

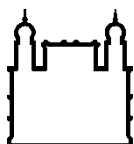
MINISTÉRIO DA SAÚDE
FUNDAÇÃO OSWALDO CRUZ
INSTITUTO OSWALDO CRUZ

Mestrado em Biologia Computacional e Sistemas

ESTUDO COMPUTACIONAL DE DESORDEM PROTEICA NOS
GENOMAS DE *Cryptococcus* spp.

GRACE SANTOS TAVARES

Rio de Janeiro
Março de 2016



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

Programa de Pós-Graduação em Biologia Computacional e Sistemas

GRACE SANTOS TAVARES

Estudo computacional de desordem proteica nos genomas de *Cryptococcus* spp.

Dissertação apresentada ao Instituto Oswaldo Cruz como parte dos requisitos para obtenção do título de Mestre em Ciências, com área de concentração em Genômica Funcional, Evolução e Filogenômica.

Orientador (es): Prof. Dr. Jeronimo Conceição Ruiz
Profª. Dra. Daniela de Melo Resende

RIO DE JANEIRO
MARÇO DE 2016

Ficha catalográfica elaborada pela
Biblioteca de Ciências Biomédicas/ ICICT / FIOCRUZ - RJ

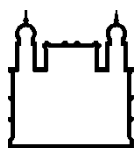
T231 Tavares, Grace Santos

Estudo computacional de desordem proteica nos genomas de
Cryptococcus spp. / Grace Santos Tavares. – Rio de Janeiro, 2016.
xvi, 126 f. : il. ; 30 cm.

Dissertação (Mestrado) – Instituto Oswaldo Cruz, Pós-Graduação em
Biologia Computacional e Sistemas, 2016.
Bibliografia: f. 69-74

1. Proteínas intrinsecamente desordenadas. 2. Regiões
intrinsecamente desordenadas. 3. *Cryptococcus*. 4. Fatores virulência. 5.
Papel biológico de IDPs e IDRs. I. Título.

CDD 572.65



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

Programa de Pós-Graduação em Biologia Computacional e Sistemas

AUTORA: GRACE SANTOS TAVARES

Estudo computacional de desordem proteica nos genomas de *Cryptococcus* spp.

**ORIENTADOR (ES): Prof. Dr. Jeronimo Conceição Ruiz
Prof^a. Dr^a. Daniela de Melo Resende**

Aprovada em: 30/03/2016

EXAMINADORES:

Prof. Dr. Marcos P. Catanho de Souza (Instituto Oswaldo Cruz - RJ)
Prof. Dr. Daniel de Assis Santos (Universidade Federal de Minas Gerais – MG)
Prof. Dr. Gabriel da Rocha Fernandes (Centro de Pesquisas René Rachou – MG)
Prof. Dr. Douglas E. Valente Pires (Centro de Pesquisas René Rachou – MG)
Prof^a. Dr^a. Cristiana F. Alves de Brito (Centro de Pesquisas René Rachou – MG)

Rio de Janeiro, 30 de março de 2016

AGRADECIMENTOS

A realização desta pesquisa contou com a colaboração de amigos, colegas de laboratório e instituições, aos quais tenho muito a agradecer.

Ao Instituto Oswaldo Cruz, ao Centro de Pesquisas René Rachou (CPqRR) e ao Programa de Pós-Graduação em Biologia Computacional e Sistemas pelos conhecimentos adquiridos durante esses dois anos de intenso estudo.

À Coordenação e aos professores do Programa de Pós-Graduação em Biologia Computacional e Sistemas, em especial ao Dr. Ernesto Caffarena e à secretária Rose Pani, por possibilitarem o acesso aos estudos de forma simultânea em duas unidades da Fundação Oswaldo Cruz.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa de Mestrado.

À minha ex, mas querida professora Luciana Márcia Oliveira, que sempre me incentivou e apoiou. Por me fazer enxergar além dos muros da Biologia Molecular e por despertar em mim a paixão que hoje tenho pela Bioinformática.

