

BUSCA POR HOMÓLOGAS À PROTEÍNA HIPOTÉTICA CNA00530 DE *Cryptococcus neoformans* JEC21 UTILIZANDO O BASIC LOCAL ALIGNMENT SEARCH TOOL – BLAST

SEARCH FOR HOMOLOGISTS OF HYPOTHETICAL PROTEIN CNA00530 OF *Cryptococcus neoformans* JEC21 USING BASIC LOCAL ALIGNMENT RESEARCH TOOL - BLAST

¹PAVAN, Andressa Aparecida; ²GARCIA, Anderson

¹Discente do Curso de Ciências Biológicas – Faculdades Integradas de Ourinhos - FIO/FEMM

²Professor Orientador - Curso de Ciências Biológicas – Faculdades Integradas de Ourinhos - FIO/FEMM

RESUMO

Cryptococcus neoformans é um fungo leveduriforme causador da criptococose, doença responsável por 10% das mortes no mundo. Embora *C. neoformans* tenha seu genoma sequenciado, grande parte seus genes e proteínas codificadas por eles ainda permanecem sem função conhecida, sendo assim anotados como hipotéticos. O presente trabalho teve como objetivo investigar a ocorrência de PROTEÍNAS homólogas ou ortólogas à proteína hipotética CNA00530 de *Cryptococcus neoformans* JEC21 com função já descrita. Para isto, foi realizada uma análise comparativa com PROTEÍNAS depositados no banco de dados do *National Center for Biotechnology Information* (NCBI), por meio da ferramenta *Basic Local Alignment Search Tool* (BLAST). Os resultados apontaram que a PROTEÍNA hipotética CNA00530 apresenta de 73 a 95% de homologia com a subunidade regulatória N3 do proteasoma 26S encontrados outras espécies do gênero. Decorrente dos dados apresentados, propõe-se que CNA00350 possa exercer a mesma função de suas homólogas encontradas na análise comparativa.

Palavras-chave: *Cryptococcus neoformans*. Proteínas hipotéticas. Proteasoma.

ABSTRACT

Cryptococcus neoformans is a yeast that causes cryptococcosis, a disease responsible for 10% of deaths in the world wide. Although *C. neoformans* has its genome sequenced, much of its genes and proteins encoded by them still remain without known function, thus being annotated as hypothetical. The present work had as objective to investigate the occurrence of homologous or orthologous proteins to the hypothetical protein CNA00530 of *Cryptococcus neoformans* JEC21 with function already described. For this, a comparative analysis was performed with proteins deposited in the National Center for Biotechnology Information (NCBI) database, using the Basic Local Alignment Search Tool (BLAST). The results indicated that the hypothesized CNA00530 protein presents 73-95% homology with the N3 regulatory subunit of the 26S proteasome found in other species of the genus. Based on the presented data, we proposed that CNA00350 can exert the same function of its homologues found in the comparative analysis.

Keywords: *Cryptococcus neoformans*. Hypothetical proteins. Proteasome.

INTRODUÇÃO

Cryptococcus consiste em um gênero de fungos leveduriformes encapsulados com mais de 70 espécies podendo ser encontradas dispersamente em ambientes úmidos. A espécie *neoformans* possui variedades A, D e AD, que podem colonizar a mucosa do papo dos pombos de várias partes do mundo e são também encontrados em suas fezes (RIBAS et al., 2011; FARIA et al., 2010; FILIÚ et al., 2002).

O *Cryptococcus neoformans* tem distribuição mundial e atuam como patógeno oportunista, sendo responsável pela criptococose, que é uma infecção que ataca primeiramente o pulmão, causando micose pulmonar, e tardiamente pode migrar para o sistema nervoso central (KON et al., 2008; REOLON et al., 2004; DAHER, 2015).

O contágio é realizado pela inalação dos basidiósporos ou leveduras desidratadas infectantes no ambiente e ingestão de alimentos contaminados com excretas de pombos (ARAUJO et al., 2015).

