

UNIVERSIDADE FEDERAL DE MINAS GERAIS

Instituto de Ciências Biológicas

Programa de Pós-graduação em Microbiologia

Ludmila Ferreira Gouveia

**ANOTAÇÃO FUNCIONAL E GENÔMICA COMPARATIVA DE VIAS METABÓLICAS
DE *Cryptococcus* spp.**

Belo Horizonte

2017

Ludmila Ferreira Gouveia

**ANOTAÇÃO FUNCIONAL E GENÔMICA COMPARATIVA DE VIAS METABÓLICAS
DE *Cryptococcus* spp.**

Dissertação apresentada ao Programa de Pós-graduação em Microbiologia da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Mestre (a) em Ciências Biológicas: Microbiologia.

Orientador: Prof. Dr. Daniel de Assis Santos

Coorientador: Dr. Jeronimo Conceição Ruiz

Belo Horizonte

2017

043

Gouveia, Ludmila Ferreira.

Anotação funcional e genômica comparativa de vias metabólicas de *Cryptococcus* spp. [manuscrito] / Ludmila Ferreira Gouveia. – 2017.
98 f. : il. ; 29,5 cm.

Orientador: Prof. Dr. Daniel de Assis Santos. Coorientador: Dr. Jeronimo Conceição Ruiz.

Dissertação (mestrado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa de Pós-Graduação em Microbiologia.

1. Microbiologia. 2. *Cryptococcus*. 3. Redes e Vias Metabólicas. 4. Virulência. 5. Bioinformática. I. Santos, Daniel de Assis. II. Ruiz, Jeronimo Conceição. III. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. IV. Título.

CDU: 579



UNIVERSIDADE FEDERAL DE MINAS GERAIS

INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM MICROBIOLOGIA

ATA DA DEFESA DE DISSERTAÇÃO DE MESTRADO

Nº REGISTRO: 2015654938

Coorientador: Dr. Jeronimo Conceição Ruiz

Relatora e Suplente: Dra. Thais Furtado Ferreira Magalhães

Às 14:00 horas do dia 22 de fevereiro de 2017, reuniu-se, no Instituto de Ciências Biológicas da UFMG, a Comissão Examinadora composta pela Dra. Ludmila de Matos Baltazar - (Departamento de Microbiologia/ICB/UFMG), pela Dra. Luciana Márcia de Oliveira (Centro Universitário UNA) e pelo Prof. Daniel de Assis Santos - Orientador, para julgar o trabalho final "Anotação funcional e genômica comparativa de vias metabólicas de *Cryptococcus* spp.", da aluna Ludmila Ferreira Gouveia, requisito final para a obtenção do Grau de **MESTRE EM CIÊNCIAS BIOLÓGICAS: MICROBIOLOGIA**. Abrindo a sessão, o Presidente da Comissão, Prof. Flávio Guimarães da Fonseca - Coordenador do Programa, após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra à candidata, para a apresentação de seu trabalho. Seguiu-se a arguição pelos Examinadores, com a respectiva defesa da candidata. Logo após, a Comissão se reuniu, sem a presença da candidata e do público, para julgamento e expedição de resultado final. A candidata foi considerada **APROVADA**. O resultado final foi comunicado publicamente à candidata pelo Presidente da Comissão. Nada mais havendo a tratar, o Presidente encerrou a reunião e lavrou a presente ata, que será assinada por todos os membros participantes da Comissão Examinadora. Belo Horizonte, 22 de fevereiro de 2017.

Dra. Luciana Márcia de Oliveira

Dra. Ludmila de Matos Baltazar

Prof. Daniel de Assis Santos (Orientador)

Prof. Flávio Guimarães da Fonseca
Coordenador

AGRADECIMENTOS

Agradeço a Deus, em primeiro lugar, por estar presente em todos os momentos de minha vida e por me capacitar a cada dia.

Ao meu orientador, Daniel de Assis Santos, por toda contribuição para o meu crescimento acadêmico, pela colaboração para a realização deste trabalho e pela amizade.

Ao meu coorientador, Jeronimo Ruiz, pelos ensinamentos e por toda a contribuição para a realização deste trabalho.

À banca examinadora composta pela Dra. Ludmila, pela Dra. Luciana e pela Dra. Thais, que gentilmente aceitou o convite.

Ao meu esposo, Isaías, por todo amor, carinho, compreensão e pela paciência. Por sempre acreditar no meu potencial e pelo incentivo.

À minha mãe, pelo amor incondicional, por sempre me incentivar e todo o apoio.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo suporte financeiro.

A todos os colegas do Laboratório de Micologia da UFMG, por todo apoio e companheirismo, pelos bons momentos e por todo o auxílio para que esse trabalho pudesse ser realizado.

Aos alunos e colaboradores do Grupo Informática de Biossistemas e Genômica da Fiocruz, em especial à Grace, Leilane, pela colaboração neste projeto. Ao Jader por todo suporte técnico.

Aos meus familiares, pelo carinho. Vocês foram essenciais para a concretização de mais uma etapa de minha carreira acadêmica.

A todos que, de alguma forma, contribuíram para a realização deste trabalho.

RESUMO

Cryptococcus neoformans e *C. gattii* são os principais agentes da criptococose, uma micose sistêmica que acomete os pulmões, com tendência a migrar para o sistema nervoso central. Essas espécies compartilham quase todos os fatores de virulência. Entretanto, diferentes perfis de hospedeiros são observados para esses micro-organismos. Uma hipótese para essa diferença poderia ser diferenças no metabolismo desses patógenos. A bioinformática representa uma área interdisciplinar de estudo que permite a análise de dados *in silico* provenientes de sequenciamento genômico e outras técnicas de Biologia Molecular. O emprego de ferramentas computacionais permite o estudo de estruturas como DNA, RNA e proteínas, bem como a predição de vias metabólicas de organismos e comparação destas com outras vias. Sendo assim, o objetivo deste trabalho foi caracterizar, *in silico*, vias metabólicas de *C. neoformans* e *C. gattii*, bem como suas funções, além de caracterizar a presença de sequências que codificam fatores de virulência desses fungos por meio da análise dos genomas disponíveis nos bancos de dados de domínio público. Três genomas de *C. neoformans* e três de *C. gattii* com anotação genômica e funcional foram utilizados. De forma a contribuir para a anotação funcional desses genomas, foi realizada a predição de RNAs ribossomais e RNAs transportadores para cada linhagem utilizada. Uma análise de similaridade de sequência entre linhagens de uma mesma espécie foi feita, sendo observado um alto grau de conservação de sequências de nucleotídeos. Os proteomas preditos foram extraídos dos genomas e o mapeamento de vias metabólicas foi realizado pelo programa Blast2GO. Das 29 vias metabólicas mapeadas, 16 foram comuns a todos os organismos. Os resultados referentes à análise de agrupamento demonstraram que praticamente todas as proteínas associadas à virulência pertencem ao *core genome*. Também foi observado que proteínas associadas à virulência estão envolvidas em diversos processos biológicos. Além disso, foi evidenciado que a linhagem *C. neoformans* H99 foi a única a possuir enzimas associadas à virulência mapeadas na via de fosforilação oxidativa, sendo essa observação um possível objeto de estudos posteriores. Tendo em vista os resultados observados, concluímos que estudos envolvendo vias metabólicas de *Cryptococcus* são de extrema importância para uma melhor compreensão acerca dos mecanismos envolvidos na virulência desses patógenos.

Palavras-chave: *Cryptococcus* spp.; vias metabólicas; virulência; bioinformática.

ABSTRACT

Cryptococcus neoformans and *C. gattii* are the main agents of Cryptococcosis, a systemic mycosis that affects lungs, tending to migrate to the Central Nervous System. These species share almost all the virulence factors. However, different patterns of hosts are observed for these microorganisms. One hypothesis for this is that differences in the metabolism of these pathogens could occur. Bioinformatics represents an area of studies that allows the analysis of *in silico* data provided from genomic sequencing and other Molecular Biology techniques. The use of computational tools allows the study of structures like DNA, RNA and proteins, as well as the prediction of metabolic pathways of organisms and the comparison of them with other pathways. Therefore, the aim of this study was to characterize, *in silico*, metabolic pathways of *C. neoformans* and *C. gattii*, as well as their functions, besides to identify the presence of sequences that code virulence factors of these organisms, by analyzing genome sequences available in public databases. Three genomes of *C. neoformans* and three of *C. gattii* with genomic and functional annotation were used. In order to contribute to the functional annotation of these genomes, it was performed the prediction of ribosomal RNAs and transfer RNAs for each strain used. A sequence similarity search between strains within the same species was performed, and it was observed a high level of sequence conservation in the level of nucleotides. The predicted proteomes were extracted from the genomes and the metabolic pathways mapping was performed by Blast2GO. From the 29 metabolic pathways mapped, 16 were common to all organisms. The results obtained from the cluster analysis demonstrated that almost all of the proteins associated with virulence belong to the core genome. It was also observed that proteins associated with virulence are involved in several biological processes. Moreover, it was evidenced that the strain *C. neoformans* H99 was the only one to pursue enzymes associated with virulence mapped in the oxidative phosphorylation pathway, being this a possible target for future studies. Based on the results obtained in this study, we concluded that studies about metabolic pathways of *Cryptococcus* spp. are extremely important for a better understanding about the mechanisms involved in the virulence of these pathogens.

Keywords: *Cryptococcus* spp.; metabolic pathways; virulence; Bioinformatics.

LISTA DE FIGURAS

Figura 1: Ciclo de vida de <i>Cryptococcus</i> e de infecção do hospedeiro.....	16
Figura 2: Diferentes sorotipos e genótipos dos dois principais agentes etiológicos da criptococose, <i>C. neoformans</i> e <i>C. gattii</i>	18
Figura 3: Resposta de <i>C. gattii</i> frente ao fluconazol.	29
Figura 4: Representação esquemática das principais ciências ômicas.	30
Figura 5: Gráfico de comparação de similaridade global entre <i>C. neoformans</i> JEC21, <i>C. neoformans</i> H99 e <i>C. neoformans</i> B-3501A	45
Figura 6 Gráfico de comparação de similaridade global entre <i>C. neoformans</i> B-3501A e <i>C. neoformans</i> JEC21.....	46
Figura 7: Gráfico de comparação de similaridade global entre <i>C. gattii</i> WM276, <i>C. gattii</i> R265 e <i>C. gattii</i> CBS 7750	48
Figura 8: Gráfico de comparação de similaridade global entre <i>C. gattii</i> R265 e <i>C. neoformans</i> H99	50
Figura 9: Distribuição de números EC após mapeamento e anotação do proteoma total de <i>C. gattii</i> WM276 pelo programa Blast2GO.	50
Figura 10: Via de biossíntese de fenilalanina, tirosina e triptofano.	54
Figura 11: Via de fosforilação oxidativa.	56
Figura 12: Diagrama de Venn referente à análise de agrupamento.....	58
Figura 13: Representação gráfica da comparação de similaridade entre trechos dos genomas de <i>C. neoformans</i> JEC21, <i>C. neoformans</i> H99 e <i>C. neoformans</i> B-3501A.....	60
Figura 14: Representação gráfica da comparação de similaridade entre trechos dos genomas de <i>C. gattii</i> R265 e <i>C. neoformans</i> H99.....	61
Figura 15: Gráfico de comparação de similaridade entre <i>C. gattii</i> R265 e <i>C. neoformans</i> H99...	62

Figura 16: Gráfico referente à análise de enriquecimento de proteínas relacionadas à virulência em <i>C. gattii</i> , linhagem R265.....	74
Figura 17: Gráfico referente à análise de enriquecimento de proteínas relacionadas à virulência em <i>C. gattii</i> , linhagem CBS7750.....	74
Figura 18: Gráfico referente à análise de enriquecimento de proteínas relacionadas à virulência em <i>C. neoformans</i> , linhagem B-3501A.....	75
Figura 19: Gráfico referente à análise de enriquecimento de proteínas relacionadas à virulência em <i>C. neoformans grubii</i> , linhagem H99	75
Figura 20: Gráfico referente à análise de enriquecimento de proteínas relacionadas à virulência em <i>C. neoformans</i> , linhagem JEC21.	76

LISTA DE TABELAS

Tabela 1: Lista de genomas de <i>Cryptococcus spp.</i> utilizados para análise.	35
Tabela 2: RNAs ribossomais e RNAs transportadores preditos.	42
Tabela 3: Relação das vias metabólicas comuns a todos os organismos estudados.	53
Tabela 4: Vias metabólicas mapeadas em <i>C. gattii</i> CBS 7750.	77
Tabela 5: Vias metabólicas mapeadas em <i>C. gattii</i> R265.	80
Tabela 6: Vias metabólicas mapeadas em <i>C. gattii</i> WM276.	82
Tabela 7: Vias metabólicas mapeadas em <i>C. neoformans</i> B-3501A.	86
Tabela 8: Vias metabólicas mapeadas em <i>C. neoformans</i> var. <i>grubii</i> H99.	90
Tabela 9: Vias metabólicas mapeadas em <i>C. neoformans</i> var. <i>neoformans</i> JEC21.	94

LISTA DE ABREVIATURAS E SIGLAS

5-FC – 5-Flucitosina

AmB – Anfotericina B

ACT – Artemis Comparison Tool

BD – Banco de dados

BLAST – Basic Alignment Search Tool

BNICE – Biochemical Network Integrated Computational Explorer

CIM – Concentração Inibitória Mínima

FLC – Fluconazol

FMM – From Metabolite to Metabolite

GO – Gene Ontology

GalXM – Galactoxilomanana

GXM – Glucuronoxilomanana

KEGG – Kyoto Encyclopedia of Genes and Genomes

NCBI – National Center for Biotechnology Information

ORFs – Open reading frames

Pb – pares de base

PCR – Polymerase Chain Reaction

RFLP – Restriction fragment length polymorphism

RID – Região intrinsecamente desordenada “IDR”

SIDA – Síndrome da Deficiência adquirida “AIDS”

SNC – Sistema Nervoso Central

SUMÁRIO

1	Introdução	15
1.1	<i>Cryptococcus</i> spp. e criptococose	15
1.2	<i>Cryptococcus neoformans</i>	19
1.3	<i>Cryptococcus gattii</i>	19
1.4	Fatores de virulência	20
1.4.1	Cápsula polissacarídica	20
1.4.2	Crescimento a 37°C	21
1.4.3	Urease	21
1.4.4	Proteases e fosfolipases	21
1.4.5	Melanina	22
1.5	Epidemiologia	22
1.6	Tratamento	24
1.6.1	Anfotericina B	25
1.6.2	5-Flucitosina	25
1.6.3	Azólicos	26
1.7	Resistência aos antifúngicos	26
1.8	Vias metabólicas de <i>Cryptococcus</i> spp.	27
1.9	Bioinformática	29
2	Justificativa	33
3	Objetivos	34

3.1	Objetivo geral	34
3.2	Objetivos específicos	34
4	Material e métodos	35
4.1	<i>Scripts</i> desenvolvidos	35
4.2	Genomas de <i>Cryptococcus</i> spp.	35
4.3	Anotação do proteoma de <i>Cryptococcus</i> spp.	36
4.3.1	Artemis	36
4.3.2	Predição de RNAs ribossomais e RNAs transportadores	36
4.4	Análise de similaridade de sequências	37
4.5	ACT	38
4.6	Predição e anotação de vias metabólicas	38
4.7	Criação de banco de dados de virulência.....	39
4.8	Análise de enriquecimento funcional	39
4.9	Mapeamento das vias metabólicas e associação aos fatores de virulência.....	40
4.10	Análise de agrupamento	40
5	Resultados e discussão	42
5.1	Seleção de genomas de <i>Cryptococcus</i> spp.	42
5.2	Predição de RNAs ribossomais e RNAs transportadores	42
5.3	Análise de similaridade de sequências	43
5.3.1	<i>Cryptococcus neoformans</i>	44
5.3.2	<i>Cryptococcus gattii</i>	47

5.3.3	<i>Cryptococcus gattii</i> vs <i>C. neoformans</i>	49
5.4	Predição de vias metabólicas.....	50
5.5	Análise de enriquecimento funcional	51
5.6	Mapeamento das vias metabólicas e associação aos fatores de virulência.....	52
5.7	Análise de agrupamento	57
6	Conclusão	64
7	Perspectivas	65
8	Referências bibliográficas	66
9	Apêndice A – gráficos relacionados à análise de enriquecimento entre as espécies/ linhagens (processo biológico)	74
10	Apêndice b - Tabelas referentes ao mapeamento de proteínas associadas à virulência de <i>Cryptococcus</i> spp. nas vias metabólicas.....	77

1 INTRODUÇÃO

1.1 *Cryptococcus* spp. e criptococose

A criptococose é uma micose sistêmica que acomete os pulmões, com tendência à migração para o Sistema Nervoso Central (SNC), por isso as manifestações clínicas mais comuns são a doença pulmonar e a meningoencefalite (MAY *et al.*, 2016). Os principais sintomas e sinais da criptococose são: tosse, produção de muco, dificuldade para respirar, perda de visão, dor no peito, febre e perda de peso (CHEN *et al.*, 2014). Essa infecção é causada por fungos do gênero *Cryptococcus*, que é constituído por aproximadamente 70 espécies, sendo *C. neoformans* e *C. gattii* as principais espécies responsáveis por essa micose (FANG *et al.*, 2014). Esse gênero pertence ao filo Basidiomycota, uma vez que esses micro-organismos possuem uma fase sexual filamentosa que resulta na produção de esporos, que são liberados de uma estrutura denominada basídio (IDNURM, 2010). A reprodução de *C. neoformans* e *C. gattii* ocorre principalmente de forma assexuada, mas essas espécies também possuem um ciclo celular sexuado, composto por dois *mating types*, MATa e MATα. *Filobasidiella neoformans* e *Filobasidiella bacillispora* são os nomes dados aos estados sexuais de *C. neoformans* e *C. gattii*, respectivamente (COGLIATI, 2013).

Atualmente, a taxonomia desse grupo tem sido discutida após HAGEN e colaboradores (2015) sugerirem a divisão de *C. gattii* e *C. neoformans* em oito espécies haploides (HAGEN *et al.*, 2015) o que tem sido endossado pelo trabalho de FERNANDES e colaboradores (2017). Essa classificação, entretanto, ainda não tem sido muito bem aceita (KWON-CHUNG *et al.*, 2017) e a divisão defendida por Kwon-Chung e colaboradores (2002) em duas espécies tem prevalecido (KWON-CHUNG *et al.*, 2002). Recentemente, sugeriu-se a utilização dos termos “complexo de espécies *C. gattii*” e “complexo de espécies *C. neoformans*” para denominar esses micro-organismos (KWON-CHUNG *et al.*, 2017).

C. gattii e *C. neoformans* podem ser encontrados como sapróbios no ambiente e, quando acometem o homem, (LIN e HEITMAN, 2006) a infecção se dá pela inalação de propágulos fúngicos presentes em fontes ambientais, causando o quadro de infecção criptocócica (Figura 1). Esses propágulos podem ser a levedura dessecada ou esporos, apesar de não haver relatos da identificação de esporos no ambiente (BOTTS e HULL, 2010). O período de incubação do fungo

varia de acordo com a espécie (CHEN *et al.*, 2014). Para *C. gattii* o período de incubação é normalmente de 6-7 meses (MACDOUGALL e FYFE, 2006); já para *C. neoformans* esse período pode chegar até 110 meses (GARCIA-HERMOSO *et al.*, 1999). Casos de meningoencefalite estão mais comumente relacionados à infecção causada por *C. neoformans*, enquanto *C. gattii* está mais relacionado com infecções pulmonares (FEDER *et al.*, 2015). Em um estudo com as linhagens *C. neoformans* H99 e *C. gattii* R265 em modelo murino de inalação, foi observado que *C. neoformans* cresceu mais rapidamente no cérebro, causando morte por meningoencefalite, enquanto *C. gattii* apresentou um crescimento mais rápido nos pulmões, provocando a morte dos animais sem apresentar um quadro fulminante de meningite criptocócica (NGAMSKULRUNGROJ *et al.*, 2012).

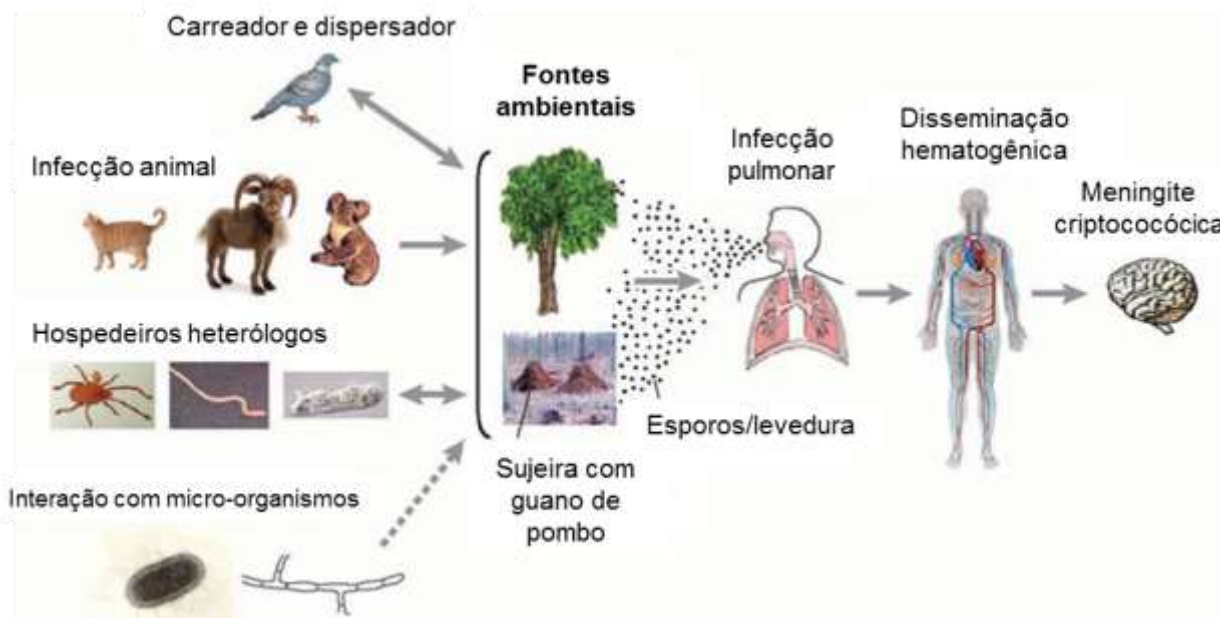


Figura 1: Ciclo de vida de *Cryptococcus* e de infecção do hospedeiro. Fonte: Adaptado de LIN e HEITMAN, 2006.

C. neoformans foi identificado por Sanfelice em 1894, após ter sido isolado em suco de pêssego (SLOAN e PARRIS, 2014; SIDRIM e ROCHA, 2004). O primeiro registro de isolamento dessa espécie como um patógeno foi feito por Busse e Buschke, que detectaram a presença dessas células em uma lesão de tíbia (SIDRIM e ROCHA, 2004). Posteriormente, *C. neoformans* foi subdividido em duas variedades, de acordo com padrões de aglutinação capsular, em *C. neoformans* var. *neoformans*, com sorotipos A e D e *C. neoformans* var. *gattii*, cujos sorotipos eram B e C (Figura 2) (FRANZOT *et al.*, 1999). *C. gattii* foi descrito pela primeira vez em 1970, ao ser isolado em um paciente africano com leucemia (VANBREUSEGHEM e TAKASHIO, 1970;

MA e MAY, 2009; ESPINEL-INGROFF e KIDD, 2015). Estima-se que os sorotipos de *C. gattii* e de *C. neoformans* divergiram um do outro há cerca de 30-40 milhões de anos (BAHN *et al.*, 2005; IDNURM *et al.*, 2005; MAY *et al.*, 2016). *C. neoformans* var. *neoformans* e *C. neoformans* var. *grubii* também podem formar híbridos por meio de recombinação, sendo os mesmos do sorotipo AD (COGLIATI, 2013).

Em 1997, BOEKHOUT e colaboradores sugeriram que essas duas variedades eram duas espécies distintas e, no ano de 2002, *C. gattii* foi definitivamente considerada como uma espécie (KWON-CHUNG *et al.*, 2002; ESPINEL-INGROFF e KIDD, 2015). Essa classificação foi feita com base em estudos filogenéticos e na ausência de recombinação dos *matings* dos isolados de cada espécie (KRONSTAD *et al.*, 2011). Outra característica que permite a diferenciação dessas duas espécies é o metabolismo de creatinina, canavanina e glicina (CHATURVEDI e CHATURVEDI, 2011). Em 1999, FRANZOT e colaboradores propuseram que somente as leveduras de sorotipo D fossem denominadas *C. neoformans* var. *neoformans*, sugerindo uma nova variedade, *C. neoformans* var. *grubii*, que é composta somente pelo sorotipo A.

A variação de sorotipos é determinada pela especificidade da estrutura da glucuronoxilomanana (GXM), sendo este o principal componente da cápsula polissacarídica dessas espécies (MA e MAY, 2009). O sorotipo mais comum de *C. neoformans* e *C. neoformans* var. *grubii* (sorotipo A), correspondendo a 82% do número de casos de criptococose no mundo, enquanto a var. *neoformans* (sorotipo D) é responsável por 20-30% dos casos dessa infecção em pacientes HIV-positivos no nordeste da Europa (FRANZOT, 1999; SHAPIRO *et al.*, 2011; SLOAN e PARRIS, 2014). Esses sorotipos divergiram há cerca de 18 milhões de anos (BAHN *et al.*, 2005; MA e MAY, 2009). Após estudos moleculares, *C. neoformans* foi classificado em quatro genótipos, VNI, VNII, VNIII e VNIV, assim como *C. gattii*, que foi classificado em VGI, VGII, VGIII e VGIV (SHUN. LI *et al.*, 2010). Dentre os genótipos de *C. neoformans*, VNI é o mais comumente isolado em pacientes, representando cerca de 80% dos casos de infecção em pacientes HIV-positivos e outros animais (NGAMSKULRUNGROJ *et al.*, 2012).

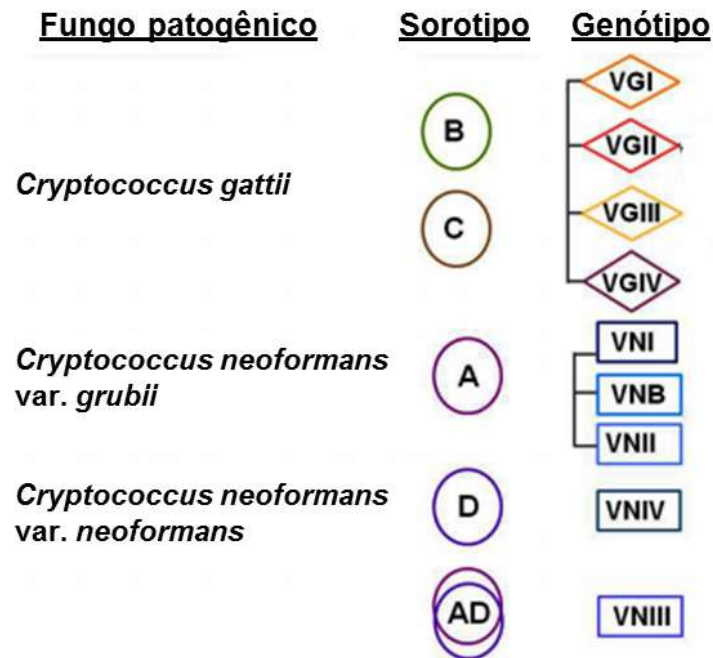


Figura 2: Diferentes sorotipos e genótipos dos dois principais agentes etiológicos da criptococose, *C. neoformans* e *C. gattii*. Fonte: Adaptado de Chaturvedi e Chaturvedi, 2011.

O aumento do número de indivíduos imunocomprometidos tem sido um dos principais fatores pelos quais a criptococose tem emergido globalmente (BAHN e JUNG, 2013). Essa é uma das três doenças oportunistas mais comuns em pacientes submetidos a tratamentos com esteroides, transplantes e aqueles com deficiência no sistema imune (FEDER *et al.*, 2015). Cerca de um milhão de casos de meningite criptocócica em pacientes com a Síndrome da Deficiência Adquirida (SIDA), a AIDS, são relatados anualmente, sendo que o número de óbitos entre esses pacientes chega a cerca de dois terços (PARK *et al.*, 2009; LEE *et al.* 2013; LIZARAZO *et al.*, 2014). Entre os anos de 1996 e 2006, a frequência de mortes causadas pela criptococose em pacientes com AIDS foi a maior registrada (PRADO *et al.*, 2009). Entretanto, outros fatores também são responsáveis pelo aumento de infecções criptocócicas. Em países desenvolvidos, onde houve uma redução no número de casos relacionados à AIDS, a medicação de pacientes com imunossupressores e o uso de corticosteroides para o transplante de órgãos ou tratamento de câncer aumentou nos últimos 30 anos, sustentando a incidência, ainda que pequena, de criptococose em pacientes HIV-negativos (SLOAN e PARRIS, 2014).

1.2 *Cryptococcus neoformans*

C. neoformans está distribuído universalmente e é encontrado primariamente em solos contaminados com guano de aves, principalmente de pombos (*Columba livia*) (GULLO *et al.*, 2013). Entretanto, essa espécie também foi isolada em outras fontes, como ar, solo e árvores (LIN e HEITMAN, 2006). A prevalência do sorotipo D é maior em países de clima temperado, em comparação às regiões de clima tropical, devido ao fato de *C. neoformans* var. *neoformans* ser mais sensível à morte em decorrência de altas temperaturas (MA e MAY, 2009).

C. neoformans é capaz de formar biofilmes em cateteres, válvulas cardíacas e outros dispositivos médicos, representando um maior risco de infecções crônicas devido sua alta resistência à resposta imune do hospedeiro e ao tratamento com antifúngicos (SANTI *et al.*, 2014). Além de causar infecções em humanos, essa espécie também é responsável por causar doenças em animais selvagens e domésticos, entre eles cães, gatos, coelhos, cavalos, ovelhas e raposas (LIN e HEITMAN, 2006).