Ao Dr. Jeronimo Conceição Ruiz, pela oportunidade oferecida, pelo incentivo ao ingresso no mundo da Bioinformática, pela confiança em mim depositada, pela paciência, pelas correções e ensinamentos, além da amizade que se consolidou.

À Dra. Daniela de Melo Resende, pelo incentivo, pelas preciosas sugestões, correções, ensinamentos e ajuda ao longo do trabalho. Pelas conversas e pela amizade formada.

A todos os demais colegas do Grupo Informática de Biosistemas, agradeço pelo carinho e amizade. Em especial, à Leilane pela grande ajuda ao longo do trabalho, pela grande amizade formada, pelas conversas e bons momentos que juntas passamos. Ao Fred por ser uma pessoa tão carismática e amiga, além da inestimável ajuda no desenvolvimento dos *scripts* e discussões científicas. Ao Jader pela enorme paciência em me ajudar nas instalações dos programas.

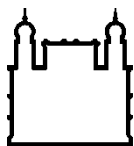
À minha grande amiga Juliana Bezerra, que apesar da distância, nunca perderemos a essência de nossa amizade.

Ao Jhonatan pelo carinho e companheirismo que sempre tivemos. Pelo inestimável apoio durante toda a minha jornada acadêmica e por todos os momentos que passamos e ainda passaremos juntos.

Aos meus pais Eder e Marise, que sempre me incentivaram a seguir meus sonhos, a torcer e vibrar pelas minhas conquistas. Por terem me dado o meu maior bem, a vida.

Às minhas queridas avós Divina e Lucy, que sempre estiveram ao meu lado, me confortando e amparando.

A todos os meus familiares pelo apoio e carinho.



Ministério da Saúde

FIOCRUZ
Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

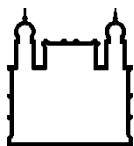
Estudo computacional de desordem proteica nos genomas de *Cryptococcus* spp.

RESUMO

DISSERTAÇÃO DE MESTRADO EM BIOLOGIA COMPUTACIONAL E SISTEMAS

Grace Santos Tavares

Desde a década de 1990, a comunidade científica tem relatado um número crescente de proteínas e trechos que não apresentam uma conformação tridimensional estável e única, mas que são biologicamente ativas, o que contrasta o paradigma estrutura-função de uma proteína. Essas proteínas podem ser denominadas como Proteínas Intrinsecamente Desordenadas (IDPs), já os trechos como Regiões Intrinsecamente Desordenadas (IDRs). As IDPs e IDRs apresentam características de composição de aminoácidos que as identificam. Além disso, como foi relatado para o protozoário *Plasmodium falciparum*, as IDPs e IDRs podem favorecer a relação parasito-hospedeiro. Causada por espécies do fungo *Cryptococcus* spp., a criptococose é uma doença que afeta principalmente pacientes com imunidade comprometida e é a principal causa de mortalidade entre os pacientes portadores do vírus HIV. Dentro deste contexto, avaliamos computacionalmente o papel biológico das IDPs e IDRs em 12 proteomas preditos de *Cryptococcus* spp. Para tanto, avaliamos as diferentes ferramentas computacionais disponíveis para a predição *ab initio* dessa interessante classe de proteínas. Parte do processo analítico envolveu a reanotação desses 12 genomas, tendo este trabalho contribuído para a anotação estrutural e funcional de oito genomas (57.680 novos genes anotados) para os quais não havia informação em banco de dados de domínio público. Dentre os 11 preditores avaliados em nossas análises, cinco foram selecionados com base em seu emprego em larga escala (DisEMBL, empregando as abordagens REM465 e HOTLOOPS, GlobPlot, IUPred e ANCHOR). A análise comparativa entre os grupos virulentos e não virulentos não revelou diferença estatisticamente significativa. Além disso, considerando as sequências consenso das predições, encontramos um conteúdo de aproximadamente 60-70% de proteínas em cada proteoma analisado, com pelo menos uma IDR de no mínimo 40 resíduos de aminoácidos consecutivos. Já os resultados referentes à análise de agrupamento revelaram que das 18.900 proteínas presentes no *core genome* dos 12 organismos, 11.409 apresentam pelo menos uma IDR. Outro ponto observado foi que muitas IDPs ou IDRs estão participando de diversos processos biológicos, estão relacionadas aos fatores de virulência do fungo e/ou atuando como enzimas em vias metabólicas distintas. Neste estudo hipotetizamos e descrevemos o papel funcional vital dessas proteínas em vários processos biológicos nos genomas de *Cryptococcus* spp.



