

UNIVERSIDADE ESTADUAL PAULISTA “Júlio de Mesquita Filho”

Instituto de Química

Programa de Pós-Graduação em Biotecnologia

Identificação de fatores de transcrição de *Neurospora crassa* envolvidos na regulação do metabolismo de glicogênio

Aluno - Rodrigo Duarte Gonçalves

Orientadora - Prof^a. Dr. Maria Célia Bertolini

2008

RODRIGO DUARTE GONÇALVES

Identificação de fatores de transcrição de *Neurospora crassa* envolvidos na regulação do metabolismo de glicogênio

Dissertação apresentada ao Instituto de Química, Universidade Estadual Paulista, Programa de Pós-Graduação em Biotecnologia, como parte dos requisitos para obtenção do título de Mestre em Biotecnologia.

Orientadora: **Prof^a. Dr. Maria Célia Bertolini**

ARARAQUARA

2008

Dados curriculares

Rodrigo Duarte Gonçalves

1. Dados pessoais

Filiação: Hermes Gonçalves e Maria Laura Duarte Gonçalves

Nascimento: 12 de agosto de 1981 – São Paulo-SP, Brasil

Documento de Identidade: 30.596.508-6 / SSP-SP

Cadastro de Pessoa Física: 304.079.778-64

Endereço Pessoal: Rua XV de Novembro, 295 – Centro

CEP: 19700-000 – Paraguaçu Paulista – SP

Endereço Profissional: Instituto de Química de Araraquara – UNESP

Departamento de Bioquímica e Tecnologia Química

Rua Prof. Francisco Degni s/n – 14800-900 Araraquara-SP

Profissão: Biólogo

2. Formação acadêmica

2006 – 2008: Mestrado em Biotecnologia, Programa de Pós-Graduação em Biotecnologia, Título: Identificação de fatores de transcrição de *Neurospora crassa* envolvidos na regulação do metabolismo de glicogênio. Orientador: Maria Célia Bertolini, Instituto de Química de Araraquara, Universidade Estadual Paulista Júlio de Mesquita Filho, UNESP, São Paulo, Brasil.

2002 – 2005: Graduação em Biologia, Faculdade de Ciências e Letras de Assis, Universidade Estadual Paulista, UNESP, São Paulo, Brasil.

3. Atividades de pesquisa

3.1. Trabalhos apresentados em congressos

GONÇALVES, R. D.; CUPERTINO, F. B.; BERTOLINI, M. C. *Neurospora crassa* transcription factors that might be involved in glycogen metabolism regulation. In: XXXVII Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular, Águas de Lindóia, 2008.

CUPERTINO, F. B.; GONÇALVES, R. D.; BERTOLINI, M. C. A systematic search for *Neurospora crassa* transcription factors involved in glycogen metabolism regulation. In: XXXVII Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular, Águas de Lindóia, 2008.

GONÇALVES, R. D.; BERTOLINI, M. C. Identificação de fatores de transcrição envolvidos no metabolismo de glicogênio em *Neurospora crassa*. In: 53º Congresso Nacional de Genética, 2007, Águas de Lindóia, 2007.

GONÇALVES, R. D.; LEVADA, P. M. ; OLIVEIRA, A. L. D. ; ALMEIDA, D. C. ; ZANOTTI-MAGALHÃES, E. M.; RIBEIRO-PAES, J. T. Análise da Expressão Gênica Diferencial do Tecido Cefalopodal em Linhagens de *Biomphalaria tenagophila* Resistentes e Suscetíveis à Infecção por *Schistosoma mansoni*. In: 51º Congresso Brasileiro de Genética, 2005, Águas de Lindóia, 2005.

GONÇALVES, R. D.; LEVADA, P. M.; OLIVEIRA, A. L. D.; ZANOTTI-MAGALHÃES, E. M ; SPADA, R. G. M.; RIBEIRO-PAES, J. T. Kinetic evaluation of hemocytes in *Biomphalaria tenagophila* strains selected and non-selected for susceptibility to *Schistosoma mansoni*. In: 10º Simpósio Internacional sobre Esquistossomose, 2005, Belo Horizonte, 2005.

GONÇALVES, R. D.; LEVADA, P. M.; OLIVEIRA, A. L. D.; ALMEIDA, D. C.; ZANOTTI-MAGALHÃES, E. M.; RIBEIRO-PAES, J. T. Análise da expressão gênica diferencial do tecido cefalopodal em linhagens de *Biomphalaria tenagophila* resistentes e suscetíveis à infecção por *Schistossoma mansoni*. In: VIII Encontro de Biociências e Biotecnologia de Assis, Assis, 2005.

LEVADA, P. M.; GONÇALVES, R. D.; OLIVEIRA, A. L. D.; RIBEIRO-PAES, J. T. Análise da variação do número de hemócitos em linhagens de *Biomphalaria tenagophila* selecionados e não selecionados geneticamente para o caráter suscetibilidade à infecção ao *Schistosoma mansoni*. In: VIII Encontro de Biociências e Biotecnologia de Assis, Assis, 2005

GONÇALVES, R. D.; LEVADA, P. M.; OLIVEIRA, A. L. D. ; ALMEIDA, D. C.; ZANOTTI-MAGALHÃES, E. M.; RIBEIRO-PAES, J. T. Comportamento de hemócitos em Linhagens de caramujos *Biomphalaria tenagophila* selecionados e não selecionados geneticamente para a suscetibilidade à infecção por *Schistosoma mansoni*. In: VI Congresso Londrinense de Biologia Aplicada à Saúde e I Simpósio Paranaense e Patologia Experimental, Londrina. 2004.

LEVADA, P. M.; GONÇALVES, R. D.; RIBEIRO-PAES, J. T.; ZANOTTI-MAGALHÃES, E. M.; OLIVEIRA, A. L. D. Comportamento de hemócitos em linhagens de caramujos *B. tenagophila* resistentes e suscetíveis à infecção por *S. mansoni*. In: XVI Congresso de Iniciação Científica, 2004, Ilha Solteira, 2004.

GONÇALVES, R. D.; ALMEIDA, D. C.; ROSSATO, D. R. ; VIANA, R. A.; LUCCAS, F. S. Distribuição das mortalidades por homicídios e acidentes de trânsito nas regiões administrativas de São Paulo. In: I Mostra de Produção Científica em Biologia, 2003, Distribuição das Mortalidades por Homicídios e Acidentes de Trânsito nas Regiões Administrativas de São Paulo, Assis, 2003.

3.2. Participação em eventos

XXXVII Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular, Águas de Lindóia, SP, 2008.

53º Congresso Brasileiro de Genética, Águas de Lindóia, SP, 2007.

II Äkta User Club Meeting, Ribeirão Preto, 2006.

51º Congresso Brasileiro de Genética, Águas de Lindóia, SP, 2005.

V Workshop de Genética, Botucatu, 2005.

IV Workshop de Genética, Botucatu, 2004.

VI Congresso Londrinense de Biologia Aplicada à Saúde e I Simpósio Paranaense de Patologia Experimental, Londrina, 2004.

49º Congresso Nacional de Genética, Águas de Lindóia, SP, 2003.

III Workshop de Genética, Botucatu, 2003.

I Mostra de Produção Científica em Biologia, Assis, 2003.

Workshop em Ciências Biológicas, Edição 2003: Biodiversidade, Tecnologia & Sociedade, Botucatu, 2003.

I Congresso Paranaense de Biologia Aplicada à Saúde, I Congresso Paranaense de Biologia Aplicada à Saúde, Londrina, 2002.

V Congresso Londrinense de Biologia Aplicada à Saúde, V Congresso Londrinense de Biologia Aplicada à Saúde, Londrina, 2002.

II Encontro Paranaense de Biomedicina, II Encontro Paranaense de Biomedicina, Londrina, 2002.

VII Encontro de Biociências de Assis, VII Encontro de Biociências de Assis, Assis, 2002.

"O sucesso nasce do querer, da determinação e persistência em se chegar a um objetivo. Mesmo não atingindo o alvo, quem busca e vence obstáculos no mínimo fará coisas admiráveis."

José de Alencar

Dedico este trabalho aos meus pais,
Hermes e Maria Laura.

AGRADECIMENTOS

Gostaria de agradecer minha orientadora, Prof^a Dr. Maria Célia Bertolini pela confiança, ensinamentos e pela oportunidade de trabalho.

A minha mãe e ao meu pai pelo carinho e por sempre me apoiarem.

A Tainara pela companhia, apoio e paciência.

Aos amigos do laboratório Luiz, Messias, Amanda, Fernanda Cupertino, Mônica e Fernanda Freitas.

Aos amigos que passaram pelo laboratório Eduardo, Elaine, Juliana, Patrícia, Margareth, Flávia, Gisele.

Aos moradores e ex-moradores da república Curva de Rio Neura, Sinistro, Bixera, Mamute, Tico, Susto, Mahatma, Novelha, Zé Ricardo, Cebinho, Batata, Francês, Babão e em especial a Dona Maria.

Aos técnicos de laboratório Fátima, Waldenir, Fernando, Zilda e especialmente ao Tarcísio.

As funcionárias da Biblioteca e da Seção de Pós-Graduação.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela bolsa concedida.

A Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pelo suporte financeiro concedido ao laboratório.

LISTA DE ABREVIATURAS

AD: Domínio protéico de ativação da transcrição
bHLH: *Basic helix-loop-helix*
BSA: Albumina bovina sérica
B-ZIP: *Basic leucine zipper*
C-terminal: região carboxi-terminal da proteína
cpm: cintilações por minuto
DBD: Domínio protéico de ligação ao DNA
DEPC: dietil pirocarbonato
DNA: ácido desoxiribonucléico
EDTA: ácido etilenodiaminotetracético
FGSC: *Fungal Genetics Stock Center*
GSN: glicogênio sintase do fungo *Neurospora crassa*
GSY2: glicogênio sintase do fungo *Saccharomyces cerevisiae*
GTF: Fatores gerais de transcrição da RNA polimerase
HAT: Histona acetil transferase
HSE: *Heat shock elements*
Id: Domínio protéico Inibidor de ligação ao DNA
kb: kilobase
kDa: kilodalton
MOPS: ácido 3-(N-morfolino) propanosulfônico
NLS: Sinais de localização nuclear
N-terminal: região amino terminal da proteína
ORF: *Open Reading Frame*
pI: Ponto isoelétrico
PMSF: fluoreto de fenilmetilsulfonil
RNA: ácido ribonucléico
RNAm: RNA mensageiro
RNAr: RNA ribossômico
RNAt: RNA transportador
rpm: rotações por minuto

SSC: Solução de citrate de sódio

SDS: Dodecil sulfato de sódio

SSPE: Solução salina fosfato-EDTA

STRE: *Stress responsive elements*

TCA: Ácido tricloroacético

TLCK: N α -p-tosil-L-lisina-clorometilcetona hidrocloreto

Tris: hidroximetilaminometano

V: volts

VM: Meio mínimo de Vogel

SUMÁRIO

Resumo	1
Abstract	3
Introdução.....	5
<i>Neurospora crassa</i> e a era pós genômica.....	6
Regulação gênica nos eucariotos	8
Famílias de fatores de transcrição	12
Sinais de localização nuclear	17
Estrutura, função e aspectos gerais do metabolismo de glicogênio	18
Objetivos.....	21
Materiais e Métodos.....	23
Linhagens utilizadas	24
Meio de cultivo para o crescimento do fungo <i>N. crassa</i>	24
Propagação das linhagens e obtenção dos conídios	25
Choque térmico	25
Preparo do extrato celular e determinação do conteúdo de glicogênio	25
Análise das linhagens mutantes pelo crescimento	26
Extração de RNA total e análise da expressão gênica por <i>Northern blot</i>	27
Análise computacional dos fatores de transcrição selecionados	28
Resultados.....	30
Determinação do conteúdo de glicogênio nas linhagens e seleção das linhagens mutantes.....	31
Caracterização bioquímica, molecular e estrutural dos fatores de transcrição selecionados.....	34
Análise do crescimento linear das linhagens selecionadas	48
Análise da expressão gênica por <i>Northern blot</i>	48
Discussão	53
Referências bibliográficas.....	63

Resumo

Este trabalho teve como objetivo identificar fatores de transcrição envolvidos na via metabólica do glicogênio no fungo *Neurospora crassa* durante o crescimento vegetativo e na condição de choque térmico. Para isso foi utilizado uma coleção de 139 linhagens mutantes do fungo, as quais possuem genes codificadores de fatores de transcrição individualmente nocauteados. O conteúdo de glicogênio de todas as linhagens foi determinado tanto a 30°C, temperatura fisiológica de crescimento, quanto a 45°C, situação de choque térmico. Do total das linhagens analisadas, 10 apresentaram um perfil distinto daquele observado na selvagem. Com o objetivo de verificar o envolvimento dos fatores de transcrição identificados na expressão do gene que codifica a enzima glicogênio sintase (*gsn*), experimentos de *Northern blot* foram realizados e revelaram que do total de linhagens analisadas, 3 apresentaram diferenças na transcrição de *gsn*, sugerindo que os fatores de transcrição nocauteados nestas linhagens atuam diretamente na regulação da transcrição do gene. As linhagens mutantes selecionadas foram avaliadas quanto ao crescimento e comparadas à linhagem selvagem. Todas as linhagens selecionadas apresentaram uma taxa de crescimento bastante semelhante à linhagem selvagem, exceto a linhagem FGSC011062 (ORF NCU09739), a qual apresentou uma taxa de crescimento muito reduzida. Ferramentas de bioinformática disponíveis na internet foram utilizadas para a dedução teórica das características bioquímicas, moleculares e estruturais dos fatores de transcrição selecionados, tais como ponto isoelétrico, massa molecular, sinais de localização nuclear clássicos e a presença de domínios conhecidos. A análise por Blastp revelou que apenas 2 das 10 ORFs selecionadas codificam fatores de transcrição com função conhecida e as demais codificam proteínas hipotéticas. Os fatores de transcrição identificados com função conhecida estão relacionados em diferentes processos biológicos: a ORF NCU08000 codifica o fator de transcrição da cutinase CTF1 α que está envolvido na indução da expressão de genes codificadores de cutinases por ácidos graxos em ascomicetos e a ORF NCU06971 codifica a proteína XlnR, fator de transcrição responsável pela ativação transcricional de genes envolvidos na degradação de xilana. Os dados obtidos sugerem que o metabolismo do glicogênio em *N. crassa* é regulado por um complexo mecanismo, no qual fatores de transcrição devem regular tanto sua síntese quanto sua degradação.

Palavras-chaves: *Neurospora crassa*, fatores de transcrição, metabolismo de glicogênio, gene *gsn*

Abstract

The objective of this work was to identify transcription factors involved in the glycogen metabolic pathway of the fungus *Neurospora crassa*. We used a collection of 139 mutant strains, in which each strain has a gene codifying for transcription factor individually knocked-out. The glycogen content was quantified during the vegetative growth (30°C), and after a stress condition such as heat shock (transfer for 30°C to 45°C). Of all analyzed strains, 10 showed a pattern of glycogen accumulation distinct of the one observed for the wild type strain. To verify the involvement of the transcription factors identified in the expression of the gene encoding glycogen synthase (*gsn*), Northern blot experiments were performed and revealed that 3 strains showed differences in gene transcription, either before or after heat shock. The results suggested that the transcription factors knocked-out in the strains might act directly in the regulation of the *gsn* gene transcription. The growth of the selected mutant strains was evaluated and compared to the wild type strain. Most of the selected mutant had a growth rate similar to the wild type strain, except the strain FGSC01162 (ORF NCU09739), which showed a reduced growth rate. Online bioinformatics tools were used to predict the biochemical parameters, such as isoelectric point and molecular weight, as well protein domain, such as classic nuclear localization signal (cNLS). Analysis by Blastp revealed that 2 from 10 selected ORFs codify transcription factors having functional roles in public databases: the ORF NCU08000 codifies the cutinase transcription factor 1 α that is involved in the induction of the expression of cutinases genes by fatty acids in ascomycetes, and the ORF NCU06971 codifies the protein XlnR a transcription factor responsible for the transcription activation of genes involved in the xylan degradation. The results described in this work show that the glycogen metabolism in *N. crassa* is regulated by a complex mechanism, in which transcription factors regulate either its synthesis or its degradation.

Keywords: *Neurospora crassa*, transcription factors, glycogen metabolism, *gsn* gene.

Introdução

***Neurospora crassa* e a era pós genômica**

Os organismos pertencentes ao reino dos fungos têm exercido um papel fundamental no avanço da compreensão dos mecanismos moleculares de organismos eucariotos devido as suas facilidades de manipulação e o conhecimento das características genéticas e bioquímicas envolvidas em seu ciclo de vida. O fungo *N. crassa*, com um legado de mais de 70 anos de intensa pesquisa, é provavelmente um dos fungos filamentosos melhor conhecidos, sendo amplamente utilizado como organismo modelo em estudos de expressão gênica, desenvolvimento e diferenciação celular, ritmo circadiano, defesa do genoma, bem como outros aspectos da biologia de eucariotos. Pesquisas utilizando *N. crassa* como modelo nas primeiras décadas do século XX determinaram o destino da genética moderna e da biologia molecular. Na década de 40, Beadle e Tatum utilizando *N. crassa* estabeleceram a relação entre genes e proteínas, formulando a hipótese de “um gene, uma enzima”.

N. crassa é um fungo filamentoso pertencente à classe dos *Eumycetes*, subclasse *Ascomyceta*, família *Sordariaceae*, subfamília *Sphaariales*. Ao contrário das leveduras, é um organismo multicelular e possui pelo menos 28 tipos distintos de células, muitos dos quais derivam das hifas, que contribuem para seu complexo ciclo de vida (Bistis *et al.*, 2003, Borkovich *et al.*, 2003). Suas hifas são tubulares, ramificadas e septadas, permitindo a passagem de organelas e núcleos entre os septos. Fusões entre as hifas formam uma rede complexa denominada de micélio que promovem a formação de heterocários, os quais múltiplos genomas contribuem para o metabolismo de um único micélio. É um organismo saprófito e não patogênico embora seu genoma esteja intimamente relacionado com fungos patogênicos. Apresenta tanto o desenvolvimento sexual quanto o assexual em seu ciclo de vida, o qual é altamente influenciado por fatores do meio ambiente, como nutrientes, luz e temperatura.

O genoma de fungo está organizado em sete cromossomos que variam entre 4 a 10 Mb de tamanho (Schutle *et al.*, 2002) e foi recentemente seqüenciado por Galagan *et al.* (2003) revelando ser constituído por cerca de 40 Mb, muito maior se comparado com outros fungos com genomas já conhecidos, como *Sacharomyces cerevisiae* e *Schizosacharomyces pombe*, os quais possuem genomas com aproximadamente 12 Mb. *N. crassa* possui um total de 10.082 genes codificadores de proteínas, que cobrem 44% do genoma. Do total de genes anotados 9.200 codificam proteínas com mais de 100 resíduos,

cerca de duas vezes o total de genes de *S. pombe* (aproximadamente 4.800 genes) e *S. cerevisiae* (aproximadamente 6.300 genes) (Galagan *et al.*, 2003). Um fato que chama a atenção é que 41% dos genes que codificam proteínas não possuem ortólogos conhecidos depositados nos bancos de dados, refletindo o baixo conhecimento dos genomas dos fungos. Além disso, 57% das proteínas de *N. crassa* não possuem relação a proteínas conhecidas de leveduras. Quando comparado com genomas conhecidos de eucariotos, 1.421 genes de *N. crassa* apresentaram as melhores correlações em relação a proteínas de plantas ou animais (Galagan *et al.*, 2003). Isso sugere uma relação mais estreita entre fungos filamentosos e eucariotos superiores, se comparado com leveduras e outros eucariotos inferiores. O fato de *N. crassa* ser um organismo multicelular faz com que uma grande quantidade de seus genes não possua homólogos em *S. cerevisiae*, sugerindo que *N. crassa* pode ser um organismo mais adequado para a compreensão dos aspectos de regulação celular dos eucariotos superiores (Dunlap *et al.*, 2007). A presença de um grande número de genes no genoma de *N. crassa* ainda com funções desconhecidas aponta este organismo como um promissor modelo para o estudo de novos mecanismos genéticos e bioquímicos não identificados até o momento, os quais podem possivelmente representar genes com novas funções ou novos genes desempenhando maneiras diferentes de realizar funções conhecidas (Dunlap *et al.*, 2007).

Em eucariotos existem dois mecanismos de recombinação gênica, a homóloga que envolve a quebra do DNA, seguida da troca entre seqüências homólogas e a recombinação não homóloga, que envolve a ligação direta das extremidades da fita de DNA, independente da homologia. A recombinação homóloga é o mecanismo predominante em *S. cerevisiae*, representando o método mais eficiente pelo qual se pode inativar um gene de interesse. Apesar de ocorrer com muita freqüência em leveduras, o mecanismo de recombinação homóloga ocorre em raríssimas ocasiões tanto em *N. crassa* como nos demais eucariotos, dependendo do gene e do tamanho da região nucleotídica homóloga que flanqueia o gene de interesse (Dunlap *et al.*, 2007). Já o mecanismo de recombinação gênica não homóloga é conservado evolutivamente em todos os reinos, sendo o mecanismo predominante em eucariotos superiores, como humanos, plantas e insetos (Ninomiya *et al.*, 2004). O conjunto de proteínas envolvidas neste mecanismo são as proteínas Ku70 e Ku80, conservadas tanto em leveduras quanto em humanos, que se ligam nas extremidades do DNA e o complexo DNA ligase IV-XRCC4.

Com a identificação de genes homólogos aos genes codificadores das proteínas Ku70 e Ku80 no genoma de *N. crassa*, *mus-51* e *mus-52*, respectivamente, Ninomiya *et al.* (2004) desenvolveram um procedimento, baseado em PCR, para a inativação dos genes homólogos de *N. crassa* e, desta maneira, construir linhagens nas quais o processo de recombinação não homóloga é inoperante. O procedimento utilizado pelos autores está sendo utilizado para a construção de um banco de linhagens de *N. crassa* mutantes nos 10.000 genes codificadores de proteínas. O nocaute de todos os genes de *N. crassa* tem como objetivo determinar a função de cada deles, pois apenas 15% dos genes são associados com o fenótipo e pouco mais de um terço destes genes não possuem ortólogos em outros organismos, não existindo evidências de suas funções (Dunlap *et al.*, 2007). Linhagens contendo genes individualmente inativados começaram a se tornar disponíveis para a comunidade científica pelo *Fungal Genetics Stock Center* (FGSC, Kansas City, Missouri, USA). Em 2006, a coleção de linhagens nocauteadas em genes que codificam fatores de transcrição foi adquirida pelo nosso laboratório e a utilização de tais linhagens tornou-se um material interessante para o início dos estudos acerca dos mecanismos regulatórios envolvidos no metabolismo do glicogênio em *N. crassa*, o qual vem sendo estudado em nosso laboratório.

Assim, o seqüenciamento do genoma de *N. crassa*, aliado aos avanços das metodologias disponíveis envolvidas na inativação de genes específicos deu início às análises da genômica funcional do fungo. A inativação de genes de interesse, juntamente com a análise de sua regulação sob diferentes condições, bem como a análise fenotípica das linhagens mutantes, tem como objetivo expandir o conhecimento do genoma funcional em fungos, permitindo a realização de comparações entre importantes processos biológicos em organismos distintos, com o intuito de revelar os mecanismos bioquímicos envolvidos nos processos de regulação das vias metabólicas dos organismos.

Regulação gênica nos eucariotos

Com o aumento da complexidade dos genomas dos organismos eucarióticos, houve a necessidade de um sistema mais elaborado de controle dos genes. Células eucarióticas possuem três classes de enzimas RNA polimerases, funcionalmente distintas compostas por 12 a 17 cadeias polipeptídicas que transcrevem diferentes classes de genes: a RNA

polimerase I sintetiza RNA ribossomal (rRNA); a RNA polimerase II transcreve os RNAs mensageiros (mRNA) e a RNA polimerase III transcreve o RNA transportador (tRNA) e o rRNA 5S. Para que a transcrição gênica ocorra, as três RNA polimerases eucarióticas requerem uma série de fatores adicionais denominados de fatores gerais de transcrição (GTF) que interagem entre si ajudando a posicionar a RNA polimerase II corretamente no promotor além de separar as fitas de DNA para que o processo de transcrição seja iniciado. Além disso, a RNA polimerase II possui um domínio adicional na extremidade C-terminal em sua subunidade maior que é constituído por uma série de repetições do heptapeptídeo Tyr – Ser – Pro – Thr – Ser – Pro – Ser (Young, 1991) que por meio de sua fosforilação reversível regula a atividade da RNA polimerase II nas distintas etapas do processo transcricional. Quando sua extremidade C-terminal se encontra hipofosforilado, a enzima é capaz de participar do complexo de pré-inicialização da transcrição. Já a sua forma hiperfosforilada confere capacidade para a elongação do transcrito (Riedl & Egly, 2000). O número total das enzimas RNA polimerases bem como sua complexidade e de seus fatores refletem a necessidade de um sistema de controle transcricional mais elaborado nos eucariotos.

