NDHM_ORYSJ. NAD(P)H-quinone oxidoreductase subunit M, chloroplastic	12; 14; 15	SCEPSD2005D11 (1E-72)	3	nd	nd	nd	nd	X	<u>EC=1.6.5.-</u>
proteólise	sp Q7NFI1 NDHO_GLOVI NAD(P)H-quinone oxidoreductase subunit O	16	SCAGAM2126G09 (1E-05)	1	X	nd	nd	nd	nd	<u>EC=1.6.5.-</u>
	sp Q93VK5 LUT5_ARATH. Protein LUTEIN DEFICIENT 5, chloroplastic	7	SCJFLR1073C07 (6E-73)	1	X	nd	nd	X	nd	<u>EC=1.14.-.-</u>
	sp O64464 PSB1_ORYSJ. Proteasome subunit beta type-1	13	SCACLR2007D12 (1E-105) / SCBGLR1002D03 (1E-122) / SCRUF1115H03.b (4E-39)	3	X	nd	nd	nd	X	<u>EC=3.4.25.1</u>
	sp P0CH04 UBI1P_PETCR Polyubiquitin	10	SCCCLB2004C08 (6E-89)	1	X	nd	nd	nd	X	
	sp P0CH05 UBI2P_PETCR. Polyubiquitin	10	SCCCST1005F12 (1E-112) / SCJFRT2055G07 (1E-111)	2	X	nd	nd	nd	X	

Tabela 6 – Proteínas nucleares identificadas em folhas de 1 mês de idade

(continuação)

Função	Identificação	Banda	Cluster	vezes	B	S	W	P	PI	E.C.
	sp P0CH32 UBQ4_ARATH. Polyubiquitin 4	10; 11; 13	SCBFLR1026D07 (0) / SCBFRZ3006E06 (1E-120) / SCCCCL3080A11 (1E-117) / SCCCCL3080A11.b (1E-117) / SCCCCL4014E10 (0) / SCCCLB1026D12 (1E-94) / SCCCLR1001C11 (0) / SCCCLR1C03C10 (1E-161) / SCCCLR1C03H09 (1E-132) / SCEZAM2059H12 (1E-114) / SCSGFL4118F01 (1E-125) / SCVPRZ3026B09 (1E-117) / SCSGAM1095C10 (0)	25	X	nd	nd	nd	X	
	sp P46466 PRS4_ORYSJ. 26S protease regulatory subunit 4 homolog	8	SCEPLB1041G09 (1E-124)	1	nd	nd	nd	X	X	
	sp P52428 PSA1_ORYSJ. Proteasome subunit alpha type-1	10		1	X	nd	nd	nd	X	<u>EC=3.4.25.1</u>
	sp P69309 UBIQP_AVEFA. Polyubiquitin	10; 11; 13	SCCCST2001G02 (1E-114) / SCJFRT1010F04 (2E-76) / SCQSLB1052H09 (1E-70) / SCQSRT2032E09 (1E-124) / SCRLLV1026A01.b (5E-48) / SCRUF1021F10.b (1E-44) / SCSFSD1065H04.b (2E-80) / SCVPRZ2038C12 (1E-168)	12	X	nd	nd	nd	X	
	sp P73689 SPPA_SYNY3. Protease 4	7	SCUTLR1037A04 (1E-13)	1	X	nd	nd	nd	nd	<u>EC=3.4.21.-</u>
	sp P85200 PRS6B_HELAN. 26S protease regulatory subunit 6B homolog	8	SCCCLR1C07H11 (0)	1	X	nd	nd	nd	X	
	sp Q10KF0 PSA2_ORYSJ. Proteasome subunit alpha type-2	13	SCQSLR1018D04 (1E-126)	1	X	nd	nd	nd	X	<u>EC=3.4.25.1</u>
	sp Q6YT00 PSA7A_ORYSJ. Proteasome subunit alpha type-7-A	11	SCEZLB1009E05 (1E-123) / SCJFRZ2010D05 (1E-121) / SCVPRZ2038H03 (3E-18)	3	X	nd	nd	nd	X	<u>EC=3.4.25.1</u>
	sp Q7DLR9 PSB4_ARATH. Proteasome subunit beta type-4	13	SCCCCL3120F07 (2E-94)	1	X	nd	X	nd	X	<u>EC=3.4.25.1</u>
	sp Q8H159 UBQ10_ARATH. Polyubiquitin 10	10; 11; 13	SCACLR1057F07 (0) / SCCCZR1002C04 (0)	6	X	nd	nd	nd	X	
	sp Q8LD27 PSB6_ARATH. Proteasome subunit beta type-6	14	SCCCRZ2C04B02 (1E-104) / SCQSAM1032E01 (5E-73)	2	nd	nd	nd	nd	X	<u>EC=3.4.25.1</u>
	sp Q8LQJ8 FTSH5_ORYSJ. ATP-dependent zinc metalloprotease FTSH 5, mitochondrial	7	SCCCLR1078F01 (0)	1	X	nd	nd	nd	X	<u>EC=3.4.24.-</u>

Tabela 6 – Proteínas nucleares identificadas em folhas de 1 mês de idade

(continuação)

Função	Identificação	Banda	Cluster	vezes	B	S	W	P	PI	E.C.
regulação da transcrição DNA-dependente regulação da transcrição RNA-dependente resposta ao estresse	sp Q9C5U3 PRS8A_ARATH. 26S protease regulatory subunit 8 homolog A	9	SCCCCL4006C07 (0)	1	nd	nd	nd	nd	X	
	sp Q9LST6 PSB2_ORYSJ. Proteasome subunit beta type-2	13; 14	SCCCLR1048G12 (1E-101)	2	nd	nd	nd	nd	X	<u>EC=3.4.25.1</u>
	sp Q9LST7 PSB3_ORYSJ. Proteasome subunit beta type-3	13	SCCCLR1C07G01 (1E-108)	1	X	nd	nd	nd	X	<u>EC=3.4.25.1</u>
	sp Q9LSU3 PSA6_ORYSJ. Proteasome subunit alpha type-6	13	SCCCLR1001B04 (1E-133)	1	X	nd	nd	nd	X	<u>EC=3.4.25.1</u>
	sp Q9MAK9 PS10B_ARATH. 26S protease regulatory subunit S10B homolog B	9	SCCCLR1022E06 (0)	1	nd	nd	nd	nd	X	
	sp Q9SEI4 PRS6B_ARATH. 26S protease regulatory subunit 6B homolog	8	SCCCLR1077D09 (0)	1	X	nd	nd	nd	X	
	tr C5XWW6 C5XWW6_SORBI Proteasome subunit beta type	13	SCCCLR1001C09 (1.0×10-145)	1	X	nd	nd	nd	X	
	tr C5YWZ8 C5YWZ8_SORBI ATP-dependent Clp protease proteolytic subunit	13	SCQSAM2100C03 (1.0×10-129)	1	nd	X	nd	nd	nd	
	tr G1L2G6 G1L2G6_AILME Uncharacterized protein (Fragment)	8	SCMCAM2081F09 (7.0×10-1)	1	X	X	X	X	X	
	sp Q8S0J7 IM30_ORYSJ. Probable membrane-associated 30 kDa protein, chloroplastic	10; 11; 12; 13	SCUTLR2015H12 (1E-151)	4	nd	nd	nd	X	nd	
	sp P09189 HSP7C_PETHY. Heat shock cognate 70 kDa protein	1	SCAGLR1043F02 (0) / SCCCCL3120G07 (0) / SCCCLR1079D09 (6E-90) / SCCCLR1C01D05 (0)	4	nd	nd	nd	X	nd	
	sp P27322 HSP72_SOLL. Heat shock cognate 70 kDa protein 2	1	SCCCRZ1002G01 (0) / SCSFST1064C10 (1E-103)	2	nd	nd	nd	X	nd	
	sp Q07078 HSP83_ORYSJ. Heat shock protein 81-3	1; 5	SCCCLR2C02F10 (1E-115) / SCRLLR1110D08 (1E-160)	3	X	nd	X	X	nd	
	sp Q08277 HSP82_MAIZE. Heat shock protein 82	1	SCJFRT2055B05 (6E-99)	1	X	nd	nd	X	nd	

Tabela 6 – Proteínas nucleares identificadas em folhas de 1 mês de idade

(continuação)

Função	Identificação	Banda	Cluster	vezes	B	S	W	P	PI	E.C.
sinalização	sp Q0J4P2 HSP81_ORYSJ. Heat shock protein 81-1 OS= Oryza sativa subsp. japonica GN=HSP81-1 PE=2 SV=2	1; 5	SCCCLR1075E09 (0) / SCEZLB1006B07 (1E-100) / SCJLLR1011E03 (0) / SCMCCL6049H03 (1E-101)	6	X	nd	nd	X	nd	
	sp Q5QWR2 HTPG_IDILO. Chaperone protein htpG	5	SCCCCL3005G07.b (1E-104)	1	X	nd	nd	nd	nd	
	tr C5X3T9 C5X3T9_SORBI Putative uncharacterized protein Sb02g028050	1; 5	SCCCLR1075H11 (7.0×10 ⁻¹)	2	X	nd	X	X	nd	
	sp P49027 GBLPA_ORYSJ. Guanine nucleotide-binding protein subunit beta-like protein A	10	SCACLR2007C03 (1E-171) / SCSFAM1076C12 (5E-80)	2	X	nd	nd	nd	nd	
	sp P49106 14331_MAIZE. 14-3-3-like protein GF14-6	7; 10; 11	SCCCLR1022D05 (1E-142) / SCRUF1017E10.b (1E-71)	5	X	nd	nd	nd	X	
	sp Q2R2W2 14334_ORYSJ. 14-3-3-like protein GF14-D	10; 11	SCCCLR1048F12 (1E-124) / SCEQRT1025D06 (3E-79) / SCMCRT2105C11 (7E-79)	6	X	nd	nd	nd	X	
	sp Q6ZKC0 14333_ORYSJ. 14-3-3-like protein GF14-C	7; 10; 11; 12	SCBFLR1026E02 (1E-131) / SCCCRZ1001D02 (1E-134) / SCEQRT1031D02 (1E-113) / SCCCCL5072E02 (1E-40)	12	X	nd	nd	nd	X	
	sp Q84J55 14331_ORYSJ. 14-3-3-like protein GF14-A	10; 11	SCEQLR1007H12 (1E-33) / SCEQRT2094B01 (6E-73) / SCMCRT2102A01 (3E-74)	6	X	nd	nd	nd	X	
	sp Q9SP07 1433_LILLO. 14-3-3-like protein	7; 10; 11	SCAGLB1069G08 (2E-76)	3	X	nd	nd	nd	X	
	tr C5Y3X7 C5Y3X7_SORBI Putative uncharacterized protein Sb05g021020	10	SCQGST1034A08 (7.0×10 ⁻¹)	1	X	nd	X	nd	X	
tradução	tr C5YJQ8 C5YJQ8_SORBI Putative uncharacterized protein Sb07g029110	10; 11	SCCCCL4011G09 (7.0×10 ⁻¹)	2	X	nd	nd	nd	X	
	sp A1E9W6 RK2_SORBI 50S ribosomal protein L2, chloroplastic	19	SCACSB1035E05 (4E-59)	1	nd	nd	nd	X	nd	
	sp A1E9W9 RR7_SORBI. 30S ribosomal protein S7, chloroplastic	14; 15; 16	SCCCLR2C01E06 (8E-81)	3	nd	nd	nd	X	nd	
	sp A3BLC3 RRFC_ORYSJ. Ribosome-recycling factor, chloroplastic	14	SCACLB1047F10 (1E-111)	1	nd	nd	nd	x	nd	
	sp A5BUU4 RSSA_VITVI. 40S ribosomal protein SA	10	SCCCLR1001D03 (1E-110) / SCCCST1001H07 (1E-118) / SCEPLR1051E04 (1E-110) / SCSGAM1094H05 (1E-90) / SCSGLR1045A06 (1E-116)	5	X	nd	nd	nd	nd	