Os atuais avanços no campo de estudo da micologia, como abordagem e aplicação genômica, prometem revolucionar os estudos sobre a patogenicidade fúngica através da busca por genes e proteínas envolvidas na infecção (SAVILLE et al., 2005).

O *Cryptococcus neoformans*, embora tenha seu genoma sequenciado, muitos de seus genes ainda permanecem sem função conhecida ou atribuída e desta forma, continuam anotados como genes hipotéticos nos bancos de dados. As proteínas oriundas destes genes são também chamadas hipotéticas e constituem 25% de qualquer genoma recém-sequenciado e atribuir função a esses genes e seus produtos, é um dos desafios da era pós-genômica (TAN et al., 2014).

O avanço das técnicas de sequenciamento de DNA tem gerado um acúmulo de dados de sequências de genes e genomas que não avançam da mesma forma como a elucidação de suas funcionalidades. A homologia de sequências contribui para predição da função do gene e através da comparação entre sequências é possível prever a sua função (SMITH, 1996; CLAVERIE, 1997; DESLER et al., 2009; DOERKS et al., 2012).

As proteínas possuem um ou mais domínios que são unidades funcionais e parte da cadeia peptídica estável. Através da identificação dos domínios da proteína pode-se procurar domínios semelhantes em outras proteínas e propor funções semelhantes para os domínios achados. (LODISH et al., 2014; SILVA, 2014).

Desta maneira, presente trabalho teve como objetivo investigar a ocorrência de proteínas homólogas ou ortólogas a proteínas hipotética CNA00530, com função já descrita, contribuindo assim como a predição da sua função *in silico*.

MATERIAL E MÉTODOS

Para a escolha da proteína alvo deste estudo, foi realizada a busca por PROTEÍNAS hipotéticas no genoma do fungo *Cryptococcus neoformans* var. *neoformans* JEC21 depositado no banco de dados genômico do *National Center for Biotechnology Information* (NCBI) e escolhida aleatoriamente a PROTEÍNA hipotética CNA00530, cujo gene está localizado no cromossomo 1.

Para a busca por proteínas homólogas à CNA00530 utilizou-se o *Basic Local Alignment Search Tool* (BLAST) no modo proteína (BLASTp). O BLASTp faz a comparação da sequência de estudo com sequências de proteínas depositadas no banco de dados e gera um arquivo comparativo entre a homologia.

Com os resultados obtidos, foi confeccionada uma tabela com as dez primeiras proteínas com alto grau de homologia encontrada e realizada a análise e discussão.

RESULTADOS E DISCUSSÃO

Na Tabela 1, estão listados os resultados gerados a partir da comparação da sequência primária de CNA00530 com as dez proteínas homólogas encontradas pelo BLASTp. É possível notar grande homologia entre a proteína de estudo com outras proteínas de fungos também do gênero *Cryptococcus* (representada pela sigla SI), variando o grau de 73 a 95%. As homólogas encontradas exercem a função de subunidades regulatórias N3 do proteasoma 26S.

Tabela 1. Comparação da homologia entre as dez primeiras proteínas encontradas pelo BLAST com a proteína hipotética CNA00530.

Número de acesso	Título/função	Organismo	*SI (%)	*SS	E-value
XP_566480.1	hypothetical protein CNA00530	<i>Cryptococcus neoformans</i> var. <i>neoformans</i> JEC21	100	100	0.0
XP_012046502.1	26S proteasome regulatory subunit N3	<i>Cryptococcus neoformans</i> var. <i>grubii</i> H99	95	100	0.0
XP_003191890.1	26S proteasome regulatory subunit N3	<i>Cryptococcus neoformans</i> var. <i>grubii</i> H99	95	100	0.0
KIR83082.1	26S proteasome regulatory subunit N3	<i>Cryptococcus tetragattii</i> IND107	94	100	0.0
KIR33625.1	26S proteasome regulatory subunit N3	<i>Cryptococcus deuterogattii</i> MMRL2647	93	100	0.0