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

Computational study of protein disorder in *Cryptococcus* spp. genomes

ABSTRACT

MASTER DISSERTATION IN COMPUTATIONAL BIOLOGY AND SYSTEMS

Grace Santos Tavares

Since the 1990s, the scientific community has reported an increasing number of proteins and protein regions that don't have a stable and unique three-dimensional conformation, but which are biologically active, which contrasts the structure-function paradigm of a protein. These proteins are called Intrinsically Disordered Proteins (IDPs) and the regions, Intrinsically Disordered Regions (IDRs). IDPs and IDRs have compositional biases that characterize them. Moreover, as has been reported for the parasite *Plasmodium falciparum*, IDPs and IDRs can favor the host-parasite relationship. Caused by species of the fungus *Cryptococcus* spp., Cryptococcosis is a disease that primarily affects patients with compromised immunity and is the leading cause of mortality among patients with HIV. Within this context, we evaluated computationally the biological role of IDPs and IDRs on 12 predicted *Cryptococcus* spp. proteomes. Therefore, we evaluated the various computer tools available for *ab initio* prediction of this interesting class of proteins. Part of the analytical process involved annotation of these 12 genomes, and this work contributed to the structural and functional annotation of eight genomes (57,680 new genes annotated) for which there was no information on the public domain databases. Among the 11 predictors evaluated in our analysis, five were selected based on their large-scale use (DisEMBL, employing REM465 and HOTLOOPS approaches, GlobPlot, IUPred and ANCHOR). The comparative analysis between virulent and non-virulent groups showed no statistically significant difference. Furthermore, considering the consensus sequences of the predictions, we found a content of approximately 60-70% of proteins on each proteome with at least one IDR with at least 40 consecutive amino acid residues. Regarding the results of the cluster analysis, we observed that, from the 18,900 proteins present in the core genome of the analyzed organisms, 11,409 have at least one IDR. Another point observed was that IDPs or IDRs are participating in many biological processes, are related to virulence factors of the fungus and/or acting as enzymes in different metabolic pathways. In this study we hypothesize and describe the vital functional role of these proteins in various biological processes in the genomes of *Cryptococcus* spp.

SUMÁRIO

RESUMO.....	VII
ABSTRACT.....	VIII
1 INTRODUÇÃO	17
1.1 Desordem estrutural e definição da nomenclatura	17
1.2 Características gerais da desordem estrutural.....	19
1.2.1 Composição de aminoácidos.....	19
1.2.2 Interações proteicas	20
1.2.3 Papel biológico	21
1.2.4 Desenvolvimento de novas drogas.....	22
1.3 Predição computacional de IDPs e IDRs.....	22
1.3.1 Características gerais	22
1.4 Organismo modelo: <i>Cryptococcus</i> spp.	24
1.4.1 <i>Cryptococcus</i> spp e epidemiologia.	24
1.4.2 Criptococose.....	27
1.4.3 Fatores de virulência.....	28
1.4.4 Resposta imune do hospedeiro	30
2 JUSTIFICATIVA	32
3 OBJETIVOS.....	33
3.1 Objetivo Geral.....	33
3.2 Objetivos Específicos	33
4 MATERIAIS E MÉTODOS	34
4.1 Scripts desenvolvidos	34
4.2 Genomas de <i>Cryptococcus</i> spp.....	34
4.3 Anotação e Reanotação dos proteomas preditos de <i>Cryptococcus</i> spp.	34
4.3.1 Artemis.....	35
4.3.2 Normalização e reanotação genômica.....	35
4.3.3 Dados genômicos sem predição gênica	36