Nas células eucarióticas, o genoma é constituído pela molécula de DNA associada a uma grande quantidade de proteínas que interagem de forma dinâmica com a cadeia nucleotídica, sendo responsável pela sua compartimentalização no núcleo bem como pelo controle da expressão gênica. Essas proteínas pertencem a duas classes distintas: histonas e proteínas cromossômicas não-histonas. Histonas são pequenas proteínas básicas que interagem com o DNA sob a forma de octâmeros, os quais são constituídos por dois pares das histonas H2A, H2B, H3 e H4, formando um *core* globular que interage em 14 pontos com a molécula de DNA (Luger *et al.*, 1997). As histonas H1 possuem a função de unir os diversos octâmeros formados. O conjunto dos octâmeros de histonas unidos pelas histonas H1 forma um complexo nucleoproteico denominado de nucleossomos (Kornberg, 1974) e representam um dos mais estáveis complexos DNA-proteína sob condições fisiológicas (Li *et al.*, 2007). Essa estrutura é extremamente conservada nos organismos eucarióticos ao longo da evolução e constitui a unidade básica da cromatina (McGhee & Felsenfeld, 1980).

As histonas estão sujeitas a um grande número de modificações pós-traducionais tanto as caudas N-terminais quanto o *core* globular formado por suas interações. Estas modificações ocorrem pela ação de proteínas que catalisam a adição de grupos funcionais presentes na cauda flexível N-terminal, como por exemplo, acetilação e metilação dos

resíduos de lisinas e fosforilação dos resíduos de serina e treonina. Sabe-se que as histonas são acetiladas em determinados resíduos de lisina, e que esta acetilação está relacionada com a expressão e silenciamento gênico. Alterações nas estruturas das histonas provocam alterações na estrutura da cromatina e, conseqüentemente, alteram o padrão de regulação gênica. Esse mecanismo recebe o nome de modelagem da cromatina e é amplamente usado para descrever alterações na estrutura durante os processos regulatórios. Genes, cujas regiões promotoras estão presentes em nucleossomos podem ter suas expressões inibidas caso os nucleossomos não estejam relaxados (Almer *et al.*, 1986; Wasylyk & Chambon, 1979). No entanto, estudos envolvendo o mapeamento de genomas revelaram que a densidade de nucleossomos nas regiões promotoras é menor se comparado com as regiões codificadoras (Bernstein *et al.*, 2004; Lee *et al.*, 2004; Sekinger *et al.*, 2005). Estes dados sugerem que o posicionamento dos nucleossomos no genoma está intimamente relacionado com a disposição das regiões codificadoras do genoma.

Outro mecanismo responsável pela regulação gênica nos eucariotos é a presença de determinadas seqüências nucleotídicas que atuam como sítios alvos para fatores de transcrição. Essas seqüências recebem o nome de elementos *cis*. Diferentemente dos organismos procarióticos que possuem apenas as regiões -35 e -10 na unidade transcricional, os eucariotos apresentam em seus genomas inúmeros elementos *cis*, que atuam como sítios alvos para a interação de diferentes proteínas regulatórias. Esses elementos *cis* estão localizados tanto *upstream* quanto *downstream* da unidade transcricional e são responsáveis tanto pela ativação quanto pela inibição da expressão gênica. Alguns elementos *cis* podem estar localizados a centenas de bases do sítio de início da transcrição, como por exemplo, os elementos *enhancers*. Proporcionalmente ao aumento da complexidade dos elementos regulatórios *cis*, 5-10% da capacidade total de codificação dos metazoários são dedicados a proteínas que regulam a transcrição (Levine & Tjian, 2003). Essas proteínas estão subdivididas em duas maiores classes: as proteínas denominadas fatores de transcrição que ligam a seqüências nucleotídicas específicas por meio de domínios de ligação ao DNA e medeiam a ativação ou repressão transcricional seletiva dos genes, e uma grande quantidade de proteínas gerais que constituem a maquinaria multiproteica da RNA polimerase II, requeridas para o reconhecimento do promotor e catálise da síntese de RNA (Levine & Tjian, 2003).

A transcrição dos genes estruturais das células eucarióticas é regulada por fatores de transcrição que interagem com os elementos *cis* presentes nas regiões promotoras por meio de um domínio de ligação específico ao DNA (DBD). Estudos envolvendo fatores de transcrição têm mostrado que essas proteínas regulatórias possuem diferentes domínios relacionados a importantes funções do processo transcricional (Latchman, 2003). Além do domínio de ligação ao DNA, os fatores de transcrição precisam também apresentar um domínio de ativação (AD), requerido para estimular a transcrição. É importante ressaltar que alguns fatores de transcrição não possuem o AD. Neste caso, o DBD e o AD estão localizados em diferentes proteínas no complexo de ligação ao DNA, sendo necessária a interação das mesmas para que possa ocorrer a transcrição (Latchman, 2003).

Para que o controle do processo transcricional seja eficiente, é necessário que ocorra uma ação combinatória de vários ativadores ou inibidores que se ligam aos elementos *cis* presente no DNA. Como consequência, um grande número de complexos regulatórios pode ser arranjado a partir de uma grande quantidade de combinações de fatores. Esta idéia ganhou força com o descobrimento de complexos multiproteicos que pode alterar a estrutura da cromatina, como por exemplo, os reguladores transcricionais globais em *S. cerevisiae*, tais como o complexo SAGA, com função da enzima histona acetiltransferase (HAT) que transfere o grupo acetil da acetil coenzima A (acetil-CoA) para os resíduos de lisina presente nas caudas das histonas H3 e/ou H4, alterando a carga deste resíduo, fazendo com que diminua sua interação ao DNA, e o complexo Swi/Snf, com capacidade de desestabilizar tanto os dímeros H2A/H2B quanto os tetrâmeros H3/H4 de histonas liberando os elementos *cis* bloqueados pelos nucleossomos (Wolffe & Hayes, 1999). É importante ressaltar que todos os reguladores de cromatina precisam se ligar a regiões específicas da cromatina para poderem efetuar suas respectivas funções. Entretanto, esses reguladores geralmente não possuem sítios de ligação no DNA e interagem com determinadas proteínas já ligadas ao DNA para poderem desempenhar suas funções.

Em suma, para que ocorra o controle da expressão gênica, é necessário que os mecanismos regulatórios estimulem ou inibam a transcrição por meio do recrutamento da maquinaria protéica requerida pela RNA polimerase II e proteínas modificadoras de cromatina para desbloquear o promotor dos genes alvos. Uma consequência da modelagem da cromatina é a exposição ou inibição dos elementos *cis* aos fatores de transcrição, que

podem estar na superfície do nucleossomo ou no interior dos mesmos, fazendo com que genes específicos sejam ativados ou reprimidos no momento específico.

Famílias de fatores de transcrição

Para que o processo de transcrição ocorra, é necessário que proteínas regulatórias possuam domínios de ligação aos elementos *cis*. Essas proteínas são agrupadas em família de fatores de transcrição de acordo com as características estruturais desses domínios. O completo seqüenciamento do genoma de *N. crassa* permitiu a anotação de 182 fatores de transcrição. A maioria desses fatores pertence à família *zinc finger*, que corresponde aos domínios de ligação Zn_2Cys_6 , C_2H_2 e GATA, mas também foram identificados no genoma fatores pertencentes às famílias bHLH, B-ZIP, entre outras (Borkivich et al., 2003). Somente 21 genes codificadores de fatores de transcrição foram caracterizados em *N. crassa* até o momento. A maioria desses fatores conhecidos desempenha funções relacionadas ao controle do metabolismo durante o desenvolvimento, como por exemplo, os genes *acu-15*, *cpc-1*, *nit-2*, *nit-4*, *pco-1*, *nuc-1*, *cys-3*, *qa-1F* e *sre*, envolvidos na utilização de nutrientes, os genes *fl*, *rca-1*, e *rco-1*, envolvidos com o controle do crescimento e macroconididação, os genes *wc-1* e *wc-2*, relacionados com resposta a luz e ritmo circadiano, os genes *asd-4*, *asm-1*, e *pp-1*, envolvidos com o desenvolvimento sexual, o gene *uvs-2*, relacionado ao reparo de DNA, e os genes *mat A-1*, *mat A-3*, e *mat a-1*, relacionados com o *mating type* (Colot et al., 2006).

Os fatores pertencentes à família *zinc finger* constituem uma das maiores famílias de fatores de transcrição em eucariotos, possuindo uma grande quantidade de estruturas secundárias distintas, sendo responsáveis pela regulação de distintos processos metabólicos. A particularidade desta família é a presença de um ou mais átomos de zinco necessários para a formação da estrutura de seu sítio ativo denominado *zinc finger* (Branden & Tooze, 1999). As cadeias polipeptídicas do domínio *zinc finger* são geralmente pequenas, com um padrão regular de resíduos de cisteína e/ou histidina ao longo de sua cadeia. Esses resíduos são responsáveis pela interação da seqüência polipeptídica com um ou mais átomos de zinco, proporcionando a formação de uma conformação independente do domínio protéico. Atualmente sabe-se que as proteínas dessa família não se ligam apenas aos sítios presentes nas moléculas de ácidos nucleicos, podendo também atuar em

processos de diversas naturezas incluindo mediação da interação proteína-proteína, remodelagem da cromatina, exercer função de proteínas chaperoninas, ligação em lipídeos e sensoriamento de zinco (Laity *et al.*, 2001).

Em eucariotos, as principais classes de proteínas *zinc fingers* responsáveis pela interação ao DNA são representadas pela classe I, que corresponde ao domínio C₂H₂, pela classe II, que corresponde ao domínio GATA e pela classe III que corresponde ao domínio Zn₂Cys₆. O clássico domínio *zinc finger*, também denominado de C₂H₂, é constituído por uma unidade de conformação independente descrita como uma proteína miniglobular, com um *core* hidrofóbico e uma cadeia lateral polar em sua superfície (figura 1a) (Branden & Tooze, 1999). Os resíduos de 1 a 10 formam duas folhas β anti-paralelas seguidas de um *hairpin*, com o zinco ligando a Cys3 na primeira folha β e a Cys6, presente na pequena volta que une as duas folhas betas. Essa estrutura é seguida por uma α -hélice com 3 voltas e meia (resíduos 12 – 24). As demais ligações com o zinco ocorrem na His19 e His23 na extremidade C-terminal. Essa hélice é denominada hélice 3₁₀, onde pontes de hidrogênio ocorrem a cada 3 resíduos. Foram descritas diversas proteínas possuindo mais de um domínio *zinc finger*, como por exemplo, a proteína de rato Zif268 que é constituída por 90 resíduos, distribuídos em 3 domínios *zinc finger* unidos por uma região *linker* de 4 resíduos (Pavletich & Pabo, 1991). Essas repetições dos domínios *zinc fingers* assumem um arranjo modular que oferece um grande número de possibilidades de combinações para reconhecimento específico de DNA, além de permitir um grande aumento na especificidade de reconhecimento dos sítios de ligação ao DNA (Branden & Tooze, 1999). Essa família representa a segunda maior classe de fatores de transcrição existente no genoma de *N. crassa*. Até o momento 43 fatores de transcrição foram identificados. Podemos citar como exemplo o gene *steA*, envolvido na formação de corpos de frutificação. Linhagens mutantes que apresentam esse gene inativado são estéreis, mas não apresentaram alterações no padrão de seu crescimento vegetativo, bem como em suas hifas (Vallim *et al.*, 2000).

A classe representada pelos fatores GATA, também denominados de C₄, recebem esta denominação devido à sua capacidade de se ligar aos sítios (A/T)GATA(A/T) presente nas regiões regulatórias do DNA. Esses fatores de transcrição foram inicialmente identificados a partir de estudos envolvendo a região promotora dos

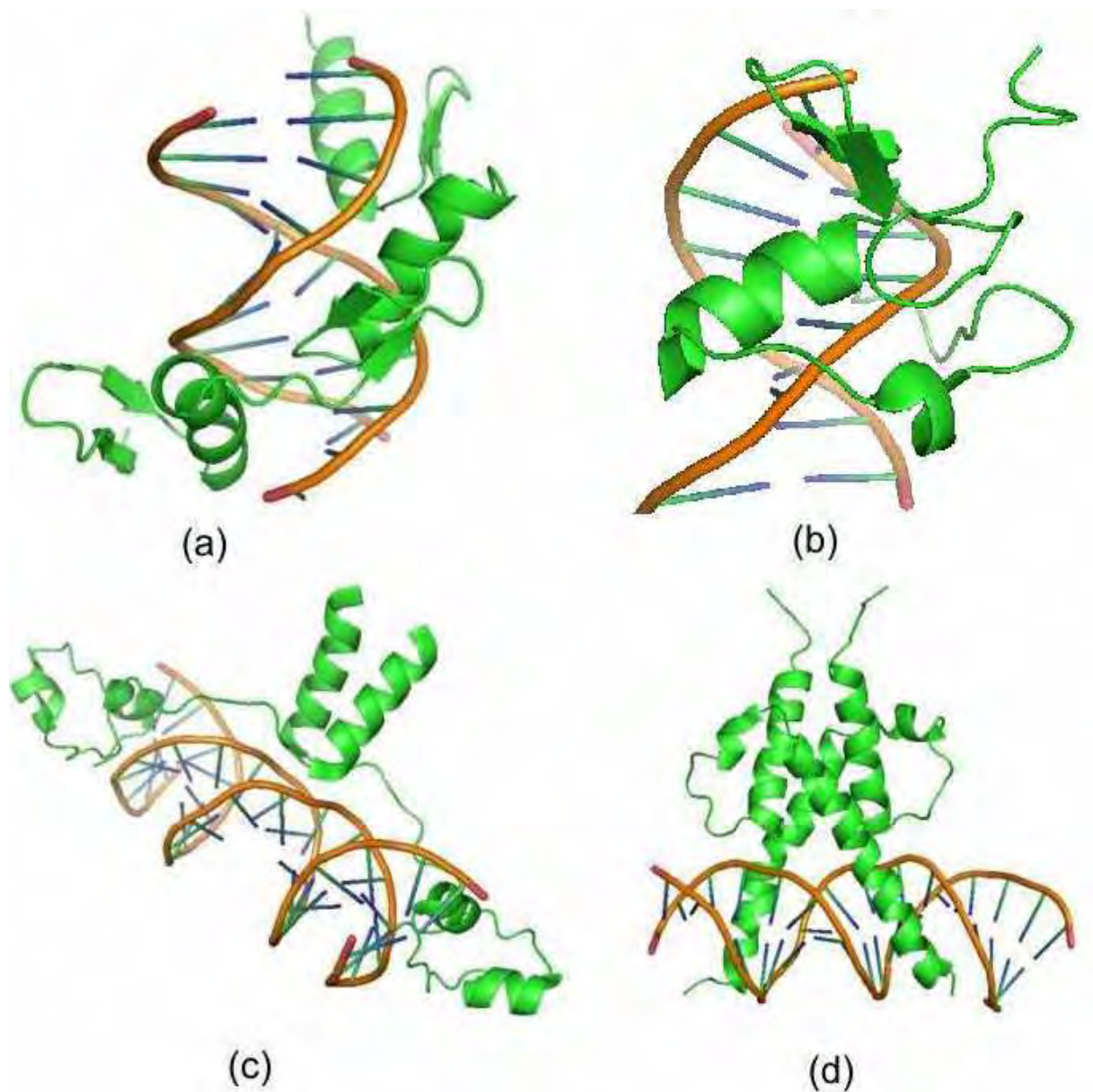


Figura 1 - Representação tridimensional dos domínios de ligação presentes nos fatores de transcrição selecionados. (a) proteína ZIF 268 formando um trímero que se liga ao DNA por meio de seus domínios C2H2 (PDB 1A1L); (b) fator de transcrição GATA-1 interagindo com o DNA (PDB 1GAT); (c) proteína GAL4 ligada ao DNA por meio de seus domínios Zn₂Cys₆ (PDB 1D66); (d) proteína PHO4 ligada ao DNA por seu domínio bHLH (PDB 1A0A); (<http://www.rcsb.org/pdb/home/home.do>).

genes da α e β globina de humanos, sendo posteriormente identificados em muitos outros organismos, como por exemplo fungos, *Caenorhabditis elegans*, *Drosophila melanogaster*, aves e mamíferos. Vertebrados possuem dois domínios *zinc fingers* com a configuração Cys-X17-Cys-X2-Cys (figura 1b). Já os invertebrados podem apresentar dois domínios de ligação ao DNA, como ocorre, por exemplo, em *C. elegans*, ou apenas um domínio de ligação, como observado em fungos. É importante ressaltar que o domínio C-terminal é requerido para a ligação do fator de transcrição ao DNA, enquanto o domínio N-terminal é responsável pelo aumento da estabilidade e especificidade da ligação com o DNA (Merika & Orkin, 1993). Podemos citar como exemplo em *N. crassa*, o fator de transcrição NIT-2, que juntamente com fator NIT-4, regulam a expressão do gene *nit-3*, o qual codifica a enzima nitrato redutase, envolvida no metabolismo de nitrogênio no fungo.

A família Zn_2Cys_6 , também conhecida como *binuclear cluster* é restrita aos fungos, sendo predominante nos organismos pertencentes à família dos ascomicetos (Todd & Adrianopoulos, 1997). O domínio *zinc finger* Zn_2Cys_6 consiste de um par de α -hélices que coordena dois átomos de zinco através de 6 resíduos de cisteína, uma α -hélice se liga ao DNA pelo sulco maior e a segunda interage com o esqueleto de fosfato (figura 1c) (Luscombe *et al.*, 2000). Seus sítios de ligação ao DNA consistem em uma trinca de nucleotídeos conservada e intercalada por uma sequência variável (CGGN_xCCG). A diferença entre as proteínas desta família é o espaçamento entre os dois domínios (clássico e Zn_2Cys_6) que se ligam ao DNA, ou seja, o que determina a especificidade de ligação de uma proteína da família Zn_2Cys_6 com seu sítio alvo no DNA é o espaçamento entre os nucleotídeos (Vashee *et al.*, 1993). Um bom representante desta família de fator de transcrição é a proteína GAL4 de *S. cerevisiae* que se liga na forma de dímero ao DNA e ativa a transcrição de genes requeridos para a quebra de galactose e melibiose (Johnston, 1987). Em seu domínio *zinc finger* dois íons zinco estão ligados a seis resíduos de cisteínas, sendo que dois resíduos estão ligados aos dois íons. O sítio de ligação ao DNA desta proteína é CCG distanciados por 11 nucleotídeos (Johnston, 1987). Os fatores de transcrição Zn_2Cys_6 representam a maior família de fatores de transcrição em *N. crassa* onde foram identificadas 77 ORFs que codificam proteínas contendo esse domínio. Destas, 20 não possuem homologia com proteínas de leveduras e 9 mostram homologia apenas com proteínas de outros fungos filamentosos (Borkovich *et al.*, 2004). Podemos citar como

exemplo em *N. crassa*, o fator de transcrição NIT-4, que como já citado, está envolvido com a regulação de genes envolvidos na assimilação de nitrogênio.

Os fatores de transcrição pertencentes à família bHLH são exclusivos de eucariotos, são altamente conservados em diversos organismos, desde leveduras até mamíferos, estão envolvidos no controle de diversas vias metabólicas e estão divididas em diferentes classes, de acordo com suas funções (Murre *et al.*, 1994). Essas proteínas regulatórias são constituídas de duas α -hélices anfipáticas, separadas por uma região espaçadora em forma de alça, que pode variar entre 5 a 23 resíduos (figura 1d) (Branden & Tooze, 1999).

Além disso, os fatores de transcrição pertencentes a esta família possuem duas regiões altamente conservadas: uma que consiste em uma região com resíduos predominantemente básicos, responsáveis pela interação dessas proteínas ao DNA, e outra formada por resíduos hidrofóbicos, responsáveis pela dimerização dos monômeros. Os dímeros são estabilizados por uma série de interações hidrofóbicas e eletrostáticas entre as α -hélices, devido à formação de um feixe de 4 hélices que mantém os monômeros unidos, constituindo uma interface de ligação ao DNA. Essa estrutura possui uma geometria precisa, que interage via o sulco maior do DNA em uma sequência consenso (CANNTG) denominada de *E BOX* (Branden & Tooze, 1999). É importante ressaltar que as proteínas que apresentam domínios de ligação bHLH não possuem a capacidade de diferenciar os sítios de ligação ao DNA sob a forma de monômeros. Assim, o mecanismo de dimerização tem a função de converter moléculas monoméricas inativas em complexos diméricos com atividade transcricional em períodos determinados durante o ciclo celular.

Alguns membros dessa família de fatores de transcrição não possuem a região básica, ficando impossibilitados de se ligarem ao DNA. Esses fatores são denominados inibidores de ligação ao DNA (Id). Os fatores Id possuem ação dominante como regulador negativo de diferentes vias metabólicas e atuam seqüestrando monômeros, inibindo a formação de dímeros ativos. Dessa forma, os fatores Id podem formar heterodímeros com outros membros da família que possuem a região básica, exercendo uma função inibitória, ou seja, inibe a interação do heterodímero ao DNA, atuando então como reguladores negativos da expressão gênica (Latchman, 2003). Em *N. crassa*, 10 proteínas foram identificadas nesta família, das quais somente a codificada pelo gene *nuc-1* foi caracterizada, estando envolvida com a regulação da utilização de fosfato (Kang & Metzenberg, 1990).

Sinais de localização nuclear

As células eucarióticas possuem seu material genético e sua maquinaria transcricional localizados no núcleo com a finalidade de aperfeiçoar seus processos biológicos. Fatores de transcrição são proteínas traduzidas no citoplasma que desempenham suas funções regulatórias no interior do envelope nuclear. Para que estas proteínas possam desempenhar suas respectivas funções, em momentos específicos, é necessário que elas atravessem a barreira física representada pelo núcleo. Conseqüentemente, é necessário que ocorra o tráfego entre moléculas do citoplasma e do núcleo. Este fluxo ocorre por meio de estruturas localizada na membrana nuclear denominada de complexo nuclear do poro (NPC) constituído por um complexo protéico, com cerca de 125 MDa em células de mamíferos e 60 MDa em células de leveduras (Chook & Blobel, 2001). Estas estruturas permitem o transporte passivo de íons, pequenas moléculas e pequenas proteínas, restringindo grandes moléculas e proteínas com massa molecular superior a 40 kDa. Proteínas com massa molecular acima de 40 kDa necessitam do auxílio de proteínas carreadoras, também conhecidas como importinas para translocar para o núcleo, via o complexo do poro.

As importinas são grupos de proteínas com capacidade de realizar o transporte de proteínas do citosol para o núcleo devido à presença de sinais de localização nuclear (NLS) na seqüência primária das proteínas alvo. Esses sinais correspondem a pequenas seqüências de aminoácidos que são requeridas para a ligação da proteína alvo às importinas. Os sinais de localização nuclear melhores conhecidos são os sinais de localização nuclear clássicos (cNLS), que podem ser de duas formas: monopartido ou bipartido. O sinal monopartido pode apresentar uma seqüência de 4 a 7 resíduos de aminoácidos básicos (lisina e arginina) seguidos ou não por um resíduo de histidina ou uma prolina (Hicks & Rainkhel, 1995). Os sinais bipartidos, descobertos a partir de estudos com *Xenopus* (Robbins *et al.*, 1991) possuem duas regiões básicas separadas por aproximadamente 10 resíduos de aminoácidos. A primeira região possui dois resíduos básicos e a segunda é composta por pelo menos 3 resíduos de aminoácidos básicos (Robbins *et al.*, 1991). De forma resumida, o mecanismo de ação das importinas em relação aos cNLS inicia no citoplasma, pela formação de um complexo importina α/β , unidas pela cauda N-terminal básica, autoinibitória da importina α . Deste modo, ocorre a formação do

complexo importina α/β , permitindo a interação dos cNLS presentes na proteína alvo com o complexo formado via domínios específicos existentes na importina α . Ao penetrarem no núcleo, via NPC, uma GTPase denominada de RanGTP se liga a importina β , provocando alterações conformacionais no complexo protéico. Estas alterações fazem com que o complexo importina α/β seja desfeito. Desta forma, a cauda autoinibitória da importina α , que estava unida à importina β fica livre, iniciando uma competição com os cNLS da proteína alvo, o que contribui para uma baixa afinidade da importina α pelos cNLS no interior do núcleo, facilitando a liberação da proteína alvo no interior do núcleo. É importante ressaltar que os cNLS não são os únicos sinais pelos quais as proteínas carreadoras reconhecem e translocam proteínas do citoplasma para o núcleo. Recentemente uma série de NLS não clássicos foram descobertos (Pemberton, *et al.*, 1998; Lee *et al.*, 2006; Fries *et al.*, 2007), onde as proteínas alvos interagem diretamente a importinas β específicas, ao contrário do mecanismo que ocorre em relação aos cNLS. Foram identificadas cerca de 14 importinas β distintas em leveduras e mais de 20 em células de mamíferos (Chook & Blobel, 2001).

Estrutura, função e aspectos gerais do metabolismo de glicogênio

O glicogênio é um polímero ramificado de glicose encontrado em células de diferentes organismos com a função de armazenamento de energia para as células. Seus resíduos de glicose estão unidos covalentemente por ligações glicosídicas lineares α -1,4 e ligações α -1,6, que ocorre a cada 4 resíduo de glicose, responsáveis pela ramificação da molécula. Sua síntese é bem conservada nos eucariotos, envolvendo as mesmas enzimas: a glicogenina, enzima auto-glicosilável responsável pelo início da biossíntese do glicogênio; a glicogênio sintase, que catalisa a formação das ligações glicosídias α -1,4 pela transferência de glicose a partir de UDP-glicose à extremidade não redutora de uma cadeia polissacarídica em crescimento e a enzima ramificadora, a enzima glicosil-(4–6)-transferase, que catalisa a formação das ligações ramificadas α -1,6 da molécula de glicogênio.