Tabela 6 – Proteínas nucleares identificadas em folhas de 1 mês de idade

(continuação)

Função	Identificação	Banda	Cluster	vezes	B	S	W	P	PI	E.C.
	sp O48557 RL17_MAIZE. 60S ribosomal protein L17	15; 16	SCBGLR1120F10 (2E-94) / SCEZRZ3095E08 (2E-26) / SCJFLR1017D02 (8E-82) / SCJFLR1017H09 (2E-94) / SCJFRZ2005H06 (6E-88)	10	nd	nd	X	nd	nd	
	sp O64937 EF1A_ORYSJ. Elongation factor 1-alpha	1; 2; 7; 8; 9; 10	SCCCLR1022B09 (0) / SCEZRZ1013H01 (0) / SCJLRT1006G01 (0) / SCVPLB1020E07 (0) / SCACLR1127E08 (3E-94) / SCRFLR2038B04 (9E-72) / SCUTSD2088D06 (8E-60)	24	nd	nd	nd	X	X	
	sp O65731 RS5_CICAR. 40S ribosomal protein S5 (Fragment)	14; 15; 16	SCACLR1127E08 (3E-94) / SCRFLR2038B04 (9E-72) / SCUTSD2088D06 (8E-60)	9	X	nd	nd	nd	nd	
	sp O80360 RK3_TOBAC. 50S ribosomal protein L3, chloroplastic (Fragment)	11; 12; 13; 14	SCEQLB1066D08 (2E-97)	4	nd	nd	X	X	nd	
	sp O82204 RL281_ARATH. 60S ribosomal protein L28-1	16	SCCCLR2002C09 (2E-52)	1	nd	nd	nd	X	nd	
	sp P19950 RS141_MAIZE. 40S ribosomal protein S14	16; 17	SCCCLR1048A02 (5E-62) / SCCCLR1C03A02 (1E-61)	4	X	nd	nd	nd	nd	
	sp P19951 RS142_MAIZE. 40S ribosomal protein S14	17	SCCCLR2002B03 (6E-63) / SCRUF13066B06.b (9E-58)	2	X	nd	nd	nd	nd	
	sp P29766 RL8_SOLLC. 60S ribosomal protein L8	11	SCCCLR2004E06 (1E-133) / SCVPLR2027H04 (1E-134)	2	X	nd	nd	nd	nd	
	sp P34788 RS18_ARATH. 40S ribosomal protein S18	16; 17	SCCCLR1001H01 (4E-67) / SCCCRZ2002C12 (1E-66)	3	nd	nd	nd	X	X	
	sp P34823 EF1A2_DAUCA. Elongation factor 1-alpha	1; 2; 7; 8; 9; 10	SCSBRT3039F07 (8E-45)	6	nd	nd	nd	nd	X	
	sp P41378 IF4A_WHEAT. Eukaryotic initiation factor 4A	8	SCRFRT3057C02 (2E-74)	1	X	nd	nd	nd	X	<u>EC=3.6.4.13</u>
	sp P42791 RL182_ARATH. 60S ribosomal protein L18-2	14; 15	SCJFRZ2015G01 (1E-61)	2	X	nd	X	X	nd	
	sp P45633 RL10_MAIZE. 60S ribosomal protein L10	13	SCBFLR1026E01 (1E-123) / SCBFRZ2016G12 (6E-98) / SCBGLR1119H10 (2E-79)	3	nd	nd	nd	X	X	
	sp P49199 RS8_ORYSJ. 40S ribosomal protein S8	11; 12	SCCCLR1048F11 (5E-94) / SCCCRZ2002D10 (5E-78) / SCQSST3114E08 (3E-41)	6	nd	nd	nd	X	nd	
	sp P49210 RL9_ORYSJ. 60S ribosomal protein L9	14; 15	SCCCLR1069A12 (7E-90) / SCBFRZ2045D04 (7E-94) / SCJFFL3C02H03.b (6E-39)	4	X	nd	nd	nd	nd	

Tabela 6 – Proteínas nucleares identificadas em folhas de 1 mês de idade

(continuação)

Função	Identificação	Banda	Cluster	vezes	B	S	W	P	PI	E.C.
	sp P49397 RS3A_ORYSJ. 40S ribosomal protein S3a	11	SCBGLR1047A12 (1E-128) / SCCCCL3005C01.b (1E-123) / SCCCCL4006A03 (1E-119) / SCCCLR2001F09 (1E-123) / SCMCLR1032F10 (3E-78) / SCRLAM1014C12 (5E-87) / SCSFST3074H02 (1E-116) / SCVPRZ2037B05 (3E-19)	8	X	nd	nd	X	X	
	sp P50883 RL121_ARATH. 60S ribosomal protein L12-1	16	SCAGLR1021E04 (2E-80) / SCCCLR1C01C09 (4E-77)	2	nd	nd	nd	nd	X	
	sp Q03033 EF1A_WHEAT. Elongation factor 1-alpha	8	SCCCLB1021F05 (3E-89) / SCJFAD1012C08 (8E-90)	2	nd	nd	nd	X	X	
	sp Q05761 RS13_MAIZE. 40S ribosomal protein S13	16	SCCCLR1024D10 (5E-79) / SCUTLR1058E09 (8E-80)	2	nd	nd	X	X	nd	
	sp Q07761 RL23A_TOBAC. 60S ribosomal protein L23a	16	SCBGFL5081F09 (1E-37) / SCCCLR1C03A11 (2E-41) / SCEQLR1091A10 (2E-41)	3	nd	nd	X	X	nd	
	sp Q08699 RS14_PODCA. 40S ribosomal protein S14	16; 17	SCSBSD2054A09 (6E-38)	2	X	nd	nd	nd	nd	
	sp Q0DK10 RL11_ORYSJ. 60S ribosomal protein L11	15; 16	SCCCRZ2001F06 (7E-88) / SCVPLR2019F01 (5E-88) / SCSGLR1045F12 (3E-87)	5	nd	nd	nd	X	nd	
	sp Q0IQF7 RS16_ORYSJ. 40S ribosomal protein S16	16; 17	SCEPLR1030C12 (3E-74)	2	nd	nd	nd	X	nd	
	sp Q0ITS8 RL101_ORYSJ. 60S ribosomal protein L10-1	13	SCEPRZ1010E07 (1E-120) / SCEZRZ1017F07 (1E-78)	2	nd	nd	nd	X	X	
	sp Q0JGY1 RL51_ORYSJ. 60S ribosomal protein L5-1	10	SCCCLR1048A09 (1E-130) / SCEQLR1091D05 (1E-141) / SCUTLR1037F12 (1E-127)	3	nd	nd	nd	X	X	
	sp Q38884 EIF3I_ARATH. Eukaryotic tradução init iation factor 3 subunit I	10	SCBFFL5074B10 (4E-53)	1	X	nd	nd	nd	X	
	sp Q40680 EF1D1_ORYSJ. Elongation factor 1-delta 1	10	SCCCRZ2002D07 (2E-76)	1	X	nd	nd	nd	nd	
	sp Q41741 IF4A_MAIZE. Eukaryotic initiation factor 4A	8; 9	SCCCLR1C07G11 (0)	2	X	X	nd	nd	X	<u>EC=3.6.4.13</u>
	sp Q41803 EF1A_MAIZE. Elongation factor 1-alpha	1; 2; 7; 8; 9; 10	SCCCCL3004F01.b (0) / SCEQLR1007G03 (0)	12	nd	nd	nd	X	X	
	sp Q5N1T5 EFP_SYNP6. Elongation factor P	14; 15	SCQSAM2099H05 (5E-44)	2	X	nd	nd	nd	nd	
	sp Q6Z2Z4 IF4A3_ORYSJ. Eukaryotic initiation facto r 4A-3	8; 9	SCCCLR1072G07 (0)	3	X	X	nd	nd	X	<u>EC=3.6.4.13</u>

Tabela 6 – Proteínas nucleares identificadas em folhas de 1 mês de idade

(continuação)