1.3 *Cryptococcus gattii*

C. gattii é encontrado principalmente em regiões tropicais e subtropicais, tendo como habitat o solo e algumas espécies arbóreas (CDC, 2015; BIELSKA e MAY, 2016). O primeiro isolamento de *C. gattii* ocorreu em árvores de eucalipto (*Eucalyptus camaldulensis*), que por muito tempo foi tido como o único hábitat dessa espécie fúngica (ELLIS e PFEIFFER, 1990; CHATURVEDI e CHATURVEDI, 2011). Entretanto, em estudos posteriores, esse patógeno foi isolado em outras espécies vegetais. Na América do Sul e na Ásia, *C. gattii* foi isolado em árvores como oiti (*Licania tomentosa*), amendoeira (*Terminalia cattapa*), Cássia-rosa (*Cassia grandis*), ficus (*Ficus microcarpa*) e jamelão (*Syzygium cumini*) (KAOCHAROEN *et al.*, 2013). No Brasil, foi relatado o isolamento de *C. gattii* em árvores urbanas (DIXIT *et al.*, 2009; FARIA *et al.*, 2010). *C. gattii* era considerado predominantemente como um patógeno tropical e subtropical até que, no ano de 1999, houve um surto de infecções na ilha de Vancouver, Canadá (KIDD *et al.*, 2007; CHATURVEDI e CHATURVEDI, 2011; ESPINEL-INGROFF e KIDD, 2015). Após esse surto de criptococose, essa espécie despertou a atenção de pesquisadores e passou a ser considerada um problema de saúde pública. Desde então diversos estudos sobre esse micro-organismo foram realizados (NGAMSKULRUNGROJ *et al.*, 2012).

C. gattii infecta predominantemente indivíduos hígidos (CHEN *et al.*, 2014). Apesar de ser um patógeno primário, os genótipos VGIII e VGIV estão mais associados a pacientes infectados pelo vírus HIV do que indivíduos imunocompetentes, sendo isolados em 24-30% de pacientes HIV-positivos em algumas regiões da América Central e do Sul e do sudeste da África (SLOAN e PARRIS, 2014; ESPINEL-INGROFF e KIDD, 2015). Em um estudo realizado com os quatro genótipos de *C. gattii*, foi observada uma perda de enolases e de domínios de famílias de transportadores de cobre nos genótipos VGIII e VGIV, o que poderia estar associado à baixa taxa de isolamento desses genótipos em indivíduos hígidos (FARRER *et al.*, 2015). Acredita-se que *C. gattii* seja mais virulento em comparação a *C. neoformans*, devido à sua capacidade de causar várias lesões nos pulmões e no cérebro, enquanto *C. neoformans* não apresenta o mesmo perfil (CHATURVEDI e CHATURVEDI, 2011). Quando comparado a *C. neoformans*, *C. gattii* causa mais criptococomas no cérebro e uma menor resposta aos antifúngicos empregados no tratamento da criptococose (ESPINEL-INGROFF e KIDD, 2015). Foi relatado que, além de humanos, outros animais apresentam a infecção ou são carreadores de *C. gattii*, indicando uma vasta gama de hospedeiros desse patógeno (CHATURVEDI e CHATURVEDI, 2011).

1.4 Fatores de virulência

C. neoformans e *C. gattii* compartilham quase todos os fatores de virulência, um fato interessante devido ao diferente perfil de hospedeiros que as duas espécies possuem, uma vez que *C. neoformans* geralmente afeta indivíduos imunocomprometidos, enquanto *C. gattii* está mais comumente associado a infecções em indivíduos hígidos (KRONSTAD *et al.*, 2011). Dentre esses fatores pode-se citar a cápsula polissacarídica, crescimento a 37 °C, produção das enzimas lacase, urease e fosfolipase e síntese de melanina (CASADEVALL *et al.*, 2000; FEDER *et al.*, 2015).

1.4.1 Cápsula polissacarídica

A cápsula polissacarídica é composta por 90%-95% de glucuronoxilomanana (GXM) e 5% de galactoxilomanana (GalXM), sendo as diferenças na estrutura da cápsula as responsáveis pela determinação dos sorotipos de *C. neoformans* e *C. gattii* (MA e MAY, 2009). Dentre as funções da cápsula estão o aumento da invasividade e patogenicidade, além de proteger a levedura contra o ambiente e o hospedeiro (GULLO *et al.*, 2013). A cápsula está presente em isolados ambientais e provavelmente protege as células da desidratação (SHAPIRO *et al.*, 2011; PRATES *et al.*, 2013).

Essa estrutura tem função antifagocítica e interfere na ativação de células dendríticas, interferindo na função das células T. Além disso, GalXM induz a apoptose de células T humanas e macrófagos têm apoptose induzida tanto por GXM quanto por GalXM (SHUN. LI *et al.*, 2010; BIELSKA e MAY, 2016). Outra função importante da cápsula é sua resistência à digestão por fagossomos (TUCKER e CASADEVALL, 2002; MA e MAY, 2009).

1.4.2 Crescimento a 37°C

A capacidade de *Cryptococcus* spp. de crescer a 37-39 °C, temperatura corporal de mamíferos, permite que essas espécies causem infecções em humanos (IDNURM *et al.*, 2005). Essa termotolerância é possibilitada por diferentes funções, dentre as quais estão vias de sinalização (KRONSTAD, et al, 2011). Apesar de ser encontrado em guano de pombos, que possuem temperatura corporal entre 40 e 42 °C, *C. neoformans* não causa infecção nesses animais, sugerindo que a temperatura de crescimento é um fator importante na patogenicidade desse micro-organismo (MA e MAY, 2009). Essa capacidade de sobreviver em um ambiente a 37 °C está relacionada à enzima calcineurina, uma proteína serina-treonina fosfatase dependente de cálcio e de calmodulina (LIU *et al.*, 1991).

1.4.3 Urease

A urease potencializa a penetração das células fúngicas no sistema nervoso central (IDNURM *et al.*, 2005). SHI *et al.* (2010) demonstraram por meio de imagens em tempo real que a migração de *C. neoformans* é dependente de urease. Para isso, foram utilizados no estudo a linhagem selvagem H99 e um *knockout* para a síntese de urease. Essa enzima catalisa a hidrólise de ureia em amônia e carbamato, aumentando a capacidade das células fúngicas de alcançar o SNC (KRONSTAD *et al.*, 2011). FEDER *et al.* (2015), observaram que linhagens de *C. gattii* mutantes para a síntese de urease ou que expressavam apoureases inativas enzimaticamente não foram capazes de multiplicar em macrófagos, indicando a importância desse fator de virulência na patogenicidade dessa espécie.

1.4.4 Proteases e fosfolipases

Lacase e fosfolipases são necessárias para a saída das leveduras dos pulmões (IDNURM *et al.*, 2005). Dentre as funções das proteinases estão a degradação de proteínas do hospedeiro, dentre

elas fibrinogênio, colágeno, imunoglobulinas, elastina e fatores complemento (MA e MAY, 2009). A fosfolipase B atua na invasão dos pulmões pelo fungo e em sua disseminação (KRONSTAD, *et al.*, 2011). A fosfolipase B1 é responsável por facilitar a aderência das células no epitélio dos pulmões, além da sobrevivência e multiplicação dentro do patógeno nos macrófagos e disseminação através do sangue e do sistema linfático (SHUN. LI *et al.*, 2010).

1.4.5 Melanina

A melanina protege *C. neoformans* contra a radiação ultravioleta, permitindo a sobrevivência desse fungo no ambiente (CASADEVALL *et al.*, 2000). Essa espécie é comumente encontrada em fezes de pombos, onde parece estar em seu estado melanizado (NOSANCHUK *et al.*, 1999). Essa substância pode ser produzida pela oxidação de catecolaminas do hospedeiro, encontradas principalmente no cérebro, protegendo o fungo contra radicais tóxicos produzidos pelo hospedeiro (IDNURM *et al.*, 2005). Acredita-se que o tropismo de *C. neoformans* pelo SNC esteja relacionado à sua capacidade em converter catecolaminas em melanina (CASADEVALL *et al.*, 2000). A alta resistência da melanina à degradação confere a essa substância a capacidade de neutralizar radicais livres (SHUN. LI *et al.*, 2010). Além disso, a melanina também está relacionada à resistência de *C. neoformans* a alguns antifúngicos, dentre eles a anfotericina B e caspofungina (DUIN *et al.*, 2002). Dois genes presentes no genoma de *C. neoformans*, LAC1 e LAC2, estão diretamente relacionados à síntese de melanina. LAC1, que codifica a enzima lacase, é essencial para a produção de melanina (IDNURM *et al.*, 2005), que confere proteção às leveduras contra o estresse oxidativo provocado pelo hospedeiro (FARRER *et al.*, 2016). LAC2 contribui para a polimerização de moléculas difenólicas para formar a melanina (IDNURM *et al.*, 2005).

1.5 Epidemiologia

Estudos epidemiológicos têm sido realizados por meio do uso de ferramentas de Biologia Molecular, tais como reação em cadeia da polimerase (PCR) para tipagem molecular, polimorfismo de fragmentos de restrição (RFLP), sequenciamento de espaços intergênicos e tipagem de sequências *multilocus* (COGLIATI, 2013; ESPINEL-INGROFF e KIDD, 2015). Embora haja uma grande diversidade de métodos para a determinação do perfil epidemiológico de *C. neoformans* e *C. gattii*, há uma dificuldade na obtenção de dados epidemiológicos mais assertivos devido ao fato de que a diferenciação de infecções causadas por esses patógenos não é feita rotineiramente

(CHATURVEDI e CHATURVEDI, 2011; ESPINEL-INGROFF e KIDD, 2015). Apesar disso, diversos estudos têm sido realizados para uma maior compreensão acerca da epidemiologia de *Cryptococcus* spp. Em um estudo recente, COGLIATI (2013) analisou e compilou os dados epidemiológicos da criptococose em todo o mundo, fornecendo uma visão sobre a distribuição global da criptococose, bem como os sorotipos e genótipos mais prevalentes em cada região.

O genótipo VNI é o mais amplamente distribuído mundialmente, sendo o mais prevalente em todo mundo com exceção da Austrália e Papua New Guiné, onde o genótipo VGI de *C. gattii* é o mais prevalente (COGLIATE, 2013). Além disso, em um estudo realizado na Índia, *C. neoformans* var. *neoformans* foi mais prevalente (LIN e HEITMAN, 2006). Em um recente estudo em Seul, na Coreia, foi relatado que todos os isolados clínicos de *C. neoformans* possuíam genótipo VNI, sendo esses isolados uma população homogênea, semelhantemente ao que é observado na China e no Japão, que apresentam VNI como o maior genótipo isolado (PARK *et al.*, 2014). Outro país no qual o VNI é o genótipo mais prevalente é a Tailândia, com aproximadamente 95% dos isolados clínicos e 98% de isolados ambientais apresentando genótipo (KAOCHAROEN *et al.*, 2013). O genótipo VNII é raramente isolado, apesar de também apresentar distribuição global. Além disso, o isolamento de VNIII parece estar diretamente relacionado com a presença de VNIV, uma vez que na Europa e nos EUA, onde o isolamento de VNIV é maior que em outras regiões, porcentagens semelhantes de VNIII são observadas (COGLIATE, 2013).

Por muito tempo acreditou-se que *C. gattii* era endêmico em regiões de clima temperado. Entretanto, diversos relatos de casos de criptococose causada por essa espécie em locais considerados não endêmicos ocorreram, endossando o conceito de que essa espécie possui distribuição global (BALTAZAR e RIBEIRO, 2008; CHATURVEDI e CHATURVEDI, 2011). Um estudo recente apontou evidência de dispersão global de *C. gattii*, sendo o genótipo VGI o mais amplamente distribuído no mundo (KIDD *et al.*; 2007; MA e MAY, 2009; PARK *et al.*, 2014). Entretanto, foi observado um perfil diferente no Canadá, onde um surto de criptococose ocorreu em 1999. Dos mais de 218 isolados na ilha de Vancouver, mais de 97% eram VGII, sendo o restante VGI (MA e MAY, 2009; BIELSKA e MAY, 2016). Atualmente, nesse país o genótipo VGII apresenta a mesma proporção de isolamento clínico e a maior parte dos isolados ambientais são VGIIa (CHEN *et al.*, 2014; ESPINEL-INGROFF e KIDD, 2015). O genótipo VGI é o mais prevalente na Austrália, e sua maior incidência é observada na população aborígine situada no

norte do país (NGAMSKULRUNGROJ *et al.*, 2012). Apesar de causar infecções principalmente em indivíduos hígidos, *C. gattii* também apresenta risco para indivíduos imunocomprometidos. Um estudo reportou a presença dessa espécie em pacientes com o sistema imune comprometido, no sudeste da Califórnia (BYRNES *et al.*, 2011). Os sorotipos VGIII e VGIV estão mais associados a infecções em pacientes imunossuprimidos, incluindo aqueles com HIV/AIDS (BYRNES *et al.*, 2011; NGAMSKULRUNGROJ *et al.*, 2012; ESPINEL-INGROFF e KIDD, 2015).

No Brasil, *Cryptococcus neoformans* é encontrado predominante nas regiões sul, sudeste e centro-oeste, enquanto *C. gattii* é endêmico no nordeste brasileiro (GULLO *et al.*, 2013). Em um estudo de sorotipagem de isolados clínicos e ambientais de *C. neoformans* e *C. gattii* no Brasil, NISHIKAWA *et al.*, 2003, observaram a prevalência do sorotipo A (*C. neoformans* var. *grubii*) no sul e sudeste, enquanto o sorotipo B (*C. gattii*) foi mais prevalente no nordeste do país. Em estudo epidemiológico reunindo estudos de caso da criptococose no mundo, COGLIATI (2013) observou que, na América do Sul e Central, o Brasil é o país que apresenta o maior número de isolados. Com 5.709 isolados relatados no estudo, o Brasil foi o quarto país do mundo com maior número de isolados de *Cryptococcus* spp., sendo os genótipos VNI e VGII os mais prevalentes. Além disso, o país foi o que apresentou o maior número de isolados VGII mundialmente (COGLIATI, 2013). Um estudo de genotipagem de 50 isolados clínicos e ambientais de *C. gattii* no Brasil mostrou que 47 dessas amostras eram VGII, enquanto apenas três eram VGI, sendo este um perfil sorológico semelhante ao observado no Canadá (MA e MAY, 2009; SILVA *et al.*, 2012). Foram isolados no país todos os genótipos de *C. gattii*, exceto VGIV (COGLIATI, 2013).

1.6 Tratamento

O tratamento de micoses é difícil devido ao fato das células fúngicas apresentarem diversas maquinarias celulares similares às das células animais (IDNURM *et al.*, 2005). Além disso, poucos são os antifúngicos disponíveis para o tratamento da criptococose em consequência à dificuldade de serem desenvolvidos fármacos que atuam em alvos conservados a uma alta gama de fungos, mas que ao mesmo tempo sejam divergentes entre fungos e células de mamíferos (CHONG *et al.*, 2012; FERREIRA *et al.*, 2013).

Estudos demonstraram que a combinação de antimicrobianos pode reduzir drasticamente a quantidade de antifúngico necessária para inibir completamente o crescimento do patógeno

(CHONG *et al.*, 2012). Essa redução na quantidade de antifúngico empregada no tratamento da criptococose representa uma diminuição da resistência do patógeno frente a esses fármacos. Além disso, a combinação entre anfotericina B e fluconazol reduzem a toxicidade, aumentando a tolerância do paciente durante o tratamento (SANTOS *et al.*, 2012). Entretanto, SANTOS e colaboradores (2012), demonstraram que a combinação entre anfotericina B e o fluconazol é dose-dependente, sendo que o aumento da concentração de fluconazol pode resultar em um efeito antagônico entre esses antifúngicos. Além disso, o sucesso da terapia com anfotericina B e fluconazol tem se mostrado mais difícil em pacientes infectados por *C. gattii*, em comparação àqueles infectados por *C. neoformans* (SILVA *et al.*, 2012).

1.6.1 Anfotericina B

No final dos anos 1950, a anfotericina B (AmB) foi introduzida como agente terapêutico no tratamento da criptococose, permanecendo até hoje como droga de escolha no tratamento dos pacientes que apresentam essa doença (SAAG *et al.*, 2000; LARSEN *et al.*, 2004). Esse antifúngico faz parte da classe dos poliênicos e seu alvo é o ergosterol, principal componente da membrana celular fúngica (FERREIRA *et al.*, 2013). A anfotericina B se liga ao ergosterol, interferindo na permeabilidade da célula em decorrência da formação de poros na membrana (RANG-DALE, 2007). O tratamento com a anfotericina B se dá pela administração intravenosa, uma vez que a administração oral resulta em uma baixa absorção desse antifúngico. Complexos com lipossomos também são formados a fim de melhorar a farmacocinética e reduzir os efeitos colaterais causados pela anfotericina B (RANG-DALE, 2007). Apesar de ser altamente eficiente, a anfotericina B é nefrotóxica, o que limita o seu uso (SANTOS *et al.*, 2012; SLOAN e PARRIS, 2014).

1.6.2 5-Flucitosina

Introduzida no tratamento da criptococose no final dos anos 1970, a 5-Flucitosina (5-FC) possui um baixo espectro de ação, sendo utilizada principalmente contra leveduras (LARSEN *et al.*, 2004; RANG-DALE, 2007). Esse antifúngico é um análogo de pirimidina, que tem como alvo a divisão celular, sendo convertido em 5-fluorouracil (5-FU) somente em células fúngicas, interferindo na síntese de DNA devido à inibição que esse antimetabólito promove na timidilato sintetase (RANG-DALE, 2007; GULLO *et al.*, 2013). Apesar da indicação da combinação entre a

anfotericina B e a 5-Flucitosina durante as fases de indução e consolidação do tratamento contra a criptococose, a 5-Flucitosina não está mais disponível para consumo no Brasil (MOREIRA, 2005).

1.6.3 Azólicos

Os azólicos foram introduzidos no tratamento da criptococose no final dos anos 1980 e início dos anos 1990, sendo o fluconazol e o itraconazol antifúngicos pertencentes a essa classe (LARSEN *et al.*, 2004). Esses fármacos atuam sobre o ergosterol, inibindo sua síntese, o que culmina na alteração da composição da membrana plasmática e, portanto, a sua integridade (RANG-DALE *et al.*, 2007). Essa inibição ocorre em decorrência da atuação do fluconazol e do itraconazol sobre a enzima codificada pelo gene *ERG11* (SHAPIRO *et al.*, 2011), a lanosterol 14- α demetilase, promovendo o acúmulo de lanosterol, que é um intermediário tóxico para as células fúngicas (CHONG *et al.*, 2012). A capacidade do fluconazol de atingir altas concentrações no líquido cefalorraquidiano e nos líquidos oculares faz desse fármaco uma boa opção para o tratamento de pacientes com meningite criptocócica (RANG-DALE *et al.*, 2007).

1.7 Resistência aos antifúngicos

O tratamento da criptococose se dá pela administração de fluconazol, quando a infecção não é severa. Nos casos mais graves, a anfotericina B é utilizada, apesar de sua nefrotoxicidade (SMITH *et al.*, 2014). Entretanto, nos últimos anos tem sido observado um aumento no isolamento de linhagens menos sensíveis e resistentes aos fármacos utilizados no tratamento da doença (CHEN *et al.*, 2015; SMITH *et al.*, 2015). *C. gattii* é menos suscetível a esse fármaco, sendo observada uma susceptibilidade menor do genótipo VGII em relação aos genótipos VGI e VGIII (CHONG *et al.*, 2012). Foi observada uma menor sensibilidade aos azólicos em linhagens de *C. gattii* isoladas no Noroeste do Pacífico, em comparação com outras cepas de *C. gattii* e *C. neoformans* (GAST *et al.*, 2013).

O uso prolongado de fluconazol pode promover a resistência do fungo a esse antifúngico (SANTOS *et al.*, 2014). Um estudo realizado por SILVA *et al.*, 2012 demonstrou altos valores de concentração inibitória mínima (CIM) de *C. gattii* frente ao fluconazol, sendo que isolados VGII apresentaram os maiores valores de CIM. A resistência aos azólicos também pode causar alterações morfológicas em *C. gattii*. Em um estudo realizado com uma linhagem com indução à resistência

ao fluconazol, foi observada a alteração da cápsula polissacarídica dessa levedura, bem como uma maior internalização dessa desse patógeno por macrófagos (SANTOS *et al.*, 2014).

Mutações ou a superexpressão do gene *ERG11*, que codifica lanosterol 14- α demetilase, bem como a superexpressão de proteínas relacionadas a bombas de efluxo de fluconazol são um dos mecanismos pelos quais *Cryptococcus* spp. se tornam resistentes a esse azólico (SHAPIRO *et al.*, 2011; CHEN *et al.*, 2014; BASSO *et al.*, 2016). Em um estudo com cinco isolados clínicos recuperados consecutivamente de um paciente com AIDS, os quatro primeiros isolados foram sensíveis ao fluconazol, sendo que o quinto isolado apresentou resistência a esse antifúngico. Após análise do gene *ERG11*, foi observada a substituição do aminoácido G484S no lanosterol 14- α demetilase, indicando que essa mutação resultou em uma menor susceptibilidade ao antifúngico (RODERO *et al.*, 2003). O efluxo de antifúngicos é um outro mecanismo regulatório encontrado em leveduras, conferindo uma maior resistência à ação desses compostos e pode ocorrer em consequência à superexpressão de genes relacionados ao transporte desses fármacos para fora das células fúngicas (CHONG *et al.*, 2012; BASSO *et al.*, 2016). Em um estudo foi observada a super expressão do mRNA responsável por codificar o gene *AFR1*, que está relacionado ao transporte do tipo ABC em *C. neoformans* (SHAPIRO *et al.*, 2011).

Um outro mecanismo de resistência de *Cryptococcus* spp. é a dissomia de cromossomos dessas espécies, conferindo um perfil de heterorresistência aos antifúngicos. Estudos de hibridização genômica demonstraram que a resistência ao fluconazol ocorreu devido à dissomia de múltiplos cromossomos do genoma fúngico, sendo esse perfil de resistência perdido após a remoção do antifúngico (heterorresistência) (SIONOV *et al.*, 2008).

1.8 Vias metabólicas de *Cryptococcus* spp.

O metabolismo é caracterizado por reações, catalisadas por enzimas, de quebra (catabolismo) e formação (anabolismo) de moléculas necessárias para a sobrevivência dos organismos. Dentre as funções do metabolismo estão a manutenção da homeostase, crescimento, produção de energia e eliminação de excretas (CHEN e HOFESTÄDT, 2006). Alterações no ambiente em que os micro-organismos vivem implicam na necessidade de estratégias metabólicas bastante desenvolvidas (FERNANDES *et al.*, 2015). Pressões ambientais podem induzir a ativação ou silenciamento das mesmas. Fatores como a disponibilidade de nutrientes podem influenciar o

ajuste dos perfis de transcrição e metabolismo dos organismos (BRILLI *et al.*, 2007). Geralmente, os processos biológicos envolvem mais de uma via, resultando em uma conexão entre essas vias metabólicas nos organismos (WERNER, 2008). Estudos metabólicos envolvendo *C. gattii* e *C. neoformans* têm sido desenvolvidos nos últimos anos, contribuindo para uma maior compreensão acerca da patogenicidade desses organismos, bem como dos mecanismos de resistência desses micro-organismos frente aos antifúngicos atualmente empregados no tratamento da criptococose.

C. neoformans é capaz de formar biofilme, que confere maior resistência a antifúngicos e defesa do hospedeiro, sugerindo uma diferenciação no metabolismo entre células planctônicas e células que formam biofilme (SANTI *et al.*, 2014). Em um estudo proteômico, SANTI *et al.* (2014) observaram uma maior atividade de alanina aminotransferase, catalase, superóxido dismutase e fosfatase em biofilme de *C. neoformans*, quando comparado às células planctônicas.

Em um estudo sobre a resposta de *C. gattii* frente ao fluconazol, realizado por CHONG *et al.*, 2012, foi observada a indução da enzima piruvato quinase, componente do complexo de síntese de ácidos graxos, sugerindo que a via de síntese de ácidos graxos ocorre preferencialmente durante o stress causado pelo tratamento com fluconazol, permitindo a recomposição da membrana após os danos causados por esse azólico. Também foi observado um aumento na expressão de proteínas relacionadas à síntese de proteínas, o que indica a produção de novas proteínas. Por meio da utilização do vocabulário funcional definido pelo *Gene Ontology*, foi possível caracterizar as interações entre as proteínas de *C. gattii* frente ao estresse provocado pelo fluconazol (Figura 3).

Apesar do aumento do uso das ciências ômicas, poucos são os estudos envolvendo o proteoma total dos patógenos do gênero *Cryptococcus*. Sendo assim, o desenvolvimento de estudos acerca dos processos metabólicos envolvidos na patogenicidade dessas espécies se faz necessário.

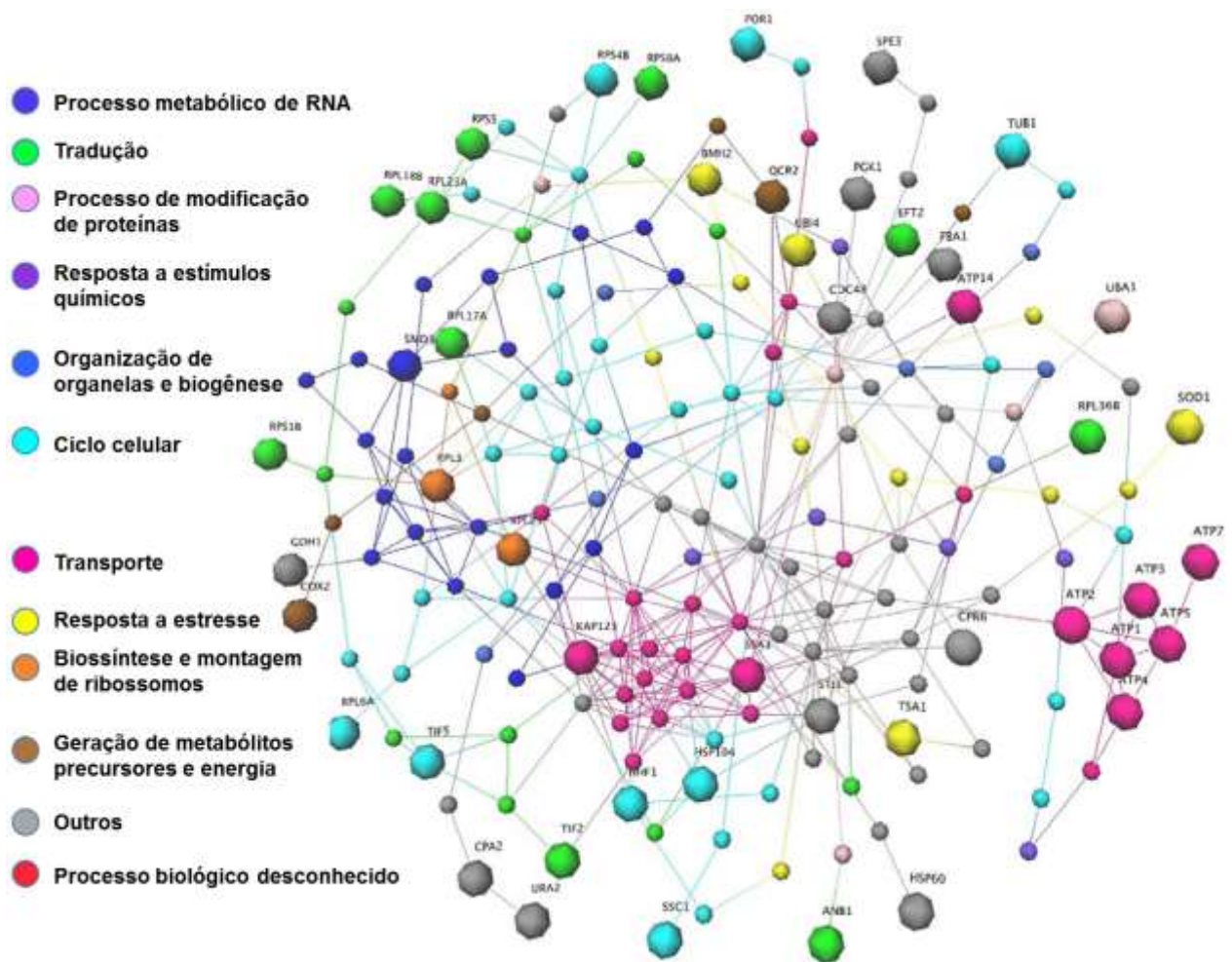


Figura 3: Resposta de *C. gattii* frente ao fluconazol. A análise de redes revela agrupamento de proteínas com funções semelhantes, e proteínas com um elevado grau de conectividade. (Fonte: Adaptado de CHONG *et al.*, 2012).

1.9 Bioinformática

A bioinformática é uma área de estudo que tem crescido rapidamente nos últimos anos, possibilitando uma análise mais rápida de dados e o compartilhamento de informações biológicas (ESPINDOLA *et al.*, 2010). Essa ciência foi originada na década de 1960, após o emprego do uso de computadores na análise de dados oriundos da Biologia Molecular (CATANHO *et al.*, 2007). A necessidade do uso da bioinformática se deve em decorrência do desenvolvimento de diversas técnicas, como o sequenciamento de última geração e os dados obtidos por meio das ciências ômicas (PINHO *et al.*, 2012). Na década de 1990, após o advento de estudos a nível do genoma, transcriptoma e proteoma, a bioinformática conquistou um maior espaço nos estudos biológicos, sendo hoje de extrema importância na análise de dados provenientes desses estudos (CATANHO

et al., 2007). Nesse contexto da era pós-genômica, na bioinformática originou-se o termo ômico, que faz referência à análise dos sistemas biológicos como um todo, sendo que cada tipo de análise possui um tipo de ferramenta, base de dados e softwares que permitem a análise desses sistemas. Dentro dessas ciências ômicas, está a Metabolômica (Figura 4), que é o estudo do metabolismo dos organismos frente a estresses causados pelo ambiente ou alterações genéticas (ESPINDOLA *et al.*, 2009).