Tanto microrganismos quanto mamíferos sintetizam e acumulam o glicogênio em períodos de abundância de nutrientes e o degrada quando as células são submetidas a certos tipos de estresse, como por exemplo, escassez de nutrientes (Harris, 1997). Assim, a variação do conteúdo de glicogênio celular em resposta a diversas situações ambientais indica que seu metabolismo é controlado por um complexo sistema regulatório.

Diversos trabalhos relatam importantes variações no conteúdo de glicogênio em *S. cerevisiae*. Foi observado que durante a situação de crescimento vegetativo, as células de leveduras acumulam glicogênio quando a cultura inicia sua fase de crescimento estacionário (Thevelein, 1984). Outro exemplo de situações onde células de *S. cerevisiae* acumulam glicogênio inclui a exposição da cultura ao choque térmico (Ni & Laporte, 1995). Sob condições de choque térmico a transcrição do gene que codifica uma das enzimas glicogênio sintase (*GSY2*) é aumentada, com conseqüente acúmulo de glicogênio. Foi mostrado que o aumento da expressão do gene *GSY2* durante o choque térmico é decorrente da presença de dois elementos regulatórios de DNA, os elementos STRE (*Stress Responsive Elements*) na região promotora do gene (Ni & LaPorte, 1995). Os STREs são elementos de transcrição *cis* presentes nos promotores de genes responsivos a várias condições estressantes (Martinez-Pastor *et al.*, 1996). A maioria dos genes codificadores de enzimas envolvidas no metabolismo de glicogênio em leveduras contém entre uma e várias cópias de STRE em seus promotores (Varela *et al.*, 1995).

Em *N. crassa*, estudos realizados em nosso laboratório mostraram a ocorrência de acúmulo de glicogênio durante a fase exponencial de crescimento vegetativo, o qual foi acompanhado por um aumento tanto da atividade glicogênio sintase quanto nos níveis do transcrito *gsn* desta enzima e que o glicogênio é degradado quando a cultura atinge a fase estacionária de crescimento (Noventa-Jordão *et al.*, 1996; de Paula *et al.*, 2002). Sob condição de choque térmico, ao contrario do que ocorre em *S. cerevisiae*, o nível do transcrito *gsn* sofre uma acentuada redução e retorna ao normal após a reincubação da cultura em sua temperatura fisiológica de crescimento (de Paula *et al.*, 2002). Estes resultados mostraram que a regulação do metabolismo de glicogênio em *N. crassa* envolve mecanismos diferentes dos descritos na levedura *S. cerevisiae* e se aproxima melhor dos mecanismos de regulação de células de mamíferos.

Com o objetivo de estudar a regulação da expressão do gene *gsn* estudos de caracterização molecular da região promotora do gene *gsn* de *N. crassa* foram realizados em nosso laboratório (Freitas & Bertolini, 2004). Uma análise detalhada da região 5'-flanqueadora do gene permitiu a identificação de vários prováveis elementos transcripcionais, dentre eles grande quantidade de HSEs e dois elementos STREs, um deles *upstream* a um TATA Box canônico (STRE1) e o outro (STRE2), localizado dentro de um *intron* presente na região 5'-UTR do gene *gsn*. Também foram identificados elementos STREs na região 3' do

gene. Os HSEs (*Heat-Shock Elements*) são elementos de transcrição encontrados em múltiplas cópias, nos promotores de genes responsivos ao choque térmico (Bienz & Pelham, 1987) e conferem uma forte indução frente ao estresse térmico, podendo agir tanto como elementos proximais ao TATA Box quanto como *enhancers* (Pelham & Bienz, 1982; Bienz & Pelham, 1986). Através de ensaios de retardamento em gel (*Electrophoretic Mobility Shift Assay*, EMSA) com duas regiões do promotor *gsn*, uma contendo elementos HSE e outra contendo a região correspondente ao STRE1, bandas de DNA de mobilidade reduzida foram observadas com extratos nucleares de células do fungo submetidas ao choque térmico. Estes resultados mostraram que proteínas ativadas nesta situação ambiental foram capazes de ligar aos fragmentos de DNA do promotor de *gsn* e demonstraram fortemente o envolvimento das regiões estudadas na regulação da expressão do gene durante a condição de estresse mencionada. Além disso, também em nosso laboratório, experimentos foram iniciados com o objetivo de isolar proteínas (fatores de transcrição ou não) que ligam à região promotora do gene *gsn* e, desta maneira, participam na regulação da expressão do gene. Esses estudos combinam experimentos de ensaios de retardamento em gel, acoplados a análise proteômica e identificação das proteínas por espectrometria de massas. Os experimentos realizados até o momento levaram à identificação de algumas proteínas do fungo, as quais podem ou não estar envolvidas na regulação do gene na condição de choque térmico (Freitas, 2007; Freitas *et al.*, 2008).

A utilização de um banco de linhagens de *N. crassa* mutantes nas ORFs codificadoras de fatores de transcrição, objetivo do presente trabalho, juntamente com as proteínas já identificadas, poderá contribuir enormemente para o conhecimento das proteínas envolvidas no controle transcricional do gene *gsn* em *N. crassa*. Isto poderá permitir a identificação de proteínas com funções ainda não descritas.

Objetivos

Este trabalho tem como objetivo principal identificar fatores de transcrição que participam na regulação do metabolismo do glicogênio no fungo *N. crassa*. Para isto, um banco de linhagens do fungo contendo ORFs codificadoras de fatores de transcrição individualmente nocauteadas foi utilizado. Especificamente, o objetivo é a busca de novos fatores de transcrição envolvidos na regulação da transcrição do gene *gsn* no fungo, ainda não caracterizados em outros organismos eucariotos.

Materiais e Métodos

Linhagens utilizadas

Linhagem selvagem: FGSC 9718 (a, $\Delta mus-51:: Hyg^r$).

Linhagens mutantes: Uma coleção de 139 linhagens mutantes de *N. crassa*, contendo genes que codificam fatores de transcrição individualmente nocauteados foi adquirida pelo laboratório. Das 139 linhagens, 40 correspondem a linhagens nocauteadas no mesmo gene, mas com fator de acasalamento distinto (A ou a), portanto 99 fatores de transcrição foram analisados neste trabalho. Todas as linhagens utilizadas foram adquiridas junto ao *Fungal Genetics Stock Center*, Kansas City, MO, USA.

Meio de cultivo para o crescimento do fungo *N. crassa*

O meio de cultivo utilizado para o crescimento e manutenção das linhagens do fungo *N. crassa* foi o meio mínimo de Vogel (VM) (Vogel, 1956). Esse meio foi preparado contendo sacarose, sais de Vogel e biotina. A solução de sais de Vogel foi preparada 50X concentrada, a partir da dissolução dos seguintes sais, na ordem indicada, em 700 ml de água destilada:

510 g de citrato de sódio.H₂O

250 g de KH₂PO₄

100 g de NH₄NO₃

10 g de MgSO₄.7H₂O

5 g de CaCl₂.2H₂O (pré-dissolvido em água destilada)

5 mL de solução de elementos traços

O volume foi completado para 1 L com água destilada e 300 µl de clorofórmio foi adicionado como agente conservante. A solução estoque de sais de Vogel foi armazenada a 4°C. A solução de elementos traços foi feita pela dissolução dos seguintes compostos em 90 mL de água:

5 g de MgSO₄.7H₂O

5 g de ácido cítrico.H₂O

5 g de ZnSO₄.7H₂O

1 g de F(NH₄)₂.(SO₄)₂.6H₂O

0,25 g de CuS₄.5H₂O

0,05 g de $\text{MnSO}_4 \cdot \text{H}_2\text{O}$

0,05 g de H_3BO_3

0,05 g de $\text{NaMoO}_4 \cdot 2\text{H}_2\text{O}$

O meio mínimo de Vogel foi preparado contendo 2% de sacarose, sais de Vogel 1X, biotina 0,1 mg/ml preparada em 50% etanol e 2% de ágar para culturas de meio sólido.

Propagação das linhagens e obtenção dos conídios

Para a obtenção dos micélios utilizados nos experimentos de choque térmico, o inóculo foi feito em garrafas de vidro contendo 50 ml de meio sólido e deixado a 30°C na estufa por 48 horas e, em seguida, expostos à luz até um total de 9 a 10 dias. Após esse período os conídios foram coletados pela adição de aproximadamente 10 ml de água destilada estéril à garrafa seguida da filtração em gaze estéril para retenção das hifas e os conídios foram contados em câmara de Neubauer.

Choque térmico

Uma suspensão de conídios (densidade de 10^7 células/ml) foi inoculada em 30 ml de meio VM (Erlenmeyers de 125 ml), em duplicata, e colocada a 30°C sob agitação constante de 250 rpm por 24 horas em um agitador New Brunswick. Após esse período, o micélio correspondente a um dos frascos de cada linhagem foi coletado por filtração a vácuo em papel filtro, rapidamente congelado em nitrogênio líquido e armazenado a -80°C (amostra controle, não submetida a choque térmico). O micélio da amostra do outro frasco também foi coletado por filtração a vácuo, e transferido para um terceiro frasco contendo VM líquido, pré-aquecido a 45°C, e deixado nesta temperatura por 30 minutos. Em seguida o micélio foi coletado por filtração, rapidamente congelado em nitrogênio líquido e armazenado a -80°C (amostra submetida a choque térmico).

Preparo do extrato celular e determinação do conteúdo de glicogênio

Para o preparo do extrato celular, aproximadamente 200 mg do micélio congelado a -80°C foram macerados em nitrogênio líquido até lise total. Em seguida, transferiu-se esse

material para um tubo eppendorf de 1,5 ml contendo 1 ml de solução tampão gelada (Tris-HCl 50 mM, pH 8,0, NaF 50 mM, EDTA 1 mM, PMSF 1 mM, benzamidina 1 mM, TLCK 2 mM, aprotinina 1 µg/ml e pepstatina A 1 µg/ml). A mistura foi fortemente agitada em *vortex*, e em seguida, foi centrifugada a 10.000 x *g* por 10 minutos em centrífuga Eppendorf de bancada, recolhendo o sobrenadante em um tubo eppendorf novo. Uma alíquota foi utilizada para a determinação das proteínas totais do extrato celular pelo método de Hartree modificado utilizando soroalbumina bovina (BSA) como padrão (Hartree, 1972).

O conteúdo de glicogênio foi determinado pela digestão do glicogênio com amiloglicosidase após precipitação com etanol absoluto como descrito por Hardy & Roach (1993), com algumas modificações. Primeiramente, foram coletados 100 µl do extrato celular bruto e adicionado 25 µl de TCA 50%. Em seguida, o material foi centrifugado a 5000 x *g* por 10 minutos a 4°C. O sobrenadante foi transferido para um tubo eppendorf limpo, sendo adicionado 500 µl de etanol absoluto e deixado 30 minutos a -80°C para a precipitação do glicogênio presente na fração solúvel. Finalmente, a solução foi centrifugada a 17000 x *g* por 10 minutos para a coleta do glicogênio, o qual foi lavado duas vezes com 500 µl de etanol 66%, levemente seco ao ar, e ressuspenso em 400 µl de solução acetato (acetato de sódio 50 mM, CaCl₂ 5 mM, pH 5,2). O glicogênio foi digerido com as enzimas α-amilase (1 µl de uma solução 10 mg/ml) e amiloglicosidase (1 µl de uma solução de 30 mg/ml) durante 12 a 15 horas. O conteúdo de glicose livre foi determinado pelo kit *Glicose PAP Liquiform* (Labtest) segundo as instruções do fabricante. O conteúdo de glicogênio presente em cada amostra foi calculado de acordo com uma curva padrão de glicogênio e expresso em relação à quantidade de proteína total de cada amostra.

A curva padrão de glicogênio foi preparada a partir de uma solução de glicogênio puro (Sigma) em solução acetato (acetato de sódio 50 mM, CaCl₂ 5 mM, pH 5,2) na concentração de 10 mg/ml.

Análise das linhagens mutantes pelo crescimento

As linhagens selecionadas foram avaliadas pela taxa de crescimento linear em meio sólido utilizando *race tubes*, os quais são longos tubos de vidro, com aproximadamente 50 cm, contendo suas extremidades inclinadas (ângulo de aproximadamente 30°). As linhagens foram rejuvenescidas e hifas (7 a 10 dias de crescimento) foram inoculadas em uma das

extremidades dos tubos. O avanço das hifas foi acompanhado a cada 24 horas até a linhagem selvagem atingir o final do tubo. A taxa de crescimento foi expressa em mm/dia.

Extração de RNA total e análise da expressão gênica por *Northern blot*

O RNA total das linhagens selecionadas foi extraído pelo método de cloreto do lítio (Sokolovsky *et al.*, 1995). Para isto, uma pequena quantidade do micélio congelado foi macerada com nitrogênio líquido e o pó resultante foi transferido para tubos eppendorf contendo uma mistura de 750 µL de tampão de lise (NaCl 0,6M, EDTA 10 mM, Tris-HCl 0,1M, pH 8,0 e SDS 4%) e 750 µL de fenol saturado em Tris-HCl 0,1 M, pH 8,0. A suspensão foi agitada durante 15 minutos à temperatura ambiente e a fase aquosa foi coletada a 10.000 rpm por 10 minutos a 4°C em centrífuga refrigerada de bancada. O RNA contido na fase aquosa foi precipitado com 0,75 volumes de uma solução de LiCl 8 M e mantido a 4°C durante uma noite. Após esta etapa, o material foi centrifugado a 13.000 rpm por 10 minutos a 4°C, o precipitado contendo os ácidos nucleicos foi ressuspenso em 300 µL de água e novamente precipitado com acetato de sódio 3 M e etanol absoluto gelado a - 80°C, durante 30 minutos. O RNA total foi coletado a 13.000 rpm, por 15 minutos a 4°C, lavado duas vezes com uma solução de etanol 70% gelada a 13.000 rpm por 5 minutos a 4°C, seco e solubilizado em 50 a 100 µL de água previamente tratada com DEPC. A concentração do RNA foi determinada por absorbância a 260 nm em espectrofotômetro Beckman, modelo DU 640 e sua pureza analisada através da relação de absorbâncias 260/280 nm.

A análise por *Northern blot* foi realizada segundo protocolo descrito por Sambrook & Russell (2001). Para isto, aproximadamente 15 µg de cada amostra de RNA total foi previamente desnaturada em tampão de desnaturação (MOPS 1X, formamida deionizada 50%, formaldeído 20 %, glicerol 10%, azul de bromofenol 0,025%, brometo de etídeo 0,05 µg/mL) a 65°C durante 10 minutos. As amostras desnaturadas foram submetidas à eletroforese em gel de agarose 1,5% contendo formaldeído 0,6 M a 65 V por aproximadamente 4 horas em tampão MOPS 1X (MOPS 20 mM, acetato de sódio 5 mM e EDTA 1 mM, pH 7,0). O RNA total, após eletroforese, foi transferido por capilaridade para membrana de nylon neutra (Hybond N, Amersham) utilizando solução tampão SSC 2X (NaCl 0,3 M e citrato de sódio 3 mM, pH 7,0) durante a noite. Após esta etapa, os RNAs

foram fixados na membrana por *crosslinking* (Ultraviolet Crosslinker UVP mod. CL1000) durante 2 minutos e, em seguida, a membrana foi lavada com solução tampão SSPE 2X (NaCl 0,3 M, NaH₂PO₄ 2 mM e EDTA 2 mM, pH 7,0) para a total remoção do formaldeído. Após as lavagens, a membrana foi pré-hibridizada a 42°C, durante 4 h com 5 mL da solução *ULTRAhyb*TM (Ambion) em forno de hibridização Hybaid. A sonda utilizada foi um fragmento do gene *gsn* de aproximadamente 1350 pb subclonado no vetor *pMOSBlue* (construção existente no laboratório). Esta sonda foi marcada radiativamente com α [³²P]dATP por *random priming* utilizando o kit *NEBlot*TM (Biolabs) e, em seguida, foi purificada por filtração em resina Sephadex G-50 (Sigma) equilibrada em tampão NT (Tris-HCl 1 mM, pH 8,0, NaCl 1 mM e EDTA 2 mM, pH 8,0) a fim de eliminar os nucleotídeos não incorporados. Em seguida, a sonda foi desnaturada a 100°C durante 10 minutos e 100 μ l da sonda desnaturada (aproximadamente 10⁸ cpm total) foi adicionada à solução de pré-hibridização *ULTRAhyb*TM (Ambion). A hibridização foi realizada a 42°C durante uma noite. Após a hibridização a membrana foi lavada em diferentes condições de estringência antes de ser exposta a filmes radiográficos (Kodak T-Mat G/RA).

Condições de lavagem:

SSPE 2x - 15 min/42°C

SSPE 0,5x; SDS 0,1% - 15 min/42°C

SSPE 0,1x; SDS 0,1% - 15 min/42°C

SSPE 0,1x; SDS 0,1% - 30 min/65°C

Análise computacional dos fatores de transcrição selecionados

As seqüências nucleotídicas das ORFs que codificam fatores de transcrição das linhagens mutantes selecionadas foram identificadas no genoma de *N. crassa* (<http://www.broad.mit.edu/annotation/genome/neurospora/Home.html>) e as seqüências polipeptídicas deduzidas e comparadas com seqüências disponíveis no banco de dados do *National Center for Biotechnology Information* (NCBI), utilizando a ferramenta Blastp (<http://www.ncbi.nlm.nih.gov/blast/>).

Para o alinhamento múltiplo das seqüências de proteínas que apresentaram similaridades relevantes após o Blastp foram utilizados os seguintes programas: ClustalW, no site do *European Bioinformatics Institute* (<http://www.ebi.ac.uk/Tools/clustalw>) e o

Boxshade, no site *Box Shade Server* (http://www.ch.embnet.org/software/BOX_form.html). Para a estimativa teórica das propriedades físico-químicas dos fatores de transcrição selecionados foi utilizado a ferramenta *ProtP* no site do servidor *ExPASy* (www.expasy.org/tools/protpar-ref.html). A detecção de domínios conhecidos presentes nos fatores de transcrição foi feita por meio das informações disponíveis no genoma do fungo (<http://www.fgsc.net/scripts/strainsearchform.asp>), pelo *SMART* (<http://smart.embl-heidelberg.de>) e pelo *Pfam 22.0* (<http://pfam.sanger.ac.uk>). A presença de sinais de localização nuclear clássicos (NLS) foram determinadas pelo *PSORTII* (<http://psort.nibb.ac.jp/form2.html>).

Resultados

Determinação do acúmulo de glicogênio nas linhagens e seleção das linhagens mutantes

Todas as linhagens mutantes foram analisadas em relação ao conteúdo de glicogênio antes e depois do choque térmico, tendo como padrão o conteúdo da linhagem selvagem, nas mesmas condições. A dosagem de glicogênio presente nos extratos celulares foi feita por meio da determinação da glicose liberada pela digestão do glicogênio com as enzimas α -amilase e amiloglicosidase em relação à quantidade de proteína presente no extrato protéico. Essa relação foi dada em microgramas de glicogênio por miligramas de proteínas, corrigindo-se os valores obtidos em função da concentração de proteína total nos extratos. O total de famílias dos fatores de transcrição, juntamente com o total de linhagens pertencentes a elas está representado na figura 2.

Em uma primeira análise, todas as linhagens mutantes foram avaliadas em relação ao perfil de acúmulo de glicogênio, sendo que 71 linhagens apresentaram um padrão de acúmulo de glicogênio distinto da linhagem selvagem após 30 minutos de choque térmico. Destas linhagens selecionadas, 50 mostraram concentrações de glicogênio após choque térmico semelhante ou até mesmo superior em comparação aos valores encontrados antes do choque térmico e 21 linhagens apresentaram um conteúdo de glicogênio muito inferior aos resultados obtidos antes da exposição ao estresse térmico. Um resultado típico desta primeira análise está mostrado na figura 3. É importante ressaltar que esta coleção, a qual possui 139 linhagens de *N. crassa*, apresenta 99 ORFs distintas nocauteadas. Ou seja, algumas linhagens distintas possuem a mesma ORFs nocauteada, se diferenciando apenas em seu fator de acasalamento.

As linhagens que apresentaram padrões distintos da linhagem selvagem foram selecionadas para a confirmação dos resultados. Para isto novos experimentos de choque térmico foram realizados em triplicata, utilizando a mesma metodologia descrita anteriormente. Nessa segunda análise, apenas 10 ORFs das 50 linhagens que apresentaram concentrações de glicogênio semelhante antes e após choque térmico confirmaram o resultado inicialmente obtido e, somente uma ORF das 21 linhagens que apresentaram conteúdo de glicogênio muito inferior em relação aos resultados obtidos antes do choque térmico confirmou o resultado. As linhagens selecionadas e as correspondentes ORFs nocauteadas que apresentaram acúmulo de glicogênio semelhante ou superior após

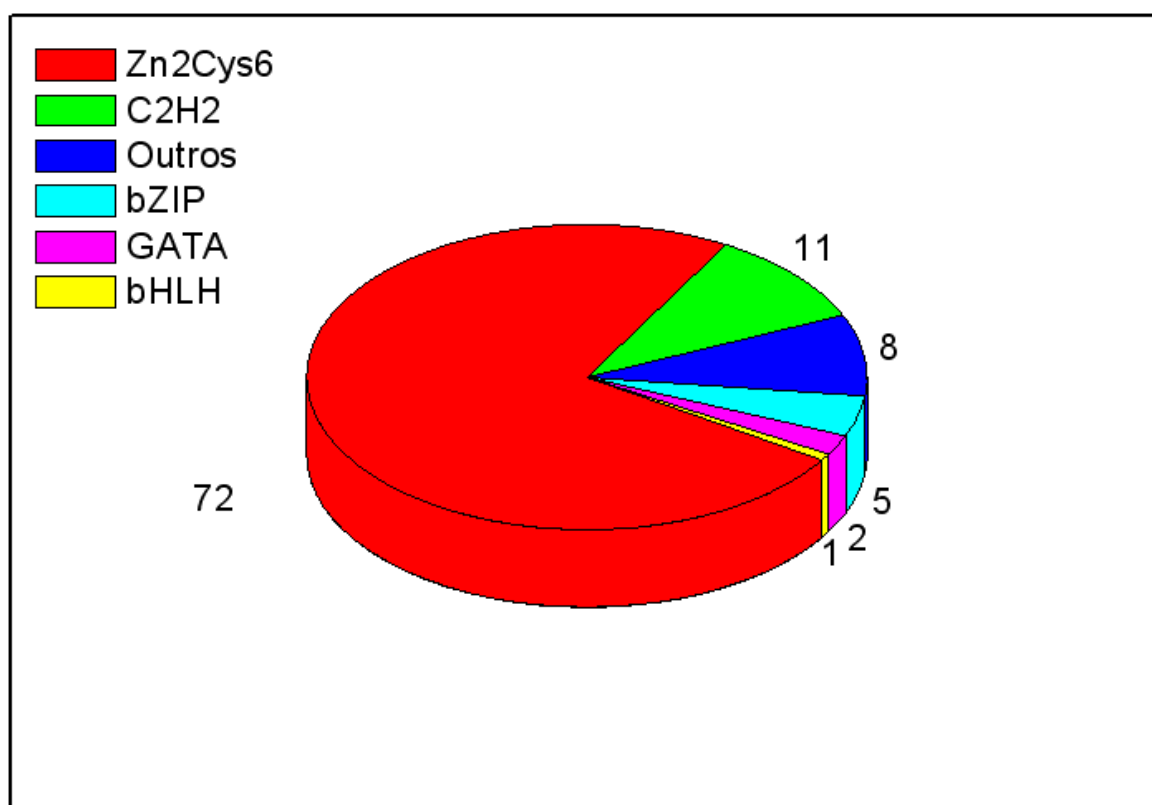


Figura 2 - Representação esquemática do número total de linhagens mutantes utilizadas no trabalho, divididas de acordo com a família do fator de transcrição cuja ORF foi nocauteada.