KIR55910.1	26S proteasome regulatory subunit N3	<i>Cryptococcus gattii</i> Ru294	94	100	0.0
KIR50126.1	26S proteasome regulatory subunit N3	<i>Cryptococcus bacillisporus</i> CA1280	93	100	0.0
KIR67104.1	26S proteasome regulatory subunit N3	<i>Cryptococcus bacillisporus</i> CA1873	93	100	0.0
KGB74654.1	26S proteasome regulatory subunit N3	<i>Cryptococcus deuterogattii</i> R265	93	100	0.0
OBR85418.1	26S proteasome regulatory subunit N3	<i>Cryptococcus dejecticola</i> CBS 10117	73	97	0.0

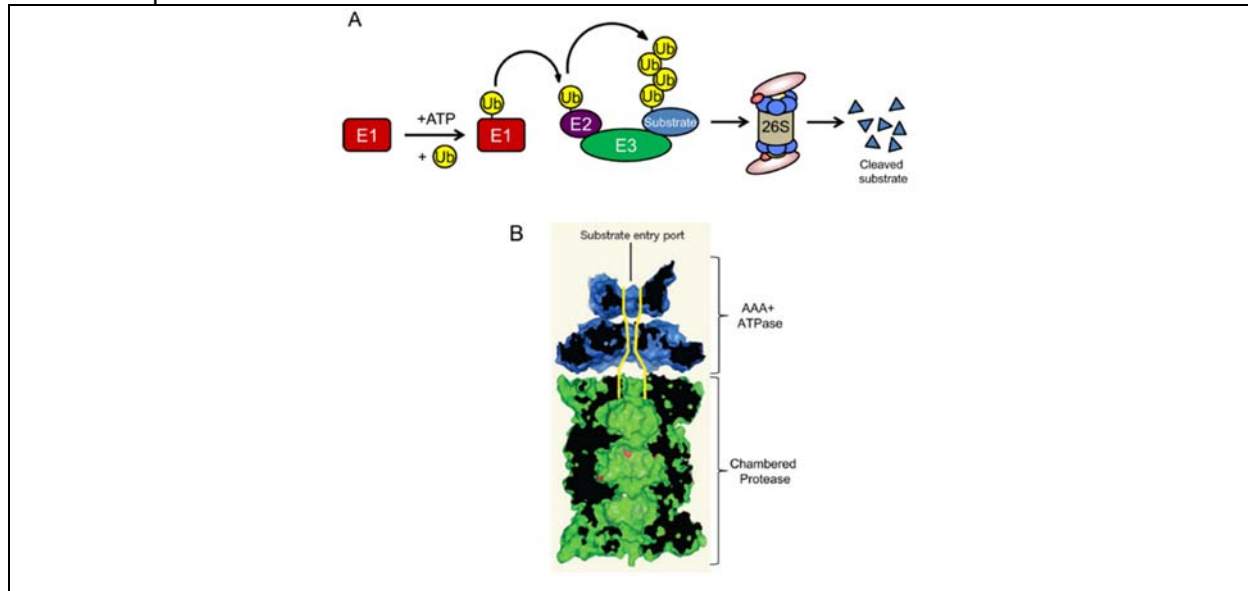
*SI = porcentagem de identidade ou homologia. **SS = Porcentagem da sequência que é coberta pelo alinhamento.
*E-value = Parâmetro de confiança.

Proteasoma é complexo proteolítico multisubunidade especialmente grande, com uma porção catalítica central (equivalente ao proteasoma 20S) e dois subcomplexos reguladores terminais, denominados PA700 ou PA28, que estão unidos a ambas as extremidades da porção central em orientações opostas para formar o complexo catalítico ativo. São gerados dois tipos de complexos proteosômicos com o mesmo núcleo catalítico e diferentes módulos de regulação, e, para isto, são necessárias cerca de 40 subunidades com tamanhos entre 20 a 110 kilodaltons (kDa). (HILT et al., 1995; HOCHSTRASSER, 1996; HERSHKO, et al., 1998; TANAKA, 1998; BAUMEISTER et al.; 1998).