4.3.4	Dados genômicos com predição gênica	37
4.4	Predição de desordem estrutural proteica.....	38
4.4.1	Seleção dos preditores computacionais de desordem estrutural proteica	38
4.4.2	Predição das IDPs e IDRs	39
4.4.3	Banco de dados relacional de IDPs e IDRs em <i>Cryptococcus</i> spp.	40
4.5.2	Sequência consenso das IDPs e IDRs preditas	40
4.4.4	Análise de Enriquecimento Funcional.....	42
4.4.5	Mapeamento das vias metabólicas e associação aos fatores de virulência.....	42
4.4.6	Análise de agrupamento	42
5	RESULTADOS E DISCUSSÃO	44
5.1	Seleção dos genomas de <i>Cryptococcus</i> spp.....	44
5.2	Anotação genômica	44
5.2.1	Anotação estrutural.....	46
5.2.2	Anotação funcional	46
5.3	Seleção dos preditores de IDPs e IDRs.....	48
5.3.1	DisEMBL.....	51
5.3.2	GlobPlot.....	53
5.3.3	IUPred.....	54
5.3.4	ANCHOR	55
5.4	Desordem proteica	56
5.4.1	Análise de enriquecimento funcional	61
5.4.2	IDPs e IDRs relacionadas aos fatores de virulência de <i>Cryptococcus</i> spp.	62
6	CONCLUSÃO	67
7	PERSPECTIVAS.....	68
8	REFERÊNCIAS BIBLIOGRAFICAS.....	69
APENDICE A – GRÁFICOS RELACIONADOS À ANÁLISE DE ENRIQUECIMENTO ENTRE AS ESPÉCIES/LINHAGENS VIRULENTAS		75

APENDICE B- GRÁFICOS RELACIONADOS À ANÁLISE DE ENRIQUECIMENTO ENTRE AS ESPÉCIES NÃO VIRULENTAS	78
APENDICE C – TABELAS RELACIONADAS AO MAPEAMENTO DAS IDPS E IDRS NAS VIAS METABÓLICAS ENTRE AS ESPÉCIES/LINHAGENS VIRULENTAS.....	81

ÍNDICE DE FIGURA

Figura 1.1: Exemplos de IDP e IDR.....	18
Figura 1.2: Exemplo de uma proteína totalmente desordenada, em seu estado nativo, adquirindo uma conformação tridimensional ao se ligar a seu alvo proteico.	21
Figura 1.3: Classificação das espécies do complexo <i>C. neoformans</i>	25
Figura 1.4: Distribuição dos sorotipos (A, B, C, D e AD) de <i>Cryptococcus</i> spp. pelas regiões geográficas do Brasil.	26
Figura 1.5: Esquematização dos processos de migração do <i>Cryptococcus</i> ssp. para o SNC.....	28
Figura 4.1: Fluxograma esquematizando as etapas analíticas do processo de anotação e reanotação genômica utilizado neste estudo.....	35
Figura 4.2: Fluxograma esquematizando as etapas analíticas do processo de predição de desordem estrutural proteica utilizado neste estudo,.....	38
Figura 4.3: Resultado gerado pelo <i>script iupred-parser-v3.pl</i> referente ao resultado do preditor IUPred.	39
Figura 4.4: Esquema do Banco de Dados Relacional <i>idp_crypto</i>	40
Figura 4.5: Diferentes regiões consenso formadas em uma mesma proteína e que atendiam ao critério de 10% de sobreposição.....	41
Figura 5.1: Exemplo de visualização pela plataforma gráfica do programa Artemis referente à anotação genômica de <i>C. neoformans</i> var. <i>neoformnas</i> , linhagem B3501A	47
Figura 5.2: Diagrama de venn referente à análise de clusterização.....	48
Figura 5.3: Gráfico referente à predição de desordem proteica por preditor e organismo.....	51
Figura 5.4: Gráficos relacionados ao conteúdo de desordem proteica identificado por cada preditor e em cada organismo.....	58
Figura 5.5: Gráfico relacionado às IDPs e IDRs obtidas pelo consenso das análises preditivas.....	60
Figura 5.6: Diagrama de venn referente à análise de clusterização.....	61
Figura 5.7: Via de Biossíntese de N-glucanos.....	66
Figura A.1: Análise de enriquecimento de desordem proteica em <i>C. gattii</i> , linhagem WM276.....	75