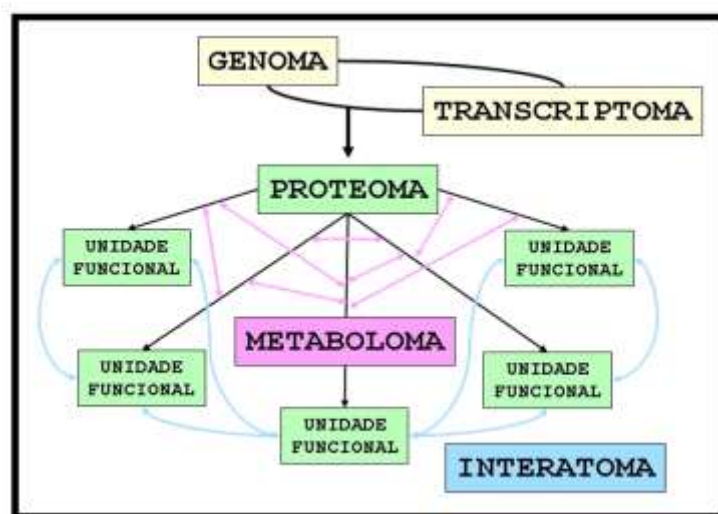


Figura 4: Representação esquemática das principais ciências ômicas. Fonte: Espindola *et al.*, 2009.

Dentre as várias aplicações da bioinformática, pode-se citar a construção e predição de rotas metabólicas a partir de sequências genômicas, que é possível pelo uso de métodos baseados na homologia entre a sequência analisada e as sequências previamente determinadas (GABALDÓN e HUYNEN, 2004). A predição de vias metabólicas pode envolver a predição de vias que são conhecidas em outros organismos ou vias que não foram previamente descritas (DALE *et al.*, 2010). Essa estratégia pode ser empregada na produção de substâncias de importância comercial como terpenoides, biodiesel, álcool e metabólitos secundários, por meio da predição de vias de interesse, seleção e inserção em um organismo de fácil cultivo. Além disso, novas vias podem ser desenvolvidas ou reconstruídas para a otimização da síntese de um determinado produto (MEDEMA *et al.*, 2012).

Diferentes algoritmos podem ser utilizados para a predição e reconstrução de vias metabólicas, dentre os quais pode-se citar o *From Metabolite to Metabolite* (FMM), que permite a reconstrução de vias metabólicas de um metabólito para outro em diferentes espécies. A partir dos

resultados obtidos pela utilização do FMM, genes específicos identificados podem ser clonados em micro-organismos de interesse, como *Escherichia coli* e *Saccharomyces cerevisiae*, sendo utilizados na fabricação de fármacos e biocombustíveis (CHOU *et al.*, 2009). Outra ferramenta utilizada na predição de vias metabólicas é o *Biochemical Network Integrated Computational Explorer* (BNICE), que faz a predição de novas vias de síntese de produtos. Dependendo da busca, esse algoritmo pode prever até 10.000 rotas diferentes de anabolismo e catabolismo para um determinado produto. Além disso, essa ferramenta também é capaz de prever vias metabólicas desconhecidas que sejam possivelmente verdadeiras (MEDEMA *et al.*, 2012).

Outra aplicação dessa ciência é a identificação de fatores de virulência, de alvos para diagnóstico e desenvolvimento de outros tipos de tratamento contra patógenos (CATANHO *et al.*, 2007). Desse modo, a disponibilidade de genomas completos e anotados permite a geração de modelos metabólicos no nível genômico (DE FILIPPO *et al.*, 2012).

A bioinformática dispõe de diversos recursos para análise de dados oriundos das ciências ômicas, dentre os quais pode-se citar o *Kyoto Encyclopedia of Genes and Genomes* (KEGG). Criado em 1995, o KEGG é uma base de dados integrada que abrange diversos tipos de dados como, por exemplo, mapas metabólicos, nomenclatura de enzimas, metabólitos e outras pequenas moléculas (KANEHISA *et al.*, 2012). Os mapas de vias KEGG são diagramas que representam interações e reações necessárias para o metabolismo, bem como o processamento de informações genéticas e processos celulares (KANEHISA *et al.*, 2012).

A bioinformática também é utilizada em estudos proteômicos de micro-organismos de interesse clínico. Um estudo proteômico em linhagem de *C. gattii* na ausência e presença de fluconazol apontou diferenças na expressão de proteínas entre as duas linhagens, sendo que 12 proteínas foram identificadas somente quando o fungo foi cultivado com o antifúngico. A maioria dessas proteínas não possuía função conhecida. Além disso, foi observada a superexpressão de enzimas relacionadas com a via de síntese de glicose, sugerindo que a resistência ao fluconazol requer um gasto energético maior (CHONG *et al.*, 2012).

Considerando-se a grande aplicabilidade da bioinformática e os recursos que essa área de estudos biológicos oferece, o emprego dessa ferramenta para a análise do comportamento de organismos patogênicos como *C. neoformans* e *C. gattii* se mostra bastante promissora no que diz

respeito a uma maior compreensão acerca dos mecanismos envolvidos na patogenicidade dessas espécies, bem como o desenvolvimento de medidas que resultem numa maior eficácia no tratamento da doença.

2 JUSTIFICATIVA

A criptococose é uma doença que afeta, mundialmente, mais de um milhão de pessoas por ano, sendo o Brasil o quarto país com maior número de isolados dos patógenos responsáveis por essa doença (COGLIATI, 2013). Ao contrário de *C. neoformans*, que é conhecido por afetar principalmente indivíduos imunocomprometidos, *C. gattii* é capaz de causar, também, infecções em indivíduos hígidos, sendo frequentemente letal (SANTOS *et al.*, 2012). Nos últimos anos, diversas técnicas foram desenvolvidas para o estudo dos organismos em nível de genoma, transcriptoma e proteoma, sendo a bioinformática uma importante ferramenta utilizada na análise dos dados obtidos (ESPINDOLA *et al.*, 2010).

Tendo em vista a importância clínica de *Cryptococcus*, a dificuldade no tratamento da doença e o aumento da resistência aos antifúngicos, uma melhor compreensão das rotas metabólicas, principalmente aquelas relacionadas à patogenicidade desse fungo, se faz necessária. A elucidação dessas vias será útil para o desenvolvimento de alvos na terapia antifúngica, bem como alternativas para o controle de infecções causadas por esses patógenos. Para isso, se propõe a utilização ferramentas de Bioinformática para a análise e o processamento de informações genéticas já depositadas em bancos de dados como, por exemplo, o NCBI (*National Center for Biotechnology Information*).

3 OBJETIVOS

3.1 Objetivo geral

Caracterizar, *in silico*, vias metabólicas de *Cryptococcus neoformans* e *Cryptococcus gattii*, com base nas sequências existentes nos bancos de dados de domínio público.

3.2 Objetivos específicos

- Realizar a normalização da anotação estrutural e funcional dos genomas de *Cryptococcus* spp. selecionados;
- realizar a predição de vias metabólicas;
- realizar a análise comparativa de vias metabólicas, bem como suas funções, além de avaliar o papel dessas vias na virulência e patogenicidade desses fungos.

4 MATERIAL E MÉTODOS

4.1 Scripts desenvolvidos

Os *scripts* desenvolvidos neste trabalho foram escritos em linguagem de programação *PERL* (<https://www.perl.org/>) e/ou em Bash (<https://www.gnu.org/software/bash/>) e em awk (<https://www.gnu.org/software/gawk/>).

4.2 Genomas de *Cryptococcus* spp.

Os genomas de *Cryptococcus* spp. utilizados no presente estudo foram obtidos a partir do banco de dados do NCBI (*National Center for Biotechnology Information*), o GenBank (em: <https://www.ncbi.nlm.nih.gov/>). Para cada espécie foram utilizados três genomas, sendo todos armazenados pelo banco de dados do NCBI (Tabela 1). Para os genomas que não possuíam predição gênica e anotação funcional, as mesmas foram realizadas pelo grupo Informática de Biosistemas do CPqRR- Fiocruz Minas, entre 2014 e 2015, seguindo a metodologia utilizada por TAVARES, 2016.

Tabela 1: Lista de genomas de *Cryptococcus* spp. utilizados para análise.

Espécie	Linhagem	Sorotipo	Genótipo	Sequenciamento
<i>C. neoformans</i> var. <i>neoformans</i>	JEC21	D	VNIV	Broad Institute
<i>C. neoformans</i> var. <i>grubii</i>	H99	A	VNI	Broad Institute/ Duke Center for Genome Technology
<i>C. neoformans</i> var. <i>neoformans</i>	B-3501A	D	VNIV	Stanford University
<i>C. gattii</i>	WM276	B	VGI	University of British Columbia
<i>C. gattii</i>	R265	B	VGII	Broad Institute
<i>C. gattii</i>	CBS7750	B	VGII	Centre for Genomic Regulation

4.3 Anotação do proteoma de *Cryptococcus* spp.

4.3.1 Artemis

Para a visualização e anotação de sequências de DNA, foi utilizada a ferramenta Artemis (<http://www.sanger.ac.uk/Software/Artemis>), a qual é um algoritmo escrito em JAVA que disponibiliza os resultados das análises de predição em sua interface, bem como nos respectivos quadros de leitura (ORFs) (RUTHERFORD *et al.*, 2000; CARVER *et al.*, 2012).

4.3.2 Predição de RNAs ribossomais e RNAs transportadores

De forma a contribuir para a anotação funcional dos genomas utilizados no presente estudo, foi realizada a predição de RNAs ribossomais e transportadores.

Os RNAs ribossomais foram preditos pelo software RNAmmer versão 1.2 (http://www.cbs.dtu.dk/cgi-bin/nph-sw_request?rnammer). Os seguintes parâmetros foram utilizados:

```
perl rnammer -S euk -m lsu,ssu,tsu -gff <arquivo-de-saída-gff> -f <arquivo-de-saída-f> -xml <arquivo-de-saída-xml> -h <arquivo-de-saída-hmm> <arquivo-de-entrada>
```

Legenda dos parâmetros:

-S: euk (predição de RNAs de eucariotos)
-m: lsu,ssu,tsu
-Formatos de saída: gff, f, xml e h

A predição de RNAs transportadores foi feita pelo tRNAscan-SE versão 1.3.1 (<http://lowelab.ucsc.edu/tRNAscan-SE/>), sendo os seguintes parâmetros utilizados:

```
tRNAscan-SE -o <output> -f <output> -m <output> <input>
```

Legenda dos parâmetros:

-o: arquivo de saída
-f: salva estruturas secundárias no arquivo de saída
-m: salva o sumário estatístico no arquivo de saída
Output format: Tabular
Searching with: tRNAscan + EufindtRNA -> Cove
Covariance model: TRNA2-euk.cm
tRNAscan parameters: Strict
EufindtRNA parameters: Relaxed (Int Cutoff= -32.1)

As informações referentes aos RNAs ribossomais e RNAs transportadores foram inseridas em cada genoma.

4.4 Análise de similaridade de sequências

Uma análise de similaridade de sequências foi realizada utilizando-se os genomas completos dos organismos do presente estudo. Para isso, um genoma de cada espécie de *Cryptococcus* foi utilizado como sequência de busca, sendo eles das linhagens *C. gattii* R265 e *C. neoformans* var. *grubii* H99. A sequência de *C. gattii* R265 foi comparada com os genomas de *C. gattii* CBS 7750 e *C. gattii* WM276, que foram definidos como banco de dados. Já o genoma da linhagem *C. neoformans* var. *grubii* H99 foi comparado aos genomas de *C. neoformans* B-3501A e *C. neoformans* JEC21, sendo estes dois utilizados como banco de dados.

Os genomas utilizados como banco de dados foram formatados pelo programa formatdb, sendo os seguintes parâmetros utilizados:

```
formatdb -i <arquivo-de-entrada> -p F -n <arquivo de saída>
```

Legenda dos parâmetros:

-i: arquivo a ser formatado
 -p: F (o arquivo possui sequências de nucleotídeos e não de proteínas)
 -n: nome do arquivo de saída

Também foi realizada uma comparação entre as duas espécies estudadas, sendo o genoma de *C. gattii* R265 comparado ao da linhagem *C. neoformans* var. *grubii* H99.

O script *big_blast.pl* foi utilizado para a análise de similaridade de sequências. Esse script executa o algoritmo BLAST e formata o *output* gerado por meio da execução do programa MSPcrunch. Esse último por sua vez gera um arquivo que serve de entrada para o programa ACT (*Artemis Comparison Tool*). O *big_blast.pl* apresenta valores padrão para o tamanho do fragmento, *score* mínimo, além de utilizar especificamente o programa *blastn* para realização das análises. O tamanho dos fragmentos escolhido foi de 40.000 pb e *score* mínimo de 150. No caso de alterações em relação a esses padrões, os parâmetros devem ser assinalados na linha de comando. Como o padrão foi utilizado na análise, a seguinte linha de comando foi utilizada:

```
big_blast.pl <banco-de-dados> <arquivo-fasta>
```

4.5 ACT

O ACT é um programa derivado do *Artemis*, também escrito em JAVA, que permite a visualização de resultados de comparações por similaridade de sequências feitas entre diferentes genomas. O programa é capaz de ler diferentes formatos de sequências, como “.embl” ou “GENBANK”, bem como “*feature tables*”, e não apresenta restrição quanto ao tamanho dos arquivos. Os arquivos de comparação no formato MSPcrunch, gerados pelo *script big_blast.pl*, foram os arquivos de entrada para a visualização da similaridade das sequências.

4.6 Predição e anotação de vias metabólicas

A predição de vias metabólicas foi realizada pela incorporação de anotação funcional com termos definidos no Gene Ontology (GO) (<http://geneontology.org/>) (The Gene Ontology Consortium, 2015). Por meio da incorporação desses termos anotadores foram correlacionados os *EC number* (*Enzyme Commission number*) de cada enzima anotada. Para isso, o programa Blast2GO PRO versão 4.0.7 foi utilizado. Essa plataforma permite a anotação funcional e mineração de dados genômicos pelo uso de vocabulários comumente utilizados em análises de dados biológicos, tais como termos do GO, vias metabólicas do KEGG e códigos de enzimas (EC) (CONESA e GÖTZ, 2008).

A primeira etapa de anotação pelo Blast2GO consistiu em uma análise de similaridade de sequências entre um banco de dados e o objeto de estudo. Para isso, uma análise de similaridade de sequências entre os proteomas preditos e o banco de dados não redundante (NR) do NCBI foi feita utilizando-se o algoritmo BLAST versão 2.5.0+ e foram utilizados os parâmetros descritos a seguir. O acesso ao banco de dados ocorreu em dezembro de 2016.

```
blastp -db <banco-de-dados> -query <arquivo-de-entrada> -out <arquivo-de-saída>
-outfmt 5 -num_threads 45 -show_gis -num_alignments 5 -evalue 0.00001
```

Legenda dos parâmetros:

```
-db: banco de dados
-query: arquivo de entrada
-out: arquivo de saída
-outfmt: 5 (formato xml)
-num_threads: 45 (número de processadores)
-show_gis: mostrar GIs do NCBI
-num_alignments: 5 (número de alinhamentos)
-evalue: 0.00001
```

Os resultados obtidos foram utilizados como arquivo de entrada para posterior anotação e mapeamento de proteínas pelo Blast2GO. O mapeamento se deu pela incorporação dos termos do GO associados aos *hits* resultantes do alinhamento. Em seguida, os números EC das enzimas e vias metabólicas do KEGG (*Kyoto Encyclopedia of Genes and Genomes*) foram inseridos na etapa de anotação.

4.7 Criação de banco de dados de virulência

Para a criação do banco de dados de proteínas relacionadas à virulência de *Cryptococcus*, foram selecionadas proteínas do Swiss-Prot, um banco não redundante altamente curado cujas proteínas são revisadas e manualmente anotadas (THE UNIPROT CONSORTIUM, 2014). Esse banco foi denominado *db_virulencia_crypto_swissprot* e os termos utilizados para a seleção das proteínas foram “*cryptococcus AND laccase*”, “*cryptococcus AND phospholipase*”, “*cryptococcus AND capsule*”, “*cryptococcus AND urease*”, “*cryptococcus AND pathogenicity*”, “*cryptococcus AND superoxide dismutase*”, “*cryptococcus AND virulence*”, “*cryptococcus AND glyceraldehyde 3 phosphate*” e “*cryptococcus AND polysaccharide*”. Proteínas repetidas foram removidas do banco de dados por meio de uma curadoria manual. O acesso ao banco de dados do Swiss-Prot foi realizado no período de 22 a 25 de novembro de 2016.

4.8 Análise de enriquecimento funcional

Visando uma maior compreensão acerca dos fatores de virulência de *Cryptococcus* spp., uma análise de enriquecimento funcional de fatores de virulência desse gênero foi realizada para cada um dos proteomas preditos. Para a identificação de proteínas relacionadas à virulência nos proteomas, foi realizada uma análise de similaridade de sequências entre os proteomas e um BD de proteínas de virulência de *Cryptococcus* spp.

A análise de similaridade de sequências dos proteomas preditos contra o banco *db_virulencia_crypto_swissprot* foi realizada por meio do algoritmo BLAST versão 2.5.0+ e os seguintes parâmetros foram utilizados:


```
blastp -db <db_virulencia_crypto_swissprot> -query <arquivo-fasta> -out
<arquivo-fasta> -outfmt 7 -num_threads 45 -show_gis -num_alignments 1 -evaluate
0.00001
```

Legenda dos parâmetros:

```
-db: banco de dados
-query: arquivo de entrada
-out: arquivo de saída
-outfmt: 7 (formato tabular com linhas comentadas)
-num_threads: 45 (número de processadores)
-show_gis: mostrar GIs do NCBI
-num_alignments: 1 (número de alinhamentos)
-evaluate: 0.00001
```

Para analisarmos o enriquecimento funcional dos fatores de virulência nos proteomas preditos, foi utilizado o programa Blast2GO PRO versão 4.0.7, que utiliza os identificadores das proteínas associadas à virulência para a análise (proteínas teste). Identificadores repetidos de proteínas foram removidos e o Teste Bilateral de Fisher com correção de Bonferroni foi empregado para a análise estatística. Valores de p menores que 0,05 foram considerados significativos.

4.9 Mapeamento das vias metabólicas e associação aos fatores de virulência

Os fatores de virulência foram mapeados nas vias metabólicas por meio dos resultados provenientes do Blast2GO, de modo que as proteínas associadas à virulência foram mapeadas em vias metabólicas presentes no KEGG (KANEHISA *et al.* 2004). Para isso, as informações referentes ao mapeamento de proteínas associadas à virulência foram extraídas dos resultados obtidos na anotação e mapeamento do proteoma total no Blast2GO. As proteínas associadas à virulência foram, então, mapeadas nas vias metabólicas presentes no KEGG.

Adicionalmente, as vias metabólicas mapeadas dos organismos foram comparadas pelo uso do *script compara_vias.pl*. Os seguintes parâmetros foram utilizados:

```
perl compara_vias.pl <arquivo-de-entrada> <arquivo-de-comparação> <error-log>
<arquivo-de-saída>
```

4.10 Análise de agrupamento

Para avaliar o conjunto de genes compartilhados entre todos os organismos estudados, ou seja, o *core genome*, bem como o conjunto de genes potencialmente linhagem-específicos, foi realizada uma análise de agrupamento utilizando o programa OrthoMCL versão 1.4 (LI *et al.*, 2003). Esse programa identifica genes ortólogos, que são aqueles que evoluíram de um ancestral

comum pelo processo de especiação. Essa identificação se faz importante em estudos de genômica comparativa e processos de anotação funcional (LI *et al.*, 2003).

Primeiramente, o OrthoMCL realiza um blastp de um proteoma em análise contra ele mesmo e contra os outros proteomas a serem analisados. As prováveis relações ortólogas e parálogas são definidas pelos pares de sequências que apresentam os melhores hits provenientes do blastp. A segunda etapa da análise consiste na conversão dessas prováveis relações em um gráfico, onde os nós representam as proteínas e as pontes indicam a relação entre essas proteínas. Essas pontes possuem pesos calculados com base nos resultados do blast e são posteriormente normalizadas. Em seguida, o algoritmo MCL (ENRIGHT *et al.*, 2002) é utilizado para selecionar apenas a clusterização de grupos ortólogos e recentes parálogos (LI *et al.*, 2003). Para a análise, os seguintes parâmetros foram utilizados:

```
./orthomcl.pl --mode 1 --fa_files <arquivo--fasta> <arquivo--fasta> <arquivo--fasta>
```

Legenda dos parâmetros:

```
OrthoMCL Mode: 1
P-value Cut-off: 1e-05
Percent Identity Cut-off: 60
Percent Match Cut-off: 60
MCL Inflation: 1.5
Maximum Weight: 316
```

O *core genome* foi avaliado pela extração dos resultados contendo 6 genes e 6 taxa, ou seja, de proteínas que apresentaram clusterização entre todos os organismos estudados. Adicionalmente, proteínas relacionadas à virulência foram identificadas para cada situação. Para isso, proteínas que apresentaram *hits* no blastp dos proteomas preditos contra o BD *db_virulencia_crypto_swissprot* foram comparadas com as proteínas linhagem-específicas e comuns a todos os organismos.

5 RESULTADOS E DISCUSSÃO

5.1 Seleção de genomas de *Cryptococcus* spp.

Seis genomas de *Cryptococcus* foram utilizados neste estudo, sendo três de *C. gattii* e três de *C. neoformans*. Diversos estudos já foram realizados com as linhagens de *C. neoformans* e duas das três de *C. gattii* (WM276 e R265), que foram utilizadas neste trabalho. Por esse motivo, essas linhagens foram selecionadas para a realização de comparações de sequências de nucleotídeos e predição e comparação de vias metabólicas.

5.2 Predição de RNAs ribossomais e RNAs transportadores

De forma a contribuir com a anotação dos genomas utilizados neste estudo, a predição de RNAs ribossomais e RNAs transportadores foi realizada, conforme seção 4.3.2 da metodologia. Em relação aos RNAs ribossomais, a linhagem *C. gattii* WM276 apresentou o maior número de RNAs preditos, com um total de 152 (**Tabela 2**). RNAs ribossomais são moléculas altamente conservadas tanto a nível de sequência quanto de estrutura. Por isso, esses RNAs podem ser utilizados em estudos filogenéticos (LAGESEN *et al.*, 2007).

Tabela 2: RNAs ribossomais e RNAs transportadores preditos.

Genoma	rRNAs preditos				tRNAs preditos
	8s	18s	28s	Total	
<i>C. gattii</i> CBS 7750	2	0	0	2	130
<i>C. gattii</i> R265	4	0	0	4	157
<i>C. gattii</i> WM276	52	50	50	152	142
<i>C. neoformans</i> B-3501A	4	2	2	8	139
<i>C. neoformans</i> H99	1	1	2	4	158
<i>C. neoformans</i> JEC21	5	2	2	9	141

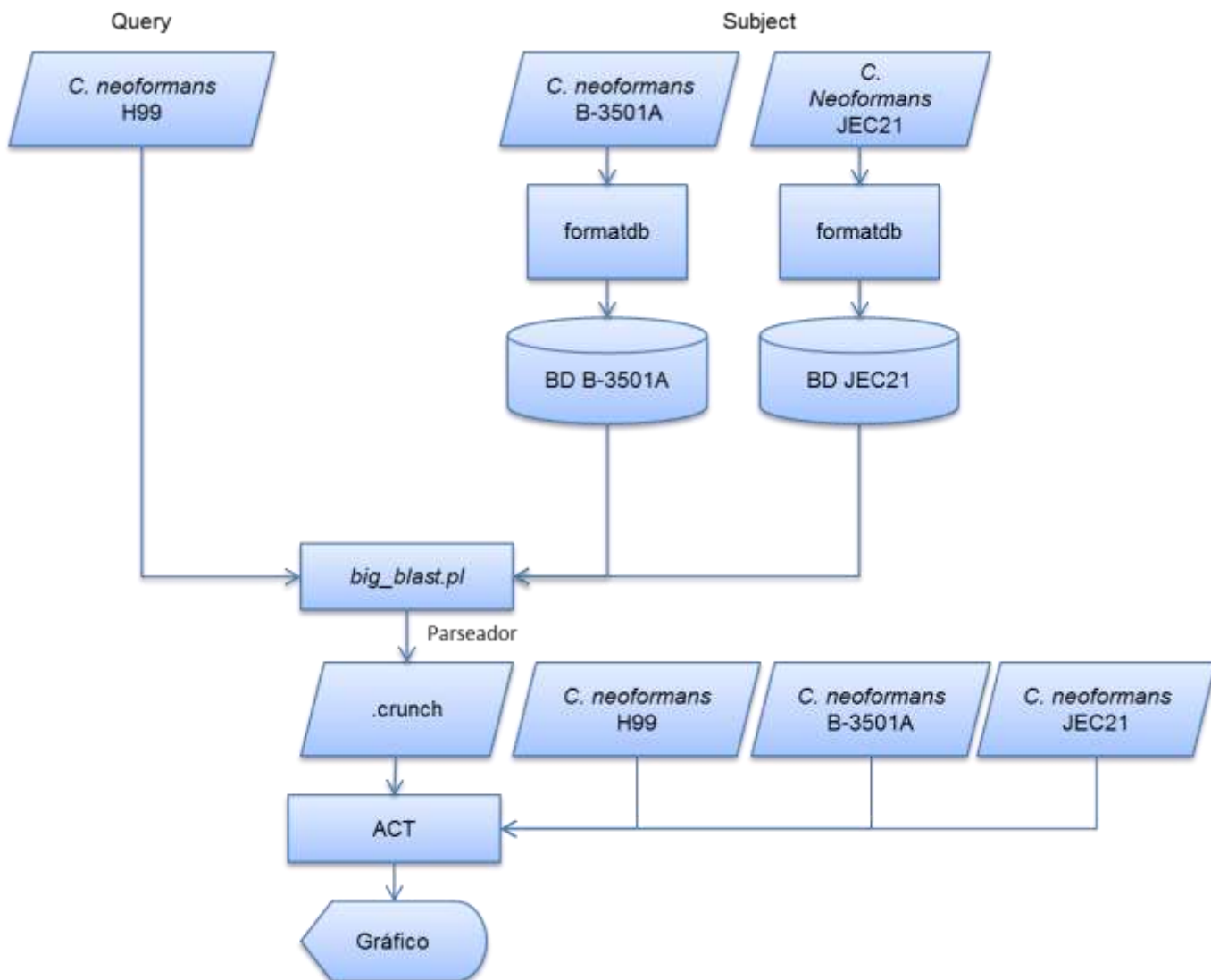
Para todos os genomas estudados houve no mínimo 130 RNAs transportadores preditos (**Tabela 2**). Os RNAs transportadores são os principais componentes necessários para a tradução de RNAs mensageiros em proteínas (LOWE e CHAN, 2016). A predição de RNAs transportadores é importante devido à influência da composição desses RNAs no viés de códons observados em

genes altamente expressos (LOWE e EDDY, 1997; SCHATTNER *et al.*, 2005). Desse modo, a anotação desses tipos de RNAs se faz necessária em genomas.

Adicionalmente, foi realizada a predição desses tipos de RNAs em outros seis genomas de organismos não virulentos de *Cryptococcus* spp. Os RNAs ribossomais e transportadores preditos foram inseridos nos genomas dos organismos utilizados, disponibilizando essas informações para futuros estudos envolvendo esses organismos (dados não mostrados).

5.3 Análise de similaridade de sequências

Com o intuito de comparar os genomas utilizados no presente estudo, uma análise de similaridade de sequências de nucleotídeos foi realizada. Para isso, o script *big_blast.pl* foi utilizado, conforme descrito na seção 4.4 da metodologia. Os resultados provenientes do alinhamento das sequências foram então utilizados para visualização no ACT. Um gráfico de comparação foi gerado para cada situação descrita na metodologia. O genoma da linhagem *C. gattii* R265 foi comparado com os outros dois genomas da mesma espécie. O mesmo foi realizado para *C. neoformans*, sendo a linhagem H99 comparada às outras duas linhagens da mesma espécie, conforme o fluxograma na página a seguir.

Fluxograma 1: metodologia utilizada para a comparação de genomas de *C. neoformans*.

Exemplo de fluxograma utilizado nas análises de similaridade de seqüências e visualização no ACT. No caso específico, o genoma de *C. neoformans* H99 (*query*) é comparado com os genomas das linhagens *C. neoformans* B-3501A e *C. neoformans* JEC21 (*subject*), por meio do script `big_blast.pl`. Os resultados gerados são reformatados por meio de um programa escrito em PERL, gerando arquivos no formato `.crunch` que permitem a visualização dos alinhamentos no `act`.

5.3.1 *Cryptococcus neoformans*

Na comparação entre as linhagens de *C. neoformans*, foram observadas regiões com orientações inversas, coloridas em azul (**Figura 5**).