O complexo regulador PA700 pode associar-se ao proteasoma 20S de uma forma dependente de ATP e assim formar o proteasoma 26S com uma massa molecular de 2000 kilodaltons (kDa) (Figura 2).

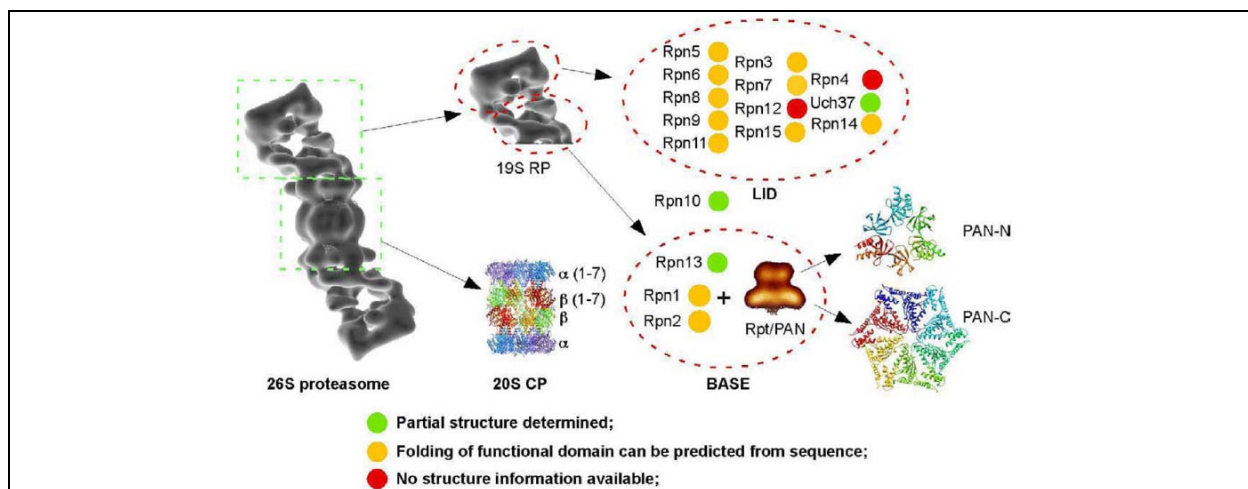
O proteasoma possui muitos substratos e sua função principal função é degradar proteínas conjugadas com ubiquitina (Figura 1). Os sítios proteolíticos do proteasoma estão dentro de uma câmara interna, que é um compartimento topológico definido pelas superfícies internas das 28 subunidades que formam a partícula nuclear do proteasoma (CP, Câmara Proteolítica também conhecida como partícula 20S). As proteínas celulares têm acesso restrito a esta câmara e aos locais proteolíticos dentro dele, diminuindo a proteólise não específica. Na holoenzima do proteasoma, o CP é complexado com a partícula reguladora de 19 subunidades (RP). A RP (proteínas de Regulação) possui receptores para a ubiquitina responsáveis pelo reconhecimento do substrato. (PETH et al., 2013; ZANET et al., 2015).

Figura 1. Esquema representativo do sistema ubiquitina-proteossoma e AAA (ATPase) + câmara proteolítica. A) Através das ações sequenciais de E1, E2 e E3, a proteína a ser degradada é modificada por poliubiquitinação, no qual serve de sinal para o reconhecimento pelo proteossoma. B) Parte do proteossomo onde é mostrado a câmara proteolítica e AAA, exibindo o caminho do substrato através de AAA e dirigindo-se a câmara proteolítica. A largura da passagem para a câmara proteolítica é mostrada nas linhas amarelas e sítio ativo proteolítico em verde.



FONTE: Tomko-Júnior e Hochstrasser (2013).

Figura 2. Representação esquemático da subunidade proteossoma 26S. Seus componentes são divididos em grupos funcionais e subunidades.