Figura A.2: Gráfico obtido a partir da análise de enriquecimento referente ao conteúdo total de desordem proteica dentro do proteoma de <i>C. gattii</i> , linhagem R265.....	75
Figura A.3: Gráfico referente à Análise de enriquecimento de desordem proteica em <i>C. neoformans</i> , linhagem B3501A.....	76
Figura A.4: Gráfico referente à análise de enriquecimento de desordem proteica em <i>C. gattii</i> , linhagem CBS7750.. ..	76
Figura A.5: Gráfico referente à Análise de enriquecimento de desordem proteica em <i>C. neoformans</i> , linhagem JEC21.....	77
Figura A.6: Gráfico referente à Análise de enriquecimento de desordem proteica em <i>C. neoformans</i> var. <i>grubii</i> , linhagem H99.. ..	77
Figura B.1: Gráfico referente à Análise de enriquecimento de desordem proteica em <i>C. dejecticola</i> , linhagem CBS10117.	78
Figura B.2: Gráfico referente à Análise de enriquecimento de desordem proteica em <i>C. bestiolae</i> , linhagem CBS10118.....	78
Figura B.3: Gráfico referente à análise de enriquecimento de desordem proteica em <i>C. heveanensis</i> , linhagem CBS569.....	79
Figura B.4: Gráfico referente à análise de enriquecimento de desordem proteica em <i>C. flavescens</i> , linhagem NRRL Y-50378.....	79
Figura B.5: Gráfico referente à análise de enriquecimento de desordem proteica em <i>C. pinus</i> , linhagem CBS10737	80
Figura B.6: Gráfico referente à análise de enriquecimento de desordem proteica em <i>C. heveanensis</i> , linhagem BCC8938.....	80

LISTA DE TABELAS

Tabela 4.1: Parâmetros solicitados pelo <i>script consenso-coord-idp-v4.pl</i> para sua execução.....	41
Tabela 5.1: Lista de genomas de <i>Cryptococcus</i> spp. selecionados.....	44
Tabela 5.2: Lista de genomas de <i>Cryptococcus</i> spp. com o número de genes encontrados pela predição <i>ab initio</i> e os provenientes da anotação original.	45
Tabela 5.3: Lista de genomas de <i>Cryptococcus</i> spp. submetidos à anotação funcional.....	47
Tabela 5.4: Relação de preditores selecionados a partir dos critérios estabelecidos	49
Tabela 5.5: Resultado do preditorDynaMine e da abordagem COILS implementada pelo DisEMBL.....	50
Tabela 5.6: Número de IDPs e IDRs preditas para as espécies <i>C. neoformans</i> e <i>C. gattii</i> identificadas entre os fatores de virulência.....	64
Tabela 5.7: Relação de vias metabólicas comum a todos os organismos virulentos	65
Tabela C.1: Vias metabólicas mapeadas em <i>C. gattii</i> CBS 7750. Todas as proteínas mapeadas referem-se às IDPs e IDRs identificadas neste estudo.....	81
Tabela C.2: Vias metabólicas mapeadas em <i>C. gattii</i> , linhagem R265. Todas as proteínas mapeadas referem-se às IDPs e IDRs identificadas neste estudo.....	95
Tabela C.3: Vias metabólicas mapeadas em <i>C. gattii</i> , linhagem WM276. Todas as proteínas mapeadas referem-se às IDPs e IDRs identificadas neste estudo.....	111
Tabela C.4: Vias metabólicas mapeadas em <i>C. neoformans</i> var. <i>grubii</i> , linhagem CBS 7750. Todas as proteínas mapeadas referem-se às IDPs e IDRs identificadas neste estudo.....	124
Tabela C.5: Vias metabólicas mapeadas em <i>C. neoformans</i> var. <i>neoformans</i> , linhagem B3501A. Todas as proteínas mapeadas referem-se às IDPs e IDRs identificadas neste estudo.....	136
Tabela C.6: Vias metabólicas mapeadas em <i>C. neoformans</i> var. <i>neoformans</i> , linhagem JEC21. Todas as proteínas mapeadas referem-se às IDPs e IDRs identificadas neste estudo.....	140