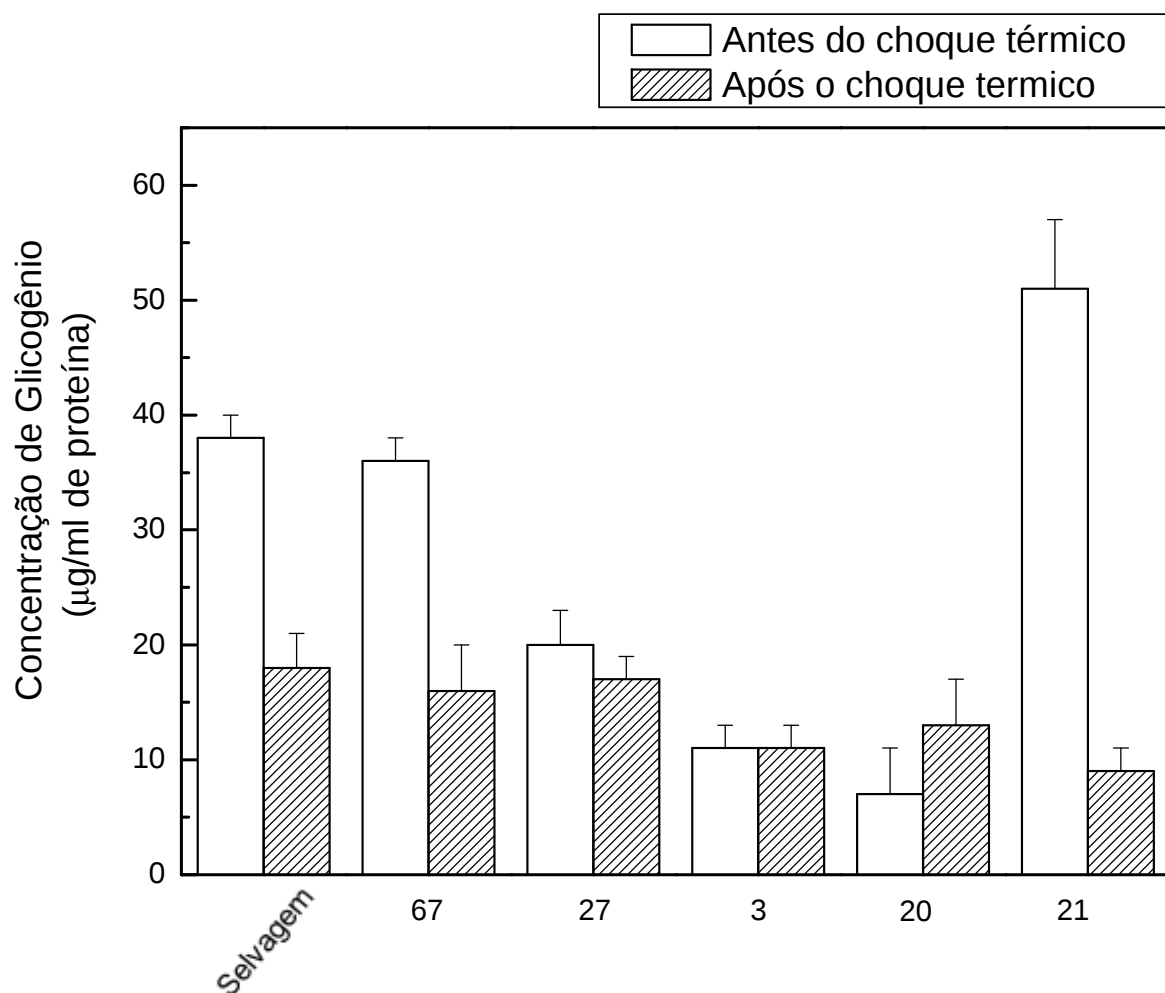


Figura 3 - Padrões de concentração de glicogênio observados nas linhagens durante a primeira análise. Os valores no eixo X correspondem ao número total das linhagens mutantes de *N. crassa* correspondentes aos padrões de acúmulo de glicogênio antes e depois do choque térmico. Células das diferentes linhagens foram submetidas ao choque térmico e o conteúdo de glicogênio foi determinado antes e após choque térmico como descrito em Materiais e Métodos.

o estresse térmico foram: FGSC# 11004 (NCU08000), FGSC#11062 (NCU09739), FGSC#11067 (NCU06971), FGSC#11089 (NCU04851), FGSC# 11102 (NCU01629), FGSC#11123 (NCU00808), FGSC#11126 (NCU01154), FGSC#11134 (NCU04390) e FGSC#11139 (NCU04731). A única linhagem que apresentou um conteúdo de acúmulo de glicogênio muito inferior à linhagem selvagem após o estresse térmico foi a FGSC# 11039 (NCU01097). Os resultados de conteúdo de glicogênio referentes à segunda análise estão mostrados na figura 4 e a classificação das linhagens selecionadas em relação às famílias de fatores de transcrição está representada na tabela 1.

Dentre as linhagens selecionadas com um maior acúmulo de glicogênio em relação à selvagem, após choque térmico, as linhagens FGSC#11004, FGSC#11062, FGSC#11189, FGSC#11067, FGSC#11123, FGSC#11126 e FGSC#11134 mostraram um acúmulo de glicogênio não tão reduzido, sendo os valores praticamente semelhantes em relação aos valores obtidos antes do estresse. Já as linhagens FGSC# 11102 e FGSC# 11139 apresentaram um padrão de acúmulo de glicogênio maior depois do choque térmico. É importante ressaltar que todas as linhagens que apresentaram alterações no acúmulo de glicogênio, tanto antes quanto após choque térmico, com exceção das linhagens FGSC#11089, FGSC#11102 e FGSC#11134, apresentaram uma concentração total de glicogênio inferior em relação à linhagem selvagem, tanto antes quanto depois do estresse térmico.

Caracterização bioquímica, molecular e estrutural dos fatores de transcrição selecionados

Após a determinação do conteúdo de glicogênio das linhagens mutantes, foram selecionadas apenas as linhagens que apresentaram um perfil de acúmulo de glicogênio distinto, tanto antes quanto após o choque térmico, em relação à linhagem selvagem. A partir das seqüências nucleotídicas das ORFs nocauteadas destas linhagens, as seqüências de aminoácidos dos fatores de transcrição foram deduzidas e analisadas. Utilizando as informações disponíveis no genoma do fungo (<http://www.fgsc.net/scripts/strainsearchform.asp>) juntamente com a utilização de ferramentas de bioinformática disponíveis na internet foi possível a caracterização bioquímica, molecular e estrutural dessas proteínas regulatórias, tais como ponto isoeletrico

Tabela 1: Classificação das linhagens selecionadas em relação às famílias de fatores de transcrição.

FGSC #	ORF	Fator de acasalamento	Família do Fator de Transcrição	Proteína Anotada
11004	NCU08000	a	Zn ₂ Cys ₆ zinc finger	Cutinase transcription factor 1 α
11039	NCU01097	A	Zn ₂ Cys ₆ zinc finger	Predicted protein
11062	NCU09739	a	Zn ₂ Cys ₆ zinc finger	Predicted protein
11067	NCU06971	A	Zn ₂ Cys ₆ zinc finger	Transcriptional activator XlnR
11089	NCU04851	A	Zn ₂ Cys ₆ zinc finger	Predicted protein
11102	NCU01629	A	C2H2 zinc finger	Predicted protein
11123	NCU00808	A	Zn ₂ Cys ₆ zinc finger	Conserved hypothetical protein
11126	NCU01154	A	GATA	Conserved hypothetical protein
11134	NCU04390	a	Zn ₂ Cys ₆ zinc finger	Conserved hypothetical protein
11139	NCU04731	a	bHLH	Conserved hypothetical protein

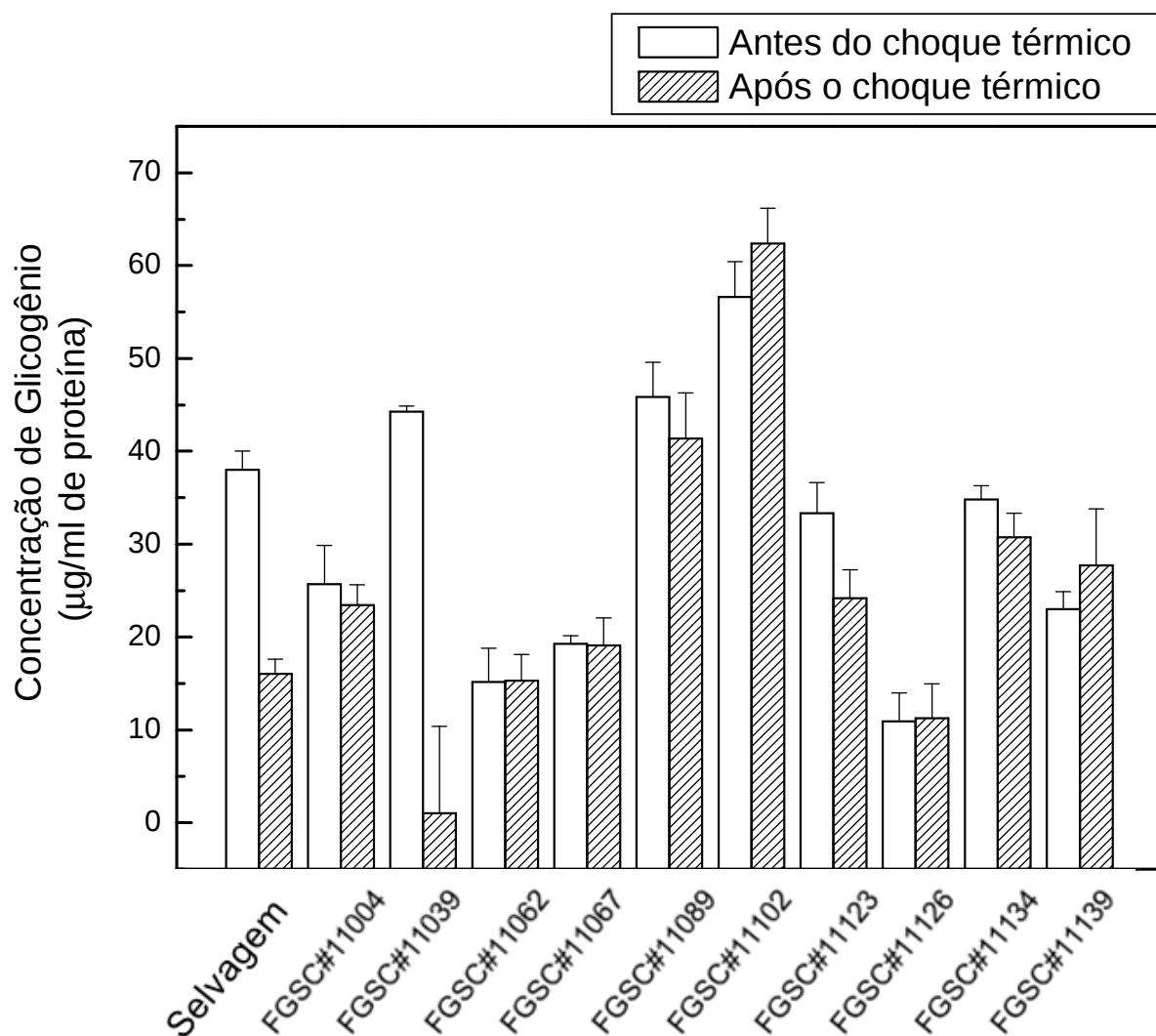


Figura 4 - Padrão de concentração de glicogênio observado nas linhagens após a segunda análise. Células das diferentes linhagens foram submetidas ao choque térmico e o conteúdo de glicogênio foi determinado antes e após choque térmico como descrito em Materiais e Métodos.

(pI), massa molecular teórica (MM), a presença de domínios conhecidos e de sinais de localização nuclear (NLS). Para a localização de domínios presentes nos fatores de transcrição utilizou-se os sites SMART (<http://smart.embl-heidelberg.de>) (Schultz *et al.*, 1998), PsortII (<http://psort.nibb.ac.jp/cgi-bin/runpsort.pl>) (Nakai & Horton, 1999) e Pfam 22.0 (<http://pfam.sanger.ac.uk/>) (Bateman *et al.*, 2004).

A figura 5 mostra uma representação esquemática das características estruturais dos fatores de transcrição selecionados e a tabela 2 mostra os dados bioquímicos obtidos por predição teórica de cada proteína selecionada. O fator de transcrição codificado pela ORF NCU08000 possui 1033 resíduos de aminoácidos com um total de 115,97 kDa de massa molecular e um ponto isoelétrico de 8,22. Esta proteína apresenta um domínio de ligação ao DNA pertencente à família Zn₂Cys₆ em sua região N-terminal, entre os resíduos 106 e 139, um domínio de ativação transcricional característico de fungos entre os resíduos 717 e 763, uma região de *coiled coil* entre os resíduos 717 e 763 e 2 sinais de localização nuclear clássicos monopartido localizados entre os resíduos 141 e 145 e entre os resíduos 706 e 710.

O fator de transcrição codificado pela ORF NCU01097 possui 817 resíduos de aminoácidos com um total de 90,23 kDa de massa molecular e um ponto elétrico de 6,34. Esta proteína possui um domínio de ligação ao DNA pertencente à família Zn₂Cys₆ entre os resíduos 44 e 73 e um sinal de localização nuclear clássico bipartido localizado entre os resíduos 36 e 52. O fator de transcrição codificado pela ORF NCU09739 possui 675 resíduos de aminoácidos com um total de 75,09 kDa de massa molecular e um ponto isoelétrico de 7,85. Esta proteína apresenta um domínio de ligação ao DNA pertencente à família Zn₂Cys₆ entre os resíduos 12 e 41 e não possui nenhum sinal de localização nuclear clássico.

O fator de transcrição codificado pela ORF NCU06971 possui 994 resíduos de aminoácidos com um total de 103,16 kDa de massa molecular e um ponto isoelétrico de 6,44. Esta proteína apresenta um domínio de ligação ao DNA pertencente à família Zn₂Cys₆ entre os resíduos 90 e 133, um domínio de ativação transcricional característico de fungos entre os resíduos 508 e 624 e dois sinais de localização nuclear monopartido entre os resíduos 87 e 93 e entre os resíduos 128 e 132. O fator de transcrição codificado pela ORF NCU04851 possui 1040 resíduos de aminoácidos com um total de 113,8 kDa de massa molecular e um ponto isoelétrico de 8,2. Esta proteína apresenta um domínio de ligação ao

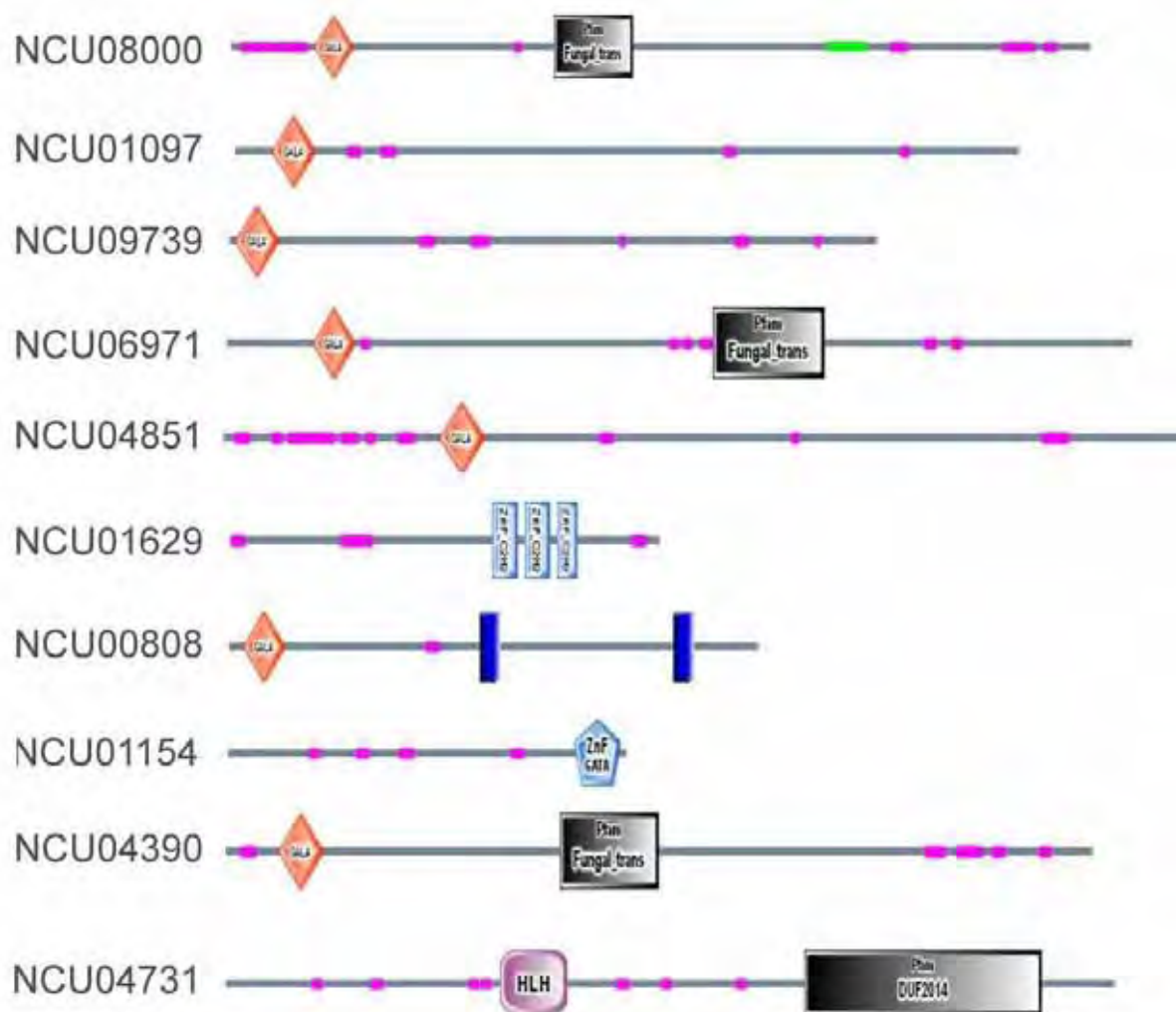


Figura 5 - Representação esquemática dos fatores de transcrição selecionados feita pelo servidor EMBL-SMART (<http://smart.embl-heidelberg.de>). As regiões em laranja representam o domínio Zn2Cys6, a em roxo representa o domínio bHLH, as barras em azul claro representam domínios C2H2, as barra em azul escuro representam domínios transmembrana, as regiões em preto representam domínios específicos, as em rosa representam seqüências com baixa complexidade e as em verde representam regiões de *coiled coil* presentes nos fatores de transcrição selecionados.

Tabela 2: Dados bioquímicos obtidos por predição teórica dos fatores de transcrição selecionados*.

ORF	Resíduos de AA	Massa molecular estimada (kDa)	Ponto isoelétrico	NLS**
NCU08000	1033	115,97	8,22	141-PKRKK (M) 706-KRHRR (M)
NCU01097	817	90,23	6,34	36-KRVKAVTQACHTCRRYK (B)
NCU09739	675	75,09	7,85	-
NCU06971	994	103,16	6,44	87-PIRRRIS (M) 128-RKKR (M)
NCU04851	1040	113,8	8,2	268-PKEKRW (M) 222-KRRNRPAVSCIPCRGRKI (B)
NCU01629	420	45,91	8,82	240-PRPKRQQ (M)
NCU00808	551	61,87	5,97	147-PRITKK (M) 11-RKRLRTSHACDLRFRKI (B)
NCU01154	445	48,68	8,91	381- KKRR (M)
NCU04390	847	93,67	6,6	89-PQPPRRRKKK (M) 94-RRKKKPKHERDLIDRLKKY (B)
NCU04731	994	108,85	6,5	276-PNSRKRK (M)

* A estimativa teórica das propriedades físico-químicas dos fatores de transcrição bem como a determinação de NLS clássicos foi realizada por meio das ferramentas ProtParam (www.expasy.org/tools/protpar-ref.html) e PsortII (<http://psort.nibb.ac.jp/cgi-bin/runpsort.pl>), respectivamente.

** M e B correspondem ao sinal de localização nuclear (NLS) monopartido (M) e bipartido (B).

DNA pertencente à família Zn_2Cys_6 entre os resíduos 225 e 271, um sinal de localização nuclear bipartido entre os resíduos 222 e 239 e um sinal monopartido entre os resíduos 268 e 274.

O fator de transcrição codificado pela ORF NCU01629 possui 420 resíduos de aminoácidos com um total de 45,91 kDa de massa molecular e um ponto isoelétrico de 8,82. Esta proteína apresenta três domínios de ligação ao DNA pertencente à família C_2H_2 em sua região C-terminal, entre os resíduos 257 e 281, entre os resíduos 289 e 314 e entre os resíduos 320 e 340 e apenas um sinal de localização nuclear clássico monopartido entre os resíduos 240 e 246.

O fator de transcrição codificado pela ORF NCU00808 possui 551 resíduos de aminoácidos com um total de 61,87 kDa de massa molecular e um ponto isoelétrico de 5,97. Esta proteína apresenta um domínio de ligação ao DNA pertencente à família Zn_2Cys_6 entre os resíduos 16 e 59, possui duas regiões transmembranas, uma entre os resíduos 262 e 284 e outra entre os resíduos 464 e 486 e possui dois sinais de localização nuclear clássicos, um monopartido localizado entre os resíduos 147 e 153 e um bipartido localizado entre os resíduos 11 e 28.

O fator de transcrição codificado pela ORF NCU01154 possui 445 resíduos de aminoácidos com um total de 48,68 kDa de massa molecular e um ponto isoelétrico de 8,91. Esta proteína apresenta um domínio de ligação ao DNA pertencente à família GATA localizado em sua região C-terminal, entre os resíduos 387 e 439 e apresenta um sinal de localização nuclear clássico monopartido entre os resíduos 381 e 384.

O fator de transcrição codificado pela ORF NCU04390 possui 847 resíduos de aminoácidos com um total de 93,67 kDa de massa molecular e um ponto isoelétrico de 6,6. Esta proteína apresenta um domínio de ligação ao DNA pertencente à família Zn_2Cys_6 localizado entre os resíduos 52 e 95, um domínio de ativação transcricional característico de fungos entre os resíduos 327 e 424, e apresenta 2 sinais de localização nuclear clássicos localizados sobrepostos, um monopartido entre os resíduos 89 e 99 e um bipartido entre os resíduos 94 e 111.

O fator de transcrição codificado pela ORF NCU04731 possui 994 resíduos de aminoácidos com um total de 108,85 kDa de massa molecular e um ponto isoelétrico de 6,5. Esta proteína possui um domínio de ligação ao DNA pertencente à família bHLH, localizado entre os resíduos 308 e 382, um domínio de fatores de transcrição ER de mamíferos, entre

os resíduos 647 e 911, envolvido no controle do metabolismo de lipídio em mamíferos e apenas um sinal de localização nuclear clássico monopartido localizado entre os resíduos 276 e 282.

As seqüências polipeptídicas foram utilizadas para a análise por Blastp para a identificação de homólogos funcionais. Apenas os fatores de transcrição codificados pelas ORFs NCU08000 e NCU06971 mostraram homologia a proteínas com funções conhecidas e depositadas em banco de dados. A análise da seqüência polipeptídica deduzida da ORF NCU8000 revelou a existência de um grande número de proteínas homólogas depositadas nos bancos de dados. Destas foram selecionados os fatores de transcrição cutinase 1 α (CTF1 α) dos ascomicetos *Fusarium oxysporum*, *Gibberella zeae*, *Nectria haematococca*, (anamorfo: *Fusarium solani*), *Ajellomyces capsulatus* e *Aspergillus terreus* por apresentarem *e-value* igual a 0.0.

As seqüências polipeptídicas destas proteínas, juntamente com a proteína de *N. crassa* foram analisadas por um alinhamento de seqüências e o resultado está mostrado na figura 6. Podemos observar que a proteína de *N. crassa* mostrou uma boa identidade às outras proteínas analisadas tendo 68%, 68%, 67%, 53% e 52% aminoácidos idênticos em relação às proteínas dos 5 microrganismos, respectivamente. O domínio de ligação ao DNA do fator de transcrição codificado pela ORF NCU08000 se mostrou extremamente conservado em relação às outras proteínas, contendo apenas uma alteração no resíduo 134, onde uma glutamina foi substituída por uma serina. Uma singularidade apresentada pela proteína codificada pela ORF NCU08000 é a presença de uma seqüência de poliglutamina (poli-Q) entre os resíduos 725 e 761, que pode formar um domínio *coiled coil* (em verde na figura 5) e não está presente nas outras proteínas analisadas.