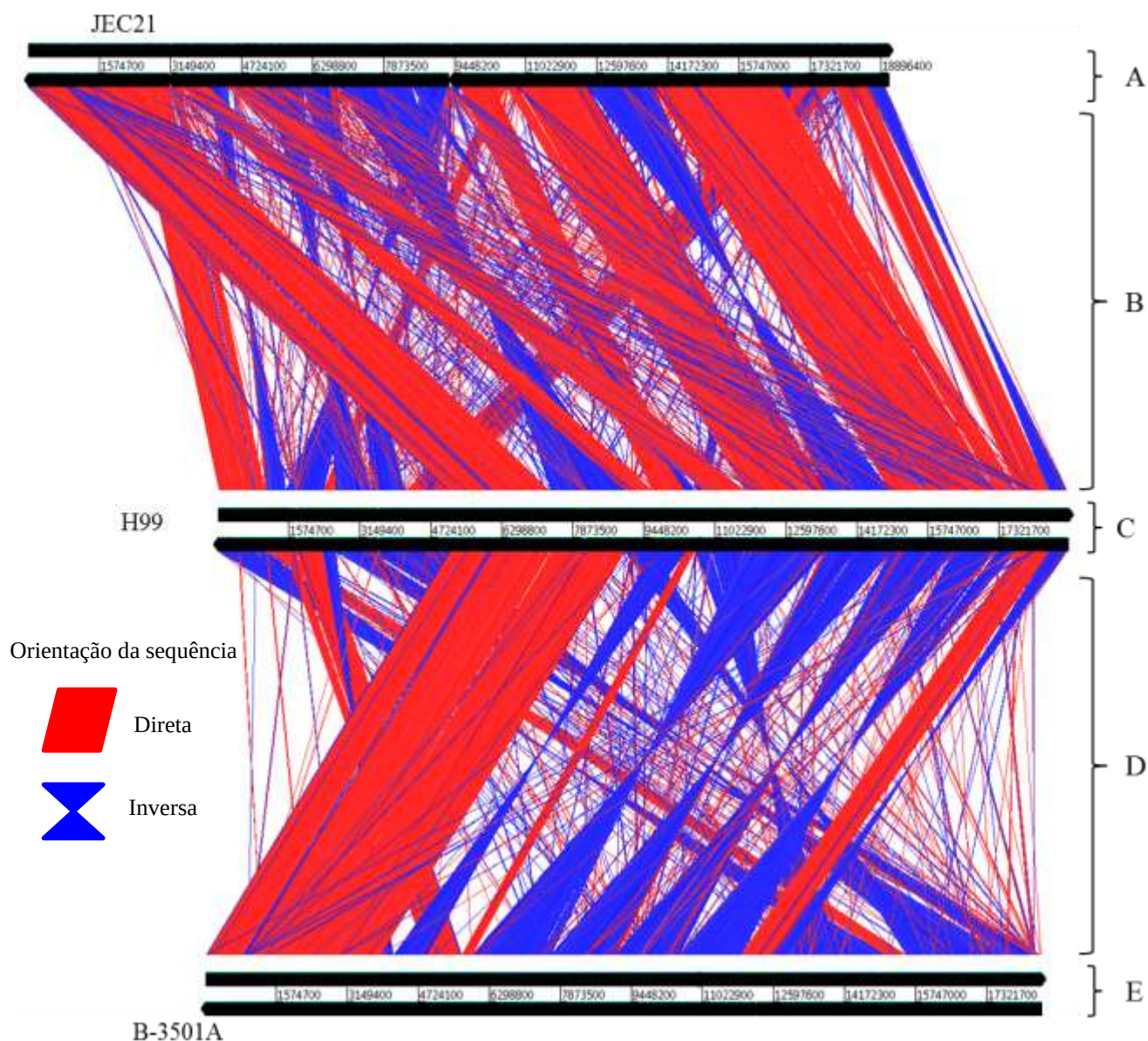
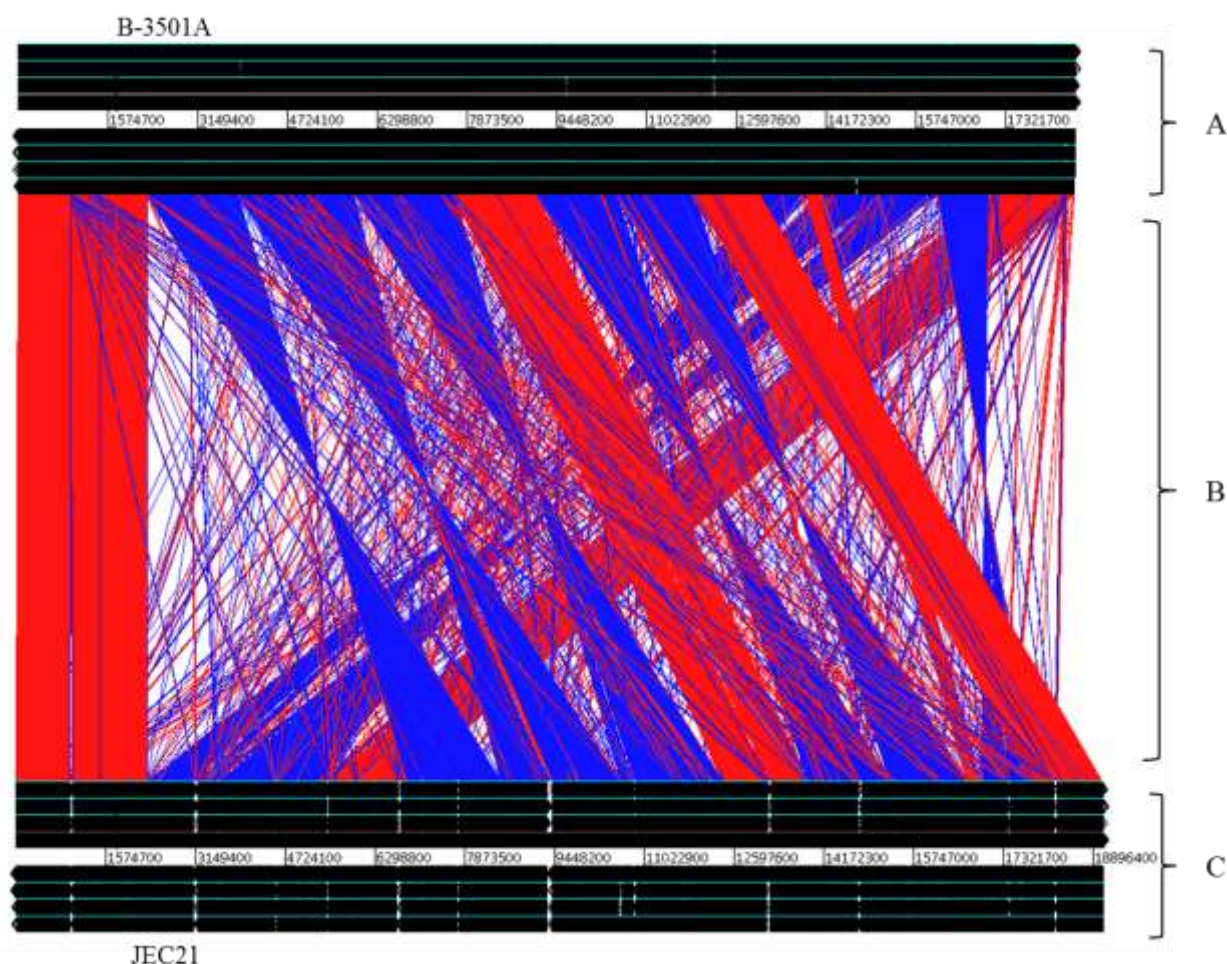


Figura 5: Gráfico de comparação de similaridade global entre *C. neoformans* JEC21, *C. neoformans* H99 e *C. neoformans* B-3501A (do topo para baixo), onde: A, C e E correspondem às sequências genômicas dos organismos em questão. As regiões similares entre os organismos estão interligadas pelos segmentos de reta indicados em B e D. Os segmentos vermelhos e azuis entre as sequências indicam se os alinhamentos são diretos ou invertidos, respectivamente. Sequências com menos de 999 pb foram removidas para facilitar a visualização.

Em uma análise filogenética de linhagens de *Cryptococcus* spp., foi observado que as linhagens *C. neoformans* B-3501A e *C. neoformans* JEC21 são mais próximas filogeneticamente entre si do que com a linhagem *C. neoformans* H99 (D'SOUZA et al., 2011). Nesse estudo, uma análise de clusterização com os proteomas preditos das linhagens H99, B-3501A, JEC21, WM276 e R265 foi realizada, por meio do programa OrthoMCL. Os alinhamentos provenientes dos grupos

contendo cópias únicas de ortólogos para todas as linhagens foram então unidos e uma análise filogenética foi realizada pelo PhyML3.0, sendo as árvores geradas pelo Seaview4 (D'SOUZA *et al.*, 2011). Por esta razão, um alinhamento entre essas duas linhagens foi realizado (**Figura 6**).



Orientação da sequência



Direta



Inversa

Figura 6 Gráfico de comparação de similaridade global entre *C. neoformans* B-3501A e *C. neoformans* JEC21 (do topo para baixo), onde A e C correspondem às sequências genômicas dos organismos em questão. As regiões similares entre os organismos estão interligadas pelos segmentos de reta indicados em B. Os segmentos vermelhos e azuis entre as sequências indicam se os alinhamentos são diretos ou invertidos, respectivamente. Sequências com menos de 999 pb foram removidas para facilitar a visualização.

Foi observado que praticamente todo o genoma da linhagem B-3501A apresenta similaridade com a sequência da linhagem JEC21, a nível de nucleotídeos. De acordo com LOFTUS e colaboradores (2005), apesar de não ser descrito metodologicamente de uma maneira clara quais estratégias foram utilizadas para obtenção da porcentagem de similaridade global entre as sequências, foi relatado que mais de 99% dos genes da linhagem B-3501A possuem regiões apresentando mais de 98% de similaridade de nucleotídeos.

5.3.2 *Cryptococcus gattii*

As linhagens *C. gattii* CBS7750 e R265 apresentaram um alto grau de similaridade (Figura 7). Ambas as linhagens possuem genótipo VGII e foram isoladas na América do Norte, enquanto a linhagem *C. gattii* linhagem WM276 é um isolado ambiental cujo genótipo é VGI. FARRER e colaboradores (2015) observaram que o genótipo VGII foi o primeiro a divergir em *C. gattii*, sendo o grupo mais isolado filogeneticamente em comparação aos outros três (VGI, VGIII e VGIV). Isso poderia explicar o fato de que praticamente todo o genoma das linhagens *C. gattii* R265 e CBS 7750 apresentar um maior número de extensas regiões similares em relação a *C. gattii* WM276.

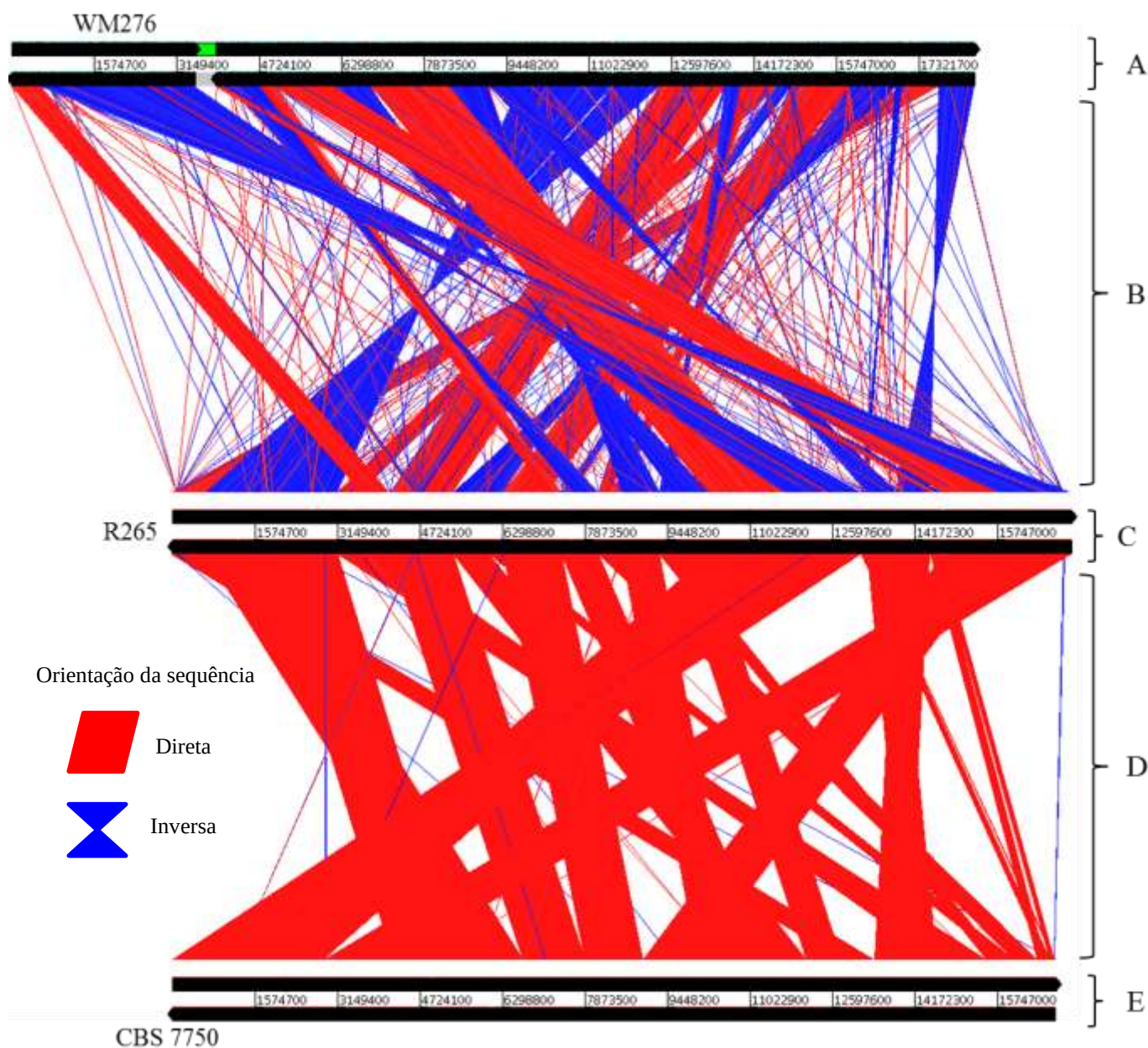


Figura 7: Gráfico de comparação de similaridade global entre *C. gattii* WM276, *C. gattii* R265 e *C. gattii* CBS 7750 (do topo para baixo), onde: A, C e E correspondem às sequências genômicas dos organismos em questão. As regiões similares entre os organismos estão interligadas pelos segmentos de reta indicados em B e D. Os segmentos vermelhos e azuis entre as sequências indicam se os alinhamentos são diretos ou invertidos, respectivamente. Sequências com menos de 999 pb foram removidas para facilitar a visualização.

As linhagens *C. gattii* WM276 e *C. gattii* R265 apresentaram diversas regiões de similaridade com orientações inversas, coloridas em azul. D'SOUZA e colaboradores (2011) observaram um diferente perfil na orientação das regiões de similaridade entre os genomas, no qual poucas inversões foram observadas para as mesmas linhagens.

D'SOUZA e colaboradores (2011) também observaram que as linhagens *C. gattii* R265 e WM276 apresentam cromossomos colineares. Contudo, existem diferenças significativas na identidade das sequências como um todo (divergência de 7,6% entre as sequências), bem como no conteúdo de genes presentes em cada linhagem.

5.3.3 *Cryptococcus gattii* vs *C. neoformans*

Uma comparação entre as duas espécies de *Cryptococcus* foi executada, sendo os genomas de *C. gattii* R265 e *C. neoformans* H99 utilizados (**Figura 8**).

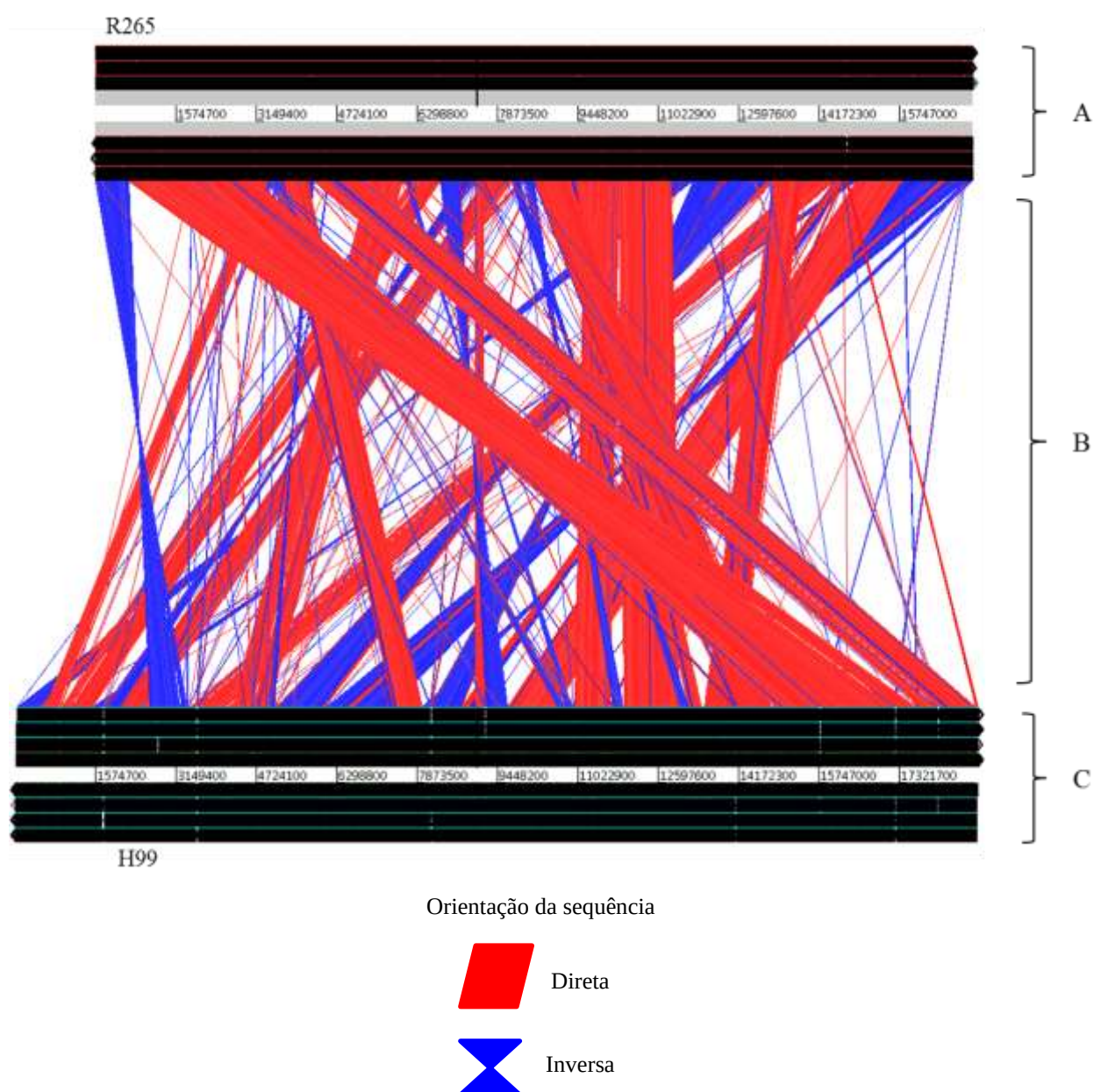


Figura 8: Gráfico de comparação de similaridade global entre *C. gattii* R265 e *C. neoformans* H99 (do topo para baixo), onde A e C correspondem às sequências genômicas dos organismos em questão. As regiões similares entre os organismos estão interligadas pelos segmentos de reta indicados em B. Os segmentos vermelhos e azuis entre as sequências indicam se os alinhamentos são diretos ou invertidos, respectivamente. Sequências com menos de 999 pb foram removidas para facilitar a visualização.

5.4 Predição de vias metabólicas

Os proteomas preditos foram extraídos do Artemis e utilizados para a predição de vias metabólicas de *Cryptococcus*. O programa Blast2GO foi utilizado para o mapeamento dos termos do GO e anotação dos números EC de enzimas, conforme descrito na seção 4.6 de Material e Métodos. Apenas as vias de proteínas associadas à virulência desses patógenos foram selecionadas para esta análise.

Em relação à anotação dos números EC, para todos os proteomas preditos a classe de enzimas que apresentou maior anotação foi a das hidrolases, enzimas que utilizam água para quebrar ligações químicas (**Figura 9**). Em um estudo de secretomas de três linhagens de *Cryptococcus*, foi observada a presença de uma provável Glicosil hidrolase em todos os secretomas analisados, que provavelmente são proteínas degradadoras presentes no secretoma basal de fungos (CAMPBELL *et al.*, 2015).

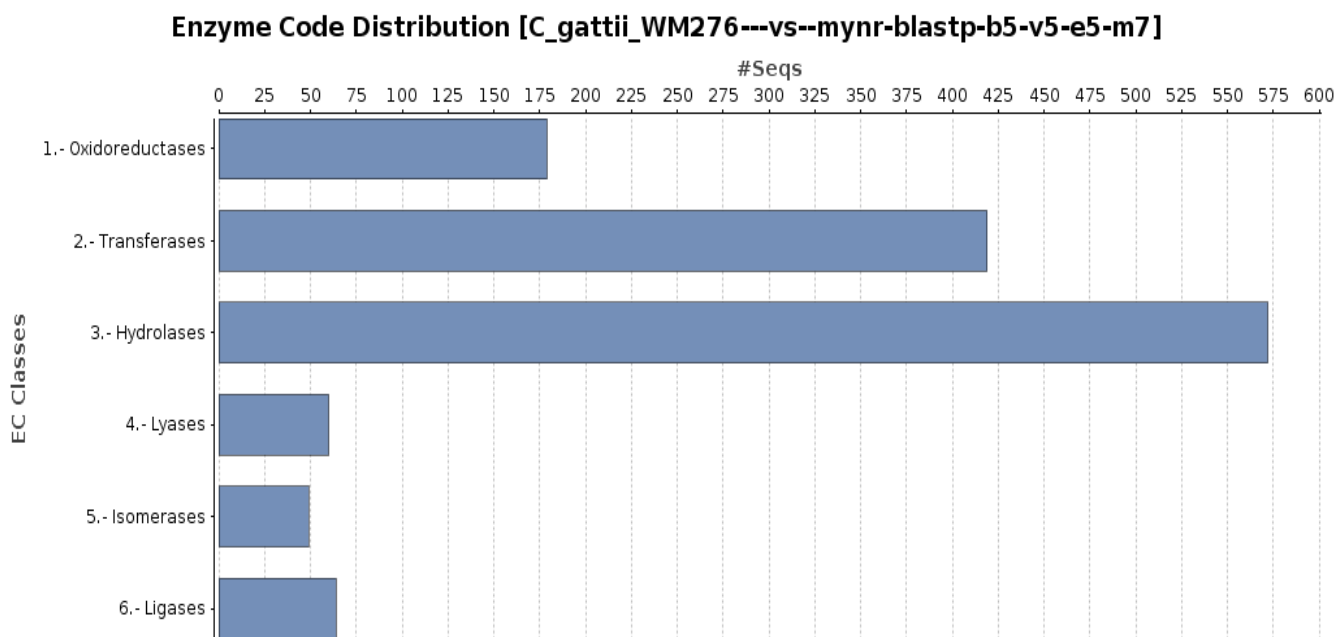


Figura 9: Distribuição de números EC após mapeamento e anotação do proteoma total de *C. gattii* WM276 pelo programa Blast2GO. O eixo x indica o número de sequências mapeadas e anotadas, enquanto o eixo y indica as classes

de enzimas. #Seqs representa o número de sequências mapeadas com os números EC e EC classes indica as seis classes de enzimas associadas ao número EC.

Em outro estudo, LEE e colaboradores (2014) demonstraram que a hidrolase Ysa1 provavelmente está envolvida na regulação da resposta de *Cryptococcus* ao estresse provocado por determinados agentes oxidantes. No estudo, mutantes de *C. neoformans* H99 para a expressão de Ysa1 e o tipo selvagem foram submetidos a diferentes testes para avaliar a resposta ao estresse oxidativo por diferentes agentes. Foi observado que as linhagens mutantes apresentaram maior susceptibilidade à diamida, indicando que a proteína Ysa1 está envolvida na resistência de *C. neoformans* a esse agente.

5.5 Análise de enriquecimento funcional

Para uma maior compreensão acerca da função das proteínas relacionadas à virulência dos organismos estudados e o papel biológico das mesmas, uma análise de enriquecimento funcional foi realizada utilizando o Teste Bilateral de Fisher com Correção de Bonferroni, abordagem presente no programa Blast2GO, de acordo com a metodologia descrita na seção 4.8 deste documento. Esse teste permite verificar se os termos do *Gene Ontology* estão enriquecidos em um grupo em relação a um grupo de referência, que no presente estudo foi o proteoma total de cada organismo.

Por meio dessa análise, foi possível observar as funções sub-representadas e super-representadas nas proteínas relacionadas à virulência de *Cryptococcus* spp. em relação ao proteoma total, permitindo uma maior compreensão sobre as funções mais presentes nesse conjunto de proteínas.

Em relação à categoria denominada “Processo Biológico” do GO, foi possível observar um enriquecimento dos termos explicitados na **Figura 16** e na **Figura 17** do apêndice deste documento em todas as linhagens de *C. gattii*, exceto a *C. gattii* WM276, uma vez que não foram obtidos resultados estatisticamente significativos na análise de enriquecimento de função para essa linhagem, com os parâmetros utilizados. Similarmente ao observado para as linhagens de *C. gattii*, a linhagem *C. neoformans* var. *grubii* H99 apresentou enriquecimento para todos os termos. Já para as linhagens *C. neoformans* B-3501A e JEC21, a única função sub-representada para as proteínas associadas à virulência foi a de transporte transmembrana, sendo todas as outras funções enriquecidas para proteínas associadas à virulência, em relação ao proteoma total.

5.6 Mapeamento das vias metabólicas e associação aos fatores de virulência

Após uma busca por proteínas no BD Swiss-prot, utilizando os termos enumerados na seção 4.7 da metodologia, foram obtidas 58 proteínas. Por meio dos resultados obtidos pelo alinhamento entre os proteomas preditos e o banco de dados *db_virulencia_crypto_swissprot* foi possível obter a lista de proteínas associadas à virulência para cada organismo estudado. Após os processos de anotação e mapeamento dos proteomas no Blast2GO, os resultados referentes a proteínas relacionadas à virulência foram selecionados e utilizados para o mapeamento de vias metabólicas do KEGG. O conjunto de vias metabólicas obtido para cada linhagem foi então comparado entre os organismos estudados, a fim de se obter a relação de vias comuns a todas as linhagens.

Observou-se que, das 29 vias metabólicas mapeadas, 16 foram comuns a todos os organismos estudados (Tabela 3). A relação de vias mapeadas para cada organismo estudado pode ser visualizada nas tabelas 4 a 9, presentes no Apêndice B deste documento.

Tabela 3: Relação das vias metabólicas comuns a todos os organismos estudados.

Via metabólica	Mapa KEGG
Aminobenzoate degradation	map00627
Amino sugar and nucleotide sugar metabolism	map00520
Biosynthesis of antibiotics	map01130
Carbon fixation in photosynthetic organisms	map00710
Drug metabolism - other enzymes	map00983
Fructose and mannose metabolism	map00051
Glycolysis / Gluconeogenesis	map00010
Monobactam biosynthesis	map00261
mTOR signaling pathway	map04150
Phenylalanine, tyrosine and tryptophan biosynthesis	map00400
Purine metabolism	map00230
Selenocompound metabolism	map00450
Sulfur metabolism	map00920
T cell receptor signaling pathway	map04660
Th1 and Th2 cell differentiation	map04658
Thiamine metabolism	map00730

Dentre as vias comuns a todos os organismos estudados, pode-se citar a via de Biossíntese de fenilalanina, tirosina e triptofano. Como exemplo, temos o mapeamento de enzimas para a linhagem R265, onde é possível observar três enzimas mapeadas na via (**Figura 10**). Um estudo sobre o papel da biossíntese de triptofano na sobrevivência de *C. neoformans* demonstrou que dois genes desta via metabólica são essenciais para *C. neoformans*, sendo a via de biossíntese de triptofano um possível alvo para o desenvolvimento de terapias contra esse patógeno (FERNANDES, *et al.*, 2015).

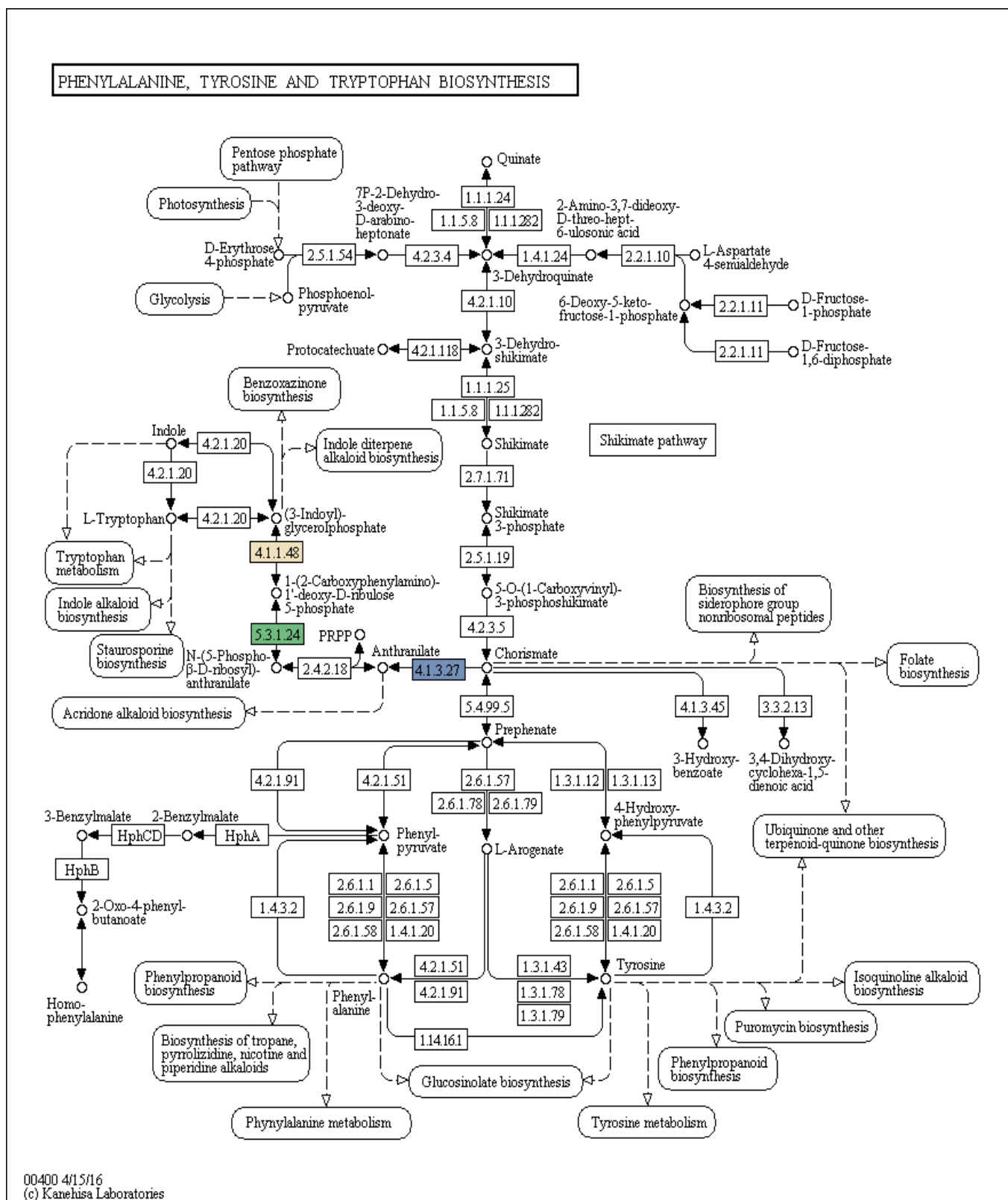


Figura 10: Via de biossíntese de fenilalanina, tirosina e triptofano. Nessa via, foram identificadas três proteínas associadas à virulência, sendo uma proteína para cada enzima do KEGG. As enzimas estão destacadas em cores diferentes no mapa, sendo elas EC: 4.1.3.27 (em azul), EC: 5.3.1.24 (em verde) e EC: 4.1.48 (em bege).