Função	Identificação	Banda	Cluster	vezes	B	S	W	P	PI	E.C.
	sp Q8LJU5 RS7_ORYSJ. 40S ribosomal protein S7	11; 12; 13; 14; 15	SCVPRZ2038F11 (5E-95)	5	X	nd	nd	X	nd	
	sp Q8VZB9 R10A1_ARATH. 60S ribosomal protein L10a- 1	11; 12; 13	SCCCLR1C05G05 (1E-93)	3	X	nd	nd	nd	nd	
	sp Q940B0 RL183_ARATH. 60S ribosomal protein L18-3	14; 15	SCBFLR1046F11 (4E-85)	2	nd	nd	X	X	nd	
	sp Q96564 RS27_HORVU. 40S ribosomal protein S27	19	SCCCCL4008F04 (4E-38)	1	nd	nd	nd	X	nd	
	sp Q9AYP4 RS10_ORYSJ. 40S ribosomal protein S10	14; 15; 16	SCCCLR2C01F11 (3E-50) / SCACLR1126F04 (5E-60) / SCMCST1051E11 (2E-42)	5	X	nd	X	nd	X	
	sp Q9C285 RL12_NEUCR. 60S ribosomal protein L12	16	SCEQRT1024D04 (7E-77)	1	nd	nd	nd	X	X	
	sp Q9C9C5 RL63_ARATH. 60S ribosomal protein L6-3	12; 13	SCCCCL3080C04 (4E-68) / SCEQLB1063E01 (1E-74)	4	nd	nd	X	X	nd	
	sp Q9FDZ9 RL212_ARATH. 60S ribosomal protein L21-2	15	SCCCCL3003D09.b (3E-83) / SCCCLR2C01F01 (3E-83) / SCJFLR1017B02 (3E-83)	3	nd	nd	nd	X	nd	
	sp Q9FJA6 RS33_ARATH. 40S ribosomal protein S3-3	13	SCAGLR1043G11 (5E-61) / SCCCCL3001F09 (1E-109) / SCCCRZ2003E04 (1E-106) / SCJFFL1C05F09.b (2E-17)	4	nd	nd	nd	X	X	
	sp Q9FKC0 R13A4_ARATH. 60S ribosomal protein L13a- 4	13; 14	SCCCCL4004E02 (4E-90) / SCRFLR1012C07 (4E-90)	4	nd	nd	nd	X	nd	
	sp Q9FLF0 RS92_ARATH. 40S ribosomal protein S9-2	14; 15	SCBFLR1083F02 (1E-91) / SCCCLR2002C01 (5E-81) / SCQGLR2017B10 (6E-91)	6	X	nd	X	X	nd	
	sp Q9FWS4 RK31_ARATH. 50S ribosomal protein L31, chloroplastic	19	SCEPLR1051F04 (2E-23)	1	nd	nd	nd	X	nd	
	sp Q9LHP1 RL74_ARATH. 60S ribosomal protein L7-4	12; 13	SCCCLR1C10E12 (5E-95) / SCCCLR2002E05 (7E-87) / SCCCLR2004C07 (7E-87) / SCCCLR1C05H09 (2E-78)	7	X	nd	nd	X	nd	
	sp Q9LZ17 RS174_ARATH. 40S ribosomal protein S17-4	15; 16; 17	SCBGLR1002D11 (8E-54) / SCCCLR1C05F01 (6E-54) / SCRLFL1011G08 (2E-48)	9	X	nd	X	X	X	
	sp Q9M339 RS32_ARATH. 40S ribosomal protein S3-2	13	SCVPRT2080E04 (1E-89)	1	nd	nd	nd	X	X	
	sp Q9SCM3 RS24_ARATH. 40S ribosomal protein S2-4	10; 11; 12	SCCCCL4015D08 (2E-99) / SCCCRZ2002H09 (1E-100) / SCEPRZ1009G03 (1E-100)	9	X	nd	nd	nd	X	
	sp Q9SF40 RL4A_ARATH. 60S ribosomal protein L4-1	7; 8; 9; 10	SCCCCL3140B12 (1E-151) / SCCCRZ1001C08 (1E-112) / SCJFRT1011A02 (2E-35)	9	X	nd	nd	X	nd	

Tabela 6 – Proteínas nucleares identificadas em folhas de 1 mês de idade

(conclusão)

Função	Identificação	Banda	Cluster	vezes	B	S	W	P	PI	E.C.
	sp Q9SS17 RS241_ARATH. 40S ribosomal protein S24-1	16	SCJLLR1054C12 (4E-55) / SCMCRT2088H11 (2E-53)	2	nd	nd	nd	X	nd	
	sp Q9XET4 RS7_SECCE. 40S ribosomal protein S7	11; 12; 13; 14; 15	SCCCCL4010B12 (8E-98) / SCVPLR2005E09 (2E-96)	10	X	nd	nd	X	nd	
	sp Q9ZUT9 RS51_ARATH. 40S ribosomal protein S5-1	14; 15; 16	SCCCCL3001E09.b (7E-93)	3	X	nd	nd	nd	nd	
	tr C5Y0E9 C5Y0E9_SORBI Putative uncharacterized protein Sb04g030890	16	SCACLR2014B04 (2E-84)	1	X	nd	nd	nd	X	
	tr C5Y3N7 C5Y3N7_SORBI Putative uncharacterized protein Sb05g001680	16; 17	SCCCLR2002A05 (1.0×10 ⁻¹⁰¹)	2	nd	nd	nd	X	nd	

ANEXO E

Tabela 7 – Proteínas nucleares identificadas em folhas de cana-de-açúcar de 4 meses de idade

(continua)

Função	Identificação	Banda	Cluster	Veze	B	S	W	P	PI	E.C.
biogênese de ribossomos ciclo celular	sp P35685 RL7A_ORYSJ. 60S ribosomal protein L7a	11	SCCCRZ2001H04 (1E-112) / SCVPLR1028C12 (1E-101)	2	X	nd	nd	X	nd	
	sp P54774 CDC48_SOYBN. Cell division cycle protein 48 homolog	4	SCVPLR1049C09 (0)	1	X	nd	nd	nd	X	
direcionamento de proteínas	sp P37107 SR54C_ARATH. Signal recognition particle 54 kDa protein, chloroplastic	8	SCQGRT1042A05 (0)	1	nd	nd	nd	X	nd	
fixação do carbono	sp P15804 CAPP3_SORBI. Phosphoenolpyruvate carboxylase 3	1; 2; 3; 4; 5; 6; 7; 8; 9; 10; 11; 12; 13; 14; 15; 16	SCRULR1020D09 (0)	16	nd	nd	nd	X	nd	<u>EC=4.1.1.31</u>
fotossíntese	sp O65107 PSAN_MAIZE. Photosystem I reaction center subunit N, chloroplastic (Fragment)	20	SCACLR1129A11 (5E-43)	1	X	nd	nd	nd	nd	
	sp P12329 CB21_MAIZE. Chlorophyll a-b binding protein 1, chloroplastic	1; 2; 5; 6; 7; 8; 9; 10; 11; 12; 13; 14; 15	SCJFLR1074A11 (1E-141)	13	X	nd	nd	nd	nd	
	sp P27524 CB4A_SOLLC. Chlorophyll a-b binding protein CP24 10A, chloroplastic	15	SCJFSB1014F03.b (4E-45)	1	X	nd	nd	nd	nd	
	sp P27525 CB4B_SOLLC. Chlorophyll a-b binding protein CP24 10B, chloroplastic	15	SCACLR2007B05 (1E-108)	1	X	nd	nd	nd	nd	
função desconhecida	sp Q8H6A5 TCTP_MAIZE. translationally-controlled tumor protein homolog	14	SCCCLR2001G09 (2E-80) / SCEZLR1031E10 (2E-80)	2	nd	nd	X	nd	nd	
	tr A2E1P4 A2E1P4_TRIVA Beige/BEACH domain containing protein	8	SCRFAM2128F04 (5.0×10 ⁻¹)	1	X	nd	nd	nd	nd	
	tr C5WRY9 C5WRY9_SORBI Putative uncharacterized protein Sb01g000680	13	SCJFRT1012A02 (3.0×10 ⁻⁷⁵)	1	X	nd	nd	nd	nd	
	tr C5YVN9 C5YVN9_SORBI Putative uncharacterized protein Sb09g010000	7; 8; 9	SCACSB1122B11 (1.0×10 ⁻¹⁰³)	3	X	nd	nd	X	X	

Tabela 7 – Proteínas nucleares identificadas em folhas de cana-de-açúcar de 4 meses de idade

(continuação)

Função	Identificação	Banda	Cluster	Veze	B	S	W	P	PI	E.C.
metabolismo de ATP	sp P05494 ATPAM_MAIZE. ATP synthase subunit alpha, mitochondrial	8	SCCCLB1004B02 (7E-91)	1	X	nd	nd	nd	nd	
	sp Q9SWE7 VATE_CITLI. V-type proton ATPase subunit E	11	SCCCLB1025B04 (5E-90)	1	X	nd	nd	nd	X	
metabolismo de proteínas	sp Q6H795 CLPD1_ORYSJ. Chaperone protein ClpD1, chloroplastic	4; 5	SCQGRT1041D10 (9E-73)	2	X	nd	nd	nd	nd	
	tr C5YF84 C5YF84_SORBI Putative uncharacterized protein Sb06g014590	4; 5	SCQGLR1062F08 (0)	2	nd	nd	nd	X	nd	
metabolismo de RNA	sp P19683 ROC4_NICSY. 31 kDa ribonucleoprotein, chloroplastic	11; 12	SCAGLR1043B12 (3E-75)	2	nd	X	nd	nd	nd	
montagem do nucleossomo	sp A2ZK26 H2A8_ORYSI. Probable histone H2A.8	16	SCJLLR1101F06 (2E-35) / SCJLLR2013B11 (2E-35)	2	nd	X	nd	nd	X	
	sp P08437 H3_VOLCA. Histone H3	17	SCQGFL1095C08 (2E-17)	1	nd	X	X	X	X	
	sp P0CG89 H4_SOYBN. Histone H4	10; 11; 18; 19	SCSGRT2062A10 (1E-38) / SCAGLB1069E02 (4E-39) / SCAGLB1070C07 (1E-39) / SCCCLR1078H02 (2E-39) / SCCCLR2001D01 (2E-35) / SCCCLR2001D12 (4E-38) / SCCCLR2002D04 (2E-39) / SCCCLR2002G09 (2E-39) / SCCCLR2003E01 (1E-39) / SCCCZR1004C05 (2E-39) / SCCCZR2001G01 (1E-36) / SCEPLR1051H12 (3E-39) / SCJFLR1013A12 (9E-37) / SCJFRZ2009F02 (2E-39) / SCJLLR1011E11 (9E-37) / SCJLLR2020B06 (1E-39) / SCRFLR2038C05 (4E-42) / SCSBFL1043F08 (2E-39) / SCSBLB1033H09 (1E-39) / SCSGHR1071E06 (2E-18) / SCUTLR2015B06 (3E-37) / SCVPRZ2035C06 (2E-35) / SCJLHR1028C12 (1E-29)	69	X	X	X	X	X	
	sp P19177 H2A_PETCR. Histone H2A	16	SCCCLR1001C03 (3E-37) / SCEZAM1082G06 (4E-38) / SCUTLR2008G03 (1E-38)	3	X	X	X	X	X	
	sp P27807 H2B1_WHEAT. Histone H2B.1	12; 15; 16; 17	SCCCLR2C02B06 (1E-46)	4	X	X	X	X	X	

Tabela 7 – Proteínas nucleares identificadas em folhas de cana-de-açúcar de 4 meses de idade

(continuação)