O outro fator de transcrição que apresentou homologia a proteínas com funções conhecidas e depositadas em banco de dados foi o codificado pela ORF NCU06971. A busca por proteínas homólogas revelou uma grande quantidade de proteínas relacionadas com a ativação da transcrição do gene codificador da proteína XlnR, envolvida na expressão de enzimas envolvidas na degradação de celulose e hemicelulose. Assim, foram selecionados para o alinhamento múltiplo de seqüência, mostrado na figura 7, as proteínas XlnR dos ascomicetos *F. oxysporum*, *Hypocrea jecorina*, *A. kawachii*, *A. níger* e *A. oryzae*, por apresentarem *e-value* igual a 0.0. A proteína de *N. crassa* mostrou identidade as outras

<i>N. crassa</i>	1	MSATTPAEKSTTASAAQQQEQQQQQQQQSQPSRQQAQPKQPPPLQPHPQFSSQSPT
<i>F. oxysporum</i>	1	-----MSSGSGTPQAQP--QQQQQQPESQLS-----
<i>G. zeae</i>	1	-----MSSGSGTPQVQPQAQQQQQQGESQISS-----
<i>N. haematococca</i>	1	-----MSSGDAPPQAQPQPHQQEQPNQRQSSTP-----
<i>A. capsulatus</i>	1	-----MSSIASDHRPDHAMTGAEDIATTSTKDAP-----
<i>A. terreus</i>	1	-----MSTTGENHTDSTSRPSPAPSATGSGTGTSG-----

<i>N. crassa</i>	61	QSRQSQSQSPAPSQAHSQGGQPPSASTATGTSTGHMSFRRQRASRACETCHARKVRCDA
<i>F. oxysporum</i>	25	-----APALPAPSTSTSTAGGVSFRRQRASRACETCHARKVRCDA
<i>G. zeae</i>	28	-----TPGPPGSIITSTISAGGVSFRRQRASRACETCHARKVRCDA
<i>N. haematococca</i>	29	-----APSAAPVPPAPSTSTSNAGGVSFRRQRASRACETCHARKVRCDA
<i>A. capsulatus</i>	30	-----SLPNTTSAPKPSVPSMQTPGHPSFRRQRASRACETCHARKVATSIH
<i>A. terreus</i>	30	-----ITVRAGPNGHMSFRRQRASRACETCHARKVRCDA

<i>N. crassa</i>	121	SLGVPCTNCVAFSIECRIPTPKRKKTAQAQSTTTTAPSKDSDSERGD-TEDRSPRQAGANN
<i>F. oxysporum</i>	67	SLGVPCTNCVAFQIECRIPNPKRKKT--QGSG-SQTNKDSDSDRGDANEDPSRPVAPSS
<i>G. zeae</i>	70	SLGVPCTNCVAFQIECRIPNPKRKKT--QGSG-SQTNKDSDSDRGDGTEDPSRPVAPPS
<i>N. haematococca</i>	75	SLGVPCTNCVAFQIECRIPNPKRKKT--QGTGGSQTNKDSDSDRGDGTEDPSRPVAP-T
<i>A. capsulatus</i>	76	DL-----HPTSVRGKESDER-----PDT
<i>A. terreus</i>	65	SLGVPCTNCVA SIECKIPTPKRKKN-----QNKSKETSGSEAGSERPPTKQEEAPAA

<i>N. crassa</i>	180	TFPAGTRPAAAYHTQEGTPTTSNVKEQAKREEYDNATLANYMNLMVKPKFTRAPITEAG
<i>F. oxysporum</i>	124	TSSLTPRAPSVYHSNNGTPPTNW-TEAQAARKEEVDSG---TYLDLVMKPKFTRAPITEAG
<i>G. zeae</i>	127	TISLTPRAPSVYHSNSGTPPTAW-TEAQAARKEEVDNG---TYLDLVMKPKFTRAPITEAG
<i>N. haematococca</i>	132	AASFAPRTPSVFHSHNGTPPTTL-TEAQAARKEEVDSG---TYLDLVMKPKFTRAPITEAG
<i>A. capsulatus</i>	94	GSGVSVDSTQGAKTVEGMPETSL-TEAQAQAQASHNAS---YAQFMKPKFARAPIKEAG
<i>A. terreus</i>	119	DGKNAFGYRSNRMAVDGMPVTTL-SDQQVAEQASQNT---YAQFMKPKFARAPIKEAG

<i>N. crassa</i>	240	RVAYMGESSNLNLLVHDRQSDSDVVHYPLPEHVRGNKARLSELDSTEIEILHQRGAFLLP
<i>F. oxysporum</i>	180	RVAYLGESSNLTLVHDRQGSADVHYPLPENVRGSRARLTELDNVEIDILHQRGAFLLP
<i>G. zeae</i>	183	RVAYLGESSNLTLVHDRQGSADVHYPLPENVRGSRARLTELDNIEIDILHQRGAFLLP
<i>N. haematococca</i>	188	RVAYLGESSNLTLVHDRQGSADVHYPLPENVRGSRARLTELDNVEIDILHQRGAFLLP
<i>A. capsulatus</i>	149	RVAYLGESSNLSLLVHDRHGTDDVVHYPLPDGIRGSRARLTE-----GAFLLP
<i>A. terreus</i>	174	EGRLPGRVVESVAPGAGSPWNRDCRALPPFHIRGSRSRLTDLNLELDILHQRGAFLLP

<i>N. crassa</i>	300	PRSLCDELIESYFQVWHPIVPVINRTKFMRYQKDPKNPPSLLLLQAMLLAGSRVCTNPQL
<i>F. oxysporum</i>	240	PRSLCDELIDAYFKVWHPIVPVINRTRFMRYQYRDPKNPPSLLLLQSVLLAGTRACNNPQL
<i>G. zeae</i>	243	PRTLCDDELIDAYFKVWHPIVPVINRTRFMRYQYRDPKNPPSLLLLQSVLLAGTRACNNHQL
<i>N. haematococca</i>	248	PRALCDELIDAYFSVWHPIVPVINRTRFMRYQYRDPKNPPSLLLLQSVLLAGTRACNNAQL
<i>A. capsulatus</i>	197	PRALCDELVDAYFRWVAPVPIINRSRFRMYRDPKNPPSLLLLQAILLAGSRVCTNPQL
<i>A. terreus</i>	233	PKPLCDELVEAYFKWVAPVPIVNRSRFRMYRDPKNPPSLLLLQAILLAGSRVCTNPQL

<i>N. crassa</i>	360	MDANGSAAPAALTFYKRAKALYEAGYEDDRVTIVQSLLLMGWYEGPEDVTKNVIFYWSRV
<i>F. oxysporum</i>	300	MDANGSTTPAALTFYKRAKALYDANYEDDRVTIVQSLLLMGWYEGPEDVTKNVIFYWSRV
<i>G. zeae</i>	303	MDANGSTTPAALTFYKRAKALYDANYEDDRVTIVQSLLLMGWYEGPEDVTKNVIFYWSRV
<i>N. haematococca</i>	308	MDANGSTTPAALTFYKRAKALYDANYEDDRVTIVQSLLLMGWYEGPEDVTKNVIFYWSRV
<i>A. capsulatus</i>	257	MDANGSTTPAAMTFYKRAKALYDANYEDDRVTIVQALVLMGWY---EDVTKNVIFYWSRV
<i>A. terreus</i>	293	MDANGSTTPAAMTFYKRAKALYDANYEDDRVTIVQALVLLMGWYEGPEDVTKNVIFYWTRV

<i>N. crassa</i>	420	ATIVAQGSGMHRSVEGSQLSKADKRLWKRIWWSLFTDRSTAVALGRPCHINLDDSDVEM
<i>F. oxysporum</i>	360	ATIVAQGSGMHRSVEQSLSRADKRLWKRIWWSLFTDRSVAVALGRPVIHNLDDADVEM
<i>G. zeae</i>	363	ATIVAQGSGMHRSVEQSLSRADKRLWKRIWWSLFTDRSVAVALGRPVIHNLDDADVEM
<i>N. haematococca</i>	368	ATIVAQGSGMHRSVEQSLSRSDKRLWKRIWWSLFTDRSVAVALGRPVIHNLDDADVEM
<i>A. capsulatus</i>	314	ATIVAQGSGMHRSVEGSQLSRDPKRLWKRIWWSLFTDRSVAVALGRPVIHNLDDSDVEM
<i>A. terreus</i>	353	AMIVAQGSGMHRSVESQLSKDPKRLWKRIWWSLFTDRSVAVALGRPVIHNLDDSDVGM
<i>N. crassa</i>	480	LTEDDFIEDEPDNPSDYPPDETHVQFFLQYVKLC EIMGLVLSQQYSVASKGKGKNAIDLT
<i>F. oxysporum</i>	420	LTEDDFIEDEVDRASEYPPDPIHVQFFLEYVKLC EIMGLVLSQQYSVASKGRQRNAIDLT
<i>G. zeae</i>	423	LTEDDFIEDEADRASEYPPDPIHVQFFLEYVKLC EIMGLVLSQQYSVASKGLKRNAIDLT
<i>N. haematococca</i>	428	LTEDDFIEDEQDRVSEYPPDPIHVQFFLEYVKLC EIMGLVLSQQYSVASKGRQRNAIDLT
<i>A. capsulatus</i>	374	LTEDDFIEDEVDSVAEYPPDPIHVQFFLQYVKLC EIMGLVLSQQYSVASKSRRTNAIDLT
<i>A. terreus</i>	413	LTEDDFIEDEIDIAAEYPPDPVHVQFFLQYVKLC EIMGLVLSQQYSVASKSRRMNAMDLT
<i>N. crassa</i>	540	HSDMALADWLQNCPKIVYWEMRNHHFW SALLHSNYYTTLCLLHRAHMPPSGSHR---WPD
<i>F. oxysporum</i>	480	HSDMALADWLQNCPKIVYWEMPRHHFW SALLHSNYYTTLCLLHRAHMPPGGSSR---LPD
<i>G. zeae</i>	483	HSDMALADWLQNCPKIVYWEMPRHHFW SALLHSNYYTTLCLLHRAHMPPGGSSR---LPD
<i>N. haematococca</i>	488	HSDMALADWLQNCPKIVYWEVPRHHFW SALLHSNYYTTLCLLHRAHMPPGGSSAR---FPD
<i>A. capsulatus</i>	434	HSDMALADWLQNCPKIVYWERPRHHFW SALLHANYYTTLCLLHRAHMPPATSIISGNYGSD
<i>A. terreus</i>	473	HSDMALADWLQNCP-----QRSLLAATAAP-----LLGRASP-----
<i>N. crassa</i>	597	DSPYPSRNIAFQAAAMITSITENLQSNHQLRYCPAYIVYSLFSALIMHVYQMRSPVASIQ
<i>F. oxysporum</i>	537	PSPYPSRNIAFQAAAMITSIVENLA AHDQLRYCPAFVVYSLFSALIMHVYQMKSPVPSIQ
<i>G. zeae</i>	540	PSPYPSRNIAFQAAAMITSIVENLA AHDQLRYCPAFVVYSLFSALIMHVYQMKSPVPSIQ
<i>N. haematococca</i>	545	PSPYPSRNIAFQAAAMITSIVENLA AHDQLRYCPAFVVYSLFSALIRHVYQMRSPVPSIQ
<i>A. capsulatus</i>	494	GMTYPSRTIAFQAAAMITSIVENLQTHGELKFTPAFI---LFSALIMHVYQMRSSVPSV
<i>A. terreus</i>	505	RKLLTHQEIRYTPAFMYVPAHQCGSSDTHQVNS---VYSLFSALIMHVYQMRSSVPTIV
<i>N. crassa</i>	657	QVTQTRIRTCMAALKDVS KVWLVGKMVFTLFESILGNKVLEERLQKAAGKRHRMQQGLS
<i>F. oxysporum</i>	597	QVTQDRLRSCMSAMKEISRVWL VGKMVYALFESIMGNKVLEERLQRAEGKRHRNLRQSL
<i>G. zeae</i>	600	QVTQDRLRSCMSAMKEISRVWL VGKMVYALFESIMGNKTLLEERLQRAEGKRHRNLRHSLT
<i>N. haematococca</i>	605	QVTQDRLRSCMSAMKEISRVWL VGKMVYALFESIIGNKVLEERLQKAAGKRHRKMRQTLT
<i>A. capsulatus</i>	551	SLSQERINICMQALKDVS KVWLVAKMVHTLFESILGNKALEERLQKAAGKRHRQVRRDQP
<i>A. terreus</i>	561	ATCQERINICMQALKDVS KVWLVAKMVHTLFESILGNKVLEERLQKAAGRRHQKVRPEAN
<i>N. crassa</i>	717	QLEQYRQHQLQQQQQQQQQQQQQQQQQQQQQNHQYQQQQQQRVHDQKRKYDEMAID
<i>F. oxysporum</i>	657	HLEQ-----QQNRQAEAPKRKYDDMAID
<i>G. zeae</i>	660	QLEQ-----QQNRQAEATKRKYDDMAID
<i>N. haematococca</i>	665	TQLE-----QHSRQQEAPKRKYDEMAID
<i>A. capsulatus</i>	611	APTIPPSS-----TTAARRPIEAPKRKFDDMEIS
<i>A. terreus</i>	621	QQQA-----LRRPDPPPKRKFDDMDIG
<i>N. crassa</i>	777	FSVNTPQPQESYERSRPQTPSL-----TTKTETTPSTMPPPVTSPNNQNGH----
<i>F. oxysporum</i>	680	FGTITPQPQESYERSRPQTPS-----AVKVESGSTMQPPPVTSNAR-----
<i>G. zeae</i>	683	FGTNTPQPQESYERSRPQTPS-----AVKTEPASSMQPPPVTSNTR-----
<i>N. haematococca</i>	688	FGTNTPQPQESYERSRPQTPS-----AVKAETSSMQPPPVSSPGAR-----
<i>A. capsulatus</i>	640	LPNGPPAPQVS YERSRPQTPAVTPSREMPPTPAGSAMAAPPSQHPHTSSTAQHQQSHQQ
<i>A. terreus</i>	643	LPNGGPTPPVSYERSRPQTPAV-----TPSRELSQPGLGVPQASPTAHR-----

<i>N. crassa</i>	823	-----RPPHDAFMGGTASRP-HTRPATP--FNPSFSVPATPPDLYLVTRN
<i>F. oxysporum</i>	722	-----QSAADTFMGGTNSRP-QTRPATP--FNPSFSVPPTPPDLYLVTRN
<i>G. zeae</i>	725	-----QSTADTFMGGTNSRP-QTRPATP--FNPSFSVPPTPPDLYLVTRN
<i>N. haematococca</i>	729	-----QSAADTFMGGTNSRP-QTRPATP--FNPSFSVPPTPPDLYLVTRN
<i>A. capsulatus</i>	700	QQQQQQQQQQQQRSINHDTFLGAAGSSDGSSTRPDSPSQFNPPFNIPATPPDLFLVTRN
<i>A. terreus</i>	687	-----DGGPGNSRANTRPTTP--FNGQFSLPATPPDLFLVTRT
<i>N. crassa</i>	865	SPNLSQNIWENFQPDQLFPSTNMP--LFPHQSP-TQQHSGLDPN-MIHMPSGLSNQPIE
<i>F. oxysporum</i>	764	SPNLSQSLWENFQPDQLFPDSAAMP--AFPNLS-PTQTHSNLDHNAMEGSPVPPNNGQGGMH
<i>G. zeae</i>	767	SPNLSQSLWENFQPDQLFPDSASMP--AFPNMSSPTQTHSNLDHNNIGSMPPGNGQSGIH
<i>N. haematococca</i>	771	SPNLSQSLWENFQPDQLFPDSAAMP--AFPNLS-PIQTHASLDHSAMGPVP-GNGQGGMP
<i>A. capsulatus</i>	760	SPNLSQSLWENFQPTQLFPDGTNMSFAGFSRSGAIDPQLQQTPTMMGSMGTPALSPGQQ
<i>A. terreus</i>	723	SPNLSQSLWENFQPDQLFPDGTATF---PELTSPQHNAVDLPQLQMQTQLQTQGIDQRM
<i>N. crassa</i>	921	YTQGVKRNLAGSPLNNPNSNNNGNLLHPGGMPGQGGQGGGYGNNQSNFNNANFDGQ
<i>F. oxysporum</i>	821	NPQ-----AG-----QYQQRGNGMMP-----QGFQG-----HSN-MWQPNMDPN
<i>G. zeae</i>	825	NPQ-----AG-----QYQNRGNGMMP-----QGFQG-----HSN-MWQPNLDPS
<i>N. haematococca</i>	827	NQQ-----AG-----QYQQRGNGILP-----QGFQG-----HTN-MWQPNLDPN
<i>A. capsulatus</i>	820	HSQLHQQLPTRG---LRDSPTMGHMHGMGPPLQQRQQQEQHPGGGWPSMQGLDGH
<i>A. terreus</i>	779	GQPVPSRSMSTGTQ---GSPPELLSSMPPG--IPLQGGQPPQMFGEHQGSWPMQGLEAT
<i>N. crassa</i>	981	IGGGGGNDGHSPSDSWSNSSVHGQSVSTLNVEDCDPSQGYMN-----LDIPELMRQL-
<i>F. oxysporum</i>	854	LP-----EGQSP-DSWSTASGPGQAVPTTLNVEDWFQFFGING-----TDPNHINLDMP
<i>G. zeae</i>	858	LP-----EGQSP-DSWSTASGPGQAVPTTLNVEDWFQFFGING-----TDPN-FNFDVP
<i>N. haematococca</i>	860	LP-----EGQSP-DSWSTSGQGQAVPTTLNVEDWFQFFGING-----TDPNHLNLDIP
<i>A. capsulatus</i>	876	MAP-TTMESSSQDDNWSTSSRGQMPIPPTLNVEDWFQFFGIN-----AGVG-----
<i>A. terreus</i>	832	LS-AATVDPANQDDNWSSSRSGPTAPTTLNVEDWFQFFGINGSLGWNAKTVHDISLKL
<i>N. crassa</i>		-----
<i>F. oxysporum</i>	902	LG-----
<i>G. zeae</i>	905	LG-----
<i>N. haematococca</i>	908	LG-----
<i>A. capsulatus</i>		-----
<i>A. terreus</i>	891	YRSNWQPPHNKQSHHLTNPPTR

Figura 6 - Alinhamento múltiplo da proteína codificada pela ORF NCU08000 com os fatores de transcrição cutinase dos ascomicetos *F. oxysporum* (ABR12478), *G. zeae* (XP_382112), *N. haematococca* (P52958), *A. capsulatus* (XP_001536925) e *A. terreus* (XP_001211421). O alinhamento de seqüências foi realizado por ClustalW e os resíduos de aminoácidos idênticos (em preto) e as trocas conservativas (em cinza) foram identificados por BoxShade. O sítio de ligação ao DNA (pontilhado) e os NLS (sublinhados) estão representados.

<i>N.crassa</i>	1	MLSNPLHRFAPYHAMPSPTLLSGGHVTASHLHAAGLDTMGP-----
<i>F.oxysporum</i>	1	MLSNPLQRFSPYQNIITSSNISPDGNVQQGTMSGTGLESLGQ-----
<i>H.jecorina</i>	1	MLSNPLRRYSAYPDISSASFDPNYHGSQSHLHSINVNTFGN-----
<i>A.kawachii</i>	1	-----
<i>A.niger</i>	1	-----
<i>A.oryzae</i>	1	MSTTSIQHFTSSFSPPSSGTQPVGMAQSQTVGLDTLAEGSQYALEQLQLSREANGASAVD
<i>N.crassa</i>	42	-----GSHYALQQLQQHVSVHNHHLARAGP-QPKHRQ-HPYGPV--TRATGAAG
<i>F.oxysporum</i>	42	-----SHQYPIQPLSQAVPLSNAHLERPGP-QVKNRQ-HPYGIH--PRNASTSG
<i>H.jecorina</i>	42	-----SHPYPMQHLAQHAELSSSRMIRASPVQPKQROGSLIAAR--KNSTGTAG
<i>A.kawachii</i>	1	-----MSHAKDQPLFDDE---RNQSAGSGFKNTLORD-PLVEARSAIRKNSSSA
<i>A.niger</i>	1	-----MSHTKDQPPFDNE---KNQSTGSGFRDALORD-PLVEARSAIRKTSSSA
<i>A.oryzae</i>	61	GGVPNPLRSSISKPPGQQLYSDSSAQHTQNAATTGFRNLPORD-QLAEARSTIRKSSNSG
<i>N.crassa</i>	87	PIRRRISRACDQCQLRTKCDGQHPCAHCIEFGLGCEYIRERKKRGKASRKDLAAQAAAA
<i>F.oxysporum</i>	87	PIRRRISRACDQCQLRTKCDGQHPCAHCIEFGLGCEYIRERKKRGKASRKELAQAAAAQ
<i>H.jecorina</i>	89	PIRRRISRACDQCQLRTKCDGLHPCAHHCIEFGLGCEYVIRERKKRGKASRKDLAAQAAA
<i>A.kawachii</i>	46	PVRRRISRACDQCQLRTKCDGQHPCAHCIEFGLTCEYARERKKRGKASKKDLAAAAAAA
<i>A.niger</i>	46	PVRRRISRACDQCQLRTKCDGQHPCAHCIEFGLTCEYARERKKRGKASKKDLAAAAAAA
<i>A.oryzae</i>	120	PVRRRISRACDQCQLRTKCDGQNPCHAHHCIEFGLTCEYARERKKRGKASKKDLAAAAAAV
<i>N.crassa</i>	147	AAAQLNGHKNPQAGENDQSPNRTTESTTATKRASSLPTEHQTTSNDK--TMSDMSGSGV
<i>F.oxysporum</i>	147	AAAANGQTLDSESTENGQSGNKGLDSSNMVLEQQSN--ERHPSTSSK--SSRDPGDVDM
<i>H.jecorina</i>	149	AAAHSGQVQDGPEDQHRKLSRQQSESSRGSALAQAHPDPPHGHIEG--SVSSFSDNGL
<i>A.kawachii</i>	106	THG-----SNGHSGQANASLMAERTSEDSRPAQDVNGRYDSTFESH--ISSQPSHMQH
<i>A.niger</i>	106	TQG-----SNGHSGQANASLMGERTSEDSRPGQDVNGTYDSAFESH--LSSQPSHMQH
<i>A.oryzae</i>	180	ANNGTAPTNSGNTSNDVSSAKRHTPSDGQSTQEVSGRYDPNFDASRNLAAGQSQLGQH
<i>N.crassa</i>	205	R-SQRTGSMDSIDLGAHQTHIASHPGAMDRDLESPAALDLS-YGNVHQEYHRQG-----M
<i>F.oxysporum</i>	203	RHTQGLEGLDPLGNISEQPHLGRSS-LDGEHIENNGGLDLNGFGSMAGHYETQG-----L
<i>H.jecorina</i>	207	SQHAAMGGMD--GLEDHHGHVGVDPALGRTQLEASSAMGLGAYGEVHPGYESPG-----M
<i>A.kawachii</i>	158	ANNAGISGLHDSQTAPSHSQPSLGTITIDAMHLGHFNTLNDSGRPAMSMSDLRSL-----
<i>A.niger</i>	158	ASTAGISGLHESQTAPSHSQPSLGTITIDAMHLNHFNTMNDSGRPAMSSIDLRSL-----
<i>A.oryzae</i>	240	SDMSGMAGMQGSQQTTP-HSQPSLGGIDAHLNHFNTLNDSNRPQMSVPDLRSLQMLHPS
<i>N.crassa</i>	258	GAHLMNGASHHTPYGSNQAAMSNYPDLPYALHTQSPTGYSS-ANTSSGFRIGASPLSAYP
<i>F.oxysporum</i>	257	EGPVLNGQSYAANGRGN---MPGYAEFPYSMQAQSPPNFA--NNPT--FRMGNSPLG-YS
<i>H.jecorina</i>	260	NGHVMVPPSYGAQ-----TTMAGYSGISYAAQAPSPATYS--SDGN--FRLTG-HIHDYP
<i>A.kawachii</i>	212	-----PPSVLPPQGLS-----SGYNASAFALVNPQEPGSP-ANQFRLGSSAENPTAPFLG
<i>A.niger</i>	212	-----PPSVLPPQGLS-----SGYNASAFALVNPQEPGSP-ANQFRLGSSAENPTAPFLG
<i>A.oryzae</i>	299	GANTRSPSGALPPQGMN---SGYNDGAYSLMNASEANHPSSINQYRLGNSAENPPAPFLG
<i>N.crassa</i>	316	MAGGSTSPGW-MNLASPPPQFAQHIPPITYSHAQLRYPVLEPLLPHLGNLMPVSLACDLI
<i>F.oxysporum</i>	309	MGKG-TSPGWGISMASPPGQYQSQVPAPAFNNSKLRYPVLEPLVPYLLNNPIPIPLACDLI
<i>H.jecorina</i>	310	LANG-SSPSWGVSLASP-----SLRYHVLRPVLLDVNRNITYPVSLACDQM
<i>A.kawachii</i>	261	LSPPGQSPGW-LPLPSPSPANFPFSFLHPFS-STLRYPVLPVLPPIASTIIPQSLACDLL
<i>A.niger</i>	261	LSPPGQSPGW-LPLPSPSPANFPFSFLHPFS-STLRYPVLPVLPPIASTIIPQSLACDLL
<i>A.oryzae</i>	355	LSPPAQSPGW-LSLPSPSPANFASFMSPPFS-STLRYPVLPVLPPIASTIIPQSLACDLL