Dentre as 16 vias comuns a todos os organismos, 6 não foram descritas na literatura, sendo elas as vias de degradação de aminobezoato, biossíntese de antibióticos, fixação de carbono em organismos fotossintéticos, metabolismo de compostos de selênio e a via de sinalização mTOR.

Em *Aspergillus flavus*, foi demonstrado que o aumento de estresse oxidativo exerce um impacto na expressão de genes relacionados à degradação de aminobenzoato (FOUNTAIN, *et al.*, 2016). Estudos relacionando o estresse oxidativo e a regulação da expressão de genes de degradação de aminobenzoatos poderiam ser realizados em linhagens de *Cryptococcus* spp.

Diversos micro-organismos são capazes de sintetizar metabólitos secundários que são utilizados como antimicrobianos. Dentre esses metabólitos pode-se citar a cefalosporina C, um composto beta-lactâmico produzido pelos fungos *P.chrysogenum* e *Ac. chrysogenum* que possui ação tanto contra bactérias gram-positivas quanto gram-negativas (KÜCK *et al.*, 2014). Em decorrência da inexistência de estudos que relacionem espécies de *Cryptococcus* à biossíntese de antibióticos, tornando a via de biossíntese dessas substâncias um possível alvo para estudos envolvendo esses micro-organismos.

A fixação de carbono é um importante processo na produção de energia por organismos autótrofos, como plantas e cianobactérias, no qual o dióxido de carbono serve como fonte para a produção carbono orgânico para o próprio fotossintetizante, bem como para organismos heterótrofos. (HUGLER e SIEVERT, 2010). Além da produção de carbono orgânico por organismos autótrofos, estudos têm demonstrado a importância de enzimas dessa via para organismos heterótrofos. Um estudo realizado com *Sclerotinia sclerotiorum*, um fungo fitopatogênico, demonstrou que enzimas associadas à via de fixação de carbono em plantas estão presentes no genoma desse fungo e que algumas dessas enzimas estão associadas à virulência desse micro-organismo (LYU *et al.*, 2015). Dessa forma, estudos envolvendo a via de assimilação de carbono pode ser um possível alvo de estudos envolvendo espécies patogênicas de *Cryptococcus*.

Além da determinação de vias comuns a todos os organismos, também foram feitas comparações entre as vias mapeadas para cada organismo a fim de se obter vias metabólicas únicas para cada linhagem. Após a comparação, foi observado que somente a linhagem H99 apresentou uma via não mapeada para os outros organismos, sendo esta a via de fosforilação oxidativa (Figura 11).

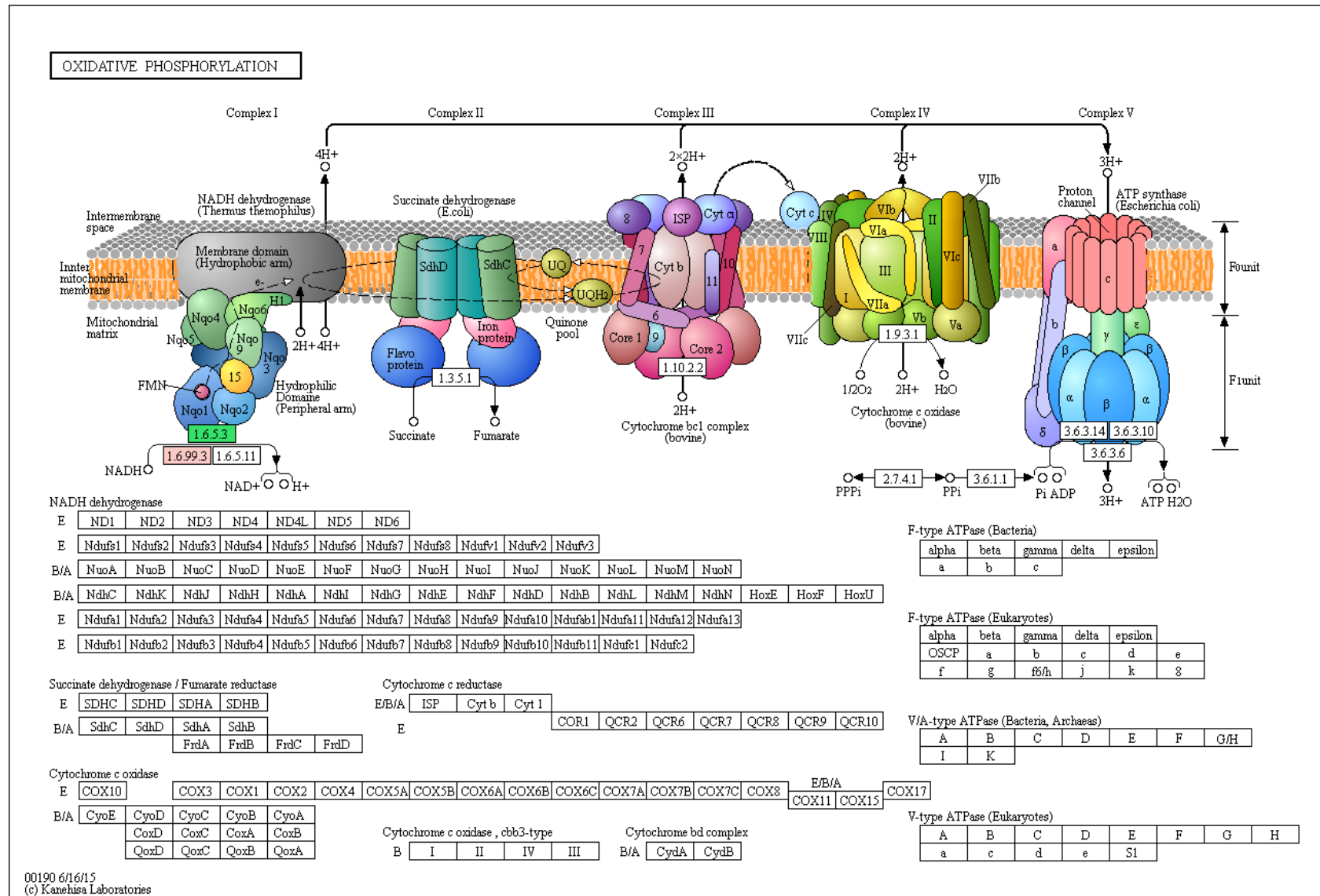


Figura 11: Via de fosforilação oxidativa. Nessa via, foram identificadas três proteínas associadas à virulência, sendo todas relacionadas a cada enzima do KEGG. As enzimas estão destacadas em cores diferentes no mapa, sendo elas EC: 1.6.5.3 (em verde) e EC: 1.6.99.3 (em rosa).

Apesar da via de fosforilação oxidativa estar presente em todos os organismos estudados, apenas a linhagem *C. neoformans* H99 apresentou proteínas associadas à virulência mapeadas na mesma. Em um estudo sobre desordem proteica em *Cryptococcus* spp. foi observado que, das seis linhagens virulentas utilizadas para as análises, somente três possuíam proteínas com pelo menos uma Região Intrinsecamente Desordenada (IDR) mapeadas na via de fosforilação oxidativa, incluindo a linhagem *C. neoformans* H99 (TAVARES, 2016). Em outro estudo, o papel da via glicolítica na virulência e persistência de *C. neoformans* no hospedeiro foi realizado. Para isso, um mutante para a síntese da enzima piruvato quinase (*pk1Δ*) foi utilizado e se observou uma redução na produção de ATP por essas leveduras. Foi constatado que a utilização de glicose é essencial para a virulência e persistência desse patógeno no SNC, sendo sugerido que inibidores de glicólise podem ser úteis no tratamento da criptococose (PRICE *et al.*, 2011).

5.7 Análise de agrupamento

Para a análise de agrupamento dos proteomas preditos, o programa OrhtoMCL foi utilizado, conforme descrito na seção 4.10 deste documento. Grupos de proteínas Ortólogas foram identificadas por meio de agrupamentos que continham uma proteína para cada organismo estudado. Os seis proteomas preditos de *Cryptococcus* spp. foram utilizados para a análise de agrupamento, totalizando 40.967 proteínas. Foram identificados 4484 grupos, representando 26.904 sequências de proteínas (**Figura 12**). A linhagem CBS 7750 de *C. gattii* apresentou o maior número de proteínas linhagem-específica (1193), enquanto a linhagem *Cryptococcus gattii* R265 apresentou o menor número de proteínas linhagem-específica (77).

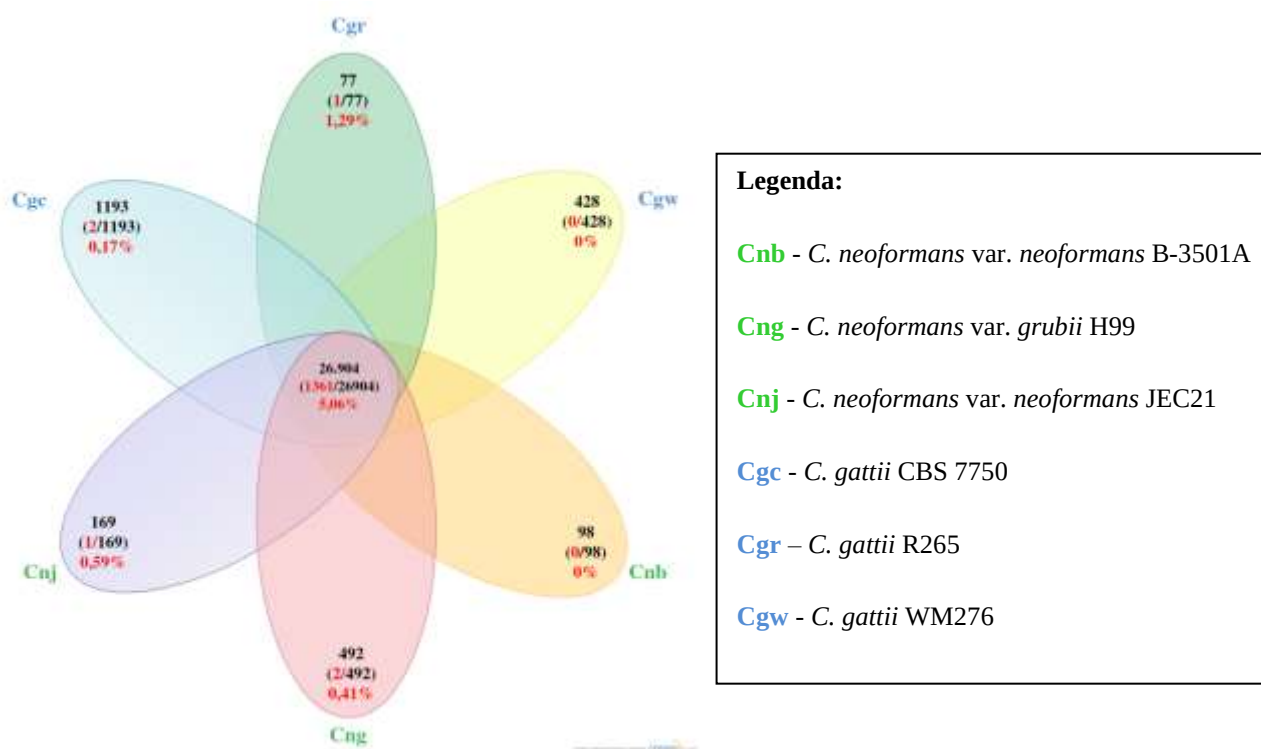


Figura 12: Diagrama de Venn referente à análise de agrupamento. No diagrama estão descritos o número de genes relacionados ao *core-genome*, bem como os genes linhagem-específicos. Em preto estão destacados o número de proteínas para cada situação e em vermelho estão representados os genes relacionados à virulência de *Cryptococcus* e a correspondente porcentagem dessas proteínas.

Em um estudo realizado com cinco das seis linhagens utilizadas no presente estudo (*C. gattii* R265, *C. gattii* WM276, *C. neoformans* var. *neoformans* B-3501A, *C. neoformans* var. *grubii* H99 e *C. neoformans* var. *neoformans* JEC21), 5.171 grupos foram identificados (D'SOUZA *et al.*, 2011). D'SOUZA e colaboradores também fizeram a predição de genes em *C. gattii* R265 e identificaram 108 genes anotados unicamente para esse genoma, sendo a maioria deles genes de proteínas hipotéticas.

Outro estudo de comparação de genomas, em que linhagens dos quatro genótipos de *C. gattii* foram sequenciadas, demonstrou que cada linhagem estudada possuía genes únicos classificados como genes putativos envolvidos na virulência como, por exemplo, genes que se ligam a ferro e cobre, bem como genes que estão envolvidos na resposta a estresses causados às leveduras (FARRER, *et al.*, 2015).

Um grande número de genes únicos enriquecidos para proteínas de choque térmico e domínios ao ferro é observado no genótipo VGII, sendo esses genes possivelmente responsáveis pela hipervirulência desse genótipo (FARRER *et al.*, 2015).

Das 492 proteínas linhagem-específicas identificadas na linhagem H99, duas estão associadas à virulência. Essas duas proteínas, cngnu.077910 e cngnu.077960, foram mapeadas na via de fosforilação oxidativa, sendo esta via mapeada somente para a linhagem H99. Outro fato relevante é que essas proteínas foram anotadas no genoma dessa linhagem como NADH desidrogenases, sendo essa mesma função atribuída às enzimas na via de fosforilação oxidativa (**Figura 11**).

Uma análise mais detalhada da região do genoma em que essas proteínas estão presentes, por meio do uso dos resultados provenientes da análise de similaridade de sequências e visualização no ACT revelou que não houve similaridade de sequências de nucleotídeos entre a extensão total dos genes cngnu.077910 e cngnu.077960 e os genomas JEC21 e B-3501A de *C. neoformans* (**Figura 13**). É possível observar nos gráficos provenientes do ACT que apenas um pequeno trecho desses genes de H99 apresenta alinhamento com os genomas das outras linhagens de *C. neoformans*. Esses resultados sugerem que os genomas dessas duas linhagens realmente não possuem essas proteínas.

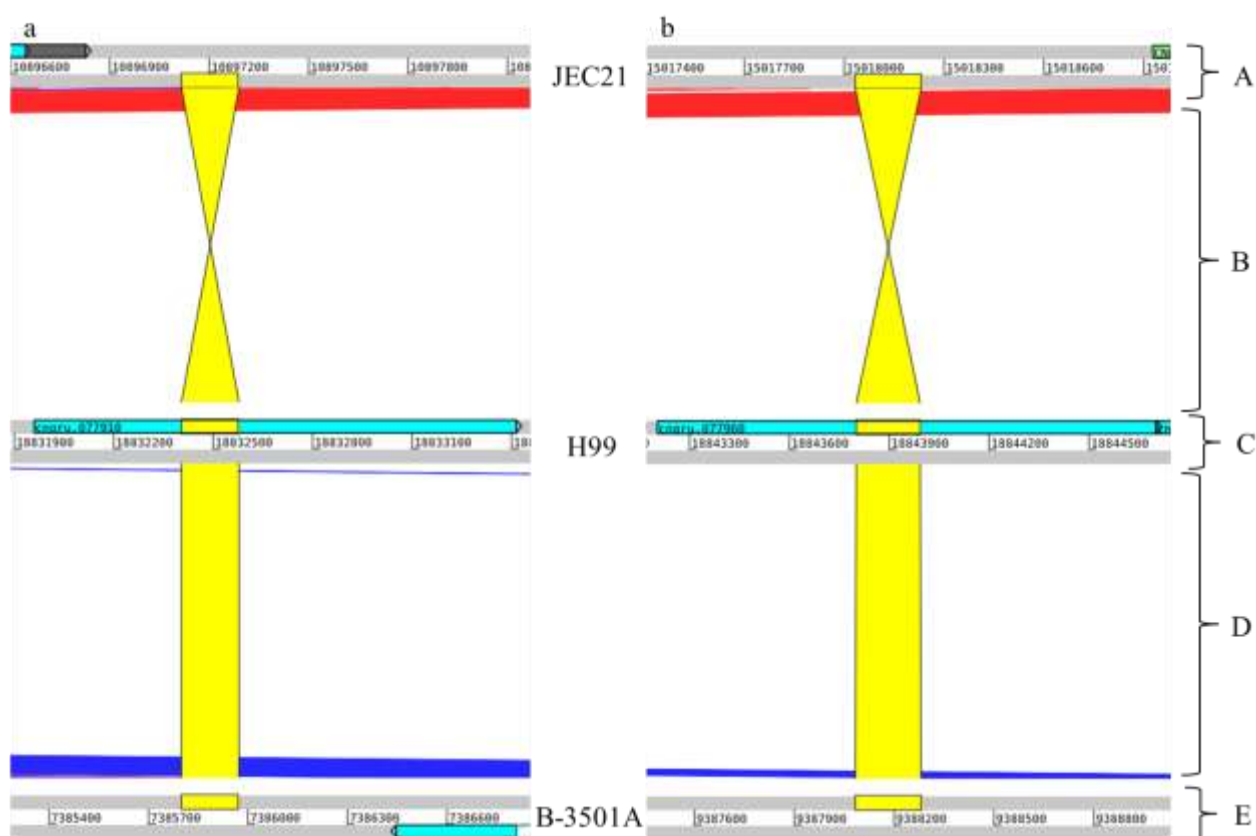


Figura 13: Representação gráfica da comparação de similaridade entre trechos dos genomas de *C. neoformans* JEC21, *C. neoformans* H99 e *C. neoformans* B-3501A (do topo para baixo), onde: A, C e E correspondem às sequências genômicas dos organismos em questão. As regiões similares entre os organismos estão interligadas pelos segmentos de reta indicados em B e D. Em C, as regiões que compreendem os genes *cngru.77910* e *cngru.77960* estão destacadas em azul claro. Os trechos dos genes que apresentam similaridade com os genomas das linhagens JEC21 e B-3501A estão destacados em amarelo.

Similarmente, uma análise mais detalhada entre essa mesma região onde os genes *cngru.077910* e *cngru.077960* estão localizados e o genoma da linhagem R265 de *C. gattii* revelou algumas que existem grandes trechos de similaridade entre genes e o genoma de R265. Quando comparado ao genoma da linhagem R265, o gene *cngru.077910* apresentou similaridade de nucleotídeos em quase toda a extensão do gene, com porcentagem de identidade de 95% (**Figura 14**). Entretanto, a porção inicial do gene *cngru.077910*, compreendendo as trincas de nucleotídeos que codificam os 7 primeiros aminoácidos do gene, incluindo a metionina, não foi alinhada com o genoma da linhagem R265. Isso poderia justificar o fato desse genoma de *C. gattii* não ter esse gene anotado ou da anotação automática ter falhado na predição desse gene.

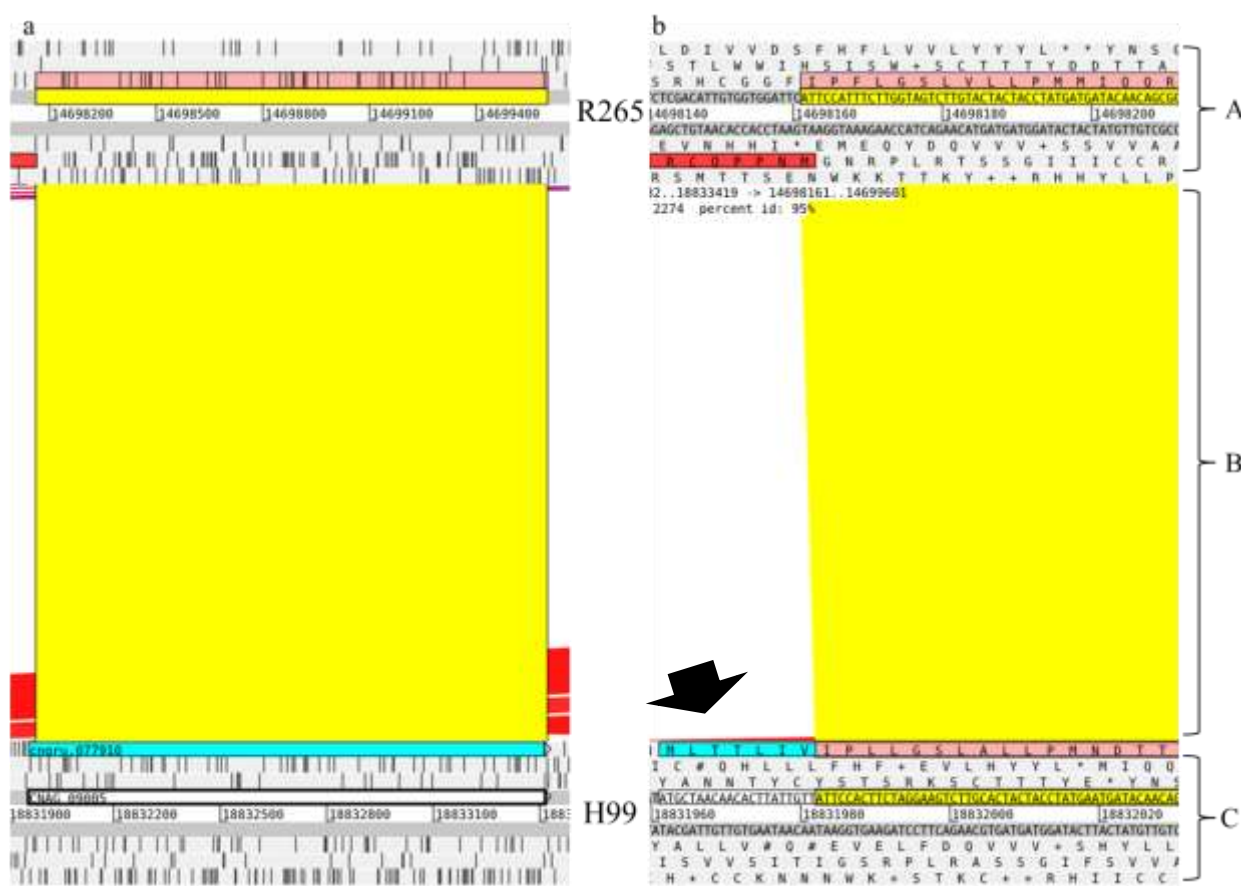


Figura 14: Representação gráfica da comparação de similaridade entre trechos dos genomas de *C. gattii* R265 e *C. neoformans* H99 (do topo para baixo), onde A e C correspondem às sequências genômicas dos organismos em questão, bem como as possíveis fases de leitura. As regiões similares entre os organismos estão interligadas pelos segmentos de reta indicados em B. Em C, a região que compreende o gene *cngru.77910* está destacada em azul. Os trechos dos genes que apresentam similaridade com o genoma da linhagem R265 estão destacados em amarelo. (a) visão global da comparação entre o gene *cngru.077910* e o genoma de R265; (b) detalhe da comparação entre os dois genomas, mostrando que o genoma da linhagem R265 não possui a mesma sequência de nucleotídeos que são traduzidos para os 7 primeiros aminoácidos da linhagem H99, indicado pela seta.

Já a análise do gráfico de alinhamento do ACT entre a região em que o gene *cngru.077960* está localizado e o genoma da linhagem R265 de *C. gattii* revelou que há similaridade de sequência de nucleotídeos entre toda a extensão do gene *cngru.077960* e a região do genoma de R265, com porcentagem de identidade de 95% (**Figura 15**). Uma análise mais detalhada do gráfico revelou que uma das fases de leitura do trecho do genoma da linhagem R265 possui a mesma sequência de aminoácidos que são codificados pelo gene *cngru.077960*. Entretanto, esse gene não está anotado no genoma da linhagem R265.

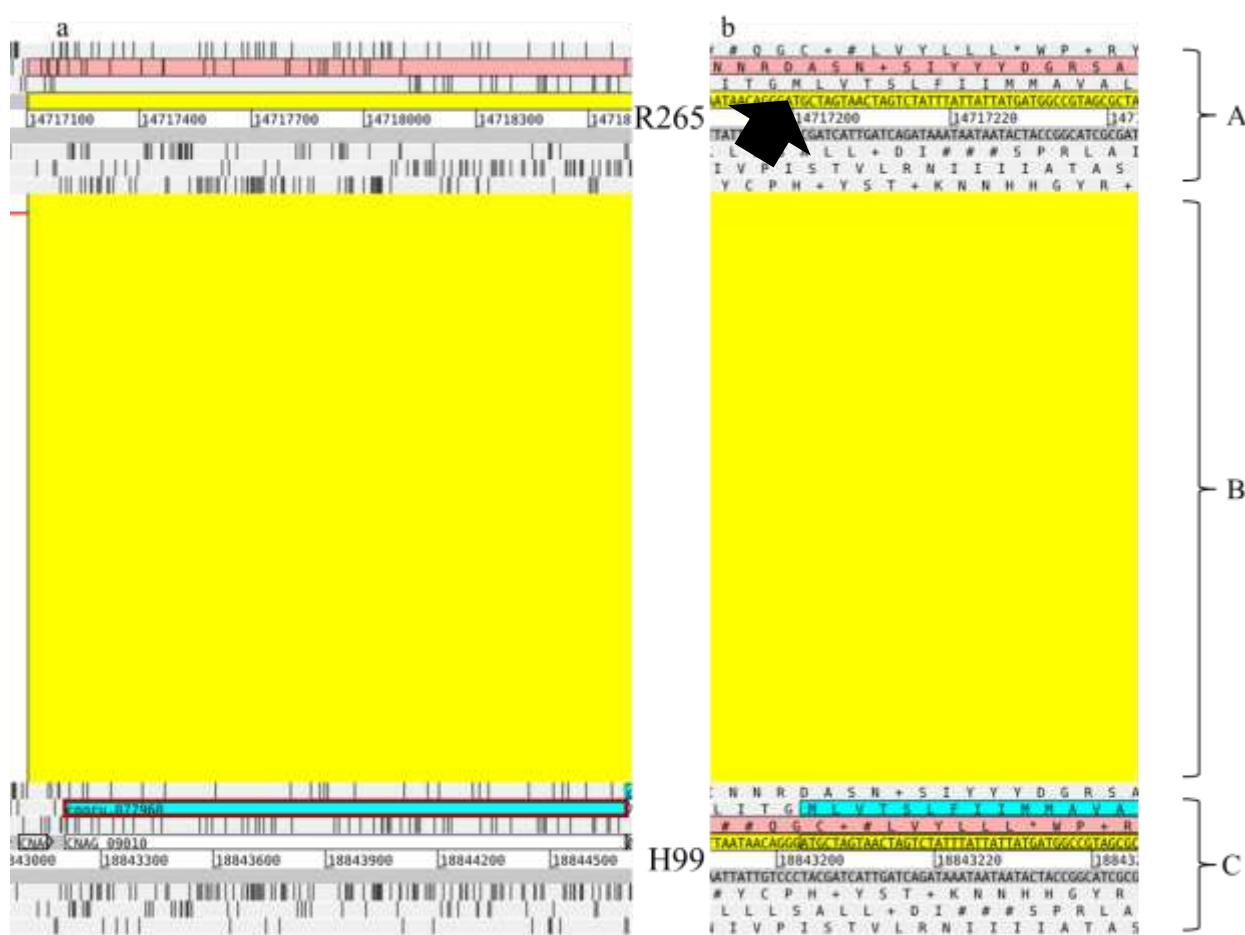


Figura 15: Gráfico de comparação de similaridade entre *C. gattii* R265 e *C. neoformans* H99 (do topo para baixo), onde A e C correspondem às sequências genômicas dos organismos em questão, bem como as possíveis fases de leitura. As regiões similares entre os organismos estão interligadas pelos segmentos de reta indicados em B. Em C, a região que compreende o gene *cngru.77910* estão assinaladas em azul. Os trechos dos genes que apresentam similaridade com o genoma da linhagem R265 estão destacados em amarelo. (a) visão global da comparação entre o gene *cngru.077960* e o genoma de R265, evidenciando que existe similaridade de sequência de nucleotídeos entre toda a extensão desse gene e um trecho do genoma de R265; (b) detalhe da comparação entre os dois genomas, mostrado a trinca de nucleotídeos que codificam o aminoácido Metionina (M), indicado pela seta.