Função	Identificação	Banda	Cluster	Veze	B	S	W	P	PI	E.C.
	sp P54348 H2B5_MAIZE. Histone H2B.5	12; 15; 16; 17	SCBGFL3095E08 (1E-36) / SCCCLR1024A09 (1E-36) / SCCCLR1048C02 (2E-36) / SCCCLR1C04B01 (1E-36) / SCCCLR2002G11 (5E-37) / SCCCRZ1001H04 (1E-36) / SCEPLB1042A02 (7E-37) / SCEQLR1050A02 (2E-38) / SCJLLR2020F01 (7E-37) / SCUTLR2015D03 (8E-37)	40	X	X	X	X	X	
	sp P59226 H32_ARATH Histone H3.2	17	SCCCLR1C04B07 (8E-59)	1	nd	X	X	X	X	
	sp P68428 H32_WHEAT. Histone H3.2	17	SCBFLR1026A06 (6E-59) / SCAGLR2026E02 (7E-59) / SCCCRZ2001G08 (4E-54) / SCVPLB1016B08 (8E-59) / SCCCLR2001D10 (2E-54) / SCEQLB1063F02 (1E-58) / SCVPLB1017H05 (5E-57) / SCAGFL8013E01 (3E-54) / SCQGLR1019F06 (8E-57) / SCCCLR1001F12 (1E-58) / SCEQLB1066C12 (2E-56) / SCQGLR1062C03 (4E-54) / SCJFRZ2009C12 (3E-54) / SCCCLR1068F12 (2E-58) / SCMCLR1032D12 (2E-54) / SCBGLR1002H11 (9E-59)	16	nd	X	X	X	X	
	sp Q28DR4 H4_XENTR. Histone H4	19	SCAGLR2026F05 (7E-27)	1	X	X	X	X	X	
	sp Q2QPG9 H2AXB_ORYSJ Probable histone H2AXB	16	SCCCLR1001G05 (5E-46) / SCCCLR1068C05 (7E-43) / SCEPAM1020H08 (3E-31)	3	X	X	X	nd	X	
	sp Q43216 H2B5_WHEAT. Histone H2B.5	15; 16; 17	SCMCLR1032A12 (7E-16)	3	X	X	X	X	X	
	sp Q6L500 H2A4_ORYSJ. Probable histone H2A.4	16	SCCCRZ2002H03 (2E-39)	1	X	X	X	X	X	
	sp Q6ZBP3 H2B2_ORYSJ. Histone H2B.2 OS=Oryza sativ a subsp. japonica GN=H2B.2 PE=3 SV=1	12; 15; 16; 17	SCBGLR1002E03 (1E-43) / SCCCLR1068G11 (5E-44) / SCCCRZ1004F04 (5E-42) / SCEZLB1014E02 (4E-43) / SCJLLR1101H08 (4E-44) / SCMCRT2108A05 (4E-43) / SCQGLR1041C10 (9E-44) / SCQGLR1062G09 (1E-43) / SCSFLL3090D03 (7E-44) / SCSGAM1096G11 (8E-44) / SCSGFL5C05H11 (2E-33)	44	X	X	X	X	X	
	sp Q6ZL42 H2A2_ORYSJ. Probable histone H2A.2	16	SCJFRZ2010D04 (9E-35) / SCJFRZ2014A08 (8E-35) / SCVPLR2019A02 (1E-34)	3	X	X	X	nd	X	

Tabela 7 – Proteínas nucleares identificadas em folhas de cana-de-açúcar de 4 meses de idade

(continuação)

Função	Identificação	Banda	Cluster	Vezes	B	S	W	P	PI	E.C.
morfogênese foliar	sp Q71H73 H33_VITVI. Histone H3.3	17; 19	SCBGFL3095F09 (1E-70) / SCJFRZ2006G04 (1E-70) / SCCCZR2002A05 (2E-70) / SCCCRL2001F08 (2E-70) / SCJFLR1013C11 (9E-71) / SCCCFL5003D06 (1E-70) / SCJFLR1035H03 (1E-70) / SCUTLR2008H04 (1E-70) / SCVPRZ2039D03 (2E-70) / SCJFRT1060F02 (2E-70) / SCBGLR1023F03 (2E-70) / SCEQLR1094C02 (2E-70) / SCSGFL4037A02 (6E-57) / SCCCCL7C02A12 (2E-65) / SCBGLR1002A05 (2E-70) / SCCCCL3120E10 (2E-70) / SCRFLR1055C06 (2E-70) / SCVPLR2027D12 (8E-40) / SCRFLR1055A01 (4E-43) / SCQSRZ2035A08 (2E-32)	37	nd	X	X	X	X	
	sp Q75L11 H2A6_ORYSJ. Probable histone H2A.6	16	SCAGLR1021A12 (5E-35) / SCCCFL3003C11 (3E-38) / SCCCBL1024F05 (7E-37) / SCCCRL1048F07 (2E-41) / SCCCRL2001C09 (4E-39) / SCCCRL2001G07 (3E-35) / SCCCRL2C02C12 (7E-40) / SCCCZR2C03C04 (2E-41) / SCEPLR1008D05 (1E-57) / SCJLLR1101A02 (4E-47)	8	X	X	X	X	X	
	sp Q84MP7 H2AV3_ORYSJ. Probable histone H2A variant 3	16; 18	SCEPLR1008D05 (1E-57) / SCJLLR1101A02 (4E-47)	4	nd	X	X	nd	X	
	sp Q8H7Y8 H2AV1_ORYSJ. Probable histone H2A variant 1	16; 18	SCCCCL4007C09 (8E-54)	2	X	X	X	nd	X	
	sp Q943L2 H2B11_ORYSJ Histone H2B.11	12; 15; 16; 17	SCVPLR1049C05 (1E-37)	4	X	X	X	X	X	
	tr B6UIQ0 B6UIQ0_MAIZE Histone H2A	16	SCRLFL1005C02 (2.0×10-59)	1	X	X	X	nd	X	
	sp O24412 PSD7_ARATH. Probable 26S proteasome non-ATPase regulatory subunit 7	10	SCJFRZ2027F12 (1E-160)	1	X	nd	nd	nd	nd	
	sp Q6AV34 ARGC_ORYSJ. Probable N-acetyl-gamma-glutamyl-phosphate reductase, chloroplastic	10	SCEPAM1015F11 (0)	1	X	nd	nd	nd	nd	EC=1.2.1.38
	sp Q9M158 Y4105_ARATH. Uncharacterized protein At4 g01050, chloroplastic	8	SCEZHR1047F05 (2E-68)	1	X	nd	nd	nd		

Tabela 7 – Proteínas nucleares identificadas em folhas de cana-de-açúcar de 4 meses de idade

(continuação)

Função	Identificação	Banda	Cluster	Veze	B	S	W	P	PI	E.C.
proteólise	sp P0CH32 UBQ4_ARATH. Polyubiquitin 4	11; 12	SCBFLR1026D07 (0) / SCBFRZ3006E06 (1E-120) / SCCCL3080A11 (1E-117) / SCCCL3080A11.b (1E-117) / SCCCL4014E10 (0) / SCCCLR1001C11 (0) / SCCCLR1C03C10 (1E-161) / SCCCLR1C03H09 (1E-132) / SCVPRZ3026B09 (1E-117)	18	X	nd	nd	nd	X	
	sp P69309 UBIQP_AVEFA. Polyubiquitin	11; 12	SCCST2001G02 (1E-114) / SCQSRT2032E09 (1E-124) / SCVPRZ2038C12 (1E-168)	6	X	nd	nd	nd	X	
	sp Q6YT00 PSA7A_ORYSJ. Proteasome subunit alpha type-7-A	11	SCEZLB1009E05 (1E-123) / SCJFRZ2010D05 (1E-121)	2	X	nd	nd	nd	X	<u>EC=3.4.25.1</u>
	sp Q8H159 UBQ10_ARATH. Polyubiquitin 10	11; 12	SCACLR1057F07 (0) / SCCCRZ1002C04 (0)	4	X	nd	nd	nd	X	
	sp Q8LQJ8 FTSH5_ORYSJ. ATP-dependent zinc metalloprotease FTSH 5, mitochondrial	6	SCCCLR1078F01 (0)	1	X	nd	nd	nd	X	<u>EC=3.4.24.-</u>
	sp Q9C5U3 PRS8A_ARATH. 26S protease regulatory subunit 8 homolog A	9	SCCCCL4006C07 (0)	1	nd	nd	nd	nd	X	
	sp Q9LT08 PSDE_ARATH. 26S proteasome non-ATPase regulatory subunit 14	11	SCCCLR1048F08 (1E-151)	1	X	nd	X	X	X	<u>EC=3.4.19.-</u>
	sp Q9MAK9 PS10B_ARATH. 26S protease regulatory sub unit S10B homolog B	9	SCCCLR1022E06 (0)	1	nd	nd	nd	nd	X	
	tr G1L2G6 G1L2G6_AILME Uncharacterized protein (Fragment)	8	SCMCAM2081F09 (7.0×10-1)	1	X	X	X	X	X	
	sp Q8S0J7 IM30_ORYSJ. Probable membrane-associated 30 kDa protein, chloroplastic	11	SCUTLR2015H12 (1E-151)	1	nd	nd	nd	X	nd	
regulação da transcrição DNA-dependente	sp P09189 HSP7C_PETHY. Heat shock cognate 70 kDa protein	5	SCAGLR1043F02 (0) / SCCCL3120G07 (0) / SCCCLR1079D09 (6E-90) / SCCCLR1C01D05 (0)	4	nd	nd	nd	X	nd	
	sp P24067 BIP2_MAIZE. Luminal-binding protein 2	5	SCJLLR1054F05 (0)	1	nd	nd	nd	nd	X	
	sp P27322 HSP72_SOLLC. Heat shock cognate 70 kDa protein 2	5	SCCCRZ1002G01 (0) / SCSFST1064C10 (1E-103)	2	nd	nd	nd	X	nd	
regulação da transcrição RNA-dependente										
resposta ao estresse										

Tabela 7 – Proteínas nucleares identificadas em folhas de cana-de-açúcar de 4 meses de idade

(continuação)