<i>N.crassa</i>	375	DL YFASSSSAQ MHPMSPYVLGFVFRKRSFLHPTKPRQCQ PALLASMLWVAAQTSDAPFLT
<i>F.oxysporum</i>	368	DL YFASSSSAQ MHPMSPYVLGFVFRKRYFLDQTRPRPCQ PALLASMLWVAAQTSDAPFLA
<i>H.jecorina</i>	353	DM YFSSSSSAQ MRPMSPYVEGFVFRKRSFLHPTDPRRCQ PALLASMLWVAAQTSEASFLT
<i>A.kawachii</i>	319	DVYFHSSSSPSHLSPPSPYVVG YIFRKQSFLHPTKPRLCSSG LLASMLWVAAQTSEAPFLT
<i>A.niger</i>	319	DVYFTSSSSSHLSPLSPYVVG YIFRKQSFLHPTKPRICSPG LLASMLWVAAQTSEAAFLT
<i>A.oryzae</i>	413	DVYFTSFSPSHLSPOSPYVVG YIFRKQSFLHPTKPRVCSPG LLASMLWVAAQTSDAAFLT
<i>N.crassa</i>	435	SVPSARGKICQKLELTVSLLKPLIHTPSE---EPSPVSSPIVDGVALGGLGVALPGSIS
<i>F.oxysporum</i>	428	STPSARAKTCQKLELTVYLLRPLIHT-----APSDAPSPVADGVALGGLGVAMPGSIS
<i>H.jecorina</i>	413	SLPSARSKVCQKLELTVGLLQPLIHTGTN---SPSPKTSPPVGAALGVLGVAMPGSLN
<i>A.kawachii</i>	379	SPPSARGRVCQKLELTIIGLLRPLVHG PATGEASPNYAANM VINGVALGGFG-----VS
<i>A.niger</i>	379	SPPSARGRVCQKLELTIIGLLRPLVHG PATGEASPNYAANM VINGVALGGFG-----VS
<i>A.oryzae</i>	473	SPPSARGRVCQKLELTVGLLRPLIHG PAPGETSPNYAANM VINGVALGGFG-----VS
<i>N.crassa</i>	492	MDALTGETGAFGAAGT LDDVVTYIHLATVVSASEYKGASLRWNA AAWSLARELKLGREIP
<i>F.oxysporum</i>	482	LDATSGESGPFGAAGSLDDVITYIHLAVVVSASEYKGASMRWNTAAWGLARELKLGRELP
<i>H.jecorina</i>	470	MDSLAGETGAFGAIGSLDDVIAYVRLATVVSASEYKGASLRWNGAAWSLARELKLGRELP
<i>A.kawachii</i>	433	MDQLGAQSS---ATGAVDDVATYVHLATVVSASEYKAASMRWNTAAWSLARELKLGRELP
<i>A.niger</i>	433	MDQLGAQSS---ATGAVDDVATYVHLATVVSASEYKAASMRWNTAAWSLARELKLGRELP
<i>A.oryzae</i>	527	MDQLGAQSS---ATGAVDDVATYVHLATVISASEYKAASMRWNTAAWSLARELKLGRELP
<i>N.crassa</i>	552	QN-SPSMQN-----SGSELDGEMGN-----IPGMITEEEREERRRIWWLVYI
<i>F.oxysporum</i>	542	PGPSPATQENM--DTDTADDGEGGTS-----GSGYVGEEREEREERRRIWWLVYI
<i>H.jecorina</i>	530	PGNPPANQEDG--EGLSEDVDEHDLNRNNTRLGRKRSAKSDAITEEEREERRRAWWLVYI
<i>A.kawachii</i>	490	PNVSHARQDGE--RDGDGEADRRHPPTLITSLGHGPGSSGINVTEEEREERRRLWVLLYA
<i>A.niger</i>	490	PNVSHARQDGE--RDGDGEADKRHPPTLITSLGHGSGSSGINVTEEEREERRRLWVLLYA
<i>A.oryzae</i>	584	PNAPQPRQDGEPEDDTDVDMSKRNLPPLITSVGGNSGSTILNVTEEEREERRRLWVLLYA
<i>N.crassa</i>	593	VDRHLALCYNRPLFLLDIECDGLLQPMDDTDYQNGNFYAY---TDPNVLASDPN---TP
<i>F.oxysporum</i>	588	VDRHLALCYNRPLFLLDIECQGLLQPMDDARWQSGDFSGHSNSTTDPNLLGTSPEGYGAD
<i>H.jecorina</i>	588	VDRHLALCYNRPLFLLDSECSLDLHPMDDIKWQAGKFRSHDAGNSSINIDSSMTDEFGDS
<i>A.kawachii</i>	548	TDRHLALCYNRPLTLLDKECGGLLQPMNDDLWQVGDFAAA-----
<i>A.niger</i>	548	TDRHLALCYNRPLTLLDKECGGLLQPMNDDLWQVGDFAAA-----
<i>A.oryzae</i>	644	TDRHLALCYNRPLTLLDKECEGLLQPMNDDLWQAGDFAGA-----
<i>N.crassa</i>	646	AARHRGPSFVCTGHSIFGYFLPLMTILGEIVDLHHAHNHPRFGVGFRSSREWDDQTAEIT
<i>F.oxysporum</i>	648	MTQAHGPQYECRGHSIFGYFLPLMTILGEIVDLHHAHNHPRFGTGFRQGHEWNAQTAEIT
<i>H.jecorina</i>	648	PRAARGAHYECRGHSIFGYFLSLMTILGEIVDVHHAHSHPRFGVGFRSARDWDEQVAEIT
<i>A.kawachii</i>	588	AYRQVGPPVECTGHSMYGYFLPLMTILGGIVDLHHAENHPRFGLAFRNSPEWERQVQDVT
<i>A.niger</i>	588	AYRQVGPPVECTGHSMYGYFLPLMTILGGIVDLHHAENHPRFGLAFRNSPEWERQVLDVT
<i>A.oryzae</i>	684	TYRQVGPPVECTGHSMFGFFLPLMTILGEIVDLQQAKEHPRFGRVFRNSADWDHQVLEIT
<i>N.crassa</i>	706	RHLEIYEESIKRFEHRNLSLSAQQA-ADEKAAEAAGVPTANDVPHDAGTPSVQSVHSHVH
<i>F.oxysporum</i>	708	RHLEIYEQSLQAFEHKNLPRPAEERVDAQNEGNERSGVPDAN-----TPSAHSHVHTNG
<i>H.jecorina</i>	708	RHLDMYEESLKRFAKHLPLSSKDKE--QHEMHDSGAVTDMQ-----SPLSVRTNA
<i>A.kawachii</i>	648	RQLDTYGRSLKEFEARYTSSLTLGTA-ENEPAVEGAHLDHTS-----PSGRSSSTVG
<i>A.niger</i>	648	RQLDTYGRSLKEFEARYTSSLTLGAT-DNEPVEGAHLDHTS-----PSGRSSSTVG
<i>A.oryzae</i>	744	RQLDTYAQSLKEFEARYTSSLALGAG-ESEAAIEGSHLDHVS-----PSGRSTSTAG

<i>N. crassa</i>	765	TTSSRMTESDIQTRIVMAYGTHVMHVLHILLTGKWDPINLLDDNDLWISSQGFITATGHA
<i>F. oxysporum</i>	761	--SNRLTESNIQTRIVIAYGTHVMHVLHILLAGKWDPINLLDDEDLWISSQGFITSTSHA
<i>M. jecorina</i>	757	--SSRMTESEIQASIVVAYSTHVMHVLHILLADKWDPINLLDDDLWISSEGFVTATSHA
<i>A. kawachii</i>	699	---SRVSGSIMHTRMVVAYGTHIMHVLHILLAGKWDPNLLEDHDLWISSESFVSAMSHA
<i>A. niger</i>	699	---SRVSESIHTRMVVAYGTHIMHVLHILLAGKWDPNLLEDHDLWISSESFVSAMSHA
<i>A. oryzae</i>	795	---SRVNESIVHTKMVVAYGTHIMHVLHVLLAGKWDPINLLEDHDLWISSESFIAAMSHA

<i>N. crassa</i>	825	VSAAEAISNILEYDPGLEFMPFFFGIYLLQGSFLLLLIADKLQVEASPSVVKACETIIRA
<i>F. oxysporum</i>	819	VAAAEAIDQILEFDPGLEFMPFFFGIYLLQGSFLLLLIADKLQSEASPSVAKACETIVRA
<i>M. jecorina</i>	815	VSAAEAISQILEFDPGLEFMPFFFGIYLLQGSFLLLLIADKLQAEASPSVIKACETIVRA
<i>A. kawachii</i>	756	VGAAEAAAETILEHDPDLSFMPFFFGIYLLQGSFLLLLAADKLQGDASPSVVRACETIVRA
<i>A. niger</i>	756	VGAAEAAAETILEYDPDLSFMPFFFGIYLLQGSFLLLLAADKLQGDASPSVVRACETIVRA
<i>A. oryzae</i>	852	VGAAADAAADILEYDPDITFMPFFFGIYLLQGSFLLLLAADKLQGDVSPSVVRACETIVRA

<i>N. crassa</i>	885	HEACVVTLNTEYQRNFSRVMSALAQVRGRVPEDLGEQHQRRELLALYRWTGDGTGLAL
<i>F. oxysporum</i>	879	HEACVVTLSLEYQRKFQKVMRSALAQVRGRVPEDLGEQQQRRELLAVYRWTKDGTGLAL
<i>M. jecorina</i>	875	HEACVVTLSLEYQRNFSKVMRSALALIRGRVPEDLAEQQQRRELLALYRWTGNGTGLAL
<i>A. kawachii</i>	816	HEACVVTLNTEYQRTFRKVMRSALAQVRGRIPEDFGEQQQRREVLALYRWSGDGSGGLAL
<i>A. niger</i>	816	HEACVVTLNTEYQRTFRKVMRSALAQVRGRIPEDFGEQQQRREVLALYRWSGDGSGGLAL
<i>A. oryzae</i>	912	HEACVVTLNTEYQRTFRKVMRSALAQVRGRMPEDFGEQQQRREVLALYRWTGDGSGLAL

Figura 7 - Alinhamento múltiplo de seqüências da proteína codificada pela ORF NCU06971 com as proteínas de *F. oxysporum* (ABN41464), *H. jecorina* (AAO33577), *A. oryzae* (BAA95967), *A. kawachii* (BAB62022) e *A. niger* (XP_001397110). O alinhamento de seqüências foi realizado por ClustalW e os resíduos de aminoácidos idênticos (em preto) e as trocas conservativas (em cinza) foram identificados por BoxShade. O sítio de ligação ao DNA (pontilhado) e os NLS (sublinhados) estão representados.

proteínas analisadas tendo 63%, 58%, 53%, 53% e 49% dos aminoácidos idênticos em relação às proteínas dos 5 microrganismos, respectivamente.

Análise do crescimento linear das linhagens selecionadas

Com o objetivo de analisar se a ausência de um determinado fator de transcrição poderia afetar o crescimento celular da linhagem mutante, o crescimento das linhagens selecionadas foi analisado. O crescimento foi avaliado pela taxa de crescimento em meio sólido utilizando *race tubes*, como mostra a figura 8. A figura 9 mostra os resultados do crescimento das linhagens mutantes e da selvagem. Podemos constatar que o crescimento das linhagens mutantes não se diferenciou muito em relação ao crescimento da linhagem selvagem, exceto o apresentado pela linhagem FGSC#11062 (NCU09739). O crescimento médio apresentado por esta linhagem comparado ao apresentado pela linhagem selvagem até o quinto dia foi de aproximadamente 50% nos experimentos realizados em meio VM a 30°C.

Análise da expressão gênica por *Northern blot*

Com o objetivo de avaliar a expressão do gene *gsn* antes e após o choque térmico nas linhagens selecionadas, o RNA total foi extraído a partir dos micélios congelados utilizados na dosagem de glicogênio das linhagens selecionadas, tanto os crescidos à 30°C quanto os submetidos ao estresse térmico por 30 minutos, tendo a linhagem selvagem FGSC# 9718 como referencia. A expressão do gene foi analisada por *Northern blot*, utilizando o transcrito do gene codificador do RNA ribossômico 18 S como controle endógeno. Os resultados são preliminares, mas algumas informações importantes foram obtidas.

A figura 10 mostra o nível do transcrito *gsn* nas seis linhagens mutantes analisadas e na linhagem selvagem antes e depois do choque térmico. Através da figura podemos ver a presença de uma única banda de hibridização com 3,0 kb, a qual corresponde ao tamanho transcrito *gsn*. Podemos notar que o nível transcricional do gene das linhagens mutantes, antes e depois do choque térmico é aproximadamente semelhante ao apresentado pela linhagem selvagem quando submetida às mesmas condições, exceto para a linhagem FGSC#11062 (NCU09739), a qual mostrou que o nível do transcrito antes do choque



Figura 8 - Avaliação do crescimento das linhagens mutantes selecionadas e da selvagem em *race tubes* em meio VM a 30°C.

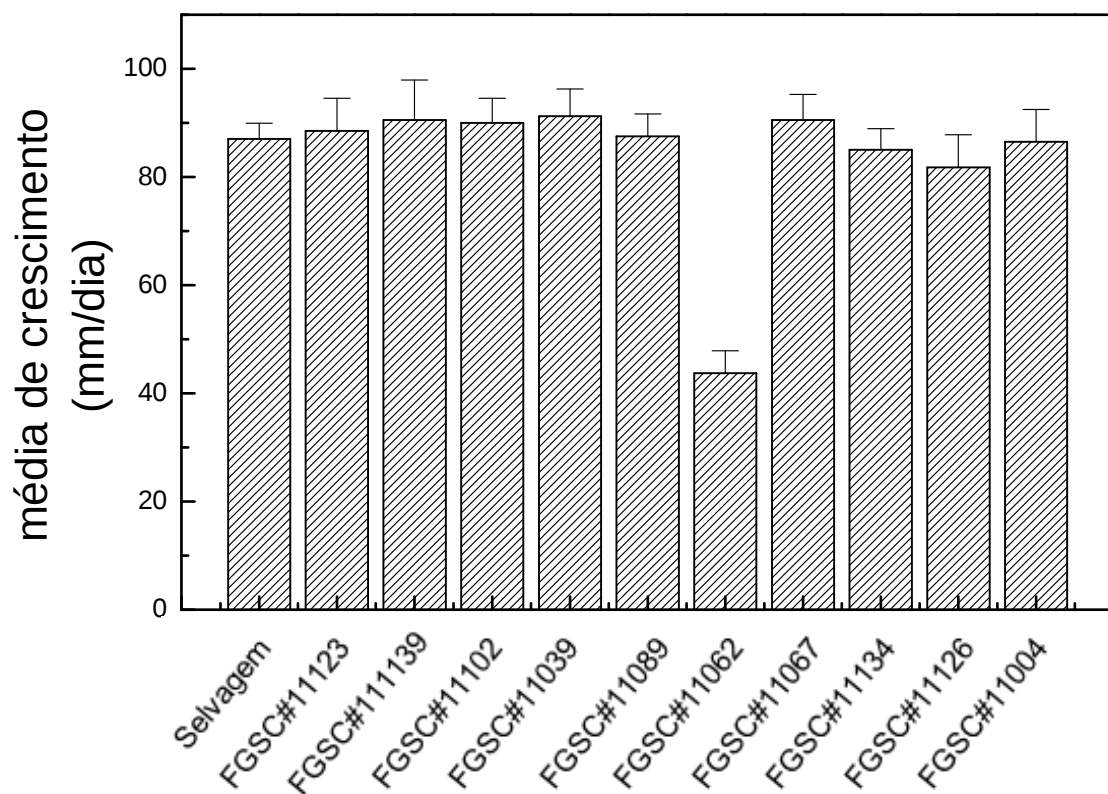


Figura 9 - Média de crescimento avaliado pela extensão das hifas aéreas em *race tubes* das linhagens mutantes selecionadas e da selvagem em meio VM contendo sacarose 2% na temperatura de 30°C ao longo de 5 dias.

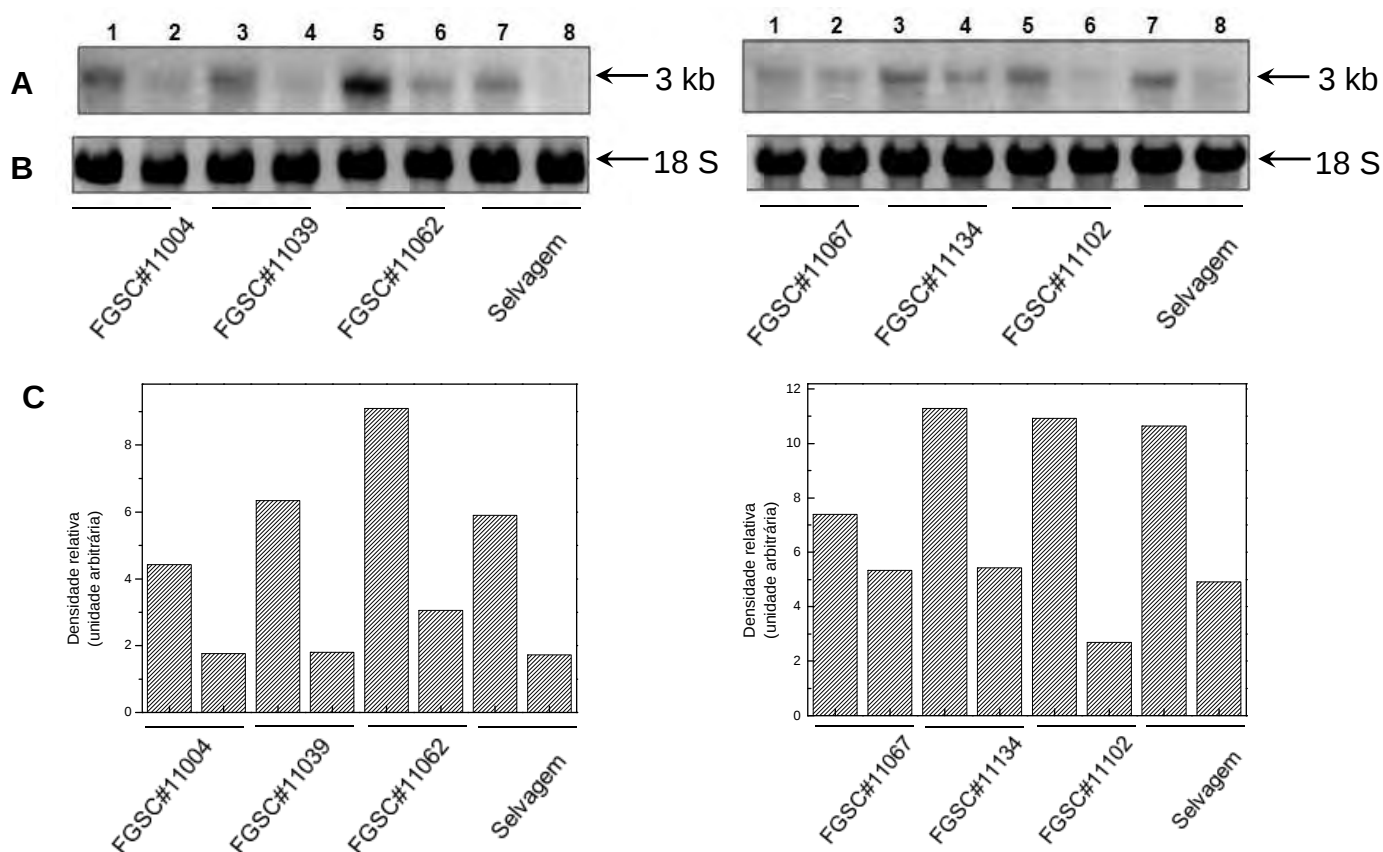


Figura 10 - Expressão do gene *gsn* antes e depois do choque térmico nas linhagens mutantes selecionadas. Amostras de RNA foram separadas por eletroforese em gel de agarose contendo formaldeído e transferidas a uma membrana de nylon neutra. Um fragmento do gene *gsn*, marcado radioativamente com $\alpha^{32}\text{P}$ -ATP foi utilizado como sonda. (A) Autoradiografia dos géis. 1, 3, 5 e 7 representam o RNA total antes do choque térmico. 2, 4, 6, 8 representam o RNA total das linhagens submetidas ao choque térmico (45°C). (B) Perfil do RNAr 18 S visualizado com brometo de etídio. (C) Análise densitométrica da imagem utilizando o software ImageJ (Abramoff *et al.*, 2004).

térmico foi muito superior antes do choque térmico quando comparado as demais linhagens. Outra linhagem mutante interessante é a linhagem FGSC#11067 (NCU06971), a qual mostrou que os níveis do transcrito, antes e após choque térmico são praticamente semelhantes, embora o nível, antes do choque térmico, tenha sido ligeiramente inferior ao apresentado pela linhagem selvagem. Outra linhagem que parece ser interessante é a linhagem FGSC#11134 (NCU04390), pois mostrou um nível do transcrito, após choque térmico ligeiramente superior ao apresentado pela linhagem selvagem. Embora os experimentos necessitam ser repetidos os resultados da expressão do gene, em nível de transcrição parecem indicar que alguns fatores de transcrição possam atuar diretamente sobre a expressão do gene. É importante ressaltar que os níveis do RNA ribossomal 18 S, utilizado como controle endógeno não mostraram alterações tanto antes quanto depois da situação de estresse térmico.

Discussão

A regulação da transcrição é um dos principais pontos de controle da atividade gênica, a qual é responsável pelo controle dos processos biológicos nos organismos eucariotos. Embora a síntese protéica possa ser controlada na célula por meio de uma variedade de mecanismos, alteração do nível transcricional do gene é o processo celular mais direto e utilizado pela célula (Wasserman & Fickett, 1998). Em contraste com os organismos procariotos, onde o processo de regulação transcricional pode ser compreendido por meio da indução por um único fator, nos organismos eucarióticos este processo é principalmente controlado por uma sofisticada interação de múltiplos fatores de transcrição, que se ligam a elementos *cis* distribuídos ao longo de todo o genoma, fornecendo um aumento do controle específico dos genes.

Embora um grande número de projetos genomas tenha sido finalizado, somente uma pequena fração dos mecanismos regulatórios da expressão gênica nos organismos eucariotos é conhecida. Este processo representa a coordenação espacial e temporal da tradução do material genético para proteínas. Para isso, uma complexa maquinaria interpreta os estímulos internos e externos transferindo funções aos fatores de transcrição. Por meio deste processo, as células eucarióticas estão mais bem preparadas para adaptarem-se às alterações ambientais, continuando a exercer suas funções vitais. O conhecimento dos mecanismos regulatórios é essencial para uma compreensão profunda da maioria dos aspectos fundamentais da vida, ou seja, decifrar a linguagem dos elementos *cis* é descrever sua influência no fator de transcrição e no controle do gene alvo. Um conhecimento da posição de elementos regulatórios individuais, bem como sua atuação na regulação gênica é um pré-requisito para a compreensão de como as células trabalham com a finalidade da manutenção da homeostase celular.

O término do seqüenciamento do genoma do fungo *N. crassa*, juntamente com avanços na metodologia de inativação gênica, permitiram o início de estudos mais aprofundados da função e regulação de genes específicos, sob determinadas condições, iniciando análises da genômica funcional do fungo. Pelo fato do genoma de *N. crassa* apresentar um grande número de genes sem ortólogos em outros organismos (Galagan *et al.*, 2003) tem sido realizada uma exaustiva busca por genes que apresentem novas funções ou novos genes desempenhando maneiras diferentes de realizar funções conhecidas (Dunlap *et al.*, 2007). Nosso laboratório vem estudando alguns aspectos da regulação do metabolismo do glicogênio no fungo, pois foi demonstrado que o conteúdo de glicogênio é

diminuído quando o fungo é submetido a uma situação de estresse térmico, fato contrário ao apresentado pela levedura *S. cerevisiae*, portanto, a regulação do metabolismo deste carboidrato se aproxima mais de células de mamíferos. O gene que codifica a proteína glicogênio sintase (GSN) foi clonado em nosso laboratório e a região 5'-flanqueadora do gene *gsn* foi determinada, revelando a existência de elementos *cis* regulatórios envolvidos no controle transcricional de genes cuja transcrição é modulada sob situação de estresse térmico, como por exemplo, vários *Heat Shock Elements* (motif NGAAN) e 2 *Stress Responsive Elements* (motif NAGGGGN) contidos nesta região (de Paula *et al.*, 2002; Freitas *et al.*, 2004). Tentando compreender o modo pelo qual o processo de regulação do gene *gsn* é desencadeado no fungo *N. crassa*, experimentos foram iniciados com o objetivo de identificar proteínas (fatores de transcrição) que se ligam à região promotora do gene *gsn* e regulam a expressão do gene.

A disponibilidade de uma coleção de linhagens de *N. crassa* constituída por mutantes contendo genes codificadores de fatores de transcrição individualmente inativados tornou possível a realização de uma triagem na busca de proteínas envolvidas na regulação de fenótipos específicos e, desta forma, foi iniciada a procura por proteínas regulatórias envolvidas no metabolismo do glicogênio. Trabalho semelhante foi desenvolvido por Wilson *et al.* (2002) no qual foi utilizado um total de 4600 mutantes homozigotos nulos *S. cerevisiae* na busca por genes envolvidos no metabolismo de glicogênio, revelando a atuação de 566 genes na via metabólica do glicogênio, onde 324 mutantes apresentaram baixo acúmulo e 242 mutantes apresentaram elevado acúmulo de glicogênio. É importante ressaltar que no trabalho de Wilson *et al.* (2002) foram utilizadas linhagens contendo genes estruturais inativados e não apenas fatores de transcrição como realizado neste trabalho. Os genes identificados pelos autores codificam proteínas que participam de diversos processos celulares.

No presente trabalho, todas as linhagens mutantes do fungo *N. crassa* (139 linhagens) foram submetidas a uma situação de estresse térmico (45°C) por 30 minutos após crescimento a 30°C por 24 horas. A dosagem de glicogênio antes e depois do choque térmico nestas linhagens revelou a existência de 10 linhagens com o perfil de acúmulo de glicogênio distinto do fenótipo apresentado pela linhagem selvagem, sob a mesma condição (figura 4). Na tentativa de identificar se os fatores de transcrição estavam regulando a transcrição do gene *gsn*, na condição ambiental estudada, experimentos de *Northern blot*

foram realizados. A análise da expressão deste gene foi utilizada, uma vez que os resultados desta metodologia foram bem conhecidos no laboratório. O resultado revelou que das 10 linhagens selecionadas, apenas as linhagens FGSC#11062, FGSC#11067 e FGSC#11134, contendo nocauteadas as ORFs NCU09739, NCU06971 e NCU04390, respectivamente, apresentaram modulação no nível de transcrição do gene *gsn*, sugerindo que estes fatores de transcrição possam estar diretamente envolvidos no controle do metabolismo do glicogênio por meio do controle transcricional do gene *gsn*. Os demais fatores de transcrição selecionados estão envolvidos na regulação do metabolismo, mas possivelmente não no controle transcricional do gene *gsn*. Pelo fato da enzima GSN ser a enzima regulatória do processo de síntese de glicogênio, mas não ser a única envolvida no processo, é possível que os outros fatores de transcrição possam atuar na regulação de outras proteínas que também participam do processo metabólico. Por exemplo, a via metabólica do AMP cíclico controla a expressão gênica de inúmeras proteínas por meio da regulação de mecanismos relacionados à fosforilação de proteínas regulatórias (Smith *et al.*, 1998; Gorner *et al.*, 1998; Hardy *et al.*, 1994). Desta forma, se esta via deixar de ser corretamente controlada, devido à ausência de fatores, ela deixará de funcionar corretamente, interferindo no controle da atividade de outras enzimas que necessitam ser fosforiladas ou desfosforiladas para desempenharem, ou não, sua respectiva função no momento adequado.