Os genes *cngru.077910* e *cngru.077960*, anotados no genoma da linhagem H99 de *C. neoformans* foram identificados como sendo linhagem-específicos, por meio dos resultados provenientes das análises de agrupamento feitas pelo OrthoMCL. Uma análise mais detalhada dos resultados dos gráficos gerados pelo ACT após a análise de similaridade de sequências pelo *big_blast.pl* entre as regiões do genoma onde esses genes estão localizados, em comparação com os genomas das linhagens B-3501A e JEC21, foi realizada com a finalidade de verificar se esses dois genomas realmente não possuíam alinhamentos entre os genes *cngru.077910* e *cngru.077960*. A presença de apenas pequenos trechos de alinhamento entre os genes das linhagens e os genomas

das outras linhagens de *C. neoformans* são um indicativo de que esses genomas não possuem esses genes anotados na linhagem H99 (**Figura 13**). Entretanto, não se pode afirmar que as linhagens B-3501A e JEC21 não possuem esses genes, tendo apenas como base o sequenciamento e anotação gênica. Para isso, ensaios laboratoriais precisam ser realizados para a validação dos resultados observados *in silico* erros durante a montagem de genomas podem ocorrer em decorrência de diferentes motivos como, por exemplo, a exclusão de fragmentos consideradas erroneamente como erros ou repetições e a inserção de pedaços em regiões ou orientações erradas (BAKER, 2012).

A análise dos gráficos de similaridade de sequências entre os genes cngru.077910 e cngru.077960 e o genoma da linhagem R265 de *C. gattii* demonstrou que, para o gene cngru.077960, existe uma região no genoma da linhagem R265 que apresenta 95% de identidade de sequência de nucleotídeos. Outra observação feita é que o genoma da linhagem R265 possui o códon de iniciação ATG, que codifica o aminoácido metionina. Esse aminoácido está presente em uma das possíveis fases de leitura, mas foram observadas substituições de aminoácidos, que poderiam justificar o fato dessa proteína não ter sido anotada no genoma. No entanto, essa alta taxa de identidade entre o gene cngru.077960 e a região do genoma da linhagem R265 poderia ser um indicativo da presença dessa proteína no genoma de *C. gattii*. Novamente, validações experimentais seriam necessárias para confirmar a presença ou não no genoma da linhagem em questão. Alternativamente podemos hipotetizar que um erro de anotação ocorreu e que na verdade houve uma deleção na região N-terminal do gene. Estas questões só podem ser confirmadas por meio de experimentação e/ou avaliação da qualidade de bases das leituras que geraram a sequência consenso presente no genoma.

Esses dados demonstram a importância da re-anotação genômica, da validação experimental de dados oriundos de análises realizadas *in silico* e da análise criteriosa dos resultados gerados.

6 CONCLUSÃO

Por meio dos resultados obtidos neste estudo, levantamos as seguintes hipóteses e conclusões:

1. As linhagens estudadas apresentaram alta taxa de similaridade de nucleotídeos.
2. Para todos os organismos utilizados no estudo, Proteínas associadas à virulência foram mapeadas em pelo menos 16 vias metabólicas descritas no KEGG, demonstrando a importância destas proteínas ligadas à virulência no metabolismo de *Cryptococcus* spp.
3. Somente *C. neoformans* var. *grubii*, linhagem H99 apresentou enzimas associadas à virulência mapeadas em uma via que não foi mapeada nas outras linhagens, sendo que essas enzimas atuam como desidrogenases e redutases na via de fosforilação oxidativa dessa linhagem.
4. Mais de 99% das proteínas associadas à virulência identificadas nos proteomas dos organismos estudados são oriundas de genes ortólogos, ou seja, genes que evoluíram de ancestral comum pelo processo de especiação.
5. Em vista dos resultados preliminares obtidos, avaliamos que a re-anotação e curadoria manual de todos os genes preditos trariam grande benefício e ganho nas análises.

7 PERSPECTIVAS

1. Realizar um teste de inibição para avaliar a possível biossíntese de antibióticos pelas linhagens H99 de *C. neoformans* e R265 de *C. gattii*.
2. Amplificar os genes cngru.077910 e cngru.077960 da linhagem H99 de *C. neoformans* e verificar a presença desses genes na linhagem R265 de *C. gattii*, bem como em outras linhagens de *C. neoformans*.
3. Realizar a re-anotação dos genomas das linhagens H99 de *C. neoformans* e R265 de *C. gattii*.

8 REFERÊNCIAS BIBLIOGRÁFICAS

BAHN, Y.-S. & Jung, K.-W. Stress Signaling Pathways for the Pathogenicity of *Cryptococcus*. **Eukaryotic Cell** 12, 1564–1577 (2013).

BAHN, Y.-S., Kojima, K., Cox, G. M. & Heitman, J. Specialization of the HOG Pathway and Its Impact on Differentiation and Virulence of *Cryptococcus neoformans*. **Mol Biol Cell** 16, 2285–2300 (2005).

BAKER, M. De novo genome assembly: what every biologist should know. **Nature Meth** [Internet]. 2012; 9(4):333–7.

BALTAZAR, L. de M. & Ribeiro, M. A. First isolation of *Cryptococcus gattii* from the environment in the State of Espírito Santo. *Revista da Sociedade Brasileira de Medicina Tropical* 41, 449–453 (2008).

BASSO, L. R., Gast, C. E., Bruzual, I. & Wong, B. Identification and properties of plasma membrane azole efflux pumps from the pathogenic fungi *Cryptococcus gattii* and *Cryptococcus neoformans*. **J Antimicrob Chemother** 70, 1396–1407 (2015).

BIELSKA, E., May, R. C. & Munro, C. What makes *Cryptococcus gattii* a pathogen? **FEMS Yeast Res** 16, (2016).

BOEKHOUT, T. et al. Molecular typing of *Cryptococcus neoformans*: taxonomic and epidemiological aspects. **Int. J. Syst. Bacteriol.** 47, 432–442 (1997).

BOTTS, M. R. & Hull, C. M. Dueling in the lung: how *Cryptococcus* spores race the host for survival. **Current Opinion in Microbiology** 13, 437–442 (2010).

BYRNES III, E. J., Bartlett, K. H., Perfect, J. R. & Heitman, J. *Cryptococcus gattii*: an emerging fungal pathogen infecting humans and animals. **Microbes and Infection** 13, 895–907 (2011).

BYRNES EJ, Li W, Ren P, Lewit Y, Voelz K, Fraser JA, et al. A Diverse Population of *Cryptococcus gattii* Molecular Type VGIII in Southern Californian HIV/AIDS Patients. **PLoS Pathog** [Internet]. 2011 Sep [cited 2017 Feb 3];7(9). Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3164645/>

CAMPBELL, L. T. et al. *Cryptococcus* Strains with Different Pathogenic Potentials Have Diverse Protein Secretomes. **Eukaryot Cell** 14, 554–563 (2015).

CARVER, T., Harris, S. R., Berriman, M., Parkhill, J. & McQuillan, J. A. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. **Bioinformatics** 28, 464–469 (2012).

CASADEVALL, A., Rosas, A. L. & Nosanchuk, J. D. Melanin and virulence in *Cryptococcus neoformans*. **Current Opinion in Microbiology** 3, 354–358 (2000).

CATANHO, M. P., Miranda, A. B. de & Degraive, W. M. S. Comparando genomas: bancos de dados e ferramentas computacionais para a análise comparativa de genomas procarióticos. **Revista Eletrônica de Comunicação, Informação & Inovação em Saúde** 1, (2007).

CDC: <http://www.cdc.gov/fungal/diseases/cryptococcosis-gattii/index.html>. Acesso em 09/06/2015.

CHATURVEDI, V. & Chaturvedi, S. *Cryptococcus gattii*: a resurgent fungal pathogen. **Trends in Microbiology** 19, 564–571 (2011).

CHEN, C.-H. et al. Epidemiological characterization and prognostic factors in patients with confirmed cerebral cryptococcosis in central Taiwan. **Journal of Venomous Animals and Toxins including Tropical Diseases** 21, 12 (2015).

CHEN, M. & Hofestädt, R. A medical bioinformatics approach for metabolic disorders: Biomedical data prediction, modeling, and systematic analysis. **Journal of Biomedical Informatics** 39, 147–159 (2006).

CHEN, S. C.-A., Meyer, W. & Sorrell, T. C. *Cryptococcus gattii* Infections. **Clin Microbiol Rev** 27, 980–1024 (2014).

CHONG, H. S. et al. Time-Course Proteome Analysis Reveals the Dynamic Response of *Cryptococcus gattii* Cells to Fluconazole. **PLOS ONE** 7, e42835 (2012).

CHOU, C.-H., Chang, W.-C., Chiu, C.-M., Huang, C.-C. & Huang, H.-D. FMM: a web server for metabolic pathway reconstruction and comparative analysis. **Nucleic Acids Res** 37, W129–W134 (2009).

COGLIATI, M. Global Molecular Epidemiology of *Cryptococcus neoformans* and *Cryptococcus gattii*: An Atlas of the Molecular Types. **Scientifica** 2013, (2013).

CONESA, A. & Götz, S. Blast2GO: A Comprehensive Suite for Functional Analysis in Plant Genomics. **Int J Plant Genomics** 2008, (2008).

D'SOUZA, C. A. et al. Genome Variation in *Cryptococcus gattii*, an Emerging Pathogen of Immunocompetent Hosts. **mBio** 2, (2011).

DALE, J. M., Popescu, L. & Karp, P. D. Machine learning methods for metabolic pathway prediction. **BMC Bioinformatics** 11, 15 (2010).

DE FILIPPO, C., Ramazzotti, M., Fontana, P. & Cavalieri, D. Bioinformatic approaches for functional annotation and pathway inference in metagenomics data. **Brief Bioinform** 13, 696–710 (2012).

DIXIT A, Carroll SF, Qureshi ST. *Cryptococcus gattii*: An Emerging Cause of Fungal Disease in North America. **Interdisciplinary Perspectives on Infectious Diseases** [Internet]. 2009 May 25; 2009:e840452.

ELLIS, D. H. & Pfeiffer, T. J. Natural habitat of *Cryptococcus neoformans* var. *gattii*. **J Clin Microbiol** 28, 1642–1644 (1990).

ENRIGHT, A. J., Van Dongen, S. & Ouzounis, C. A. An efficient algorithm for large-scale detection of protein families. **Nucleic Acids Res** 30, 1575–1584 (2002).

ESPINDOLA, F. S. et al. Recursos de bioinformática aplicados às ciências ômicas como genômica, transcriptômica, proteômica, interatômica e metabolômica = Bioinformatic resources applied on the omic sciences as genomic, transcriptomic, proteomic, interatomic and metabolomic. **Bioscience Journal** 26, (2010).

ESPINEL-INGROFF, A. & Kidd, S. E. Current trends in the prevalence of *Cryptococcus gattii* in the United States and Canada. **Infect Drug Resist** 8, 89–97 (2015).

FANG, W., Fa, Z. & Liao, W. Epidemiology of *Cryptococcus* and cryptococcosis in China. **Fungal Genetics and Biology** 78, 7–15 (2015).

FARRER, R. A. *et al.* Genome Evolution and Innovation across the Four Major Lineages of *Cryptococcus gattii*. **mBio** 6, e00868–15 (2015).

FARIA RO de, Nascente P da S, Meinerz ARM, Cleff MB, Antunes T de A, Silveira É da S, et al. Occurrence of *Cryptococcus neoformans* in pigeon excrement in the city of Pelotas, State of Rio Grande do Sul. **Revista da Sociedade Brasileira de Medicina Tropical** [Internet]. 2010 Apr; 43(2):198–200.

FARRER, R. A. *et al.* Microevolutionary traits and comparative population genomics of the emerging pathogenic fungus *Cryptococcus gattii*. **Philos Trans R Soc Lond B Biol Sci** 371, (2016).

FEDER, V. *et al.* *Cryptococcus gattii* urease as a virulence factor and the relevance of enzymatic activity in cryptococcosis pathogenesis. **FEBS J** 282, 1406–1418 (2015).

FERNANDES, J. D. S., Martho, K., Tofik, V., Vallim, M. A. & Pascon, R. C. The Role of Amino Acid Permeases and Tryptophan Biosynthesis in *Cryptococcus neoformans* Survival. **PLoS One** 10, (2015).

FERNANDES, K. E., Dwyer, C., Campbell, L. T. & Carter, D. A. Species in the *Cryptococcus gattii* Complex Differ in Capsule and Cell Size following Growth under Capsule-Inducing Conditions. **mSphere** 1, (2016).

FERREIRA, G. F. *et al.* The role of oxidative and nitrosative bursts caused by azoles and amphotericin B against the fungal pathogen *Cryptococcus gattii*. **J. Antimicrob. Chemother.** 68, 1801–1811 (2013).

FOUNTAIN JC, Bajaj P, Nayak SN, Yang L, Pandey MK, Kumar V, et al. Responses of *Aspergillus flavus* to Oxidative Stress Are Related to Fungal Development Regulator, Antioxidant Enzyme, and Secondary Metabolite Biosynthetic Gene Expression. **Front Microbiol** [Internet]. 2016 Dec 21; 7.

FRANZOT, S. P., Salkin, I. F. & Casadevall, A. *Cryptococcus neoformans* var. *grubii*: Separate Varietal Status for *Cryptococcus neoformans* Serotype A Isolates. **J Clin Microbiol** 37, 838–840 (1999).

GABALDÓN, T. & Huynen, M. A. Prediction of protein function and pathways in the genome era. **CMLS, Cell. Mol. Life Sci.** 61, 930–944 (2004).

GARCIA-HERMOSO, D., Janbon, G. & Dromer, F. Epidemiological Evidence for Dormant *Cryptococcus neoformans* Infection. **J Clin Microbiol** 37, 3204–3209 (1999).

GAST, C. E., Basso, L. R., Bruzual, I. & Wong, B. Azole Resistance in *Cryptococcus gattii* from the Pacific Northwest: Investigation of the Role of ERG11. **Antimicrob Agents Chemother** 57, 5478–5485 (2013).

GULLO, F. P. *et al.* Cryptococcosis: epidemiology, fungal resistance, and new alternatives for treatment. **Eur J Clin Microbiol Infect Dis** 32, 1377–1391 (2013).

HAGEN, F. *et al.* Recognition of seven species in the *Cryptococcus gattii*/*Cryptococcus neoformans* species complex. **Fungal Genetics and Biology** 78, 16–48 (2015).

HUGLER, M.; Sievert, S.M. Beyond the Calvin Cycle: Autotrophic Carbon Fixation in the Ocean. **Annual Review of Marine Science** [Internet]. 2011;3(1):261–89.

IDNURM, A. A Tetrad Analysis of the Basidiomycete Fungus *Cryptococcus neoformans*. **Genetics** 185, 153–163 (2010).

IDNURM, A. *et al.* Deciphering the Model Pathogenic Fungus *Cryptococcus Neoformans*. **Nat Rev Micro** 3, 753–764 (2005).

KANEHISA, M., Goto, S., Sato, Y., Furumichi, M. & Tanabe, M. KEGG for integration and interpretation of large-scale molecular data sets. **Nucl. Acids Res.** 40, D109–D114 (2012).

KAOCHAROEN, S. *et al.* Molecular Epidemiology Reveals Genetic Diversity amongst Isolates of the *Cryptococcus neoformans*/*C. gattii* Species Complex in Thailand. **PLOS Neglected Tropical Diseases** 7, e2297 (2013).

KIDD, S. E. *et al.* *Cryptococcus gattii* Dispersal Mechanisms, British Columbia, Canada. **Emerg Infect Dis** **13**, 51–57 (2007).

KRONSTAD, J. W. *et al.* Expanding fungal pathogenesis: *Cryptococcus* breaks out of the opportunistic box. **Nat Rev Micro** **9**, 193–203 (2011).

KÜCK U, Bloemendal S, Teichert I. Putting Fungi to Work: Harvesting a Cornucopia of Drugs, Toxins, and Antibiotics. **PLoS Pathog** [Internet]. 2014 Mar 13 ; 10(3).

KWON-CHUNG, K. J. *et al.* The Case for Adopting the ‘Species Complex’ Nomenclature for the Etiologic Agents of Cryptococcosis. **mSphere** **2**, (2017).

KWON-CHUNG, K. J., Boekhout, T., Fell, J. W. & Diaz, M. (1557) Proposal to Conserve the Name *Cryptococcus gattii* against *C. hondurians* and *C. bacillisporus* (Basidiomycota, Hymenomycetes, Tremellomycetidae). **Taxon** **51**, 804–806 (2002).

LAGESEN, K. *et al.* RNAmmer: consistent and rapid annotation of ribosomal RNA genes. **Nucleic Acids Res** **35**, 3100–3108 (2007).

LARSEN, R. A., Bauer, M., Thomas, A. M. & Graybill, J. R. Amphotericin B and Fluconazole, a Potent Combination Therapy for Cryptococcal Meningitis. **Antimicrob Agents Chemother** **48**, 985–991 (2004).

LEE, K.-T. *et al.* Systematic functional analysis of kinases in the fungal pathogen *Cryptococcus neoformans*. *Nat Commun* **7**, (2016).

LEE, K.-T., Kwon, H., Lee, D. & Bahn, Y.-S. A Nudix Hydrolase Protein, Ysa1, Regulates Oxidative Stress Response and Antifungal Drug Susceptibility in *Cryptococcus neoformans*. **Mycobiology** **42**, 52–58 (2014).

LI, L., Stoeckert, C. J. & Roos, D. S. OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes. **Genome Res** **13**, 2178–2189 (2003).

LI, S. S. & Mody, C. H. *Cryptococcus*. *Proc Am Thorac Soc* **7**, 186–196 (2010).

Lin, X. & Heitman, J. The Biology of the *Cryptococcus neoformans* Species Complex. **Annual Review of Microbiology** **60**, 69–105 (2006).

LIN. X., HEITMAN, J., The Complex Biology of the *Cryptococcus neoformans* Species. **Annual Review of Microbiology**, 60:69–105. 2006.

LIU, J. *et al.* Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes. **Cell** **66**, 807–815 (1991).

LIZARAZO, J., Escandón, P., Agudelo, C. I. & Castañeda, E. Cryptococcosis in Colombian children and literature review. **Mem Inst Oswaldo Cruz** **109**, 797–804 (2014).

LOWE, T. M. & Chan, P. P. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. **Nucleic Acids Res** 44, W54–W57 (2016).

LOWE, T. M. & Eddy, S. R. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. **Nucleic Acids Res** 25, 955–964 (1997).

LYU X, Shen C, Xie J, Fu Y, Jiang D, Hu Z, et al. A “footprint” of plant carbon fixation cycle functions during the development of a heterotrophic fungus. **Sci Rep** [Internet]. 2015 Aug 11; 5

MA, H. & MAY, R. C. in (ed. Microbiology, B.-A. in A.) 67, 131–190 (**Academic Press**, 2009).

MACDOUGALL, L. & Fyfe, M. Emergence of *Cryptococcus gattii* in a Novel Environment Provides Clues to Its Incubation Period. **J Clin Microbiol** 44, 1851–1852 (2006).

MAY, R. C., Stone, N. R. H., Wiesner, D. L., Bicanic, T. & Nielsen, K. *Cryptococcus*: from environmental saprophyte to global pathogen. **Nat Rev Micro** 14, 106–117 (2016).

MEDEMA, M. H., van Raaphorst, R., Takano, E. & Breitling, R. Computational tools for the synthetic design of biochemical pathways. **Nat Rev Micro** 10, 191–202 (2012).

MOREIRA, M. E. L. Controversies about the management of invasive fungal infections in very low birth weight infants. **Jornal de Pediatria** 81, S52–S58 (2005).

NGAMSKULRUNGROJ, P., Chang, Y., Sionov, E. & Kwon-Chung, K. J. The Primary Target Organ of *Cryptococcus gattii* Is Different from That of *Cryptococcus neoformans* in a Murine Model. **mBio** 3, (2012).

NISHIKAWA, M. M. et al. Serotyping of 467 *Cryptococcus neoformans* Isolates from Clinical and Environmental Sources in Brazil: Analysis of Host and Regional Patterns. **J. Clin. Microbiol.** 41, 73–77 (2003).

NOSANCHUK, J. D., Rudolph, J., Rosas, A. L. & Casadevall, A. Evidence That *Cryptococcus neoformans* Is Melanized in Pigeon Excreta: Implications for Pathogenesis. **Infect Immun** 67, 5477–5479 (1999).

PARK, B. J. et al. Estimation of the current global burden of cryptococcal meningitis among persons living with HIV/AIDS. **AIDS** 23, 525–530 (2009).

PARK, S. H., Kim, M., Joo, S. I. & Hwang, S. M. Molecular Epidemiology of Clinical *Cryptococcus neoformans* Isolates in Seoul, Korea. **Mycobiology** 42, 73–78 (2014).

PRADO, M., Silva, M. B. da, Laurenti, R., Travassos, L. R. & Taborda, C. P. Mortality due to systemic mycoses as a primary cause of death or in association with AIDS in Brazil: a review from 1996 to 2006. **Memórias do Instituto Oswaldo Cruz** 104, 513–521 (2009).

- PRATES, R. A. et al. Effect of Virulence Factors on the Photodynamic Inactivation of *Cryptococcus neoformans*. **PLOS ONE** 8, e54387 (2013).
- PRICE, M. S. et al. *Cryptococcus neoformans* Requires a Functional Glycolytic Pathway for Disease but Not Persistence in the Host. **mBio** 2, (2011).
- RANG, H. P.; DALE, M. M.; Ritter, J. M.; Flower, J. R. **Farmacologia**. Tradução: Santos, R. R. et al Rio de Janeiro: Elsevier, 2007.
- RODERO, L. et al. G484S Amino Acid Substitution in Lanosterol 14- α Demethylase (ERG11) Is Related to Fluconazole Resistance in a Recurrent *Cryptococcus neoformans* Clinical Isolate. *Antimicrob. Agents Chemother.* 47, 3653–3656 (2003).
- RUTHERFORD, K. et al. Artemis: sequence visualization and annotation. **Bioinformatics** 16, 944–945 (2000).
- SAAG, M. S. et al. Practice Guidelines for the Management of Cryptococcal Disease. **Clin Infect Dis** 30, 710–718 (2000).
- SANTI, L. et al. Proteomic Profile of *Cryptococcus neoformans* Biofilm Reveals Changes in Metabolic Processes. **J. Proteome Res.** 13, 1545–1559 (2014).
- SANTOS, J. R. A. et al. Dynamic interaction between fluconazole and amphotericin B against *Cryptococcus gattii*. *Antimicrob. Agents Chemother.* AAC.06098–11 (2012). doi:10.1128/AAC.06098-11
- SANTOS, J. R. A. et al. Fluconazole Alters the Polysaccharide Capsule of *Cryptococcus gattii* and Leads to Distinct Behaviors in Murine Cryptococcosis. **PLOS ONE** 9, e112669 (2014).
- SCHATTNER, P., Brooks, A. N. & Lowe, T. M. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. **Nucleic Acids Res** 33, W686–W689 (2005).
- SHAPIRO, R. S., Robbins, N. & Cowen, L. E. Regulatory Circuitry Governing Fungal Development, Drug Resistance, and Disease. **Microbiol Mol Biol Rev** 75, 213–267 (2011).
- SHI, M. et al. Real-time imaging of trapping and urease-dependent transmigration of *Cryptococcus neoformans* in mouse brain. **J Clin Invest** 120, 1683–1693 (2010).
- SIDRIM, J. J. C.; Rocha, M. F. G. **Micologia médica à luz de autores contemporâneos**. Rio de Janeiro: Guanabara Koogan, 2004.
- SILVA, D. C. et al. Susceptibility to antifungal agents and genotypes of Brazilian clinical and environmental *Cryptococcus gattii* strains. **Diagnostic Microbiology and Infectious Disease** 72, 332–339 (2012).

SIONOV, E., Lee, H., Chang, Y. C. & Kwon-Chung, K. J. *Cryptococcus neoformans* Overcomes Stress of Azole Drugs by Formation of Disomy in Specific Multiple Chromosomes. **PLoS Pathog** 6, (2010).

SLOAN, D. & Parris, V. Cryptococcal meningitis: epidemiology and therapeutic options. **Clinical Epidemiology** (2014). Available at: <https://www.dovepress.com/cryptococcal-meningitis-epidemiology-and-therapeutic-options-peer-reviewed-article-CLEP>.

SMITH, R. M. et al. Treatment and Outcomes among Patients with *Cryptococcus gattii* Infections in the United States Pacific Northwest. **PLoS One** 9, (2014).

TAVARES, GS. Estudo computacional de desordem proteica nos genomas de *Cryptococcus* spp. [Rio de Janeiro – Rio de Janeiro]: Instituto Oswaldo Cruz – IOC, 2016.

THE GENE ONTOLOGY CONSORTIUM. Gene Ontology: going forward. **Nucleic Acids Res** 43, D1049–D1056 (2015).

THE UNIPROT CONSORTIUM. UniProt: a hub for protein information. **Nucleic Acids Res** 43, D204–D212 (2015).

TUCKER, S. C. & Casadevall, A. Replication of *Cryptococcus neoformans* in macrophages is accompanied by phagosomal permeabilization and accumulation of vesicles containing polysaccharide in the cytoplasm. **Proc Natl Acad Sci U S A** 99, 3165–3170 (2002).

VAN DUIN, D., Casadevall, A. & Nosanchuk, J. D. Melanization of *Cryptococcus neoformans* and *Histoplasma capsulatum* Reduces Their Susceptibilities to Amphotericin B and Caspofungin. **Antimicrob Agents Chemother** 46, 3394–3400 (2002).

VANBREUSEGHEM, R. & Takashio, M. An atypical strain of *Cryptococcus neoformans* (San Felice) Vuillemin 1894. II. *Cryptococcus neoformans* var. *gattii* var. nov. *Ann Soc Belges Med Trop Parasitol Mycol* 50, 695–702 (1970).

9 APÊNDICE A – GRÁFICOS RELACIONADOS À ANÁLISE DE ENRIQUECIMENTO ENTRE AS ESPÉCIES/ LINHAGENS (PROCESSO BIOLÓGICO)

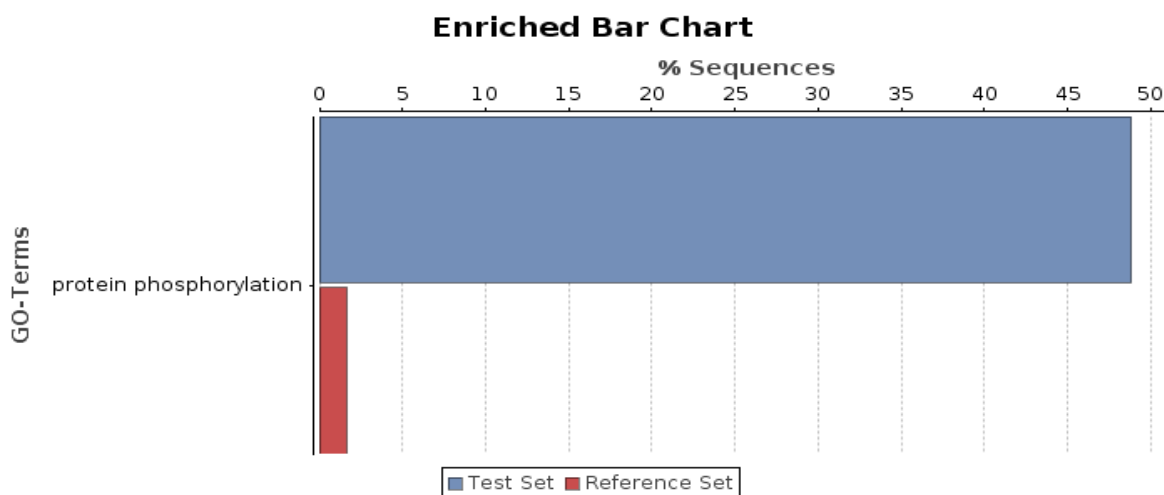


Figura 16: Gráfico referente à análise de enriquecimento de proteínas relacionadas à virulência em *C. gattii*, linhagem R265. O conteúdo total de proteínas está representado em vermelho. Já o conteúdo de proteínas relacionadas à virulência está representado em azul.

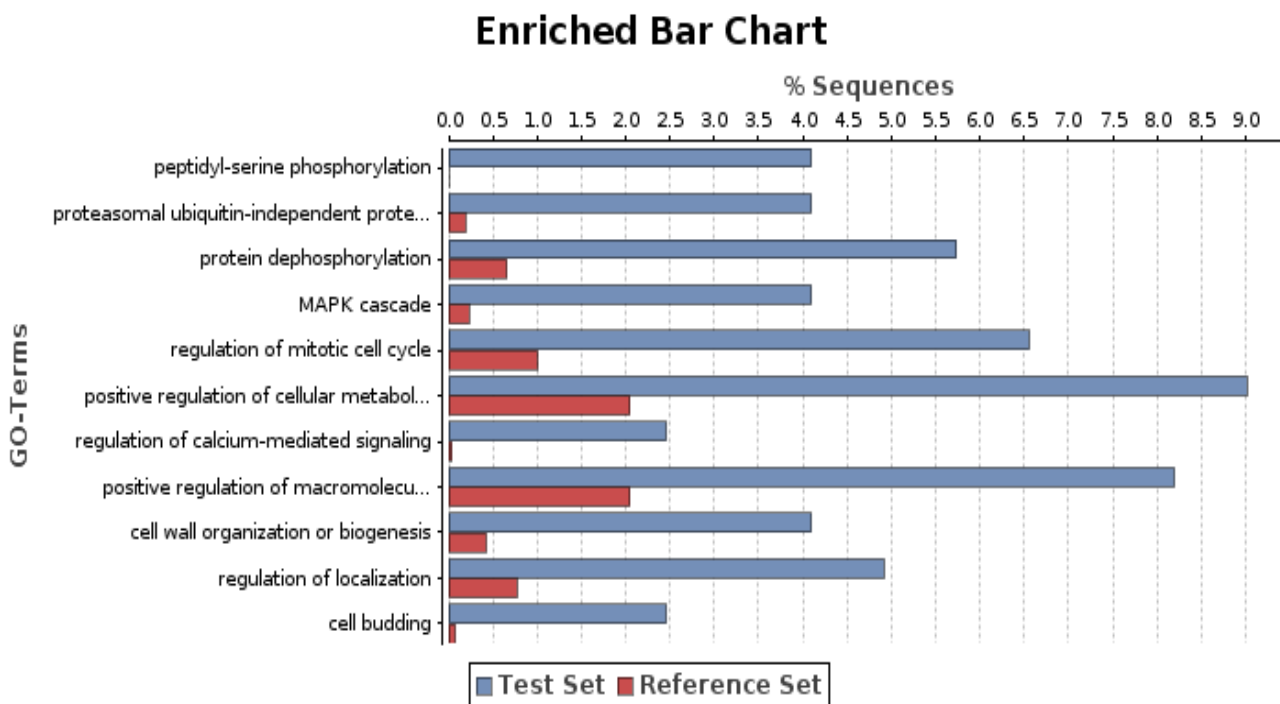


Figura 17: Gráfico referente à análise de enriquecimento de proteínas relacionadas à virulência em *C. gattii*, linhagem CBS7750. O conteúdo total de proteínas está representado em vermelho. Já o conteúdo de proteínas relacionadas à virulência está representado em azul.