Função	Identificação	Banda	Cluster	Veze	B	S	W	P	PI	E.C.
sinalização	sp Q07078 HSP83_ORYSJ. Heat shock protein 81-3	5	SCCCLR2C02F10 (1E-115) / SCRLLR1110D08 (1E-160)	2	X	nd	X	X	nd	
	sp Q08277 HSP82_MAIZE. Heat shock protein 82	5	SCJFRT2055B05 (6E-99)	1	X	nd	nd	X	nd	
	sp Q0J4P2 HSP81_ORYSJ. Heat shock protein 81-1	1; 5	SCCCLR1075E09 (0) / SCJLLR1011E03 (0) / SCEZLB1006B07 (1E-100) / SCMCCL6049H03 (1E-101)	6	X	nd	nd	X	nd	
	sp Q5QWR2 HTPG_IDILO. Chaperone protein htpG	4	SCCCCL3005G07.b (1E-104)	1	X	nd	nd	nd	nd	
	tr C5X3T9 C5X3T9_SORBI Putative uncharacterized protein Sb02g028050	5	SCCCLR1075H11 (7.0×10 ⁻¹)	1	X	nd	X	X	nd	
	sp P49027 GBLPA_ORYSJ. Guanine nucleotide-binding protein subunit beta-like protein A	10; 11	SCACLR2007C03 (1E-171) / SCSFAM1076C12 (5E-80)	4	X	nd	nd	nd	nd	
	sp P49106 14331_MAIZE. 14-3-3-like protein GF14-6	10; 11	SCCCLR1022D05 (1E-142) / SCRUF11017E10.b (1E-71)	4	X	nd	nd	nd	X	
	sp Q2R2W2 14334_ORYSJ. 14-3-3-like protein GF14-D	10; 11	SCCCLR1048F12 (1E-124) / SCEQRT1025D06 (3E-79) / SCMCRT2105C11 (7E-79)	4	X	nd	nd	nd	X	
	sp Q6ZKC0 14333_ORYSJ. 14-3-3-like protein GF14-C	10; 11; 12	SCBFLR1026E02 (1E-131) / SCCCRZ1001D02 (1E-134) / SCEQRT1031D02 (1E-113) / SCCCCL5072E02 (1E-40)	10	X	nd	nd	nd	X	
	sp Q84J55 14331_ORYSJ. 14-3-3-like protein GF14-A	11	SCEQLR1007H12 (1E-33) / SCEQRT2094B01 (6E-73) / SCMCRT2102A01 (3E-74)	3	X	nd	nd	nd	X	
	sp Q9SP07 1433_LILLO. 14-3-3-like protein	10; 11	SCAGLB1069G08 (2E-76)	2	X	nd	nd	nd	X	
	tr C5Y3X7 C5Y3X7_SORBI Putative uncharacterized protein Sb05g021020	11	SCQGST1034A08 (7.0×10 ⁻¹)	1	X	nd	X	nd	X	
tradução	tr C5YJQ8 C5YJQ8_SORBI Putative uncharacterized protein Sb07g029110	10; 11	SCCCCL4011G09 (7.0×10 ⁻¹)	2	X	nd	nd	nd	X	
	sp A1E9W6 RK2_SORBI 50S ribosomal protein L2, chloroplastic	20	SCACSB1035E05 (4E-59)	1	nd	nd	nd	X	nd	
	sp A1E9W9 RR7_SORBI. 30S ribosomal protein S7, chloroplastic	16	SCCCLR2C01E06 (8E-81)	1	nd	nd	nd	X	nd	
	sp A5BUU4 RSSA_VITVI. 40S ribosomal protein SA	9; 10	SCCCLR1001D03 (1E-110) / SCCST1001H07 (1E-118) / SCEPLR1051E04 (1E-110) / SCSGAM1094H05 (1E-90) / SCSGLR1045A06 (1E-116)	10	X	nd	nd	nd	nd	

Tabela 7 – Proteínas nucleares identificadas em folhas de cana-de-açúcar de 4 meses de idade

(continuação)

Função	Identificação	Banda	Cluster	Veze	B	S	W	P	PI	E.C.
	sp O64937 EF1A_ORYSJ. Elongation factor 1-alpha	1; 7; 8; 9; 10; 11	SCCCLR1022B09 (0) / SCEZRZ1013H01 (0) / SCJLRT1006G01 (0) / SCVPLB1020E07 (0)	24	nd	nd	nd	X	X	
	sp O65731 RS5_CICAR. 40S ribosomal protein S5 (Fragment)	15	SCACLR1127E08 (3E-94) / SCRFLR2038B04 (9E-72) / SCUTSD2088D06 (8E-60)	3	X	nd	nd	nd	nd	
	sp O80360 RK3_TOBAC. 50S ribosomal protein L3, chloroplastic (Fragment)	12; 13	SCEQLB1066D08 (2E-97)	2	nd	nd	X	X	nd	
	sp P29766 RL8_SOLLC. 60S ribosomal protein L8	11	SCCCLR2004E06 (1E-133) / SCVPLR2027H04 (1E-134)	2	X	nd	nd	nd	nd	
	sp P34823 EF1A2_DAUCA. Elongation factor 1-alpha	1; 7; 8; 9; 10; 11	SCSBRT3039F07 (8E-45)	6	nd	nd	nd	nd	X	
	sp P41129 RL132_BRANA. 60S ribosomal protein L13-2	13	SCAGLR2033E09 (6E-76) / SCCCL3003G05.b (4E-81) / SCJFRZ2013D07 (8E-80)	3	nd	nd	nd	X	nd	
	sp P41378 IF4A_WHEAT. Eukaryotic initiation factor 4A	8	SCRFR3057C02 (2E-74)	1	X	nd	nd	nd	X	<u>EC=3.6.4.13</u>
	sp P45633 RL10_MAIZE. 60S ribosomal protein L10	13; 14	SCBFLR1026E01 (1E-123) / SCBFRZ2016G12 (6E-98)	4	nd	nd	nd	X	X	
	sp P49199 RS8_ORYSJ. 40S ribosomal protein S8	13	SCCCLR1048F11 (5E-94) / SCCCRZ2002D10 (5E-78) / SCQSST3114E08 (3E-41)	3	nd	nd	nd	X	nd	
	sp P49397 RS3A_ORYSJ. 40S ribosomal protein S3a	11	SCBGLR1047A12 (1E-128) / SCCCL3005C01.b (1E-123) / SCCCL4006A03 (1E-119) / SCCCLR2001F09 (1E-123) / SCMCLR1032F10 (3E-78) / SCRLAM1014C12 (5E-87) / SCSFST3074H02 (1E-116) / SCVPRZ2037B05 (3E-19)	8	X	nd	nd	X	X	
	sp P50883 RL121_ARATH. 60S ribosomal protein L12-1	16	SCAGLR1021E04 (2E-80) / SCCCLR1C01C09 (4E-77)	2	nd	nd	nd	nd	X	
	sp Q03033 EF1A_WHEAT. Elongation factor 1-alpha	8	SCCCLB1021F05 (3E-89) / SCJFAD1012C08 (8E-90)	2	nd	nd	nd	X	X	
	sp Q05761 RS13_MAIZE. 40S ribosomal protein S13	16	SCCCLR1024D10 (5E-79) / SCUTLR1058E09 (8E-80)	2	nd	nd	X	X	nd	
	sp Q0DK10 RL11_ORYSJ. 60S ribosomal protein L11	16	SCCCRZ2001F06 (7E-88) / SCVPLR2019F01 (5E-88)	2	nd	nd	nd	X	nd	
	sp Q0ITS8 RL101_ORYSJ. 60S ribosomal protein L10-1	13; 14	SCEPRZ1010E07 (1E-120)	2	nd	nd	nd	X	X	

Tabela 7 – Proteínas nucleares identificadas em folhas de cana-de-açúcar de 4 meses de idade

(continuação)

Função	Identificação	Banda	Cluster	Veze	B	S	W	P	PI	E.C.
	sp Q0JGY1 RL51_ORYSJ. 60S ribosomal protein L5-1	10	SCCCLR1048A09 (1E-130) / SCEQLR1091D05 (1E-141) / SCUTLR1037F12 (1E-127)	3	nd	nd	nd	X	X	
	sp Q41741 IF4A_MAIZE. Eukaryotic initiation factor 4A	8	SCCCLR1C07G11 (0)	1	X	X	nd	nd	X	<u>EC=3.6.4.13</u>
	sp Q41803 EF1A_MAIZE. Elongation factor 1-alpha	1; 7; 8; 9; 10; 11	SCCCCL3004F01.b (0) / SCEQLR1007G03 (0)	12	nd	nd	nd	X	X	
	sp Q6Z2Z4 IF4A3_ORYSJ. Eukaryotic initiation factor 4A-3	8	SCCCLR1072G07 (0) / SCSSGSB1005E04 (7E-40)	2	X	X	nd	nd	X	<u>EC=3.6.4.13</u>
	sp Q8LJU5 RS7_ORYSJ. 40S ribosomal protein S7	11; 14	SCVPRZ2038F11 (5E-95)	2	X	nd	nd	X	nd	
	sp Q8VZB9 R10A1_ARATH. 60S ribosomal protein L10a-1	12	SCCCLR1C05G05 (1E-93)	1	X	nd	nd	nd	nd	
	sp Q9AYP4 RS10_ORYSJ. 40S ribosomal protein S10	16	SCACLR1126F04 (5E-60) / SCCCLR2C01F11 (3E-50) / SCMCST1051E11 (2E-42)	3	X	nd	X	nd	X	
	sp Q9C285 RL12_NEUCR. 60S ribosomal protein L12	16	SCEQRT1024D04 (7E-77)	1	nd	nd	nd	X	X	
	sp Q9C9C5 RL63_ARATH. 60S ribosomal protein L6-3	13	SCCCCL3080C04 (4E-68) / SCEQLB1063E01 (1E-74)	2	nd	nd	X	X	nd	
	sp Q9FJA6 RS33_ARATH. 40S ribosomal protein S3-3	13	SCAGLR1043G11 (5E-61) / SCCL3001F09 (1E-109) / SCCCRZ2003E04 (1E-106) / SCJFFL1C05F09.b (2E-17)	4	nd	nd	nd	X	X	
	sp Q9FKC0 R13A4_ARATH. 60S ribosomal protein L13a-4	14	SCCCCL4004E02 (4E-90) / SCRFLR1012C07 (4E-90)	2	nd	nd	nd	X	nd	
	sp Q9LHP1 RL74_ARATH. 60S ribosomal protein L7-4	12; 13	SCCCLR2002E05 (7E-87) / SCCCLR2004C07 (7E-87) / SCCCLR1C05H09 (2E-78) / SCCCLR1C10E12 (5E-95)	6	X	nd	nd	X	nd	
	sp Q9M339 RS32_ARATH. 40S ribosomal protein S3-2	13	SCVPRT2080E04 (1E-89)	1	nd	nd	nd	X	X	
	sp Q9SCM3 RS24_ARATH. 40S ribosomal protein S2-4	11	SCCCCL4015D08 (2E-99) / SCCCRZ2002H09 (1E-100) / SCEPRZ1009G03 (1E-100)	3	X	nd	nd	nd	X	
	sp Q9SF40 RL4A_ARATH. 60S ribosomal protein L4-1	8; 9; 10	SCCCCL3140B12 (1E-151) / SCCCRZ1001C08 (1E-112) / SCJFRT1011A02 (2E-35)	7	X	nd	nd	X	nd	
	sp Q9XET4 RS7_SECCE. 40S ribosomal protein S7	11; 14; 15	SCCCCL4010B12 (8E-98) / SCVPLR2005E09 (2E-96)	5	X	nd	nd	X	nd	
	sp Q9ZUT9 RS51_ARATH. 40S ribosomal protein S5-1	15	SCCCCL3001E09.b (7E-93)	1	X	nd	nd	nd	nd	
	tr C5Y0E9 C5Y0E9_SORBI Putative uncharacterized protein Sb04g030890	16	SCACLR2014B04 (2E-84)	1	X	nd	nd	nd	X	