A linhagem FGSC#11062 que possui a ORF NCU09739 nocauteada apresentou acúmulo de glicogênio, tanto antes, quanto depois da situação de estresse térmico muito semelhante, ou seja, em ambas as situações esta linhagem apresentou aproximadamente 15 µg de glicogênio/mg de proteína (figura 4). Análise transcricional do gene *gsn* revelou uma superexpressão do transcrito quando crescida a 30°C seguida de uma diminuição quando submetida à situação de estresse térmico, chegando próximo aos níveis apresentados pela linhagem selvagem quando submetida às mesmas condições. Esses resultados sugerem que a proteína codificada pela ORF NCU09739 possa atuar como repressor transcricional do gene *gsn* em situação de crescimento a 30°C, uma vez que em sua ausência, o nível transcricional do gene foi muito maior que o observado na linhagem selvagem. Entretanto, esta correlação não foi observada em relação ao conteúdo de glicogênio nas células crescidas a 30°C, quando o nível de glicogênio foi menor do que o apresentado pela linhagem selvagem na mesma condição. É importante ressaltar que esta

linhagem mutante foi a única que apresentou crescimento vegetativo diferente da linhagem selvagem. É possível que este fator de transcrição atue no controle do metabolismo de glicogênio e em outros processos metabólicos como, por exemplo, crescimento aéreo das hifas e até mesmo na conidiação, pois esta linhagem produz pouquíssimos conídios quando comparado com a linhagem selvagem. Uma vez que a ORF codificadora desta proteína regulatória está nocauteada, os processos nos quais este fator de transcrição está envolvido perderão parte de seus mecanismos regulatórios, podendo apresentar fenótipos muito diferentes dos apresentados pela linhagem selvagem.

Outra linhagem que apresentou padrão transcricional do gene *gsn* distinto da linhagem selvagem foi a FGSC#1134, cuja ORF nocauteada é a NCU04390. Na figura 10, podemos observar que esta linhagem mutante apresentou transcrição semelhante à apresentada pela linhagem selvagem, quando crescida a 30°C, mas após choque térmico houve um aumento da transcrição. Este resultado sugere que o fator de transcrição possa ser um repressor transcricional do gene *gsn* na situação de choque térmico.

A linhagem FGSC#11039 (NCU01097) foi a única que, antes do choque térmico apresentou um hiperacúmulo de glicogênio e, após o choque térmico, apresentou um acúmulo pouco inferior ao apresentado pela linhagem selvagem. Um fato curioso é que os níveis transcricionais do gene *gsn* nesta linhagem, tanto antes quanto depois de submetido à situação de estresse térmico, se assemelham ao apresentado pela linhagem selvagem. Uma hipótese é que este fator de transcrição esteja atuando como ativador transcricional de enzimas responsáveis pela via de degradação do glicogênio, como por exemplo, a glicogênio fosforilase, e não na via de síntese, como analisado neste trabalho, uma vez que sua ausência no fungo provoque um acúmulo anormal de glicogênio pelo fato de que o polímero de glicose não é degradado na célula.

De todos os fatores de transcrição selecionados, apenas dois apresentaram ortólogos funcionais depositados em banco de dados. O primeiro, codificado pela ORF NCU08000 (linhagem FGSC#11004), corresponde ao fator de transcrição da cutinase 1 α (CTF1 α), encontrado em ascomicetos fitopatógenos. A habilidade de parasitar plantas é amplamente difundida entre os fungos ascomicetos. Cutina é um poliéster constituído por ácidos graxos dihidroxi e epóxi ácidos graxos dos tipos *n*-C₁₆ e *n*-C₁₈ (Li *et al.*, 2002; Kolattukudy *et al.*, 1995) e representa o principal revestimento do limbo foliar, sendo responsável pela proteção da epiderme dos vegetais. Para que fungos fitopatógenos

possam atravessar a barreira cuticular e infectar as plantas, é necessário que ocorra a secreção de cutinase, enzima responsável pela hidrólise das ligações ésteres dos ácidos graxos. Em *N. haematococca*, ascomiceto fitopatógeno no qual o mecanismo de ativação do gene da cutinase é bem conhecido, existem 3 genes codificadores de cutinases: *cut1*, *cut2* e *cut3* (Li *et al.*, 2002). Os genes *cut2* e *cut3*, devido à ação do fator de transcrição de cutinase 1 β (CTF1 β) apresentam transcrição constitutiva sob níveis basais (Li *et al.*, 2002). A hidrólise da cutícula da planta ocorrida pela ação do produto dos genes *cut2* e *cut3* do fungo libera monômeros de cutina, fazendo com que a transcrição do gene *cut1* seja fortemente ativada pelo fator de transcrição cutinase 1 α (Li & Kolattukudy, 1997; Li *et al.*, 2002), o qual se liga ao elemento *cis* localizado na posição -159 de sua região promotora (Kämper *et al.*, 1994). A figura 11 representa o mecanismo de ação dos genes envolvidos na via metabólica da cutinase em *F. solani*.

Em *A. nidulans*, as proteínas FarA e FarB, ortólogas às proteínas CTF1 α e CTF1 β , regulam a expressão de genes envolvidos no metabolismo de ácidos graxos (Hynes *et al.*, 2006). Ensaios *in vitro* demonstraram que FarA e FarB interagem com o elemento *cis* CCGAGG presente na região 5' flanqueadora de genes induzidos por ácidos graxos, o mesmo elemento que se liga o fator CTF1 α (Hynes *et al.*, 2006). Isto sugere que o controle transcricional de enzimas envolvidas no metabolismo de ácido graxos possa ser amplamente conservado entre os fungos (Rocha *et al.*, 2008).

O seqüenciamento do genoma de *N. crassa* revelou que, embora *N. crassa* seja um fungo saprófito, seu genoma contém genes similares aos genes requeridos pelos fungos fitopatógenos. Esses genes identificados não possuem funções conhecidas em outros organismos, exceto nos fitopatógenos. *N. crassa* também possui um amplo arsenal de enzimas capaz de digerir polímeros de parede celular, como por exemplo, a cutinase. Além disso, *N. crassa* também possui enzimas relacionadas ao citocromo p450, importante para a desintoxicação por compostos antifúngicos produzidos pelas plantas em alguns sistemas parasita-hopedeiro (Galagan *et al.*, 2003).

Uma particularidade do fator de transcrição codificado pela ORF NCU08000 é a presença de uma inserção de resíduos de glutamina (região poli-Q) que se estende entre os resíduos 728 e 749, região de *coiled coil*, em sua região C-terminal. Regiões poli-Q são extremamente comuns em proteínas relacionadas a doenças neurodegenerativas, como por exemplo, a doença de Huntington (Shao & Diamond, 2007). Estudos envolvendo ativação da

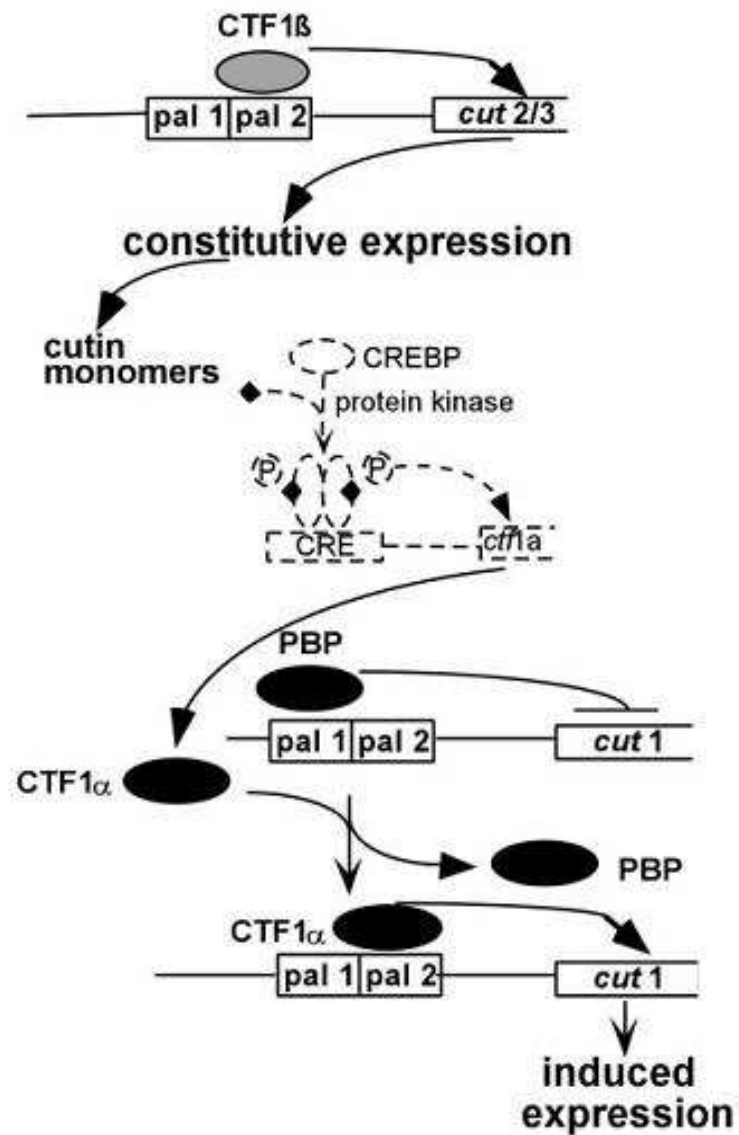


Figura 11 – Modelo esquemático pelo qual ocorre expressão dos genes *cut1*, *cut2* e *cut3* em *N. haematococca*. Os genes *cut2* e *cut3* que codificam enzimas envolvidas na degradação da cutina, possuem transcrição constitutiva a níveis basais, devido à ação do fator de transcrição CTF1β, que se liga em regiões palindrômicas na região promotora dos genes. A presença de monômeros de cutina, resultante da ação de enzimas codificadas por *cut2* e *cut3* sobre a planta, ativam a expressão do gene *cut1*, devido à ação do fator de transcrição CTF1α, elevando a transcrição de *cut1* (Li *et al.*, 2002).

expressão gênica por fatores de transcrição em células de mamíferos demonstraram alguns ativadores transcrpcionais, como por exemplo, o fator de transcrição Sp1, que apresenta 4 regiões que contribuem para a ativação transcricional, sendo que as duas regiões com maior força necessitam de pelo menos 25% de resíduos de glutamina e a minoria de seus resíduos de aminoácidos possuindo carga negativa (Courey & Tjian, 1988; Mitchell & Tjian, 1989).

É importante ressaltar, para efeito comparativo, que o fator de transcrição NIT-4, o qual está relacionado com a expressão dos genes nitrato e nitrito redutases, pertencentes ao metabolismo de nitrogênio no fungo *N.crassa*, também apresenta uma longa seqüência poliQ em sua cadeia polipeptídica, em sua região N-terminal. Estudos demonstraram que quando esta região não está presente na cadeia polipeptídica a proteína NIT-4 perde sua função, demonstrando que esta região possui importante papel no mecanismo de ação desta proteína (Yuan *et al.*, 1991).

Em relação ao conteúdo de glicogênio, esta linhagem mutante (FGSC#11004) apresentou acúmulo, tanto antes quanto depois do choque térmico, praticamente semelhantes. Como o resultado de *Northern blot* mostrou que a transcrição do gene *gsn* está de acordo com o padrão apresentado pela linhagem selvagem sob as mesmas condições, é provável que o fator de transcrição codificado pela ORF nocauteado NCU08000 atue na via de degradação de glicogênio, e não na via de síntese. É possível que este fator de transcrição exerça a função de repressor transcricional da via de degradação do glicogênio, pois em sua ausência o conteúdo de glicogênio linhagem mutante ficou muito próximo do resultado apresentado pela linhagem selvagem sob condições de choque térmico.

Outro fator de transcrição que apresentou ortólogo funcional depositado em banco de dados foi o codificado pela ORF NCU06971, o qual corresponde ao ativador transcricional da xilanase (XlnR), relacionado ao metabolismo de xilana em ascomicetos. A xilana é o principal polissacarídeo constituinte do complexo hemicelulósico das plantas e consiste de uma cadeia principal formada por resíduos de xilopiranosil unidos por ligações β -1,4-glicosídicas (Biely, 1993). As xilanases possuem um destaque especial nas indústrias de papel e celulose, pois diminui a utilização de produtos organoclorados na etapa de branqueamento da polpa de papel ou ainda podem ser utilizadas na reciclagem do papel (Georis *et al.*, 2000; Biley, 1993). Em escala industrial, as xilanases são produzidas,

principalmente, por *Aspergillus* e *Trichoderma spp* (Haltrich *et al.*, 1996). O fungo filamentoso *H. jecorina* apresenta um dos repertórios enzimáticos envolvidos na via metabólica da xilanase mais bem conhecidos (Rauscher *et al.*, 2006). Estudos revelaram que o ativador transcricional XlnR é o controlador central da expressão de mais de 10 genes envolvidos não somente na degradação de xilana, mas também relacionados ao metabolismo de xilose e degradação de celulose (Gielken *et al.*, 1999; Hasper *et al.*, 2000; Hazell *et al.*, 2000). Em *A. oryzae*, este ativador está envolvido na regulação da expressão de pelo menos 4 genes relacionados à degradação de xilana e 4 genes relacionados à degradação de celulose (Marui *et al.*, 2002a; 2002b).

Análise por *Northern blot* da linhagem FGSC#11067, a qual continha a ORF codificadora da proteína XlnR nocauteada (NCU06971), apresentou uma diferença muito sutil do nível transcricional após situação de estresse térmico se comparado com sua expressão quando crescida a 30°C. Parece que a linhagem mutante perdeu a capacidade de regular a transcrição do gene após choque térmico. Isso sugere que o ativador transcricional da xilanase atue no metabolismo de glicogênio diretamente na transcrição do gene *gsn*, uma vez que em sua ausência, a transcrição do gene *gsn*, tanto antes quanto depois da situação de estresse térmico permanece muito semelhante. É importante ressaltar que seu acúmulo de glicogênio tanto antes quanto após o choque térmico foi de aproximadamente o mesmo. Esses resultados sugerem que a proteína codificada pela ORF NCU06971, além de estar relacionada à ativação transcricional de genes envolvidos na degradação de xilana, possivelmente atue como regulador transcricional do gene *gsn*. Foi realizada uma busca por possíveis elementos *cis* para o *XlnR* presentes na região promotora do gene *gsn* de *N. crassa* por meio da ferramenta MatInspector (<http://www.genomatix.de/products/MatInspector/index.html>) o que revelou a presença de um elemento consenso GGCTA/GA localizado a 315 e 320, o qual possivelmente representa seu sítio de interação ao DNA.

Em relação às análises computacionais, foi observado que a massa molecular estimada de todos os dos fatores de transcrição selecionados são superior 40 kDa. Desta forma, todas as proteínas regulatórias selecionadas neste trabalho necessitam do auxílio das importinas para poderem exercer suas funções regulatórias no núcleo. Utilizando a ferramenta PSORTII, uma série de cNLS foram detectados nos fatores de transcrição selecionados, tanto monopartidos quanto bipartidos. Até o momento não existe nenhuma

ferramenta de bioinformática capaz de localizar os NLS não clássicos. Um fato interessante é que na representação esquemática dos fatores de transcrição na figura 5, na qual podemos observar a presença de inúmeros domínios determinados pelo servido SMART-EMBL, pode-se observar regiões em rosa, as quais representam seqüências com estrutura secundária desorganizada. Essas regiões podem ser constituídas por agrupamento de resíduos como, por exemplo, um homopolímero ou até mesmo pequenas repetições, podendo ser aperiódicas, de alguns tipos de resíduos fazendo com que essas regiões apresentem uma baixa complexidade estrutural.

Em humanos e leveduras, os NLSs não clássicos PY-NLS, foram localizados nestas regiões (Lee, *et al.*, 2006; Lange *et al.*, 2008). Além de não possuírem uma estrutura secundária organizada, esses NLS necessitam de uma grande quantidade de resíduos básicos e devem estar localizados *upstream* de um domínio R/K/H-X₍₂₋₅₎-P-Y (X representa qualquer resíduo de aminoácido), localizado na região C-terminal. Nenhum dos fatores de transcrição selecionado de *N. crassa* apresentou PY-NLS propriamente dito, mas a proteína codificada pela ORF NCU09739, a qual não apresentou nenhum NLS não clássico, possui uma grande quantidade de resíduos básicos além de 3 regiões com estrutura secundária desorganizada na região C-terminal. Apesar dos fatores de transcrição identificados neste trabalho apresentarem vários cNLS, é necessário a realização de ensaios de localização celular *in vivo* para confirmar o possível envolvimento dos *motifs* identificados por ensaios *in silico* nos mecanismos de translocação destas proteínas para o interior do núcleo.

Desta forma, o uso de linhagens mutantes contendo genes específicos inativados tem se mostrado um material de extrema importância para o aprofundamento de estudos da genômica funcional do fungo *N. crassa*, principalmente para a compreensão dos mecanismos bioquímicos envolvidos no controle das diversas vias metabólicas no fungo.

Referências bibliográficas

ABRAMOFF, M. D.; MAGELHAES, P. J.; RAM, S. J. Image processing with IMAGEJ. **Biophot. Int.**, v. 11, n. 7, p. 36-42, 2004.

ALMER, A.; RUDOLPH, H.; HINNEN, A.; HÖRZ, W. Removal of positioned nucleosomes from the yeast PHO5 promoter upon PHO5 induction releases additional upstream activating DNA elements. **EMBO J.**, n. 10, p. 2689-2696, 1986.

ALTSCHUL, S. F.; GISH, W.; MILLER, W.; MYERS, E. W.; LIPMAN, D. J. Basic local alignment search tool. **J. Mol. Biol.**, v. 3, n. 215, p. 403-410, 1990.

BATEMMAN, A.; COIN, L.; DURNIN, R.; FINN, R. D.; HOLLICH, V.; GRIFFITHS-JONES, S.; KHANNA, A.; MARSHALL, M.; MOXON, S.; SONNOHAMMER, E. L.; STUDHOLME, D. J.; YEATS, C.; EDDY, S. R. The Pfam protein families database. **Nucleic Acids Res.**, n. 32, p. 138-141, 2004.

BERNSTEIN, B. E.; LIU, C. L.; HUMPHREY, E. L.; PERLSTEIN, E. O.; SCHREIBER, S. L. Global nucleosome occupancy in yeast. **Genome Biol.**, v. 5, n. 9, p. r62, 2004.

BILEY, P. Biochemical aspects of the production of microbial hemicellulases. In: COUGHLAN, M. P.; HAZLEWOOD, G. P. **Hemicellulose and hemicellulases**. London: Portland Press, 1993. p. 29-52.

BIENZ, M.; PELHAM, H. R. Heat shock regulatory elements function as an inducible enhancer in the *Xenopus* hsp70 gene and when linked to a heterologous promoter. **Cell**, v. 5, n. 45, p. 753-760, 1986.

BIENZ, M.; PELHAM, H. R. Mechanisms of the heat shock gene activation in higher eukaryotes. **Adv. Genet.**, n. 24, p. 31-72, 1987.

BISTIS, G. N.; PERKINS, D. D.; READ, N. D. Different cell types in *Neurospora crassa*. **Fungal Genet. Newls.**, n. 50, p. 17-19, 2003.

BORKOVICH, K. A.; ALEX, L. A.; YARDEN, O.; FREITAG, M.; TURNER, G. E.; READ, N. D.; SEILER, S.; BELL-PEDERSEN, D.; PAIETTA, J.; PLESOFKY, N.; PLAMANN, M.; GOODRICH-TANRIKULU, M.; SCHULTE, U.; MANNHAUPT, G.; NARGANG, F. E.; RADFORD, A.; SELITRENNIKOFF, C.; GALAGAN, J. E.; DUNLAP, J. C.; LOROS, J. J.; CATCHESIDE, D.; INOUE, H.; ARAMAYO, R.; POLYMENIS, M.; SELKER, E. U.; SACHS, M. S.; MARZLUF, G. A.; PAULSEN, I.; DAVIS, R.; EBBOLE, D. J.; ZELTER, A.; KALKMAN, E. R.; O'ROURKE, R.; BOWRING, F.; YEADON, J.; ISHII, C.; SUZUKI, K.; SAKAI, W.; PRATT, R. Lessons from the genome sequence of *Neurospora crassa*: tracing the path from genomic blueprint to multicellular organism. **Microbiol. Mol. Biol. Rev.**, v. 1, n. 68, p. 1-108, 2004.

BRANDEN, C.; TOOZE, J. **Introduction to protein structure**. 2nd ed. New York: Garland Publishing, 1999. 410 p.

CHOOK, Y. M.; BLOBEL, G. Karyopherins and nuclear import. **Curr. Opin. Struct. Biol.**, n. 11, p. 703-715, 2001.

COLOT, H.; PARK, G.; JONES, J.; TURNER, G.; BORKOVICH, K.; DUNLAP, J. C. High throughput knockout of transcription factors in *Neurospora* reveals diverse phenotypes. **Proc. Natl. Acad. Sci. USA**, n. 103, p. 10352–10357, 2006.

COUREY, A. J.; TJIAN, R. Analysis of Sp1 *in vivo* reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. **Cell**, n. 55, p. 8878-8898, 1988.

DE PAULA, R. M.; DE PINHO, C. A.; TEREZI, H. F.; BERTOLINI, M. C. Molecular and biochemical characterization of *Neurospora crassa* glycogen synthase encoded by gsn cDNA. **Mol. Genet. Genom.**, v. 2, n. 267, p. 241-253, 2002.

DUNLAP, J. C.; BORKOVICH, K. A.; HENN, M. R.; TURNER, G. E.; SACHS, M. S.; GLASS, N. L.; McCLUSKEY, K.; PLAMANN, M.; GALAGAN, J. E.; BIRREN, B. W.; WEISS, R. L.; TOWNSEND, J. P.; LOROS, J. J.; NELSON, M. A.; LAMBREGHTS, R.; COLOT, H. V.; PARK, G.; COLLOPY, P.; RINGELBERG, C.; CREW, C.; LITVINKOVA, L.; DECAPRIO, D.; HOOD, H. M.; CURILLA, S.; SHI, M.; CEAWFORD, M.; KOERHSEN, M.; MONTGOMERY, P.; LARSON, L.; PEARSON, M.; KASUGA, T.; TIAN, C.; BASTÜRKMEN, M.; ALTAMIANO, L.; XU, J. Enabling a community to dissect an organism: overview of the *Neurospora* functional genomics project. **Adv. Genet.**, n. 57, p. 49-96, 2007.

FREITAS, F. Z.; BERTOLINI, M. C. Genomic organization of the *Neurospora crassa* gsn gene: possible involvement of the STRE and HSE elements in the modulation of transcription during heat shock. **Mol. Gen. Genomics**, v. 5, n. 272, p. 550-561, 2004.

FREITAS, F. Z. **Caracterização funcional dos elementos de transcrição do gene da glicogênio sintase (gsn) de *Neurospora crassa***. 2007. 177 f. Tese (Doutorado em biotecnologia) - Instituto de Química, Universidade Estadual Paulista, Araraquara, 2007.

FREITAS, F. Z.; CHAPEAUROUGE, A.; PERALES, J.; BERTOLINI, M. C. A systematic approach to identify STRE-binding proteins of the gsn glycogen synthase gene promoter in *Neurospora crassa*. **Proteomics**, v. 10, n. 8, p. 2052-2061, 2008.

FRIES, T.; BETZ, C.; SOHN, K.; CAESAR, S.; SCHLENSTEDT, G.; BAILER, S. M. A novel conserved nuclear localization signal is recognized by a group of yeast importins. **J. Biol. Chem.**, v. 27, n. 282, p. 19292-19301, 2007.

GALAGAN, J. E.; CALVO, S. E.; BORKOVICH, K. A.; SELKER, E. U.; READ, N. D.; JAFFE, D.; FITZHUGH, W.; MA, L. J.; SMIRNOV, S.; PURCELL, S.; REHMAN, B.; ELKINS, T.; ENGELS, R.; WANG, S.; NIELSEN, C. B.; BUTLER, J.; ENDRIZZI, M.; QUI, D.; IANAKIEV, P.; BELL-PEDERSEN, D.; NELSON, M. A.; WERNER-WASHBURNE, M.; SELITRENNIKOFF, C. P.; KINSEY, J. A.; BRAUN, E. L.; ZELTER, A.; SCHULTE, U.; KOTHE, G. O.; JEDD, G.; MEWES, W.; STABEN, C.; MARCOTTE, E.; GREENBERG, D.; ROY, A.; FOLEY, K.; NAYLOR, J.; STANGE-THOMANN, N.; BARRETT, R.; GNERRE, S.;

KAMAL, M.; KAMVYSSELIS, M.; MAUCELI, E.; BIELKE, C.; RUDD, S.; FRISHMAN, D.; KRYSTOFOVA, S.; RASMUSSEN, C.; METZENBERG, R. L.; PERKINS, D. D.; KROKEN, S.; COGONI, C.; MACINO, G.; CATCHESIDE, D.; LI, W.; PRATT, R. J.; OSMANI, S. A.; DESOUZA, C. P.; GLASS, L.; ORBACH, M. J.; BERGLUND, J. A.; VOELKER, R.; YARDEN, O.; PLAMANN, M.; SEILER, S.; DUNLAP, J.; RADFORD, A.; ARAMAYO, R.; NATVIG, D. O.; ALEX, L. A.; MANNHAUPT, G.; EBBOLE, D. J.; FREITAG, M.; PAULSEN, I.; SACHS, M. S.; LANDER, E. S.; NUSBAUM, C.; BIRREN, B. The genome sequence of the filamentous fungus *Neurospora crassa*. **Nature**, v. 422, n. 6934, p. 859-868, 2003.