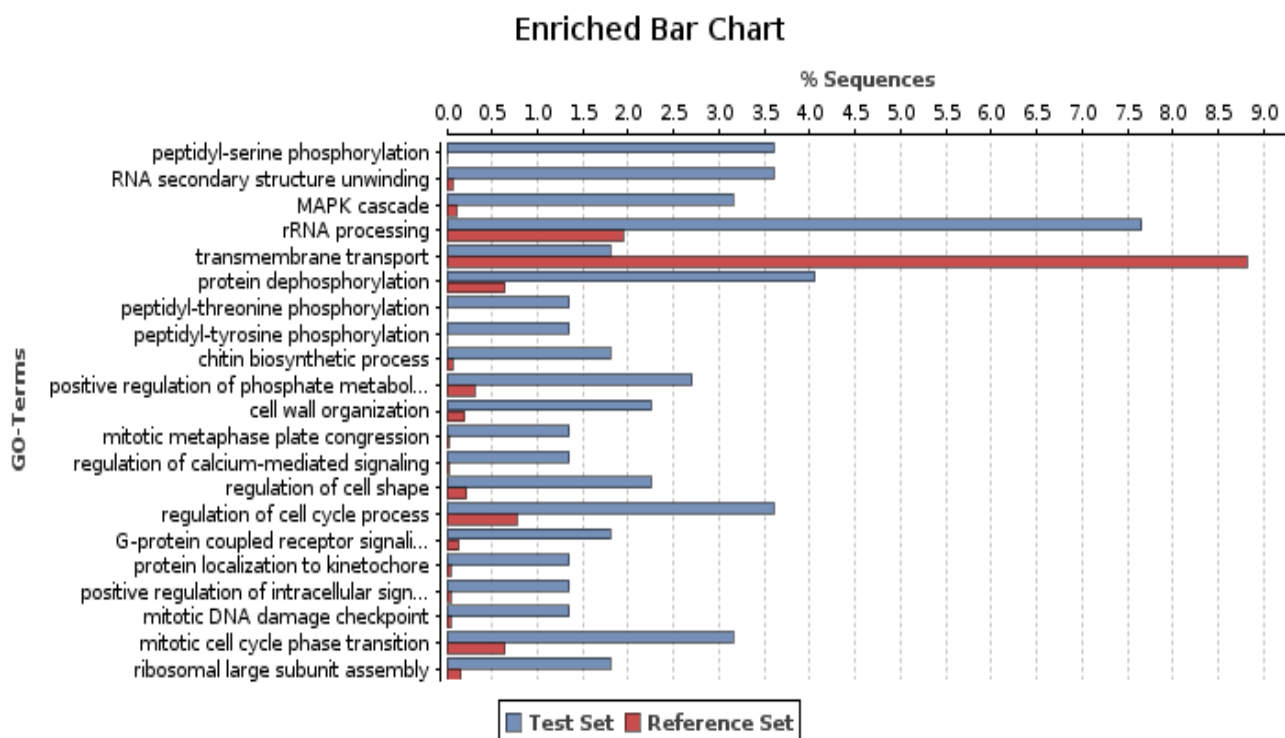


Figura 18: Gráfico referente à análise de enriquecimento de proteínas relacionadas à virulência em *C. neoformans*, linhagem B-3501A. O conteúdo total de proteínas está representado em vermelho. Já o conteúdo de proteínas relacionadas à virulência está representado em azul.

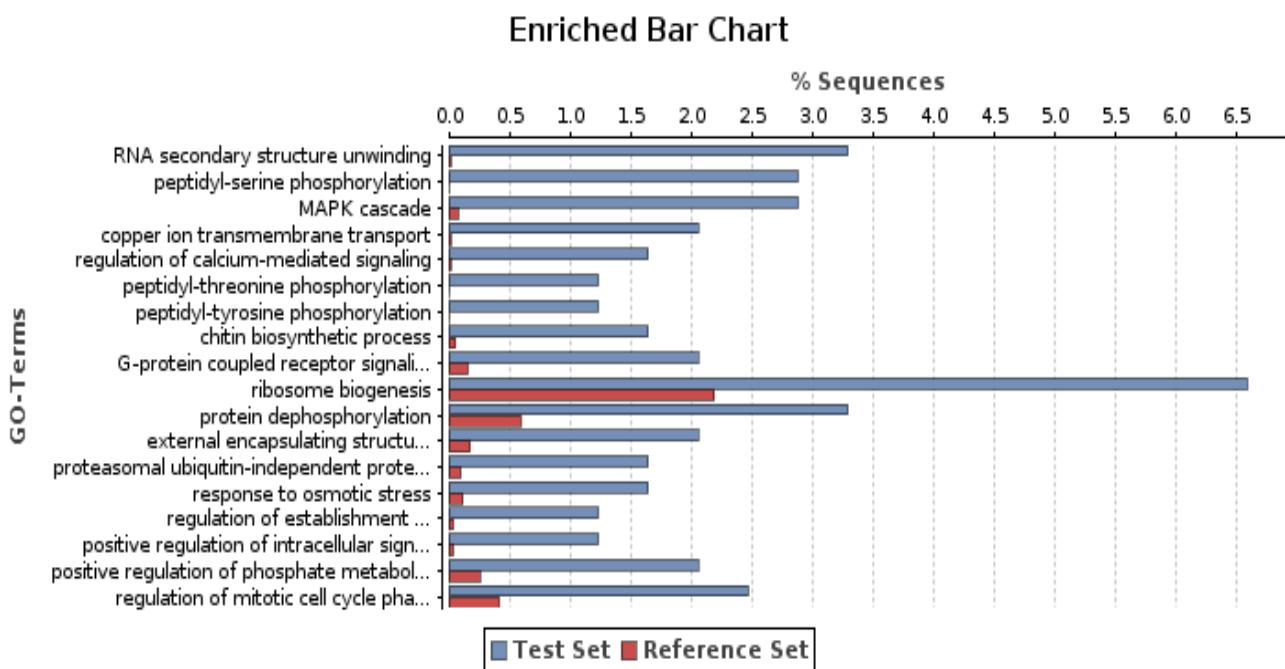


Figura 19: Gráfico referente à análise de enriquecimento de proteínas relacionadas à virulência em *C. neoformans grubii*, linhagem H99. O conteúdo total de proteínas está representado em vermelho. Já o conteúdo de proteínas relacionadas à virulência está representado em azul.

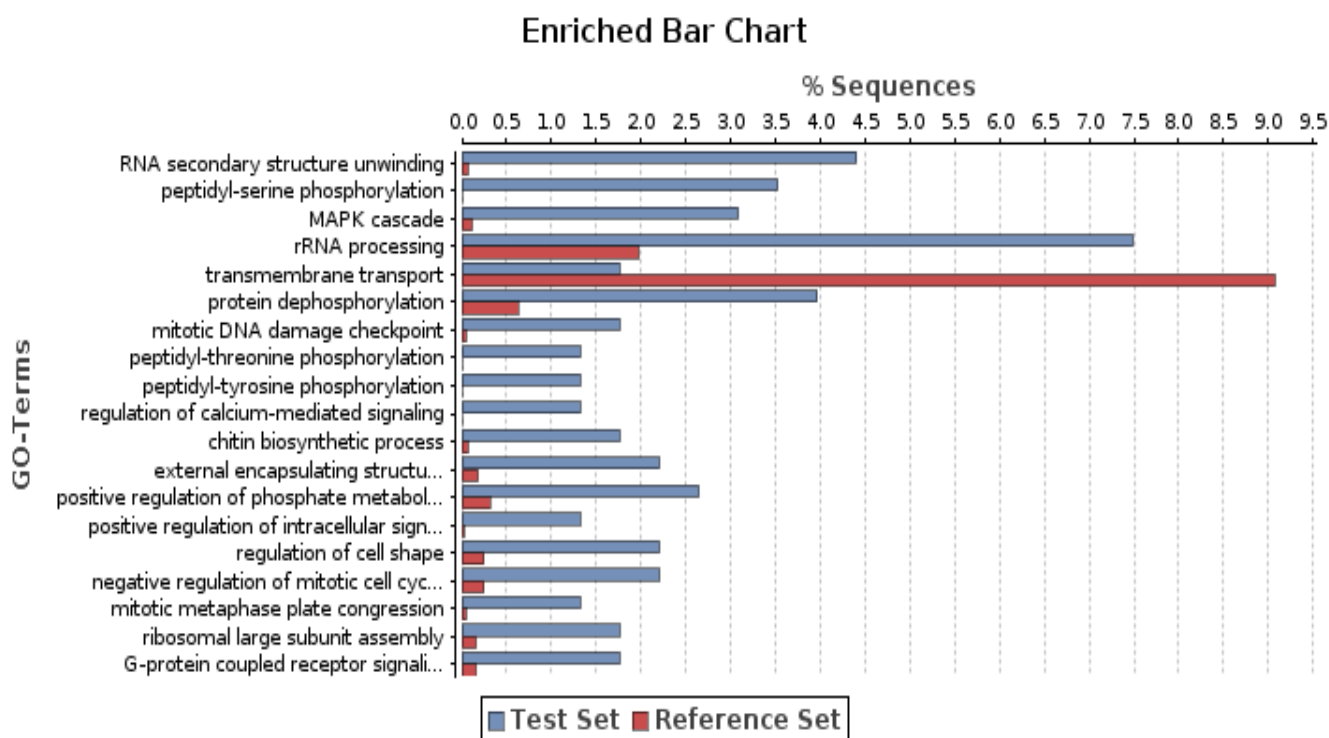


Figura 20: Gráfico referente à análise de enriquecimento de proteínas relacionadas à virulência em *C. neoformans*, linhagem JEC21. O conteúdo total de proteínas está representado em vermelho. Já o conteúdo de proteínas relacionadas à virulência está representado em azul.

10 APÊNDICE B - TABELAS REFERENTES AO MAPEAMENTO DE PROTEÍNAS ASSOCIADAS À VIRULÊNCIA DE *CRYPTOCOCCUS* SPP. NAS VIAS METABÓLICAS

Tabela 4: Vias metabólicas mapeadas em *C. gattii* CBS 7750.

Via metabólica	Mapa KEGG	Número de enzimas na via	Enzima	EC number	Número de sequências enzimáticas	Proteínas
mTOR signaling pathway	map04150	1	protein kinase	ec:2.7.11.24	3	Cgcbs.049220, Cgcbs.032670, Cgcbs.030160
Phenylpropanoid biosynthesis	map00940	1	lactoperoxidase	ec:1.11.1.7	2	Cgcbs.008010, Cgcbs.062410
Glycolysis / Gluconeogenesis	map00010	2	dehydrogenase (NAD(P)+) (phosphorylating)	ec:1.2.1.59	2	Cgcbs.021750, Cgcbs.030060
Glycolysis / Gluconeogenesis	map00010	2	dehydrogenase (phosphorylating)	ec:1.2.1.12	2	Cgcbs.021750, Cgcbs.030060
C5-Branched dibasic acid metabolism	map00660	1	synthase	ec:2.2.1.6	1	Cgcbs.003380
Aminobenzoate degradation	map00627	1	nitrophenyl phosphatase	ec:3.1.3.41	7	Cgcbs.066550, Cgcbs.070120, Cgcbs.063480, Cgcbs.001220, Cgcbs.049550, Cgcbs.062780, Cgcbs.017420
Butanoate metabolism	map00650	1	synthase	ec:2.2.1.6	1	Cgcbs.003380
Phenylalanine, tyrosine and tryptophan biosynthesis	map00400	3	synthase,	ec:4.1.3.27	1	Cgcbs.030280
Phenylalanine, tyrosine and tryptophan biosynthesis	map00400	3	isomerase,	ec:5.3.1.24	1	Cgcbs.030280
Phenylalanine, tyrosine and tryptophan biosynthesis	map00400	3	synthase	ec:4.1.1.48	1	Cgcbs.030280
Sulfur metabolism	map00920	2	adenylyltransferase	ec:2.7.7.4	1	Cgcbs.033450
Sulfur metabolism	map00920	2	kinase	ec:2.7.1.25	1	Cgcbs.063090
Pantothenate and CoA biosynthesis	map00770	1	synthase	ec:2.2.1.6	1	Cgcbs.003380

Via metabólica	Mapa KEGG	Número de enzimas na via	Enzima	EC number	Número de sequências enzimáticas	Proteínas
Biosynthesis of antibiotics	map01130	7	synthase	ec:2.2.1.6	1	Cgcbs.003380
Biosynthesis of antibiotics	map01130	7	adenylyltransferase	ec:2.7.7.4	1	Cgcbs.033450
Biosynthesis of antibiotics	map01130	7	synthase	ec:4.1.3.27	1	Cgcbs.030280
Biosynthesis of antibiotics	map01130	7	isomerase	ec:5.3.1.24	1	Cgcbs.030280
Biosynthesis of antibiotics	map01130	7	synthase	ec:4.1.1.48	1	Cgcbs.030280
Biosynthesis of antibiotics	map01130	7	dehydrogenase (NAD(P)+) (phosphorylating)	ec:1.2.1.59	2	Cgcbs.021750, Cgcbs.030060
Biosynthesis of antibiotics	map01130	7	dehydrogenase (phosphorylating)	ec:1.2.1.12	2	Cgcbs.021750, Cgcbs.030060
Thiamine metabolism	map00730	1	phosphatase	ec:3.6.1.15	17	Cgcbs.046520, Cgcbs.014430, Cgcbs.019720, Cgcbs.013240, Cgcbs.050530, Cgcbs.033650, Cgcbs.034200, Cgcbs.013800, Cgcbs.049080, Cgcbs.000090, Cgcbs.004250, Cgcbs.004380, Cgcbs.057870, Cgcbs.040930, Cgcbs.061780, Cgcbs.001150, Cgcbs.048950
Selenocompound metabolism	map00450	1	adenylyltransferase	ec:2.7.7.4	1	Cgcbs.033450
Monobactam biosynthesis	map00261	1	adenylyltransferase	ec:2.7.7.4	1	Cgcbs.033450
Purine metabolism	map00230	6	synthase (glutamine-hydrolysing),	ec:6.3.5.2	1	Cgcbs.036200
Purine metabolism	map00230	6	adenylpyrophosphatase	ec:3.6.1.3	7	Cgcbs.013800, Cgcbs.046520, Cgcbs.000090, Cgcbs.040930, Cgcbs.019720, Cgcbs.033650, Cgcbs.034200
Purine metabolism	map00230	6	phosphatase	ec:3.6.1.15	17	Cgcbs.046520, Cgcbs.014430, Cgcbs.019720, Cgcbs.013240, Cgcbs.050530, Cgcbs.033650, Cgcbs.034200, Cgcbs.013800, Cgcbs.049080, Cgcbs.000090, Cgcbs.004250, Cgcbs.004380, Cgcbs.057870, Cgcbs.040930, Cgcbs.061780, Cgcbs.001150, Cgcbs.048950

Via metabólica	Mapa KEGG	Número de enzimas na via	Enzima	EC number	Número de sequências enzimáticas	Proteínas
Purine metabolism	map00230	6	adenylyltransferase	ec:2.7.7.4	1	Cgcbs.033450
Purine metabolism	map00230	6	carboxylase	ec:4.1.1.21	1	Cgcbs.044330
Purine metabolism	map00230	6	kinase	ec:2.7.1.25	1	Cgcbs.063090
Valine, leucine and isoleucine biosynthesis	map00290	1	synthase	ec:2.2.1.6	1	Cgcbs.003380
Glycerolipid metabolism	map00561	1	lipase	ec:3.1.1.3	1	Cgcbs.044550
Starch and sucrose metabolism	map00500	1	synthase	ec:2.4.1.34	1	Cgcbs.058520
Carbon fixation in photosynthetic organisms	map00710	2	dehydrogenase (NAD(P)+) (phosphorylating)	ec:1.2.1.59	2	Cgcbs.021750, Cgcbs.030060
Carbon fixation in photosynthetic organisms	map00710	2	dehydrogenase (phosphorylating)	ec:1.2.1.12	2	Cgcbs.021750, Cgcbs.030060
Drug metabolism - other enzymes	map00983	2	synthase (glutamine-hydrolysing)	ec:6.3.5.2	1	Cgcbs.036200
Drug metabolism - other enzymes	map00983	2	ali-esterase	ec:3.1.1.1	1	Cgcbs.044550
Th1 and Th2 cell differentiation	map04658	1	phosphatase	ec:3.1.3.16	7	Cgcbs.066550, Cgcbs.070120, Cgcbs.063480, Cgcbs.001220, Cgcbs.049550, Cgcbs.062780, Cgcbs.017420
T cell receptor signaling pathway	map04660	1	phosphatase	ec:3.1.3.16	7	Cgcbs.066550, Cgcbs.070120, Cgcbs.063480, Cgcbs.001220, Cgcbs.049550, Cgcbs.062780, Cgcbs.017420
Amino sugar and nucleotide sugar metabolism	map00520	2	synthase	ec:2.4.1.16	1	Cgcbs.058710
Amino sugar and nucleotide sugar metabolism	map00520	2	isomerase	ec:5.3.1.8	1	Cgcbs.032360
Fructose and mannose metabolism	map00051	1	isomerase	ec:5.3.1.8	1	Cgcbs.032360

Tabela 5: Vias metabólicas mapeadas em *C. gattii* R265.

Via metabólica	Mapa KEGG	Número de enzimas na via	Enzima	EC number	Número de sequências enzimáticas	Proteínas
Selenocompound metabolism	map00450	1	adenylyltransferase	ec:2.7.7.4	1	Cgr26.018470
mTOR signaling pathway	map04150	1	protein kinase	ec:2.7.11.24	2	Cgr26.017780, Cgr26.015610
Monobactam biosynthesis	map00261	1	adenylyltransferase	ec:2.7.7.4	1	Cgr26.018470
Glycolysis / Gluconeogenesis	map00010	2	dehydrogenase (NAD(P)+) (phosphorylating)	ec:1.2.1.59	2	Cgr26.009620, Cgr26.015520
Glycolysis / Gluconeogenesis	map00010	2	dehydrogenase (phosphorylating)	ec:1.2.1.12	2	Cgr26.009620, Cgr26.015520
Purine metabolism	map00230	4	synthase (glutamine-hydrolysing)	ec:6.3.5.2	1	Cgr26.024290
Purine metabolism	map00230	4	adenylyltransferase	ec:2.7.7.4	1	Cgr26.018470
Purine metabolism	map00230	4	adenylpyrophosphatase	ec:3.6.1.3	3	Cgr26.019080, Cgr26.007870, Cgr26.018630
Purine metabolism	map00230	4	phosphatase	ec:3.6.1.15	6	Cgr26.019890, Cgr26.001700, Cgr26.019080, Cgr26.016060, Cgr26.007870, Cgr26.018630
Aminobenzoate degradation	map00627	1	nitrophenyl phosphatase	ec:3.1.3.41	1	Cgr26.011540
Drug metabolism - other enzymes	map00983	1	synthase (glutamine-hydrolysing)	ec:6.3.5.2	1	Cgr26.024290
Carbon fixation in photosynthetic organisms	map00710	2	dehydrogenase (NAD(P)+) (phosphorylating)	ec:1.2.1.59	2	Cgr26.009620, Cgr26.015520
Carbon fixation in photosynthetic organisms	map00710	2	dehydrogenase (phosphorylating)	ec:1.2.1.12	2	Cgr26.009620, Cgr26.015520
Th1 and Th2 cell differentiation	map04658	1	phosphatase	ec:3.1.3.16	1	Cgr26.011540
T cell receptor signaling pathway	map04660	1	phosphatase	ec:3.1.3.16	1	Cgr26.011540
Phenylalanine, tyrosine and tryptophan biosynthesis	map00400	3	synthase	ec:4.1.3.27	1	Cgr26.015720
Phenylalanine, tyrosine and tryptophan biosynthesis	map00400	3	isomerase	ec:5.3.1.24	1	Cgr26.015720

Via metabólica	Mapa KEGG	Número de enzimas na via	Enzima	EC number	Número de sequências enzimáticas	Proteínas
Phenylalanine, tyrosine and tryptophan biosynthesis	map00400	3	synthase	ec:4.1.1.48	1	Cgr26.015720
Sulfur metabolism	map00920	1	adenylyltransferase	ec:2.7.7.4	1	Cgr26.018470
Amino sugar and nucleotide sugar metabolism	map00520	1	isomerase	ec:5.3.1.8	1	Cgr26.017480
Fructose and mannose metabolism	map00051	1	isomerase	ec:5.3.1.8	1	Cgr26.017480
Biosynthesis of antibiotics	map01130	6	adenylyltransferase	ec:2.7.7.4	1	Cgr26.018470
Biosynthesis of antibiotics	map01130	6	synthase	ec:4.1.3.27	1	Cgr26.015720
Biosynthesis of antibiotics	map01130	6	isomerase	ec:5.3.1.24	1	Cgr26.015720
Biosynthesis of antibiotics	map01130	6	synthase	ec:4.1.1.48	1	Cgr26.015720
Biosynthesis of antibiotics	map01130	6	dehydrogenase (NAD(P)+) (phosphorylating)	ec:1.2.1.59	2	Cgr26.009620, Cgr26.015520
Biosynthesis of antibiotics	map01130	6	dehydrogenase (phosphorylating)	ec:1.2.1.12	2	Cgr26.009620, Cgr26.015520
Thiamine metabolism	map00730	1	phosphatase	ec:3.6.1.15	6	Cgr26.019890, Cgr26.001700, Cgr26.019080, Cgr26.016060, Cgr26.007870, Cgr26.018630

Tabela 6: Vias metabólicas mapeadas em *C. gattii* WM276

Via metabólica	Mapa KEGG	Número de enzimas na via	Enzima	EC number	Número de sequências enzimáticas	Proteínas
Atrazine degradation	map00791	1	urease	ec:3.5.1.5	1	Cgwm2_059730
mTOR signaling pathway	map04150	2	protein kinase	ec:2.7.11.11	1	Cgwm2_002980
mTOR signaling pathway	map04150	2	protein kinase	ec:2.7.11.24	5	Cgwm2_045100, Cgwm2_029870, Cgwm2_030030, Cgwm2_014830, Cgwm2_047290
Phenylpropanoid biosynthesis	map00940	1	lactoperoxidase	ec:1.11.1.7	2	Cgwm2_025720, Cgwm2_033410
Glycolysis / Gluconeogenesis	map00010	2	dehydrogenase (NAD(P)+) (phosphorylating)	ec:1.2.1.59	2	Cgwm2_045000, Cgwm2_036350
Glycolysis / Gluconeogenesis	map00010	2	dehydrogenase (phosphorylating)	ec:1.2.1.12	2	Cgwm2_045000, Cgwm2_036350
C5-Branched dibasic acid metabolism	map00660	1	synthase	ec:2.2.1.6	1	Cgwm2_010000
Aminobenzoate degradation	map00627	1	nitrophenyl phosphatase	ec:3.1.3.41	8	Cgwm2_012110, Cgwm2_012440, Cgwm2_032560, Cgwm2_014580, Cgwm2_050010, Cgwm2_033140, Cgwm2_051280, Cgwm2_003100
Butanoate metabolism	map00650	1	synthase	ec:2.2.1.6	1	Cgwm2_010000
Phenylalanine, tyrosine and tryptophan biosynthesis	map00400	3	synthase	ec:4.1.3.27	1	Cgwm2_045230
Phenylalanine, tyrosine and tryptophan biosynthesis	map00400	3	isomerase	ec:5.3.1.24	1	Cgwm2_045230
Phenylalanine, tyrosine and tryptophan biosynthesis	map00400	3	synthase	ec:4.1.1.48	1	Cgwm2_045230
Sulfur metabolism	map00920	2	adenylyltransferase	ec:2.7.7.4	1	Cgwm2_047910
Sulfur metabolism	map00920	2	kinase	ec:2.7.1.25		Cgwm2_032900
Pantothenate and CoA biosynthesis	map00770	1	synthase	ec:2.2.1.6	1	Cgwm2_010000

Glycerophospholipid metabolism	map00564	1	lecithinase B	ec:3.1.1.5	1	Cgwm2_027320
Via metabólica	Mapa KEGG	Número de enzimas na via	Enzima	EC number	Número de sequências enzimáticas	Proteínas
Phosphatidylinositol signaling system	map04070	1	kinase C	ec:2.7.11.13	1	Cgwm2_017850
Biosynthesis of antibiotics	map01130	7	synthase	ec:2.2.1.6	1	Cgwm2_010000,
Biosynthesis of antibiotics	map01130	7	adenylyltransferase	ec:2.7.7.4	1	Cgwm2_047910,
Biosynthesis of antibiotics	map01130	7	synthase	ec:4.1.3.27	1	Cgwm2_045230,
Biosynthesis of antibiotics	map01130	7	isomerase	ec:5.3.1.24	1	Cgwm2_045230,
Biosynthesis of antibiotics	map01130	7	synthase	ec:4.1.1.48	1	Cgwm2_045230,
Biosynthesis of antibiotics	map01130	7	dehydrogenase (NAD(P)+) (phosphorylating)	ec:1.2.1.59	2	Cgwm2_045000, Cgwm2_036350
Biosynthesis of antibiotics	map01130	7	dehydrogenase (phosphorylating)	ec:1.2.1.12	2	Cgwm2_045000, Cgwm2_036350
Thiamine metabolism	map00730	1	phosphatase	ec:3.6.1.15	37	Cgwm2_060590, Cgwm2_037790, Cgwm2_012170, Cgwm2_009240, Cgwm2_033970, Cgwm2_012630, Cgwm2_025740, Cgwm2_050640, Cgwm2_018020, Cgwm2_048100, Cgwm2_006420, Cgwm2_006860, Cgwm2_044620, Cgwm2_006660, Cgwm2_015070, Cgwm2_062930, Cgwm2_042600, Cgwm2_014960, Cgwm2_056160, Cgwm2_013110, Cgwm2_013550, Cgwm2_009110, Cgwm2_038130, Cgwm2_040570, Cgwm2_011910, Cgwm2_030130, Cgwm2_029360, Cgwm2_062480, Cgwm2_045190, Cgwm2_017730, Cgwm2_060180, Cgwm2_048630, Cgwm2_021390, Cgwm2_015160, Cgwm2_007100, Cgwm2_046510, Cgwm2_005860
Selenocompound metabolism	map00450	1	adenylyltransferase	ec:2.7.7.4	1	Cgwm2_047910
Monobactam biosynthesis	map00261	1	adenylyltransferase	ec:2.7.7.4	1	Cgwm2_047910
Purine metabolism	map00230	7	synthase (glutamine-hydrolysing)	ec:6.3.5.2	1	Cgwm2_055810
Purine metabolism	map00230	7	adenylpyrophosphatase	ec:3.6.1.3	12	Cgwm2_056160, Cgwm2_013110, Cgwm2_013550, Cgwm2_038130, Cgwm2_048630, Cgwm2_021390,

Cgwm2_048100, Cgwm2_006420, Cgwm2_011910,
Cgwm2_030130, Cgwm2_006660, Cgwm2_012630

Via metabólica	Mapa KEGG	Número de enzimas na via	Enzima	EC number	Número de sequências enzimáticas	Proteínas
Purine metabolism	map00230	7	phosphatase	ec:3.6.1.15	37	Cgwm2_060590, Cgwm2_037790, Cgwm2_012170, Cgwm2_009240, Cgwm2_033970, Cgwm2_012630, Cgwm2_025740, Cgwm2_050640, Cgwm2_018020, Cgwm2_048100, Cgwm2_006420, Cgwm2_006860, Cgwm2_044620, Cgwm2_006660, Cgwm2_015070, Cgwm2_062930, Cgwm2_042600, Cgwm2_014960, Cgwm2_056160, Cgwm2_013110, Cgwm2_013550, Cgwm2_009110, Cgwm2_038130, Cgwm2_040570, Cgwm2_011910, Cgwm2_030130, Cgwm2_029360, Cgwm2_062480, Cgwm2_045190, Cgwm2_017730, Cgwm2_060180, Cgwm2_048630, Cgwm2_021390, Cgwm2_015160, Cgwm2_007100, Cgwm2_046510, Cgwm2_005860
Purine metabolism	map00230	7	adenylyltransferase	ec:2.7.7.4	1	Cgwm2_047910
Purine metabolism	map00230	7	carboxylase	ec:4.1.1.21	1	Cgwm2_032130
Purine metabolism	map00230	7	urease	ec:3.5.1.5	1	Cgwm2_059730
Purine metabolism	map00230	7	kinase	ec:2.7.1.25	1	Cgwm2_032900
Glycerolipid metabolism	map00561	1	lipase	ec:3.1.1.3	1	Cgwm2_031960
Valine, leucine and isoleucine biosynthesis	map00290	1	synthase	ec:2.2.1.6	1	Cgwm2_010000
Starch and sucrose metabolism	map00500	1	synthase	ec:2.4.1.34	1	Cgwm2_065610
Drug metabolism - other enzymes	map00983	3	synthase (glutamine-hydrolysing)	ec:6.3.5.2	1	Cgwm2_055810
Drug metabolism - other enzymes	map00983	3	ali-esterase	ec:3.1.1.1	2	Cgwm2_027320, Cgwm2_031960
Drug metabolism - other enzymes	map00983	3	phosphoribosyltransferase	ec:2.4.2.10	1	Cgwm2_040490

Carbon fixation in photosynthetic organisms	map00710	2	dehydrogenase (NAD(P)+) (phosphorylating)	ec:1.2.1.59	2	Cgwm2_045000, Cgwm2_036350
---	----------	---	---	-------------	---	----------------------------