Tabela 7 – Proteínas nucleares identificadas em folhas de cana-de-açúcar de 4 meses de idade

(conclusão)

Função	Identificação	Banda	Cluster	Vezes	<i>B</i>	<i>S</i>	<i>W</i>	<i>P</i>	<i>PI</i>	E.C.
transporte	sp Q9SJ12 ATP7_ARATH. Probable ATP synthase 24 kDa subunit, mitochondrial	12	SCEQRT2098D07 (2E-67) / SCEZLR1052F04 (1E-66)	2	nd	nd	nd	X	nd	

ANEXO F

Tabela 8 – Proteínas nucleares encontradas apenas em folhas de cana-de-açúcar de 1 mês

(continua)

Função	Identificação	Banda	Cluster
amino acid metabolism fotossíntese	sp Q9SJ05 AROK_ARATH. Shikimate kinase, chloroplastic	13	SCQSST3114G03
	sp O65107 PSAN_MAIZE. Photosystem I reaction center subunit N, chloroplastic (Fragment)	19	SCACLR1129A11
	sp O80634 PPL2_ARATH. PsbP-like protein 2, chloroplastic	16	SCJLLR1102C07
	sp P12329 CB21_MAIZE. Chlorophyll a-b binding protein 1, chloroplastic	3;4;16;19	SCJFLR1074A11
	sp P12331 CB22_ORYSJ. Chlorophyll a-b binding protein 2, chloroplastic	3;4;15;16;19	SCCCLR1066F06/SCJFRZ1005F03
	sp P27524 CB4A_SOLLC. Chlorophyll a-b binding protein CP24 10A, chloroplastic	1;9;10;13;14	SCJFSB1014F03.b
	sp P27525 CB4B_SOLLC. Chlorophyll a-b binding protein CP24 10B, chloroplastic	1;8;9;10;13;14	SCACLR2007B05
	sp Q41048 PSBQ1_MAIZE. Oxygen-evolving enhancer protein 3-1, chloroplastic	9;10	SCBGLR1114G09
	sp P94063 HAL3B_ARATH. Probable phosphopantothienoylcysteine decarboxylase	7	SCQSRT2031G03
	sp Q8H6A5 TCTP_MAIZE. Translationally-controlled tumor protein homolog	13	SCCCLR2001G09/SCEZLR1031E10
função desconhecida	tr C0PAK4 C0PAK4_MAIZE Putative uncharacterized protein	13;14;15;19	SCRUAD1133C09.b
	tr C5WRY9 C5WRY9_SORBI Putative uncharacterized protein Sb01g000680	12	SCJFRT1012A02
	tr C5WVY3 C5WVY3_SORBI Putative uncharacterized protein Sb01g045750	16	SCEPSD1004E09
	tr C5XIC1 C5XIC1_SORBI Putative uncharacterized protein Sb03g033020	1	SCBFST3135C09.b
	tr C5YBP8 C5YBP8_SORBI Putative uncharacterized protein Sb06g000920	8	SCCCST2002G06
	tr C5YJN5 C5YJN5_SORBI Putative uncharacterized protein Sb07g028880	12	SCCCRT1C06B08

Tabela 8 – Proteínas nucleares encontradas apenas em folhas de cana-de-açúcar de 1 mês

(continuação)

Função	Identificação	Banda	Cluster
glicólise	tr C5YTS7 C5YTS7_SORBI Putative uncharacterized protein Sb08g007230	10	SCAGFL3023D11
	sp Q40677 ALFC_ORYSJ. Fructose-bisphosphate aldolase, chloroplastic	10	SCMCLV1030C12/SCRULR1020D11
homeostase celular redox	sp O64628 TXND9_ARATH. Thioredoxin domain-containing protein 9 homolog	13;19	SCEZRZ1012C04
metabolismo de ATP	sp Q43433 VATB2_GOSHI. V-type proton ATPase subunit B 2 (Fragment)	8	SCQSLR1089G05
	sp Q9SWE7 VATE_CITLI. V-type proton ATPase subunit E	10	SCCCLB1025B04/SCEPLB1042D04
	sp Q9XGM1 VATD_ARATH. V-type proton ATPase subunit D	11	SCEPLB1044F03
metabolismo de carboidratos	sp B0UV30 TAL_HAES2. Transaldolase	10	SCACLR2022D05
metabolismo de membrana	sp Q6H7M7 OEP80_ORYSJ. Outer envelope protein of 80 kDa	7	SCEPAM2014E05
metabolismo de nucleotídeos	sp P82277 RRP2_SPIOL. 30S ribosomal protein 2, chloroplastic	14	SCRRLRZ3039A08
metabolismo de proteínas	sp Q6H795 CLPD1_ORYSJ. Chaperone protein ClpD1, chloroplastic	1	SCQGRT1041D10
	tr C5YF84 C5YF84_SORBI Putative uncharacterized protein Sb06g014590	1;6	SCQGLR1062F08
metabolismo de RNA	sp Q9ZUU4 ROC1_ARATH. Ribonucleoprotein At2g37220, chloroplastic	13	SCCCST1008B08
metabolismo do DNA	sp Q9M7J4 MFP1_TOBAC. MAR-binding filament-like protein 1-1	5	SCCCFL5093B04
montagem de nucleossomos	sp A2ZK26 H2A8_ORYSI. Probable histone H2A.8	15	SCJLLR1101F06/SCJLLR2013B11
	sp P0CG89 H4_SOYBN. Histone H4	1;10;15;16	SCAGLB1069E02/SCAGLB1070C07/SCCCLR1078H02/SCCCLR2001D01/SCCCLR2001D12/SCCCLR2002D04/SCCCLR2002G09/SCCCLR2003E01/SCCCRZ1004C05/SCCCRZ2001G01/SCEPLR1051H12/SCJFLR1013A12/SCJFRZ2009F02/SCJLHR1028C12/SCJLLR1011E11/SCJLLR2020B06/SCRFLR2038C05/SCSBFL1043F08/SCSBLB1033H09/SCSGHR1071E06/SCSGRT2062A10/SCUTLR2015B06/SCVPRZ2035C06/SCAGLB1069E02
	sp P19177 H2A_PETCR. Histone H2A	10;15;17;18	SCEZAM1082G06/SCCCLR1001C03/SCUTLR2008G03

Tabela 8 – Proteínas nucleares encontradas apenas em folhas de cana-de-açúcar de 1 mês

(continuação)

Função	Identificação	Banda	Cluster
processo de oxidação-redução	sp P27807 H2B1_WHEAT. Histone H2B.1	10;13;1 4	SCCCLR2C02B06
	sp P54348 H2B5_MAIZE. Histone H2B.5	9;10;13; 14	SCBGFL3095E08/SCCCLR1024A09/SCCCLR1048C02/SCCCLR1C04B01/SCCCLR2002G11/SCCCRZ1001H04/SCEPLB1042A02/SCEQLR1050A02/SCJLLR2020F01/SCUTLR2015D03/SCEQLR1050A02
	sp Q28DR4 H4_XENTR. Histone H4	18	SCAGLR2026F05
	sp Q2QPG9 H2AXB_ORYSJ Probable histone H2AXb	15;10;6	SCCCLR1001G05/SCCCLR1068C05/SCEPAM1020H08/SCBGLR1117A10
	sp Q43216 H2B5_WHEAT. Histone H2B.5	10;13;1 4	SCMCLR1032A12
	sp Q6L500 H2A4_ORYSJ. Probable histone H2A.4	10;15;1 7;18	SCCCRZ2002H03
	sp Q6ZBP3.1 H2B2_ORYSJ. Histone H2B.2 RecName: Full=Histone H2B.2	10;13;1 4;	SCBGLR1002E03/SCCCLR1068G11/SCCCRZ1004F04/SCEZLB1014E02/SCJLLR1101H08/SCMCRT2108A05/SCQGLR1041C10/SCQGLR1062G09/SCSFFL3090D03/SCSGAM1096G11/SCSGFL5C05H11/
	sp Q6ZL42 H2A2_ORYSJ. Probable histone H2A.2	15	SCJFRZ2010D04/SCJFRZ2014A08/SCVPLR2019A02
	sp Q71H73 H33_VITVI. Histone H3.3	16	SCBGFL3095F09/SCJFRZ2006G04/SCBGLR1002A05/SCBGLR1023F03/SCCCCL3120E10/SCCCCL7C02A12/SCCCFL5003D06/SCCCLR2001F08/SCCCRZ2002A05/SCEQLR1094C02/SCJFLR1013C11/SCJFLR1035H03/SCJFRT1060F02/SCRFLR1055C06/SCSGFL4037A02/SCUTLR2008H04/SCVPLR2027D12/SCVPRZ2039D03
	sp Q75L11 H2A6_ORYSJ. Probable histone H2A.6	10;15;1 7;18	SCAGLR1021A12/SCCCFL3003C11/SCCCLB1024F05/SCCCLR1048F07/SCCCLR2001C09/SCCCLR2001G07/SCCCLR2C02C12/SCCCRZ2C03C04
	sp Q84MP7 H2AV3_ORYSJ. Probable histone H2A variant 3	10;14;1 5;11;12	SCEPLR1008D05/SCJLLR1101A02/SCQSLR1089C08/SCUTRZ3072F02
	sp Q8H7Y8 H2AV1_ORYSJ. Probable histone H2A variant 1	14;15	SCCCCL4007C09
	sp Q943L2 H2B11_ORYSJ Histone H2B.11	10;13;1 4	SCVPLR1049C05
	tr B6UIQ0 B6UIQ0_MAIZE Histone H2A	15	SCRLFL1005C02
	sp O04130 SERA_ARATH. D-3-phosphoglycerate dehydrogenase, chloroplastic	7	SCCCLR2002G02/SCJFRT1005D01
	sp P18123 CATA3_MAIZE. Catalase isozyme 3	7;8	SCCCLR1048H09/SCCCAD1004H02

Tabela 8 – Proteínas nucleares encontradas apenas em folhas de cana-de-açúcar de 1 mês

(continuação)