FONTE: Kin et al. (2011)

CONCLUSÃO

Verificou - se através da análise realizada que a proteína hipotética CNA00530 de *Cryptococcus neoformans* JEC21 apresenta de 73 a 95% de homologia a

subunidade regulatória N3 do proteasoma 26S encontrados outras espécies do gênero. Diante dos dados apresentados, propõe-se que a proteína hipotética CNA00530, pode exercer a função de subunidade regulatória N3 do proteasoma 26S, no entanto, para concluir a completa elucidação da função, torna-se necessário aliar os dados obtidos *in silico* com estudos realizados *in vitro*.

REFERÊNCIAS

- ARAUJO, E.J.C.; TAPARO, C.V.; UCHIDA, C.Y.; MARINHO, M. Cryptococcus: isolamento ambiental e caracterização bioquímica. **Arq. Bras. Med. Vet. Zootec**, Belo Horizonte, v. 67, n.4, p.1003-1008, 2015.
- BAUMEISTER, W.; WALZ, J.; ZÜHL, F.; AND SEEMÜLLER, E. The Proteasome: Paradigm of a Self-Compartmentalizing Protease. **Cell**, Martinsried, Alemanha v.92, p. 367–380, 1998.
- CLAVERIE, J. M. Computational methods for the identification of genes in vertebrate genomic sequences. **Hum. Mol. Genet**, Marselha, França, v.6, n. 10, p.1735–1744, 1997.
- DAHER, F.E.; NASSERALA, L.C.J; JUNIOR, S.B.G; OLIVEIRA, V.R.A; NETO, M.U.J; SOUSA, Q. Case report fatal disseminated cryptococcosis with renal involvement in an hiv-infected patient. **Rev. Inst. Med. Trop. São Paulo**, Sao Paulo, v. 57, n.4, p.366-367, 2015.
- DESLER, C.; SURAVAJHALA, P.; SANDERHOFF, M.; RASMUSSEN, M.; RASMUSSEN, L. J. IN Silico screening for functional candidates amongst hypothetical proteins. Copenhagen N- Denmark. **BMC Bioinformatics**, Copenhagen Dinamarca, v.10, n. 289, p.1-9, 2009.
- DOERKS, T.; NOORT, V.V.; MINGUEZ, P.; BORK, P. Annotation of the M. Tuberculosis Hypothetical Orfeome: Adding Functional Information to More than Half of the Uncharacterized Proteins. **PLoS ONE**, Heidelberg, Alemanha, v.7, n. 4, p.1-6, 2012.
- FARIA, R.O.; NASCENTE, P.S.; MEIREZ, A.R.M.; CLEFF, M.B.; ANTUNES, T.A.; SILVEIRA, E.S.; NOBRE, M.O; MEIRELES, M.C.A; MELLO, J.R.B. Ocorrência de Cryptococcus neoformans em excretas de pombos na Cidade de Pelotas, Estado do Rio Grande do Sul. **Revista da Sociedade Brasileira de Medicina Tropical**. Uberaba, MG, v.43, n.2, p198-200, 2010.
- FILIÚ, W.F.O.; WANKE, B.; AGÜENA, S.M.; VILELA, V.O.; MACEDO, R.C.L.; LAZERA, M. Cativeiro de aves como fonte de Cryptococcus neoformans na cidade de Campo Grande, Mato Grosso do Sul, Brasil. **Rev. Soc. Bras. Med. Trop**. Uberaba, MA, v.35, n.6, p.591-595, 2002.

HERSHKO, A.; CIECHANOVER, A. The ubiquitin system. **Rev. Biochem**, Haifa, Israel, v.67, p. 425-79,1998.

HILT, W.; WOLF, D. H. Proteasomes: destruction as a programme. **Trends Biochem. Sci.** Stuttgart, Alemanha, v.21, p.96-102, 1996.

HOCHSTRASSER, M. Ubiquitin-dependent protein degradation **Rev. Genet.** Chicago, Estados Unidos. v.30, p. 405–439, 1996.