Via metabólica	Mapa KEGG	Número de enzimas na via	Enzima	EC number	Número de sequências enzimáticas	Proteínas
Carbon fixation in photosynthetic organisms	map00710	2	dehydrogenase (phosphorylating)	ec:1.2.1.12	2	Cgwm2_045000, Cgwm2_036350
Arginine biosynthesis	map00220	1	urease	ec:3.5.1.5	1	Cgwm2_059730
Th1 and Th2 cell differentiation	map04658	1	phosphatase	ec:3.1.3.16	8	Cgwm2_012110, Cgwm2_012440, Cgwm2_032560, Cgwm2_014580, Cgwm2_050010, Cgwm2_033140, Cgwm2_051280, Cgwm2_003100
T cell receptor signaling pathway	map04660	1	phosphatase	ec:3.1.3.16	8	Cgwm2_012110, Cgwm2_012440, Cgwm2_032560, Cgwm2_014580, Cgwm2_050010, Cgwm2_033140, Cgwm2_051280, Cgwm2_003100
Amino sugar and nucleotide sugar metabolism	map00520	2	synthase	ec:2.4.1.16	4	Cgwm2_065400, Cgwm2_032760, Cgwm2_021440, Cgwm2_041740
Amino sugar and nucleotide sugar metabolism	map00520	2	isomerase	ec:5.3.1.8	2	Cgwm2_047010, Cgwm2_050800
Pyrimidine metabolism	map00240	1	phosphoribosyltransferase	ec:2.4.2.10	1	Cgwm2_040490
Fructose and mannose metabolism	map00051	1	isomerase	ec:5.3.1.8	2	Cgwm2_047010, Cgwm2_050800

Tabela 7: Vias metabólicas mapeadas em *C. neoformans* B-3501A

Via metabólica	Mapa KEGG	Número de enzimas na via	Enzima	EC number	Número de sequências enzimáticas	Proteínas
Atrazine degradation	map00791	1	urease	ec:3.5.1.5	1	Cnb35.059520
mTOR signaling pathway	map04150	1	protein kinase	ec:2.7.11.24	5	Cnb35.028050, Cnb35.045160, Cnb35.014680, Cnb35.027870, Cnb35.042980
Phenylpropanoid biosynthesis	map00940	1	lactoperoxidase	ec:1.11.1.7	2	Cnb35.031530, Cnb35.025150
Glycolysis / Gluconeogenesis	map00010	2	dehydrogenase (NAD(P)+) (phosphorylating)	ec:1.2.1.59	2	Cnb35.042900, Cnb35.034720
Glycolysis / Gluconeogenesis	map00010	2	dehydrogenase (phosphorylating)	ec:1.2.1.12	2	Cnb35.042900, Cnb35.034720
C5-Branched dibasic acid metabolism	map00660	1	synthase	ec:2.2.1.6	1	Cnb35.002480
Aminobenzoate degradation	map00627	1	nitrophenyl phosphatase	ec:3.1.3.41	9	Cnb35.012020, Cnb35.012350, Cnb35.003840, Cnb35.022250, Cnb35.012030, Cnb35.051880, Cnb35.014420, Cnb35.030690, Cnb35.031260
Butanoate metabolism	map00650	1	synthase	ec:2.2.1.6	1	Cnb35.002480
Phenylalanine, tyrosine and tryptophan biosynthesis	map00400	3	synthase	ec:4.1.3.27	1	Cnb35.043110
Phenylalanine, tyrosine and tryptophan biosynthesis	map00400	3	isomerase	ec:5.3.1.24	1	Cnb35.043110
Phenylalanine, tyrosine and tryptophan biosynthesis	map00400	3	synthase	ec:4.1.1.48	1	Cnb35.043110
Sulfur metabolism	map00920	1	adenylyltransferase	ec:2.7.7.4	1	Cnb35.046030
Pantothenate and CoA biosynthesis	map00770	1	synthase	ec:2.2.1.6	1	Cnb35.002480
Glycerophospholipid metabolism	map00564	1	lecithinase B	ec:3.1.1.5	2	Cnb35.023550, Cnb35.061580

Phosphatidylinositol signaling system	map04070	1	kinase C	ec:2.7.11.13	1	Cnb35.018010
---------------------------------------	----------	---	----------	--------------	---	--------------

Via metabólica	Mapa KEGG	Número de enzimas na via	Enzima	EC number	Número de sequências enzimáticas	Proteínas
Biosynthesis of antibiotics	map01130	8	synthase	ec:6.3.2.6	1	Cnb35.030170
Biosynthesis of antibiotics	map01130	8	synthase	ec:2.2.1.6	1	Cnb35.002480
Biosynthesis of antibiotics	map01130	8	adenylyltransferase	ec:2.7.7.4	1	Cnb35.046030
Biosynthesis of antibiotics	map01130	8	synthase	ec:4.1.3.27	1	Cnb35.043110
Biosynthesis of antibiotics	map01130	8	isomerase	ec:5.3.1.24	1	Cnb35.043110
Biosynthesis of antibiotics	map01130	8	synthase	ec:4.1.1.48	1	Cnb35.043110
Biosynthesis of antibiotics	map01130	8	dehydrogenase (NAD(P)+) (phosphorylating)	ec:1.2.1.59	2	Cnb35.042900, Cnb35.034720
Biosynthesis of antibiotics	map01130	8	dehydrogenase (phosphorylating)	ec:1.2.1.12	2	Cnb35.042900, Cnb35.034720
Thiamine metabolism	map00730	1	phosphatase	ec:3.6.1.15	38	Cnb35.036570, Cnb35.062320, Cnb35.015010, Cnb35.016880, Cnb35.012520, Cnb35.017850, Cnb35.011830, Cnb35.014820, Cnb35.025140, Cnb35.032810, Cnb35.006540, Cnb35.052500, Cnb35.032110, Cnb35.036190, Cnb35.060370, Cnb35.013000, Cnb35.013440, Cnb35.036560, Cnb35.040980, Cnb35.038940, Cnb35.020210, Cnb35.062710, Cnb35.007100, Cnb35.007560, Cnb35.014930, Cnb35.007360, Cnb35.028140, Cnb35.021280, Cnb35.044360, Cnb35.045350, Cnb35.001560, Cnb35.001680, Cnb35.007820, Cnb35.043070, Cnb35.053920, Cnb35.018130, Cnb35.012090, Cnb35.057860
Selenocompound metabolism	map00450	1	adenylyltransferase	ec:2.7.7.4	1	Cnb35.046030
Monobactam biosynthesis	map00261	1	adenylyltransferase	ec:2.7.7.4	1	Cnb35.046030
Purine metabolism	map00230	7	synthase (glutamine-hydrolysing)	ec:6.3.5.2	1	Cnb35.055470
Purine metabolism	map00230	7	synthase	ec:6.3.2.6	1	Cnb35.030170

Purine metabolism	map00230	7	adenylpyrophosphatase	ec:3.6.1.3	15	Cnb35.013440, Cnb35.036560, Cnb35.016880, Cnb35.038940, Cnb35.012520, Cnb35.007100, Cnb35.011830, Cnb35.028140, Cnb35.021280, Cnb35.044360, Cnb35.045350, Cnb35.001560, Cnb35.032810, Cnb35.053920, Cnb35.057860
-------------------	----------	---	-----------------------	------------	----	--

Via metabólica	Mapa KEGG	Número de enzimas na via	Enzima	EC number	Número de sequências enzimáticas	Proteínas
Purine metabolism	map00230	7	phosphatase	ec:3.6.1.15	38	Cnb35.036570, Cnb35.062320, Cnb35.015010, Cnb35.016880, Cnb35.012520, Cnb35.017850, Cnb35.011830, Cnb35.014820, Cnb35.025140, Cnb35.032810, Cnb35.006540, Cnb35.052500, Cnb35.032110, Cnb35.036190, Cnb35.060370, Cnb35.013000, Cnb35.013440, Cnb35.036560, Cnb35.040980, Cnb35.038940, Cnb35.020210, Cnb35.062710, Cnb35.007100, Cnb35.007560, Cnb35.014930, Cnb35.007360, Cnb35.028140, Cnb35.021280, Cnb35.044360, Cnb35.045350, Cnb35.001560, Cnb35.001680, Cnb35.007820, Cnb35.043070, Cnb35.053920, Cnb35.018130, Cnb35.012090, Cnb35.057860
Purine metabolism	map00230	7	adenyltransferase	ec:2.7.7.4	1	Cnb35.046030
Purine metabolism	map00230	7	carboxylase	ec:4.1.1.21	1	Cnb35.030170
Purine metabolism	map00230	7	urease	ec:3.5.1.5	1	Cnb35.059520
Glycerolipid metabolism	map00561	1	lipase	ec:3.1.1.3	1	Cnb35.029990
Valine, leucine and isoleucine biosynthesis	map00290	1	synthase	ec:2.2.1.6	1	Cnb35.002480
Starch and sucrose metabolism	map00500	1	synthase	ec:2.4.1.34	1	Cnb35.065660
Drug metabolism - other enzymes	map00983	3	synthase (glutamine-hydrolysing)	ec:6.3.5.2	1	Cnb35.055470
Drug metabolism - other enzymes	map00983	3	ali-esterase,	ec:3.1.1.1	3	Cnb35.023550, Cnb35.029990, Cnb35.061580

Drug metabolism - other enzymes	map00983	3	phosphoribosyltransferase	ec:2.4.2.10	1	Cnb35.038850
Carbon fixation in photosynthetic organisms	map00710	2	dehydrogenase (NAD(P)+) (phosphorylating)	ec:1.2.1.59	2	Cnb35.042900, Cnb35.034720

Via metabólica	Mapa KEGG	Número de enzimas na via	Enzima	EC number	Número de sequências enzimáticas	Proteínas
Carbon fixation in photosynthetic organisms	map00710	2	dehydrogenase (phosphorylating)	ec:1.2.1.12	2	Cnb35.042900, Cnb35.034720
Arginine biosynthesis	map00220	1	urease	ec:3.5.1.5	1	Cnb35.059520
Th1 and Th2 cell differentiation	map04658	1	phosphatase	ec:3.1.3.16	9	Cnb35.012020, Cnb35.012350, Cnb35.003840, Cnb35.022250, Cnb35.012030, Cnb35.051880, Cnb35.014420, Cnb35.030690, Cnb35.031260
T cell receptor signaling pathway	map04660	1	phosphatase	ec:3.1.3.16	9	Cnb35.012020, Cnb35.012350, Cnb35.003840, Cnb35.022250, Cnb35.012030, Cnb35.051880, Cnb35.014420, Cnb35.030690, Cnb35.031260
Amino sugar and nucleotide sugar metabolism	map00520	3	synthase	ec:2.4.1.16	6	Cnb35.037890, Cnb35.040100, Cnb35.021330, Cnb35.030890, Cnb35.037900, Cnb35.065450
Amino sugar and nucleotide sugar metabolism	map00520	3	deacetylase	ec:3.5.1.41	1	Cnb35.024180
Amino sugar and nucleotide sugar metabolism	map00520	3	isomerase	ec:5.3.1.8	2	Cnb35.052350, Cnb35.044870
Pyrimidine metabolism	map00240	1	phosphoribosyltransferase	ec:2.4.2.10	1	Cnb35.038850
Fructose and mannose metabolism	map00051	1	isomerase	ec:5.3.1.8	2	Cnb35.052350, Cnb35.044870

Tabela 8: Vias metabólicas mapeadas em *C. neoformans* var. *grubii* H99

Via metabólica	Mapa KEGG	Número de enzimas na via	Enzima	EC number	Número de sequências enzimáticas	Proteínas
Atrazine degradation	map00791	1	urease	ec:3.5.1.5	1	cngru.018050
mTOR signaling pathway	map04150	1	protein kinase	ec:2.7.11.24	5	cngru.009790, cngru.076140, cngru.073740, cngru.061230, cngru.061470
Phenylpropanoid biosynthesis	map00940	1	lactoperoxidase	ec:1.11.1.7	2	cngru.051200, cngru.057290
Glycolysis / Gluconeogenesis	map00010	2	dehydrogenase (NAD(P)+) (phosphorylating)	ec:1.2.1.59	2	cngru.073640, cngru.065380
Glycolysis / Gluconeogenesis	map00010	2	dehydrogenase (phosphorylating)	ec:1.2.1.12	2	cngru.073640, cngru.065380
C5-Branched dibasic acid metabolism	map00660	1	synthase	ec:2.2.1.6	1	cngru.023570
Aminobenzoate degradation	map00627	1	nitrophenyl phosphatase	ec:3.1.3.41	8	cngru.010090, cngru.054540, cngru.057630, cngru.034850, cngru.002730, cngru.025120, cngru.034490, cngru.058310
Butanoate metabolism	map00650	1	synthase	ec:2.2.1.6	1	cngru.023570
Oxidative phosphorylation	map00190	2	dehydrogenase	ec:1.6.99.3	3	cngru.077910, cngru.077920, cngru.077960
Oxidative phosphorylation	map00190	2	reductase (H ⁺ -translocating)	ec:1.6.5.3	3	cngru.077910, cngru.077920, cngru.077960
Phenylalanine, tyrosine and tryptophan biosynthesis	map00400	3	synthase	ec:4.1.3.27	1	cngru.073880
Phenylalanine, tyrosine and tryptophan biosynthesis	map00400	3	synthase	ec:4.1.1.48	1	cngru.073880
Phenylalanine, tyrosine and tryptophan biosynthesis	map00400	3	isomerase	ec:5.3.1.24	1	cngru.073880
Sulfur metabolism	map00920	2	adenylyltransferase	ec:2.7.7.4	1	cngru.076860
Sulfur metabolism	map00920	2	kinase	ec:2.7.1.25	1	cngru.057880
Pantothenate and CoA biosynthesis	map00770	1	synthase	ec:2.2.1.6	1	cngru.023570

Via metabólica	Mapa KEGG	Número de enzimas na via	Enzima	EC number	Número de sequências enzimáticas	Proteínas
Glycerophospholipid metabolism	map00564	1	lecithinase B	ec:3.1.1.5	2	cngru.012600, cngru.053010
Phosphatidylinositol signaling system	map04070	1	kinase C	ec:2.7.11.13	1	cngru.006210
Biosynthesis of antibiotics	map01130	7	xsynthase	ec:2.7.11.13	1	cngru.023570
Biosynthesis of antibiotics	map01130	7	synthase	ec:4.1.3.27	1	cngru.073880
Biosynthesis of antibiotics	map01130	7	synthase	ec:4.1.1.48	1	cngru.073880
Biosynthesis of antibiotics	map01130	7	adenylyltransferase	ec:2.7.7.4	1	cngru.076860
Biosynthesis of antibiotics	map01130	7	isomerase	ec:5.3.1.24	1	cngru.073880
Biosynthesis of antibiotics	map01130	7	dehydrogenase (NAD(P)+) (phosphorylating)	ec:1.2.1.59	2	cngru.073640, cngru.065380
Biosynthesis of antibiotics	map01130	7	dehydrogenase (phosphorylating)	ec:1.2.1.12	2	cngru.073640, cngru.065380
Thiamine metabolism	map00730	1	phosphatase	ec:3.6.1.15	41	cngru.075260, cngru.029500, cngru.000360, cngru.009530, cngru.056700, cngru.008880, cngru.009650, cngru.011630, cngru.028170, cngru.061110, cngru.022570, cngru.069420, cngru.029040, cngru.063330, cngru.020170, cngru.040480, cngru.036170, cngru.035060, cngru.034570, cngru.017090, cngru.055890, cngru.011170, cngru.051220, cngru.000350, cngru.035610, cngru.022440, cngru.022580, cngru.006350, cngru.008890, cngru.009440, cngru.029250, cngru.063680, cngru.043320, cngru.039390, cngru.073830, cngru.028800, cngru.071770, cngru.001960, cngru.077610, cngru.034280, cngru.055880
Selenocompound metabolism	map00450	1	adenylyltransferase	ec:2.7.7.4	1	cngru.076860
Monobactam biosynthesis	map00261	1	adenylyltransferase	ec:2.7.7.4	1	cngru.076860
Purine metabolism	map00230	7	synthase (glutamine-hydrolysing)	ec:6.3.5.2	1	cngru.005880

Via metabólica	Mapa KEGG	Número de enzimas na via	Enzima	EC number	Número de sequências enzimáticas	Proteínas
Purine metabolism	map00230	7	phosphatase	ec:3.6.1.15	41	cngru.075260, cngru.029500, cngru.000360, cngru.009530, cngru.056700, cngru.008880, cngru.009650, cngru.011630, cngru.028170, cngru.061110, cngru.022570, cngru.069420, cngru.029040, cngru.063330, cngru.020170, cngru.040480, cngru.036170, cngru.035060, cngru.034570, cngru.017090, cngru.055890, cngru.011170, cngru.051220, cngru.000350, cngru.035610, cngru.022440, cngru.022580, cngru.006350, cngru.008890, cngru.009440, cngru.029250, cngru.063680, cngru.043320, cngru.039390, cngru.073830, cngru.028800, cngru.071770, cngru.001960, cngru.077610, cngru.034280, cngru.055880
Purine metabolism	map00230	7	adenylpyrophosphatase	ec:3.6.1.3	15	cngru.075260, cngru.000350, cngru.035610, cngru.000360, cngru.022440, cngru.061110, cngru.063330, cngru.040480, cngru.036170, cngru.035060, cngru.028800, cngru.071770, cngru.055890, cngru.034280, cngru.055880
Purine metabolism	map00230	7	adenyltransferase	ec:2.7.7.4	1	cngru.076860
Purine metabolism	map00230	7	carboxylase	ec:4.1.1.21	1	cngru.058890
Purine metabolism	map00230	7	urease	ec:3.5.1.5	1	cngru.018050
Purine metabolism	map00230	7	kinase	ec:2.7.1.25	1	cngru.057880
Glycerolipid metabolism	map00561	1	lipase	ec:3.1.1.3	1	cngru.059080
Valine, leucine and isoleucine biosynthesis	map00290	1	synthase	ec:2.2.1.6	1	cngru.023570
Starch and sucrose metabolism	map00500	1	synthase	ec:2.4.1.34	1	cngru.013830
Drug metabolism - other enzymes	map00983	3	synthase (glutamine-hydrolysing)	ec:6.3.5.2	1	cngru.005880
Drug metabolism - other enzymes	map00983	3	ali-esterase	ec:3.1.1.1	3	cngru.059080, cngru.012600, cngru.053010

Via metabólica	Mapa KEGG	Número de enzimas na via	Enzima	EC number	Número de sequências enzimáticas	Proteínas
Drug metabolism - other enzymes	map00983	3	phosphoribosyltransferase	ec:2.4.2.10	1	cngru.071870
Arginine biosynthesis	map00220	1	urease	ec:3.5.1.5	1	cngru.018050
Carbon fixation in photosynthetic organisms	map00710	2	dehydrogenase (NAD(P)+) (phosphorylating)	ec:1.2.1.59	2	cngru.073640, cngru.065380
Carbon fixation in photosynthetic organisms	map00710	2	dehydrogenase (phosphorylating)	ec:1.2.1.12	2	cngru.073640, cngru.065380
Th1 and Th2 cell differentiation	map04658	1	phosphatase	ec:3.1.3.16	8	cngru.010090, cngru.054540, cngru.057630, cngru.034850, cngru.002730, cngru.025120, cngru.034490, cngru.058310
T cell receptor signaling pathway	map04660	1	phosphatase	ec:3.1.3.16	8	cngru.010090, cngru.054540, cngru.057630, cngru.034850, cngru.002730, cngru.025120, cngru.034490, cngru.058310
Amino sugar and nucleotide sugar metabolism	map00520	3	deacetylase	ec:3.5.1.41	1	cngru.052270
Amino sugar and nucleotide sugar metabolism	map00520	3	isomerase	ec:5.3.1.8	2	cngru.075810, cngru.002110
Amino sugar and nucleotide sugar metabolism	map00520	3	synthase	ec:2.4.1.16	5	cngru.072980, cngru.014040, cngru.058070, cngru.043370, cngru.070470
Pyrimidine metabolism	map00240	1	phosphoribosyltransferase	ec:2.4.2.10	1	cngru.071870
Fructose and mannose metabolism	map00051	1	isomerase	ec:5.3.1.8	2	cngru.075810, cngru.002110

Tabela 9: Vias metabólicas mapeadas em *C. neoformans* var. *neoformans* JEC21.

Via metabólica	Mapa KEGG	Número de enzimas na via	Enzima	EC number	Número de sequências enzimáticas	Proteínas
Atrazine degradation	map00791	1	urease	ec:3.5.1.5	1	Cnj21.058560
mTOR signaling pathway	map04150	1	protein kinase	ec:2.7.11.24	5	Cnj21.042170, Cnj21.063020, Cnj21.042380, Cnj21.034940, Cnj21.060780
Phenylpropanoid biosynthesis	map00940	1	lactoperoxidase	ec:1.11.1.7	2	Cnj21.038070, Cnj21.045660
Glycolysis / Gluconeogenesis	map00010	2	dehydrogenase (NAD(P)+) (phosphorylating)	ec:1.2.1.59	2	Cnj21.050200, Cnj21.060700
Glycolysis / Gluconeogenesis	map00010	2	dehydrogenase (phosphorylating)	ec:1.2.1.12	2	Cnj21.050200, Cnj21.060700
C5-Branched dibasic acid metabolism	map00660	1	synthase	ec:2.2.1.6	1	Cnj21.002560
Aminobenzoate degradation	map00627	1	nitrophenyl phosphatase	ec:3.1.3.41	9	Cnj21.024720, Cnj21.040900, Cnj21.024410, Cnj21.024710, Cnj21.035190, Cnj21.045390, Cnj21.044830, Cnj21.003790, Cnj21.010260
Butanoate metabolism	map00650	1	synthase	ec:2.2.1.6	1	Cnj21.002560
Phenylalanine, tyrosine and tryptophan biosynthesis	map00400	3	synthase	ec:4.1.3.27	1	Cnj21.060900
Phenylalanine, tyrosine and tryptophan biosynthesis	map00400	3	synthase	ec:4.1.1.48	1	Cnj21.060900
Phenylalanine, tyrosine and tryptophan biosynthesis	map00400	3	isomerase	ec:5.3.1.24	1	Cnj21.060900
Sulfur metabolism	map00920	2	adenylyltransferase	ec:2.7.7.4	1	Cnj21.063910
Sulfur metabolism	map00920	2	kinase	ec:2.7.1.25	1	Cnj21.045160
Pantothenate and CoA biosynthesis	map00770	1	synthase	ec:2.2.1.6	1	Cnj21.002560
Glycerophospholipid metabolism	map00564	1	lecithinase B	ec:3.1.1.5	2	Cnj21.039590, Cnj21.018700

Via metabólica	Mapa KEGG	Número de enzimas na via	Enzima	EC number	Número de sequências enzimáticas	Proteínas
Phosphatidylinositol signaling system	map04070	1	kinase C	ec:2.7.11.13	1	Cnj21.031690
Biosynthesis of antibiotics	map01130	8	synthase	ec:2.2.1.6	1	Cnj21.002560
Biosynthesis of antibiotics	map01130	8	synthase	ec:4.1.3.27	1	Cnj21.060900
Biosynthesis of antibiotics	map01130	8	synthase	ec:4.1.1.48	1	Cnj21.060900
Biosynthesis of antibiotics	map01130	8	synthase	ec:6.3.2.6	1	Cnj21.044360
Biosynthesis of antibiotics	map01130	8	adenylyltransferase	ec:2.7.7.4	1	Cnj21.063910
Biosynthesis of antibiotics	map01130	8	isomerase	ec:5.3.1.24	1	Cnj21.060900
Biosynthesis of antibiotics	map01130	8	dehydrogenase (NAD(P)+) (phosphorylating)	ec:1.2.1.59	2	Cnj21.050200, Cnj21.060700
Biosynthesis of antibiotics	map01130	8	dehydrogenase (phosphorylating)	ec:1.2.1.12	2	Cnj21.050200, Cnj21.060700
Thiamine metabolism	map00730	1	phosphatase	ec:3.6.1.15	39	Cnj21.008320, Cnj21.007650, Cnj21.055590, Cnj21.006980, Cnj21.007410, Cnj21.048360, Cnj21.024650, Cnj21.029560, Cnj21.042450, Cnj21.038090, Cnj21.019470, Cnj21.023760, Cnj21.024910, Cnj21.001730, Cnj21.062180, Cnj21.031570, Cnj21.007220, Cnj21.006990, Cnj21.034620, Cnj21.006430, Cnj21.031850, Cnj21.034800, Cnj21.034700, Cnj21.048370, Cnj21.028500, Cnj21.023330, Cnj21.046230, Cnj21.060130, Cnj21.060870, Cnj21.009660, Cnj21.063200, Cnj21.024240, Cnj21.048750, Cnj21.001600, Cnj21.062170, Cnj21.046910, Cnj21.019860, Cnj21.053540, Cnj21.057740
Selenocompound metabolism	map00450	1	adenylyltransferase	ec:2.7.7.4	1	Cnj21.063910
Monobactam biosynthesis	map00261	1	adenylyltransferase	ec:2.7.7.4	1	Cnj21.063910
Purine metabolism	map00230	8	synthase (glutamine-hydrolysing)	ec:6.3.5.2	1	Cnj21.013490
Purine metabolism	map00230	8	phosphatase	ec:3.6.1.15	39	Cnj21.008320, Cnj21.007650, Cnj21.055590, Cnj21.006980, Cnj21.007410, Cnj21.048360,

Cnj21.024650, Cnj21.029560, Cnj21.042450, Cnj21.038090, Cnj21.019470, Cnj21.023760, Cnj21.024910, Cnj21.001730, Cnj21.062180, Cnj21.031570, Cnj21.007220, Cnj21.006990, Cnj21.034620, Cnj21.006430, Cnj21.031850, Cnj21.034800, Cnj21.034700, Cnj21.048370, Cnj21.028500, Cnj21.023330, Cnj21.046230, Cnj21.060130, Cnj21.060870, Cnj21.009660, Cnj21.063200, Cnj21.024240, Cnj21.048750, Cnj21.001600, Cnj21.062170, Cnj21.046910, Cnj21.019860, Cnj21.053540, Cnj21.057740

Via metabólica	Mapa KEGG	Número de enzimas na via	Enzima	EC number	Número de sequências enzimáticas	Proteínas
Purine metabolism	map00230	8	synthase	ec:6.3.2.6	1	Cnj21.044360
Purine metabolism	map00230	8	adenylpyrophosphatase	ec:3.6.1.3	17	Cnj21.008320, Cnj21.006990, Cnj21.055590, Cnj21.006980, Cnj21.048360, Cnj21.048370, Cnj21.028500, Cnj21.023330, Cnj21.060130, Cnj21.063200, Cnj21.024240, Cnj21.042450, Cnj21.001600, Cnj21.062170, Cnj21.024910, Cnj21.046910, Cnj21.062180
Purine metabolism	map00230	8	adenylyltransferase	ec:2.7.7.4	1	Cnj21.063910
Purine metabolism	map00230	8	xcarboxylase	ec:2.7.7.4	1	Cnj21.044360
Purine metabolism	map00230	8	urease	ec:3.5.1.5	1	Cnj21.058560
Purine metabolism	map00230	8	kinase	ec:2.7.1.25	1	Cnj21.045160
Valine, leucine and isoleucine biosynthesis	map00290	1	synthase	ec:2.2.1.6	1	Cnj21.002560
Glycerolipid metabolism	map00561	1	lipase	ec:3.1.1.3	1	Cnj21.044200
Starch and sucrose metabolism	map00500	1	synthase	ec:2.4.1.34	1	Cnj21.022610
Drug metabolism - other enzymes	map00983	3	synthase (glutamine-hydrolysing)	ec:6.3.5.2	1	Cnj21.013490
Drug metabolism - other enzymes	map00983	3	ali-esterase	ec:3.1.1.1	3	Cnj21.039590, Cnj21.044200, Cnj21.018700

Via metabólica	Mapa KEGG	Número de enzimas na via	Enzima	EC number	Número de sequências enzimáticas	Proteínas
Drug metabolism - other enzymes	map00983	3	phosphoribosyltransferase	ec:2.4.2.10	1	Cnj21.055660
Arginine biosynthesis	map00220	1	urease	ec:3.5.1.5	1	Cnj21.058560
Carbon fixation in photosynthetic organisms	map00710	2	dehydrogenase (NAD(P)+) (phosphorylating)	ec:1.2.1.59	2	Cnj21.050200, Cnj21.060700
Carbon fixation in photosynthetic organisms	map00710	2	dehydrogenase (phosphorylating)	ec:1.2.1.12	2	Cnj21.050200, Cnj21.060700
Th1 and Th2 cell differentiation	map04658	1	phosphatase	ec:3.1.3.16	9	Cnj21.024720, Cnj21.040900, Cnj21.024410, Cnj21.024710, Cnj21.035190, Cnj21.045390, Cnj21.044830, Cnj21.003790, Cnj21.010260
T cell receptor signaling pathway	map04660	1	phosphatase	ec:3.1.3.16	9	Cnj21.024720, Cnj21.040900, Cnj21.024410, Cnj21.024710, Cnj21.035190, Cnj21.045390, Cnj21.044830, Cnj21.003790, Cnj21.010260
Amino sugar and nucleotide sugar metabolism	map00520	3	deacetylase	ec:3.5.1.41	1	Cnj21.038990
Amino sugar and nucleotide sugar metabolism	map00520	3	isomerase	ec:5.3.1.8	2	Cnj21.009800, Cnj21.062720
Amino sugar and nucleotide sugar metabolism	map00520	3	synthase	ec:2.4.1.16	5	Cnj21.056500, Cnj21.045020, Cnj21.022400, Cnj21.028440, Cnj21.054430
Pyrimidine metabolism	map00240	1	phosphoribosyltransferase	ec:2.4.2.10	1	Cnj21.055660
Fructose and mannose metabolism	map00051	1	isomerase	ec:5.3.1.8	2	Cnj21.009800, Cnj21.062720