LISTA DE SIGLAS E ABREVIATURAS

PDB – *Protein Data Bank*

IDP – Proteína Intrinsecamente Desordenada

IDR – Região Intrinsecamente Desordenada

Ala – Alanina

Arg – Arginina

Gly – Glicina

Gln – Glutamina

Ser – Serina

Pro – Prolina

Glu - Ácido Glutâmico

Lys – Lisina

Trp – Triptofano

Cys – Cisteína

Phe – Fenilalanina

Ile – Isoleucina

Tyr – Tirosina

Val – Valina

Leu – Leucina

Asn – Asparagina

His – Histidina

Met – Metionina

Thr – Treonina

Asp – Ácido Aspártico

CASP – *Critical Assessment of Structure Prediction*

ΔS – Variação de entropia

ΔH – Variação de entalpia

TLR - Receptor do tipo *Toll*

TLR-4 – Receptor do tipo *Toll* 4

TLR-9 – Receptor do tipo *Toll* 9

IL – Interleucina

IL1 α – Interleucina 1 alfa

IL-1 β - Interleucina 1 beta

IL-6 – Interleucina 6
IL-17 – Interleucina 17
IL-22 – Interleucina 22
GXM – Glucuronoxilomanana
GalXM – Galactoxilomanana
MIP-2 α - Proteína Inflamatória de Macrófagos 2 alfa
TNF α - Fator de necrose tumoral alfa
IFN- γ - Interferon gama
BD – Banco de Dados
NR – Não redundante
CDD - Banco de Dados de Domínios Conservados
Pfam –*Protein Families Database*
NCBI –*National Center for Biotechnology Information*
Å –Angstroms
CDS–*Coding Sequence*
SNC –Sistema Nervoso Central
HIV– Vírus da Imunodeficiência Humana
DSSP – Dicionário de Estruturas Secundárias de Proteínas
SCOP–*Structural Classification of Proteins*
 ΔG – Energia Livre de *Gibbs*
KEGG –Enciclopédia Kyoto de Genes e Genomas
BLAST - *Basic Aligment Search Tool*
UniProt - *Universal Protein Resource*
ABC – *ATP Binding Cassette*
ORF– Fase aberta de leitura / quadros abertos de leitura

1 INTRODUÇÃO

1.1 Desordem estrutural e definição da nomenclatura

Por mais de um século, o conceito mais aceitável sobre função proteica estava intimamente ligado ao paradigma estrutura-função, ou seja, uma proteína só poderia desempenhar sua função se apresentasse uma estrutura tridimensional estável e bem definida em condições fisiológicas. Logo, diante deste conceito, evidências experimentais, como por exemplo a ligação entre uma enzima e seu substrato, que é baseado na hipótese de “chave-fechadura” formulada por (Fischer 1894), e com o sucesso das técnicas experimentais que possibilitaram a resolução e depósito no *Protein Data Bank* (PDB) de inúmeras estruturas proteicas tridimensionais, confirmaram e deram suporte a esta visão (Bernstein et al. 1977; Torrieri 2010; Habchi et al. 2014).

Contudo, a partir da década de 50, devido alguns trabalhos, tais como a descoberta da plasticidade da Albumina bovina por (Karush 1950) e do modelo “ajuste induzido”, proposto por (Koshland 1958), o qual afirma que uma enzima pode atuar em substratos de diferentes formatos, esta visão começou a ser questionada (Dunker et al. 2001; Habchi et al. 2014).