Função	Identificação	Banda	Cluster
processos biossintéticos	sp Q69TY4 PR2E1_ORYSJ. Peroxiredoxin-2E-1, chloroplastic	15;16	SCEZRZ1014F10
	sp Q7FB12 NDHM_ORYSJ. NAD(P)H-quinone oxidoreductase subunit M, chloroplastic	12;14;15	SCEPSD2005D11
	sp Q7NF11 NDHO_GLOVI NAD(P)H-quinone oxidoreductase subunit O	16	SCAGAM2126G09
	sp Q93VK5 LUT5_ARATH. Protein LUTEIN DEFICIENT 5, chloroplastic	7	SCJFLR1073C07
	tr B6TCZ3 B6TCZ3_MAIZE Malate dehydrogenase	10	SCCCLR1072A12
	sp O04974 LEU1B_SOLPN. 2-isopropylmalate synthase B	7	SCQGLB1039D07
	sp P34106 ALA2_PANMI. Alanine aminotransferase 2	8;9	SCCCRZ1002E01
	sp P72580 PREA_SYNY3. Prenyl transferase	10	SCCCAM1002F08
processos celulares	sp P04907 GSTF3_MAIZE. Glutathione S-transferase 3	13	SCCCLR1001A07/SCEZRT3070A09/SCJFLR1073H08
	sp P24203 YJIA_ECOLI. Uncharacterized GTP-binding protein YjiA	1;8	SCRULB1060G07
	sp Q9FM19 HIR1_ARATH. Hypersensitive-induced response protein 1	10;11	SCCCRZ1002F04/SCJFRT1008C05/SCRUAD1061C12
	sp Q9M158 Y4105_ARATH. Uncharacterized protein At4 g01050, chloroplastic	7;9	SCEZHR1047F05
proteólise	sp O64464 PSB1_ORYSJ. Proteasome subunit beta type-1	13	SCACLR2007D12/SCBGLR1002D03/SCRUFL1115H03.b
	sp P0CH04 UBI1P_PETCR Polyubiquitin	10	SCCCLB2004C08
	sp P0CH05 UBI2P_PETCR. Polyubiquitin	10	SCCCST1005F12/SCJFRT2055G07
	sp P0CH32 UBQ4_ARATH. Polyubiquitin 4	10;13	SCBFLR1026D07/SCBFRZ3006E06/SCCCCL3080A11/SCCCCL3080A11.b/SCCCCL4014E10/SCCCLB1026D12SCCCLR1001C11/SCCCLR1C03C10/SCCCLR1C03H09/SCEZAM2059H12/SCSGFL4118F01/SCVPRZ3026B09/SCSGAM1095C10
	sp P46466 PRS4_ORYSJ. 26S protease regulatory subunit 4 homolog	8	SCSGAM1095C10

Tabela 8 – Proteínas nucleares encontradas apenas em folhas de cana-de-açúcar de 1 mês

(continuação)

Função	Identificação	Banda	Cluster
resposta ao estresse	sp P52428 PSA1_ORYSJ. Proteasome subunit alpha type-1	10	SCEPLB1041G09
	sp P69309 UBIQP_AVEFA. Polyubiquitin	10;13	SCCCST2001G02/SCJFRT1010F04/SCQSLB1052H09/SCQSRT2032E09/SCRLLV1026A01.b/SCRUFL1021F10.b/SCSFSD1065H04.b/SCVPRZ2038C12
	sp P73689 SPPA_SYNY3. Protease 4	7	SCUTLR1037A04
	sp P85200 PRS6B_HELAN. 26S protease regulatory subunit 6B homolog	8	SCCCLR1C07H11
	sp Q10KF0 PSA2_ORYSJ. Proteasome subunit alpha type-2	13	SCQSLR1018D04
	sp Q6YT00 PSA7A_ORYSJ. Proteasome subunit alpha type-7-A	11	SCVPRZ2038H03
	sp Q7DLR9 PSB4_ARATH. Proteasome subunit beta type-4	13	SCCCCL3120F07
	sp Q8H159 UBQ10_ARATH. Polyubiquitin 10	10;13	SCACLR1057F07/SCCCRZ1002C04/
	sp Q8LD27 PSB6_ARATH. Proteasome subunit beta type-6	14	SCCCRZ2C04B02/SCQSAM1032E01
	sp Q8LQJ8 FTSH5_ORYSJ. ATP-dependent zinc metalloprotease FTSH 5, mitochondrial	7	SCCCLR1078F01
	sp Q9LST6 PSB2_ORYSJ. Proteasome subunit beta type-2	13;14	SCCCLR1048G12
	sp Q9LST7 PSB3_ORYSJ. Proteasome subunit beta type-3	13	SCCCLR1C07G01
	sp Q9LSU3 PSA6_ORYSJ. Proteasome subunit alpha type-6	13	SCCCLR1001B04
	sp Q9SEI4 PRS6B_ARATH. 26S protease regulatory subunit 6B homolog	8	SCCCLR1077D09
	tr C5XWW6 C5XWW6_SORBI Proteasome subunit beta type	13	SCCCLR1001C09
	tr C5YWZ8 C5YWZ8_SORBI ATP-dependent Clp protease proteolytic subunit	13	SCQSAM2100C03
	sp P09189 HSP7C_PETHY. Heat shock cognate 70 kDa protein	1	SCAGLR1043F02/SCCCCL3120G07/SCCCLR1079D09/SCCCLR1C01D05
	sp P27322 HSP72_SOLLC. Heat shock cognate 70 kDa protein 2	1	SCCCRZ1002G01/SCSFST1064C10

Tabela 8 – Proteínas nucleares encontradas apenas em folhas de cana-de-açúcar de 1 mês

(continuação)

Função	Identificação	Banda	Cluster
sinalização	sp Q07078 HSP83_ORYSJ. Heat shock protein 81-3	1	SCCCLR2C02F10/SCRLLR1110D08
	sp Q08277 HSP82_MAIZE. Heat shock protein 82	1	SCJFRT2055B05
	sp Q0J4P2 HSP81_ORYSJ. Heat shock protein 81-1	1	SCEZLB1006B07/SCMCCL6049H03
	sp Q5QWR2 HTPG_IDILO. Chaperone protein htpG	5	SCCCCL3005G07.b
	tr C5X3T9 C5X3T9_SORBI Putative uncharacterized protein Sb02g028050	1	SCCCLR1075H11
	sp P49106 14331_MAIZE. 14-3-3-like protein GF14-6	7	SCCCLR1022D05
	sp Q2R2W2 14334_ORYSJ. 14-3-3-like protein GF14-D	10	SCEQRT1025D06/SCMCRT2105C11
	sp Q6ZKC0 14333_ORYSJ. 14-3-3-like protein GF14-C	7;10	SCBFLR1026E02/SCCCRZ1001D02/SCEQRT1031D02/SCCCCL5072E02
	sp Q84J55 14331_ORYSJ. 14-3-3-like protein GF14-A	10	SCEQLR1007H12/SCEQRT2094B01/SCMCRT2102A01
	sp Q9SP07 1433_LILLO. 14-3-3-like protein	7	SCAGLB1069G08
tradução	tr C5Y3X7 C5Y3X7_SORBI Putative uncharacterized protein Sb05g021020	10	SCQGST1034A08
	sp A1E9W6 RK2_SORBI 50S ribosomal protein L2, chloroplastic	19	SCACSB1035E05
	sp A1E9W9 RR7_SORBI. 30S ribosomal protein S7, chloroplastic	14;15	SCCCLR2C01E06
	sp A3BLC3 RRFC_ORYSJ. Ribosome-recycling factor, chloroplastic	14	SCACLB1047F10
	sp O48557 RL17_MAIZE. 60S ribosomal protein L17	15;16	SCBGLR1120F10/SCEZRZ3095E08/SCJFLR1017D02/SCJFLR1017H09/SCJFRZ2005H06/SCBGLR1120F10
	sp O64937 EF1A_ORYSJ. Elongation factor 1-alpha	2	SCJLRT1006G01/SCCCLR1022B09/SCVPLB1020E07/SCEZRZ1013H01
	sp O65731 RS5_CICAR. 40S ribosomal protein S5 (Fragment)	14;16	SCACLR1127E08/SCRFLR2038B04/SCUTSD2088D06
	sp O80360 RK3_TOBAC. 50S ribosomal protein L3, chloroplastic (Fragment)	11;14	SCEQLB1066D08
	sp O82204 RL281_ARATH. 60S ribosomal protein L28-1	16	SCCCLR2002C09

Tabela 8 – Proteínas nucleares encontradas apenas em folhas de cana-de-açúcar de 1 mês

(continuação)

Função	Identificação	Banda	Cluster
	sp P19950 RS141_MAIZE. 40S ribosomal protein S14	16;17	SCCCLR1048A02/SCCCLR1C03A02
	sp P19951 RS142_MAIZE. 40S ribosomal protein S14	17	SCCCRT2002B03/SCRUFL3066B06.b
	sp P34788 RS18_ARATH. 40S ribosomal protein S18	16;17	SCCCLR1001H01/SCCCRZ2002C12
	sp P34823 EF1A2_DAUCA. Elongation factor 1-alpha	2	SCSBRT3039F07
	sp P42791 RL182_ARATH. 60S ribosomal protein L18-2	14;15	SCJFRZ2015G01
	sp P45633 RL10_MAIZE. 60S ribosomal protein L10	13	SCBGLR1119H10
	sp P49199 RS8_ORYSJ. 40S ribosomal protein S8	11;12	SCCCLR1048F11/SCCCRZ2002D10/SCQSST3114E08
	sp P49210 RL9_ORYSJ. 60S ribosomal protein L9	14;15	SCCCLR1069A12/SCBFRZ2045D04/SCJFFL3C02H03.b
	sp Q07761 RL23A_TOBAC. 60S ribosomal protein L23a	16	SCBGFL5081F09/SCCCLR1C03A11/SCEQLR1091A10
	sp Q08699 RS14_PODCA. 40S ribosomal protein S14	16;17	SCSBSD2054A09
	sp Q0DK10 RL11_ORYSJ. 60S ribosomal protein L11	15;16	SCCCRZ2001F06/SCVPLR2019F01/SCSGLR1045F12
	sp Q0IQF7 RS16_ORYSJ. 40S ribosomal protein S16	16;17	SCEPLR1030C12
	sp Q0ITS8 RL101_ORYSJ. 60S ribosomal protein L10-1	13	SCEZRZ1017F07
	sp Q40680 EF1D1_ORYSJ. Elongation factor 1-delta 1	10	SCCCRZ2002D07
	sp Q41741 IF4A_MAIZE. Eukaryotic initiation factor 4A	9	SCCCLR1C07G11
	sp Q41803 EF1A_MAIZE. Elongation factor 1-alpha	2	SCCCCL3004F01.b/SCEQLR1007G03
	sp Q5N1T5 EFP_SYNP6. Elongation factor P	14;15	SCQSAM2099H05
	sp Q8LJU5 RS7_ORYSJ. 40S ribosomal protein S7	12;13;15	SCVPRZ2038F11

Tabela 8 – Proteínas nucleares encontradas apenas em folhas de cana-de-açúcar de 1 mês

(conclusão)