GEORIS, J.; GIANNOTTA, F.; DE BUYL, E.; GRANIER, B.; FRÈRE, J-M. Purification and properties of three endo- β -1,4-xylanases produced by *Streptomyces* sp. Strain S38 which differ in their ability to enhance the bleaching of kraft pulps. **Enzyme Microb. Tech.**, v. 26, p. 78-186, 2000.

GIELKENS, M. M. C.; DEKKERS, E.; VISSER, J.; DE GRAAFF, L. H. Two Cellobiohydrolase-Encoding genes from *Aspergillus niger* require D-Xylose and the xylanolytic transcriptional activator XlnR for their expression. **Appl. Environ. Microbiol.**, n. 65, p. 4340-4345, 1999.

GORNER, W.; DURCHSCHLAG, E.; MARTINEZ-PASTOR, M. T.; ESTRUCH, F., AMMERER, G.; HAMILTON, B.; RUIS, H.; SCHULLER, C. Nuclear localization of the C2H2 zinc finger protein Msn2p is regulated by stress and protein kinase A activity. **Genes Dev.**, n. 12, p. 586–597, 1998.

HALTRICH, D.; NIDETZKY, B.; KULBE, K. D.; STEINER, W.; SILVIA, Z. Production of fungal xylanases. **Biores. Technol.**, n. 2, p.137-161, 1996.

HARDY, T. A.; HUANG, D.; ROACH, P. J. Interactions between cAMP-dependent and SNF1 protein kinases in the control of glycogen accumulation in *Saccharomyces cerevisiae*. **J. Biol. Chem.**, n. 269, p. 27907–27913, 1994.

HARDY, T. A.; ROACH, P. J. Control of yeast glycogen synthase-2 by COOH-terminal phosphorylation. **J. Biol. Chem.**, v. 32, n. 268, p. 23799-23805, 1993.

HARRIS, R. A. Carbohydrate metabolism I: major metabolic pathways and their control. In: DEVLIN, T. M. **Textbook of biochemistry with clinical correlations**. 4th ed. New York: Wiley-Liss, 1997. p. 267-331.

HARTREE, E. F. Determination of protein: a modification of the Lowry method that gives a linear photometric response. **Anal. Biochem.**, v. 2, n. 48, p. 422-427, 1972.

HASPER, A. A.; VISSER, J.; DE GRAAFF, L. H. The *Aspergillus niger* transcriptional activator XlnR, which is involved in the degradation of the polysaccharides xylan and cellulose, also regulates D-xylose reductase gene expression. **Mol. Microbiol.**, n. 36, p. 193-200, 2000.

HAZELL, B. W.; TE´O, V. S.; BRADNER, J. R.; BERGGUIST, P. L.; NEVALAINEN, K. M. Rapid transformation of high cellulase-producing mutant strains of *Trichoderma reesei* by microprojectile bombardment. **Lett. Appl. Microbiol.**, n. 30, p. 282-286, 2000.

HICKS, G. R.; RAIKHEL, N. V. Protein import into the nucleus: an integrated view. **Annu. Rev. Cell. Dev. Biol.**, n. 11, p. 155-188, 1995.

HYNES, M. J.; MURRAY, S. L.; DUNCAN, A.; KHEW, G. S.; DAVIES, M. A. Regulatory genes controlling fatty acid catabolism and peroxisomal functions in the filamentous fungus *Aspergillus nidulans*. **Eukaryot. Cell.**, v. 5, n. 5, p. 794-805, 2006.

JOHNSTON, M. A model fungal gene regulatory mechanism: the GAL genes of *Saccharomyces cerevisiae*. **Microbiol. Rev.**, v. 4, n. 51, p. 458–476, 1987.

KÄMPER, J. T.; KÄMPER, U.; ROGERS, L. M.; KOLATTUKUDY, P. E. Identification of regulatory elements in the cutinase promoter from *Fusarium solani* f. sp. *pisi* (*Nectria haematococca*). **J. Biol. Chem.**, n. 269, p. 9195-9204, 1994.

KANG, S.; METZENBERG, R. L. Molecular analysis of *nuc-1+*, a gene controlling phosphorus acquisition in *Neurospora crassa*. **Mol. Cell. Biol.**, n. 10, p. 5839–5848, 1990.

KOLATTUKUDY, P. E.; ROGERS, L. M.; LI, D. X.; HWANG, C. S.; FLAISHMAN, M. A. Surface signaling in pathogenesis. **Proc. Natl. Acad. Sci. USA**, n. 92, p. 4080-4087, 1995.

KORNBERG, R. Chromatin structure: a repeating unit of histones and DNA. **Science**, n. 184, p. 868-871, 1974.

LAITY, J. H.; LEE, B. M.; WRIGHT, P. E. Zinc finger proteins: new insights into structural and functional diversity. **Curr. Opin. Struct. Biol.**, n. 11, p. 39–46, 2001.

LANGE, A.; MILS, R. E.; DEVINE, S. E.; CORBETT, A. H. A PY-NLS nuclear targeting signal is required for nuclear localization and function of the *Saccharomyces cerevisiae* mRNA-binding protein Hrp1. **J. Biol. Chem.**, n. 283, p. 12926-12934, 2008.

LATCHMAN D. S. **Eukaryotic transcription factors**. 4th ed. New York: Elsevier, 2003. 384p.

LEE, B. J.; CANSIZOGLU, A. E.; SÜEL, K. E.; LOUIS, T. H.; ZHANG, Z.; CHOOK. Y. M. Rules for nuclear localization sequence recognition by karyopherin β 2. **Cell**, n. 126, p. 543-558, 2006.

LEE, C. K.; SHIBATA, Y.; RAO, B.; STRAHL, B. D.; LIEB, J. D. Evidence for nucleosome depletion at active regulatory regions genome-wide. **Nat. Genet.**, v. 8, n. 36, p. 900-1005, 2004.

LEVINE, M.; TJIAN, R. Transcription regulation and animal diversity. **Nature**, v. 6945, n. 424, p. 147-151, 2003.

LI, B.; CAREY, M.; WORKMAN, J. L. The role of chromatin during transcription **Cell**, v. 4, n. 128, p. 707-719, 2007.

LI, D.; KOLATTUKUDY, P. E. Cloning of cutinase transcription factor 1, a transactivating protein containing Cys6Zn2 binuclear cluster DNA-binding motif. **J. Biol. Chem.**, v. 19, n. 272, p. 12462-12467, 1997.

LI, D.; SIRAKOVA, T.; ROGERS, L.; ETTINGER, W. F.; KOLATTUKUDY, P. E. Regulation of constitutively expressed and induced cutinase genes by different zinc finger transcription factors in *Fusarium solani* f. sp. *pisi* (*Nectria haematococca*). **J. Biol. Chem.**, n. 277, p. 7905-7912, 2002.

LUSCOMBE, N. M.; AUSTIN, S. E.; BERMAN, H. M.; THORNTON, J. M. An overview of the structures of protein-DNA complexes. **Genome Biol.**, n. 1, p. 1-37, 2000.

LUGER, K.; MÄDER, A. W.; RICHMOND, R. K.; SARGENT, D. F.; RICHMOND, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. **Nature**, v. 6648, n. 389, p. 251-260, 1997.

MARTINEZ-PASTOR, M.; MARCHLER, G.; SCHULLER, C.; MARCHLER, B. A.; RUIS, H.; ESTRUCH, F. The *Saccharomyces cerevisiae* zinc finger proteins Msn2p and Msn4p are required for transcriptional induction through the stress response element (STRE). **EMBO J.**, v. 9, n. 15, p. 2227-2235, 1996.

MARUI, J.; KITAMOTO, N.; KATO, M.; KOBAYASHI, T.; TSUKAGOSHI, N. Transcriptional activator, AoXlnR, mediates cellulose-inductive expression of the xylanolytic and cellulolytic genes in *Aspergillus oryzae*. **FEBS Lett.**, n. 528, p. 279-282, 2002.

MARUI, J.; TANAKA, A.; MIMURA, S.; DE GRAFF, L. H.; VISSER, J.; KITAMOTO, N.; KATO, M.; KOBAYASHI, T.; TSUKAGOSHI, N. A transcriptional activator, AoXlnR, controls the expression of genes encoding xylanolytic enzymes in *Aspergillus oryzae*. **Fungal Genet. Biol.**, n. 35, p. 157-169, 2002.

McGHEE, J. D.; FELSENFELD, G. Nucleosome structure. **Annu. Rev. Biochem.**, n. 49, p. 1115-1156, 1980.

MERIKA, M.; ORKIN, S. H. DNA-binding specificity of GATA family transcription factors. **Mol. Cell. Biol.** v. 7, n. 13, p. 3999-4010, 1993.

MITCHELL, P. J.; TJIAN, R. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. **Science**, n. 245, p. 371-378, 1989.

MURRE, C.; BAIN, G.; VAN DIJK, M. A.; ENGEL, I.; FURNARI, B. A.; MASARI, M. E.; MATTHEWS, J. R.; QUONG, M. W.; RIVERA, R. R.; STUIVER, M. H. Structure and function of helix-loop-helix proteins. **Biochim. Biophys. Acta.**, n. 1218, p. 129-135, 1994.

NAKAI, K.; HORTON, P. PSORT: a program for detecting the sorting signals of proteins and predicting their subcellular localization. **Trends Biochem. Sci.**, v. 1, n. 24, p. 34-35, 1999.

NI, H. T.; LAPORTE, D. C. Response of a yeast glycogen synthase gene to stress. **Mol. Microbiol.**, v. 6, n. 16, p. 1197-1205, 1995.

NINOMIYA, Y.; SUZUKI, K.; ISHII, C.; INOUE, H. Highly efficient gene replacements in *Neurospora* strains deficient for nonhomologous end-joining. **Proc. Natl. Acad. Sci. U S A.**, n. 101, p. 12248-12253, 2004.

NOVENTA-JORDÃO, M. A.; POLIZELI, M. L. T. M.; BONINI, B. M.; JORGE, J. A.; TERENCE, H. F. Effects of temperature shifts on the activities of *Neurospora crassa* glycogen synthase, glycogen phosphorylase and trehalose-6-phosphate synthase. **FEBS Lett.**, n. 378, p. 32-36, 1996.

PAVLETICH, N. P.; PABO, C. O. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. **Science**, v. 5007, n. 252, p. 809-817, 1991.

PELHAM, H. B. R.; BIENZ, M. A synthetic heat shock promoter element confers heat inducibility on the herpes simplex virus *tk* gene. **EMBO J.**, v. 11, n. 1, p. 1473-1477, 1992.

PEMBERTON, L. F.; BLOBEL, G.; ROSENBLUM, J. S. Transport routes through the nuclear pore complex. **Curr. Opin. Cell Biol.**, n. 10, p. 392-399, 1998.

RAUSCHER, R.; WÜRLEITNER, E.; WACENOVSKY, C.; ARO, N.; STRICKER, A. R.; ZEILINGER, S.; KUBICEK, C. P.; PENTTILÄ, M.; MACH, R. L. Transcriptional regulation of *xyn1*, encoding xylanase I, in *Hypocrea jecorina*. **Eukaryot. Cell**, n. 5, p. 447-56, 2006.

RIEDL, T.; EGLY, J. M. Phosphorylation in transcription: the CTD and more. **Gene Expr.**, v. 1-2, n. 9, p. 3-13, 2000.

ROBBINS, J.; DILWORTH, S. M.; LASKEY, R. A.; DINGWALL, C. Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. **Cell**, n. 64, p. 615-623, 1991.

ROCHA, A. L. M.; DI PIETRO, A.; RUIZ-ROLDÁN, C.; RONCERO, I. G. Ctf1, a transcriptional activator of cutinase and lipase genes in *Fusarium oxysporum* is dispensable for virulence **Mol. Plant Pathol.**, v. 3, n. 9, p. 293-304, 2008.

SAMBROOK, J.; RUSSELL, D. W. **Molecular cloning, a laboratory manual**. Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 2001.

SCHULTE, U.; BECKER, I.; MEWES, H. W.; MANNHAUPT, G. Large scale analysis of sequences from *Neurospora crassa*. **J. Biotechnol.**, n. 94, p. 3-13, 2002.

SCHULTZ, J.; MILPETZ, F.; BORK, P.; PONTING, C. P. SMART, a simple modular architecture research tool: identification of signaling domains. **Proc. Natl. Acad. Sci. U S A.**, v. 11, n. 95, p. 5857-5864, 1998.

SEKINGER, E. A.; MOGTADERI, Z.; STRUHL, K. Intrinsic histone-DNA interactions and low nucleosome density are important for preferential accessibility of promoter regions in yeast. **Mol. Cell**, v. 6, n. 18, p. 735-748, 2005.

SHAO, J.; DIAMOND, M. I. Polyglutamine diseases: emerging concepts in pathogenesis and therapy. **Hum. Mol. Genet.**, n. 2, p. 115-23, 2007.

SMITH, A.; WARD, M. P.; GARRETT, S. Yeast PKA represses Msn2p/Msn4p-dependent gene expression to regulate growth, stress response and glycogen accumulation. **EMBO J.**, n. 17, p. 3556–3564, 1998.

SOKOLOVSKY, V.; KALDENHOFF, R.; RICCI, M.; RUSSO, V. E. A. Fast and reliable mini-prep RNA extraction from *Neurospora crassa*. **Fungal Genet. Newls.**, v. 67, p. 41–43, 1995.

THEVELEIN, J. M. Regulation of trehalose mobilization in fungi. **Microbiol. Reviews**, n. 48, p. 42-59, 1984.

TODD, R. B.; AANDRIOPOULOS, A. Evolution of a fungal regulatory gene family: the Zn(II)₂Cys₆ binuclear cluster DNA binding motif. **Fungal Genet. Biol.**, n. 21, p. 388-405, 1997.

VALLIM, M. A.; MILLER, K. Y.; MILLER, B. L. *Aspergillus* SteA (Sterile 12-like) is a homeodomain-C2/H2-Zn²⁺ finger transcription factor required for sexual reproduction. **Mol. Microbiol.**, n. 36, p. 290–301, 2000.

VARELA, J. C.; PRAEKELT, U. M.; MEACOCK, P. A.; PLANTA, R. J.; MAGER, W. H. The *Saccharomyces cerevisiae* HSP12 gene is activated by the high-osmolarity glycerol pathway and negatively regulated by protein kinase A. **Mol. Cell. Biol.**, v. 11, n. 15, p. 6232-6245, 1995.

VASHEE, S.; XU, H.; JOHNSTON, S. A.; KODADEK, T. How do "Zn2cys6" proteins distinguish between similar upstream activation sites? Comparison of the DNA-binding specificity of the GAL4 protein *in vitro* and *in vivo*. **J. Biol. Chem.**, v. 33, n. 268, p. 24699-24706, 1993.

VOGEL, H. J. A convenient medium for *Neurospora crassa* (N medium). **Microbiol. Genet. Bulletin**, v. 13, p. 42-43, 1956.

WASSERMAN, W.; FICKETT, J. Identification of regulatory regions which confer muscle-specific gene expression. **J. Mol. Biol.**, n. 278, p. 167-181, 1998.

WASYLYK, B.; CHAMBON, P. Transcription by eucaryotic RNA polymerases A and B of chromatin assembled *in vitro*. **Eur. J. Biochem.** n. 98, p. 317-327, 1979.

WILSON, W. A.; WANG, Z.; ROACH, P. J. Analysis of respiratory mutants reveals new aspects of the control of glycogen accumulation by the cyclin-dependent protein kinase Pho85p. **FEBS Lett.**, n. 515, p. 104-108, 2002.

WOLFFE, A. P.; HAYES, J. J. Chromatin disruption and modification. **Nucleic Acids Res.**, v. 3, n. 27, p. 711-720, 1999.

YUAN, G. F.; FU, Y. H.; MARZLUF, G. A. nit-4, a pathway-specific regulatory gene of *Neurospora crassa*, encodes a protein with a putative binuclear zinc DNA-binding domain. **Mol. Cell Biol.**, v. 11, n. 11, p. 5735-5745, 1991.

YOUNG, R. A. RNA polymerase II. **Annu. Rev. Biochem.**, n. 60, p. 689-715, 1991.

Anexos

Tabela 3: Total de linhagens mutantes do fungo *N. crassa* utilizadas no presente trabalho.

FGSC#	ORF K.O.	Fator de acasalamento	Família do fator de transcrição nocauteado
11001	NCU06799.2	a	Zn ₂ Cys ₆
11002	NCU01478.2	a	Zn ₂ Cys ₆
11003	NCU02934.2	a	Zn ₂ Cys ₆
11004	NCU08000.2	a	Zn ₂ Cys ₆
11005	NCU08000.2	A	Zn ₂ Cys ₆
11006	NCU07007.2	a	Zn ₂ Cys ₆
11007	NCU08294.2	A	Zn ₂ Cys ₆
11008	NCU08294.2	a	Zn ₂ Cys ₆
11009	NCU08652.2	a	Zn ₂ Cys ₆
11010	NCU08652.2	A	Zn ₂ Cys ₆
11011	NCU08651.2	A	Zn ₂ Cys ₆
11012	NCU08651.2	a	Zn ₂ Cys ₆
11013	NCU06656.2	A	Zn ₂ Cys ₆
11014	NCU06656.2	a	Zn ₂ Cys ₆
11015	NCU02752.2	a	Zn ₂ Cys ₆
11016	NCU07374.2	a	Zn ₂ Cys ₆
11017	NCU06407.2	a	Zn ₂ Cys ₆
11018	NCU06407.2	A	Zn ₂ Cys ₆
11019	NCU05383.2	a	Zn ₂ Cys ₆
11020	NCU00217.1	A	Zn ₂ Cys ₆
11021	NCU00217.1	a	Zn ₂ Cys ₆
11022	NCU04866.2	a	Zn ₂ Cys ₆
11023	NCU04866.2	A	Zn ₂ Cys ₆

11024	NCU03110.2	a	Zn ₂ Cys ₆
11025	NCU05994.2	a	Zn ₂ Cys ₆
11026	NCU05994.2	A	Zn ₂ Cys ₆
11027	NCU05536.2	a	Zn ₂ Cys ₆
11028	NCU07705.2	A	Zn ₂ Cys ₆
11029	NCU07705.2	a	Zn ₂ Cys ₆
11030	NCU07788.2	a	Zn ₂ Cys ₆
11031	NCU07788.2	A	Zn ₂ Cys ₆
11032	NCU06990.2	A	Zn ₂ Cys ₆
11033	NCU06990.2	a	Zn ₂ Cys ₆
11034	NCU06028.2	a	Zn ₂ Cys ₆
11035	NCU06028.2	A	Zn ₂ Cys ₆
11036	NCU03120.2	a	Zn ₂ Cys ₆
11037	NCU03120.2	A	Zn ₂ Cys ₆
11038	NCU01097.2	a	Zn ₂ Cys ₆
11039	NCU01097.2	A	Zn ₂ Cys ₆
11040	NCU05411.2	a	Zn ₂ Cys ₆
11041	NCU07392.2	a	Zn ₂ Cys ₆
11042	NCU07392.2	A	Zn ₂ Cys ₆
11043	NCU08848.2	a	Zn ₂ Cys ₆
11044	NCU08726.2	a	Zn ₂ Cys ₆
11045	NCU04359.2	a	Zn ₂ Cys ₆
11046	NCU04359.2	A	Zn ₂ Cys ₆
11047	NCU08049.2	a	Zn ₂ Cys ₆
11048	NCU08899.2	a	Zn ₂ Cys ₆
11049	NCU03643.2	A	Zn ₂ Cys ₆

11050	NCU03643.2	a	Zn ₂ Cys ₆
11051	NCU05767.1	A	Zn ₂ Cys ₆
11052	NCU05767.1	a	Zn ₂ Cys ₆
11053	NCU03931.2	a	Zn ₂ Cys ₆
11054	NCU02307.2	a	Zn ₂ Cys ₆
11055	NCU07139.2	a	Zn ₂ Cys ₆
11056	NCU07945.2	A	Zn ₂ Cys ₆
11057	NCU07945.2	a	Zn ₂ Cys ₆
11058	NCU03320.2	a	Zn ₂ Cys ₆
11059	NCU08658.1	a	Zn ₂ Cys ₆
11060	NCU02094.2	a	Zn ₂ Cys ₆
11061	NCU08443.2	a	Zn ₂ Cys ₆
11062	NCU09739.2	a	Zn ₂ Cys ₆
11063	NCU06068.2	a	Zn ₂ Cys ₆
11064	NCU00945.2	a	Zn ₂ Cys ₆
11065	NCU09829.2	a	Zn ₂ Cys ₆
11066	NCU06971.2	a	Zn ₂ Cys ₆
11067	NCU06971.2	A	Zn ₂ Cys ₆
11068	NCU02214.2	A	Zn ₂ Cys ₆
11069	NCU02214.2	a	Zn ₂ Cys ₆
11070	NCU02896.2	a	Zn ₂ Cys ₆
11071	NCU02142.2	a	Zn ₂ Cys ₆
11072	NCU02576.2	a	Zn ₂ Cys ₆
11073	NCU04001.2	a	Zn ₂ Cys ₆
11074	NCU05294.2	a	Zn ₂ Cys ₆
11075	NCU00017.2	a	Zn ₂ Cys ₆

11076	NCU03686.2	a	Zn ₂ Cys ₆
11077	NCU03686.2	A	Zn ₂ Cys ₆
11078	NCU05993.2	A	Zn ₂ Cys ₆
11079	NCU05993.2	a	Zn ₂ Cys ₆
11080	NCU09804.2	a	Zn ₂ Cys ₆
11081	NCU09804.2	A	Zn ₂ Cys ₆
11082	NCU03417.1	a	Zn ₂ Cys ₆
11083	NCU03417.1	A	Zn ₂ Cys ₆
11084	NCU09549.2	a	Zn ₂ Cys ₆
11085	NCU00289.2	a	Zn ₂ Cys ₆
11086	NCU00289.2	A	Zn ₂ Cys ₆
11087	NCU08901.2	A	Zn ₂ Cys ₆
11088	NCU08901.2	a	Zn ₂ Cys ₆
11089	NCU04851.2	a	Zn ₂ Cys ₆
11090	NCU02768.2	A	Zn ₂ Cys ₆
11091	NCU02768.2	a	Zn ₂ Cys ₆
11092	NCU08063.2	A	Zn ₂ Cys ₆
11093	NCU08063.2	a	Zn ₂ Cys ₆
11094	NCU07535.2	A	Zn ₂ Cys ₆
11095	NCU03489.2	a	Zn ₂ Cys ₆
11096	NCU09205.2	a	Zn ₂ Cys ₆
11097	NCU05051.2	a	Zn ₂ Cys ₆
11098	NCU09529.2	A	Zn ₂ Cys ₆
11099	NCU09529.2	a	Zn ₂ Cys ₆
11100	NCU08289.2	a	Zn ₂ Cys ₆
11101	NCU01629.2	A	C ₂ H ₂ zinc finger

11102	NCU01629.2	a	C ₂ H ₂ zinc finger
11103	NCU00694.2	a	C ₂ H ₂ zinc finger
11104	NCU05909.2	a	C ₂ H ₂ zinc finger
11105	NCU06919.2	A	C ₂ H ₂ zinc finger
11106	NCU06919.2	a	C ₂ H ₂ zinc finger
11107	NCU03073.2	a	Outros
11108	NCU02017.1	a	Outros
11109	NCU03070.1	a	Outros
11110	NCU00097.2	a	Outros
11111	NCU00097.2	A	Outros
11112	NCU03962.2	A	Outros
11113	NCU03962.2	a	Outros
11114	NCU07561.2	a	Outros
11115	NCU07561.2	A	Outros
11116	NCU06411.2	a	Outros
11117	NCU00233.2	a	B-ZIP
11118	NCU00285.2	a	C ₂ H ₂ zinc finger
11119	NCU00329.2	a	B-ZIP
11120	NCU00499.2	A	B-ZIP
11121	NCU00499.2	a	B-ZIP
11122	NCU00808.2	a	Zn ₂ Cys ₆
11123	NCU00808.2	A	Zn ₂ Cys ₆
11124	NCU00902.2	a	GATA
11125	NCU01122.2	a	C ₂ H ₂ zinc finger
11126	NCU01154.2	A	GATA
11127	NCU01154.2	a	GATA

11128	NCU03356.2	a	Outros
11129	NCU03593.2	a	Outros
11130	NCU03699.2	a	C C ₂ H ₂ zinc finger
11131	NCU03905.2	a	B-ZIP
11132	NCU04179.2	a	C ₂ H ₂ zinc finger
11133	NCU04211.2	a	B-ZIP
11134	NCU04390.2	a	Zn ₂ Cys ₆
11135	NCU04561.2	A	C ₂ H ₂ zinc finger
11136	NCU04561.2	a	C ₂ H ₂ zinc finger
11137	NCU04619.1	a	C ₂ H ₂ zinc finger
11138	NCU04628.1	A	C ₂ H ₂ zinc finger
11139	NCU04731.2	a	bHLH