Assim, estudos posteriores demonstram a existência de proteínas que não apresentavam elementos estruturais secundários e terciários estáveis e únicos em condições fisiológicas ou livres em solução, mas que podiam desempenhar função biológica (Wright e Dyson 1999; Dunker et al. 2001, 2013; Dyson e Wright 2005; Russell e Gibson 2008; Huang et al. 2013).

Sendo inicialmente observadas de forma independente, estas proteínas ficaram alheias aos olhos de muitos pesquisadores, pois os autores frequentemente as descreviam sob diferentes nomes, como por exemplo “*malleable protein*”, “*flexible*”, “*vulnerable*”, “*mobile*”, “*intrinsically denatured*”, “*natively unfolded protein*”, “*intrinsically unstructured protein*” dentre outros termos ou combinações. Além disso, nenhum destes termos são adequados, pois podem evidenciar características muito específicas ou não transmitem a ideia geral por trás do fenômeno da desordem intrínseca (Dunker et al. 2013; Uversky 2013; Oldfield e Dunker 2014).

Diante deste fato, em 2013 Dunker e colaboradores afirmaram que os termos Intrinsecamente Desordenado (do inglês *Intrinsically Disordered*) são os mais indicados para se descrever a desordem proteica. Para os autores, a escolha do termo “Intrínseco” transmite a ideia de que a cadeia polipeptídica é definida exclusivamente pela sequência de aminoácidos que a compõe e que em condições fisiológicas pode mostrar uma preferência por estados não estruturados, ou seja, refere-se às tendências inatas dessas proteínas. Já a escolha do termo “Desordem”, foi devido ao fato de muitas vezes ser utilizado para comparar os diferentes estados conformacionais que uma proteína pode adotar, portanto descreve um conjunto de proteínas estruturalmente heterogêneas.

Logo, para este estudo, os termos Região(ões) Intrinsecamente Desordenada(s) (do inglês *Intrinsically Disordered Region – IDR*) serão utilizados para caracterizar os segmentos peptídicos que são incapazes de formar estruturas tridimensionais em proteínas que, ao mesmo tempo, apresentam segmentos estruturados (Figura 1.1B). Já os termos Proteína(s) Intrinsecamente Desordenada(s) (do inglês *Intrinsically Disordered Protein – IDP*), serão adotados para caracterizar as proteínas que não formam nenhuma conformação tridimensional estável e bem definida (Figura 1.1A) (van der Lee et al. 2014).