KON, A.S.; GRUMACH, A.S.; COLOMBO, A.L.; PENALVA, A.C.O.; WANKE, B.; TELLES, F.Q.; SEVERO, L.C.; ARANHA, L.F.; LAZÉRA, M.S.; RESENDE, M.R.; SALMITO, M.A.; SHIKANAI-YASUDA, M.A.; MORETTI, M.L.; FERREIRA, M.S.; SILVA-VERGARA, M.L.; ANDRADE, N.M.P.; TRABASSO, P.; MENDES, R.P.; MARTINEZ, R.; PONZIO, V. Consenso em Criptococose – 2008. **Revista da Sociedade Brasileira de Medicina Tropical**, Uberaba, MG, v.41, n.5, p.524-544, 2008.

LODISH, H.; BERK, A.; KAISER, C. A.; KRIEGER, M.; BRETSCHER, A.; PLOEGH, H.; AMON, A.; SCOTT, M. P.; **Biologia Celular e Molecular**. 7. ed. Porto Alegre: Artmed, 2014.1210p.

PETH, A.; NATHAN, J.A.; GOLDBERG, A.L. The ATP costs and time required to degrade ubiquitinated proteins by the 26 S proteasome. **J Biol Chem**, Cambridge, Reino Unido, v. 288, p. 29215–29222, 2013.

REOLON, A.; PEREZ, L.R.R.; MEZZARI, A. Prevalência de *Cryptococcus neoformans* nos pombos urbanos da Cidade de Porto Alegre, Rio Grande do Sul. **Jornal Brasileiro de Patologia e Medicina Laboratorial**, Rio de Janeiro, RJ, v.40, n.5, p.293-298, 2004.

RIBAS, R. C.; BAEZA, L. C.; RIBEIRO, F. H. M. Isolation of *cryptococcus* spp. In excrements of pigeons (*columba* sp.) In the Maringá city, Pr, Brazil. **Arq. Ciênc. Saúde UNIPAR**, Umuarama, v. 15, n. 1, p. 45-50, 2011.

SAVILLE, S. P.; TOMAS, D. P.; LOPÉZ-RIBOT J. L. Use of genome information for the study of the pathogenesis of fungal infections and the development of diagnostic tools. **Revista Iberoamericana de micologia**, San Antonio, Texas, Estados Unidos, v. 22, p. 238-241, 2005.

SILVA, P.F.F. **Caracterização Funcional de PROTEÍNAS Hipotéticas do Fungo Patogênico Humano *Paracoccidioides* sp.**133F. Dissertação [Mestrado]. Goiânia: Universidade Federal de Goiás, Instituto de Ciências Biológicas. 2014.

SMITH, R. F. Perspectives: sequence data base searching in the era of large-scale genomic sequencing. **Genome Res**, Houston, Texas, USA, v.6, p.653–660, 1996.

TANAKA, K. Proteasomes: structure and biology. **J Biochem**, Tóquio, Japão, v.123, p.195–204, 1998.

TANAKA, K. Molecular Biology of the Proteasome. **Biochemical and biophysical research communications**. Tóquio, Japão, v.247, p. 537–541, 1998.

TAN, S. H.; NORMI, Y. M.; LEOW, A. T. C.; SALLEH, A. B.; KARJIBAN, R. A.; MURAD, A. M. A.; MAHADI, N. M.; RAHMAN, M. B. A. A Sco protein among the hypothetical proteins of *Bacillus lehensis* G1: Its 3D macromolecular structure and association with Cytochrome C Oxidase. **BMC Structural Biology**, v. 14, p. 1-12, 2014.

ZANET, J.; BENRABAH, E.; LI, T.; PÉLISSIER-MONIER, A.; CHANUT-DELALANDE, H.; RONSIN, B.; BELLEN, H.J.; PAYRE, F.; PLAZA, S. Pri sORF peptides induce selective proteasome-mediated protein processing. **Science**, Washington Estados Unidos, v. 349, p. 1356–1358, 2015.