Função	Identificação	Banda	Cluster
	sp Q8VZB9 R10A1_ARATH. 60S ribosomal protein L10a- 1	11;13	SCCCLR1C05G05
	sp Q940B0 RL183_ARATH. 60S ribosomal protein L18-3	14;15	SCBFLR1046F11
	sp Q96564 RS27_HORVU. 40S ribosomal protein S27	19	SCCCCL4008F04
	sp Q9AYP4 RS10_ORYSJ. 40S ribosomal protein S10	14;15	SCCCLR2C01F11
	sp Q9C9C5 RL63_ARATH. 60S ribosomal protein L6-3	12	SCCCCL3080C04/SCEQLB1063E01
	sp Q9FDZ9 RL212_ARATH. 60S ribosomal protein L21-2	15;15	SCCCCL3003D09.b/SCCCLR2C01F01/SCJFLR1017B02
	sp Q9FKC0 R13A4_ARATH. 60S ribosomal protein L13a- 4	13	SCCCCL4004E02/SCRFLR1012C07
	sp Q9FLF0 RS92_ARATH. 40S ribosomal protein S9-2	14;15	SCBFLR1083F02/SCCCLR2002C01/SCQGLR2017B10
	sp Q9FWS4 RK31_ARATH. 50S ribosomal protein L31, chloroplastic	19	SCEPLR1051F04
	sp Q9LHP1 RL74_ARATH. 60S ribosomal protein L7-4	12	SCCCLR1C10E12
	sp Q9LZ17 RS174_ARATH. 40S ribosomal protein S17-4	15;16;17	SCBGLR1002D11/SCCCLR1C05F01/SCRLFL1011G08
	sp Q9SCM3 RS24_ARATH. 40S ribosomal protein S2-4	10;12	SCCCCL4015D08/SCCCRZ2002H09/SCEPRZ1009G03
	sp Q9SF40 RL4A_ARATH. 60S ribosomal protein L4-1	7	SCCCCL3140B12/SCCCRZ1001C08
	sp Q9SS17 RS241_ARATH. 40S ribosomal protein S24-1	16	SCJLLR1054C12/SCMCRT2088H11
	sp Q9XET4 RS7_SECCE. 40S ribosomal protein S7	12;13;15	SCCCCL4010B12/SCVPLR2005E09
	sp Q9ZUT9 RS51_ARATH. 40S ribosomal protein S5-1	14;16	SCCCCL3001E09.b
	tr C5Y3N7 C5Y3N7_SORBI Putative uncharacterized protein Sb05g001680	16;17	SCCCLR2002A05

ANEXO G

Tabela 9 – Proteínas nucleares encontradas apenas em folhas de cana-de-açúcar de 4 meses

(continua)

Função	Identificação	Banda	Cluster
biogênese de ribossomos ciclo celular	sp P35685 RL7A_ORYSJ. 60S ribosomal protein L7a	11	SCCCRZ2001H04
	sp P54774 CDC48_SOYBN. Cell division cycle protein 48 homolog	4	SCVPLR1049C09
fixação de carbono função desconhecida metabolismo de proteínas	sp P15804 CAPP3_SORBI. Phosphoenolpyruvate carboxylase 3	16	SCRULR1020D09
	tr F2CR05 F2CR05_HORVD Predicted protein	13	SCVPFL3047D06
	sp Q6H795 CLPD1_ORYSJ. Chaperone protein ClpD1, chloroplastic	4	SCQGRT1041D10
	tr C5YF84 C5YF84_SORBI Putative uncharacterized protein Sb06g014590	4	SCQGLR1062F08
metabolismo do RNA montagem de nucleossomos	sp P19683 ROC4_NICSY. 31 kDa ribonucleoprotein, chloroplastic	12	SCAGLR1043B12
	sp P08437 H3_VOLCA. Histone H3	17	SCQGFL1095C08
	sp P0CG89 H4_SOYBN. Histone H4	11	SCAGLB1069E02/SCAGLB1070C07/SCCCLR1078H02/SCCCLR2001D01/SCCCLR2001D12/SCCCLR2002D04/SCCCLR2002G09/SCCCLR2003E01/SCCCRZ1004C05/SCCCRZ2001G01/SCEPLR1051H12/SCJFLR1013A12/SCJFRZ2009F02/SCJLLR1011E11/SCJLLR2020B06/SCRFLR2038C05/SCSBFL1043F08/SCSBLB1033H09/SCSGHR1071E06/SCSGHR1071E06/SCUTLR2015B06/SCVPRZ2035C06
	sp P27807 H2B1_WHEAT. Histone H2B.1	12	SCCCLR2C02B06
	sp P54348 H2B5_MAIZE. Histone H2B.5	12	SCBGFL3095E08/SCCCLR1024A09/SCCCLR1048C02/SCCCLR1C04B01/SCCCLR2002G11/SCCCRZ1001H04/SCEPLB1042A02/SCEQLR1050A02/SCJLLR2020F01/SCUTLR2015D03
	sp Q6ZBP3.1 H2B2_ORYSJ. Histone H2B.2	12	SCBGLR1002E03/SCCCLR1068G11/SCCCRZ1004F04/SCEZLB1014E02/SCJLLR1101H08/SCMCRT2108A05/SCQGLR1041C10/SCQGLR1062G09/SCSFFL3090D03/SCSGAM1096G11/SCSGFL5C05H11

Tabela 9 – Proteínas nucleares encontradas apenas em folhas de cana-de-açúcar de 4 meses

(continuação)

Função	Identificação	Banda	Cluster
processo de oxidação redução	sp Q71H73 H33_VITVI. Histone H3.3	17;19	SCBGFL3095F09/SCJFRZ2006G04/SCCCRZ2002A05/SCCCLR2001F08/SCJFLR1013C11/SCCCFL5003D06/SCJFLR1035H03/SCUTLR2008H04/SCVPRZ2039D03/SCJFRT1060F02/SCBGLR1023F03/SCEQLR1094C02/SCSGFL4037A02/SCCCCL7C02A12/SCBGLR1002A05/SCCCCL3120E10/SCRFLR1055C06/SCVPLR2027D12/SCCCFL5003D06/SCUTLR2008H04/SCCCCL3120E10/SCVPRZ2039D03/SCBGLR1023F03/SCJFLR1035H03/SCBGLR1002A05/SCJFLR1013C11/SCCCCL7C02A12/SCSGFL4037A02/SCRFLR1055C06/SCCCRZ2002A05/SCCCLR2001F08/SCEQLR1094C02/SCRFLR1055A01/SCQSRT2035A08/SCJFRT1060F02
	sp Q943L2 H2B11_ORYSJ Histone H2B.11	12	SCVPLR1049C05
	sp Q6AV34 ARGC_ORYSJ. Probable N-acetyl-gamma-glutamyl-phosphate reductase, chloroplastic	10	SCEPAM1015F11
	sp P0CH32 UBQ4_ARATH. Polyubiquitin 4	11;12	SCBFRZ3006E06/SCBFLR1026D07/SCCCCL3080A11/SCCCCL3080A11.b/SCCCCL4014E10/SCCCLR1001C11/SCCCLR1C03C10/SCCCLR1C03H09/SCVPRZ3026B09
	sp P69309 UBIQP_AVEFA. Polyubiquitin	12	SCCCST2001G02/SCQSRT2032E09/SCVPRZ2038C12
	sp Q8H159 UBQ10_ARATH. Polyubiquitin 10	12	SCACLR1057F07/SCCCRZ1002C04
	sp Q8LQJ8 FTSH5_ORYSJ. ATP-dependent zinc metalloprotease FTSH 5, mitochondrial	6	SCCCLR1078F01
	sp Q9LT08 PSDE_ARATH. 26S proteasome non-ATPase regulatory subunit 14	11	SCCCLR1048F08
	sp P09189 HSP7C_PETHY. Heat shock cognate 70 kDa protein	5	SCAGLR1043F02/SCCCCL3120G07/SCCCLR1079D09/SCCCLR1C01D05
	sp P24067 BIP2_MAIZE. Luminal-binding protein 2	5	SCJLLR1054F05
resposta ao estresse	sp P27322 HSP72_SOLLC. Heat shock cognate 70 kDa protein 2	5	SCCCRZ1002G01/SCSFST1064C10
	sp Q07078 HSP83_ORYSJ. Heat shock protein 81-3	5	SCCCLR2C02F10
	sp Q08277 HSP82_MAIZE. Heat shock protein 82	5	SCJFRT2055B05
	sp Q0J4P2 HSP81_ORYSJ. Heat shock protein 81-1	5	SCEZLB1006B07/SCMCCL6049H03
	sp Q5QWR2 HTPG_IDILO. Chaperone protein htpG	4	SCCCCL3005G07.b

Tabela 9 – Proteínas nucleares encontradas apenas em folhas de cana-de-açúcar de 4 meses

(conclusão)

Função	Identificação	Banda	Cluster
sinalização	sp P49027 GBLPA_ORYSJ. Guanine nucleotide-binding protein subunit beta-like protein A	11	SCACLR2007C03/SCSFAM1076C12
	sp Q6ZKC0 14333_ORYSJ. 14-3-3-like protein GF14-C	12	SCCCRZ1001D02/SCEQRT1031D02
	tr C5Y3X7 C5Y3X7_SORBI Putative uncharacterized protein Sb05g021020	11	SCQGST1034A08
tradução	sp A1E9W6 RK2_SORBI 50S ribosomal protein L2, chloroplastic	20	SCACSB1035E05
	sp A5BUU4 RSSA_VITVI. 40S ribosomal protein SA	9	SCCCLR1001D03/SCCCST1001H07/SCEPLR1051E04/SCSGAM1094H05/SCSGLR1045A06
	sp O64937 EF1A_ORYSJ. Elongation factor 1-alpha	11	SCCCLR1022B09/SCEZRZ1013H01/SCJLRT1006G01/SCVPLB1020E07
	sp P34823 EF1A2_DAUCA. Elongation factor 1-alpha	11	SCSBRT3039F07
	sp P41129 RL132_BRANA. 60S ribosomal protein L13-2	13	SCAGLR2033E09/SCCCCL3003G05.b/SCJFRZ2013D07
	sp P45633 RL10_MAIZE. 60S ribosomal protein L10	14	SCBFLR1026E01/SCBFRZ2016G12
	sp P49199 RS8_ORYSJ. 40S ribosomal protein S8	13	SCCCLR1048F11/SCCCRZ2002D10/SCQSST3114E08
	sp Q0ITS8 RL101_ORYSJ. 60S ribosomal protein L10-1	14	SCEPRZ1010E07
	sp Q41803 EF1A_MAIZE. Elongation factor 1-alpha	11	SCCCCL3004F01.b/SCEQLR1007G03
	sp Q9SJ12 ATP7_ARATH. Probable ATP synthase 24 kDa subunit, mitochondrial	12	SCEQRT2098D07/SCEZLR1052F04
transporte			