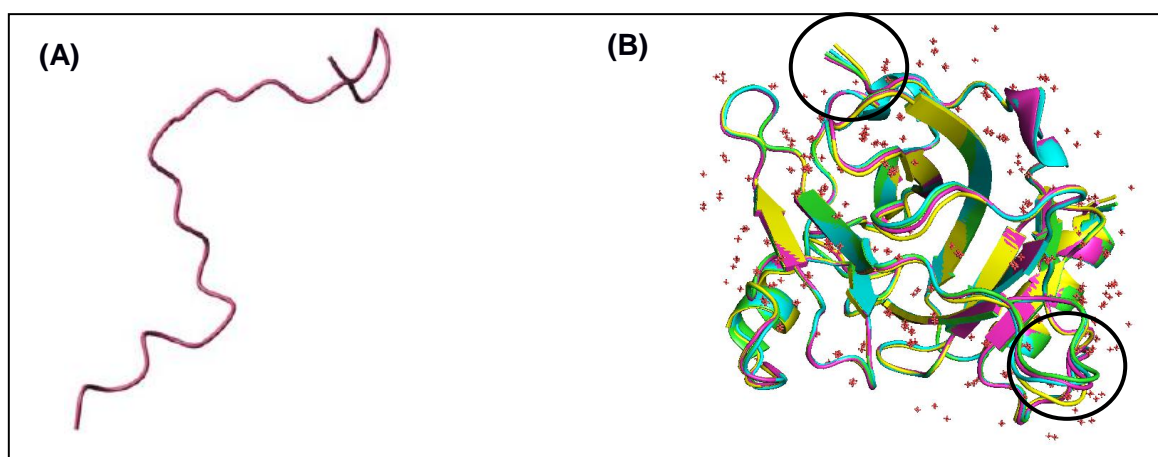


Figura 1.1: Exemplos de IDP e IDR. (A) Proteína conhecida como fator de transcrição, CREB, em seu estado nativo não apresenta nenhuma conformação tridimensional. **(B)** Proteína p53 humana, em seu estado nativo, apresenta tanto regiões ordenadas, quanto IDRs (circuladas) ao longo de sua estrutura. **Fonte: (A)** Dyson & Wright (2005) - adaptada e **(B)** Acervo pessoal.

1.2 Características gerais da desordem estrutural

1.2.1 Composição de aminoácidos

As IDPs e IDRs possuem várias características primordiais em comum que as identificam e o primeiro nível da caracterização estrutural está relacionado à estrutura primária, ou seja, à sequência de aminoácidos que formam as proteínas. Essas proteínas são enriquecidas com resíduos que apresentam elevados índices de flexibilidade, polares e carregados, presença de regiões repetidas, além de pouco ou nenhum conteúdo de aminoácidos hidrofóbicos (Tompa 2010).

Apesar disso, o conteúdo de resíduos encontrado em IDRs e IDPs varia na literatura de acordo com os critérios metodológicos utilizados por cada estudo. Assim, os aminoácidos Glutamina (Gln), Serina (Ser), Prolina (Pro), Ácido Glutâmico (Glu) e Lisina (Lys) são comumente encontrados nas sequências de IDPs e IDRs. Já os aminoácidos Triptofano (Trp), Cisteína (Cys), Fenilalanina (Phe), Isoleucina (Ile), Tirosina (Tyr), Valina (Val) e Leucina (Leu) são normalmente encontrados em proteínas globulares, sendo, portanto, relacionados à estabilidade estrutural da proteína (Dunker et al. 2001; Dyson e Wright 2005; Tompa 2010).

Em contrapartida, em relação aos resíduos Histidina (His), Treonina (Thr), Metionina (Met) e Ácido Aspártico (Asp) não há um consenso quanto à influência deles na estabilidade da molécula, pois podem ser encontrados em sequências que podem apresentar desordem estrutural ou não. Dunker e colaboradores (2001), por exemplo, classificaram estes quatro aminoácidos como neutros (não influenciam a conformação da molécula), enquanto que Linding e colaboradores. (2003a) classificaram a Met e a Thr como promotores de desordem, Dunker e colaboradores (2008) classificaram a His e a Thr como promotores de ordem e a Met e o Asp como aminoácidos neutros.

Portanto, pode-se perceber que a frequência de aminoácidos em IDPs e IDRs é significativamente diferente quando comparada com proteínas globulares, pois a composição e as características físico-químicas dos aminoácidos promotores de desordem impedem a formação de um núcleo hidrofóbico e aumentam a instabilidade estrutural da molécula (Dunker et al. 2001; Hansen et al. 2006; Tompa 2010).

1.2.2 Interações proteicas

Para que uma proteína globular consiga exercer determinada função, é necessário que ela consiga interagir com seu alvo e, portanto, o conhecimento da sua estrutura tridimensional é de fundamental importância para entender sua função. Entretanto, as proteínas são moléculas dinâmicas que podem mudar sua conformação tridimensional ao se ligarem aos seus alvos proteicos para se adaptarem ao complexo em formação.

Partindo deste princípio, estudos demonstraram que assim como as proteínas globulares, as IDPs e/ou IDRs sofrem uma mudança conformacional ao se ligarem a outras moléculas (Figura 1.2), pois a interação provoca uma alteração química que leva a uma menor densidade de carga no complexo e, assim, permite a formação de domínios estruturados. Esta característica promove um aumento significativo na capacidade da mesma proteína interagir com múltiplos alvos proteicos e de desempenhar funções distintas, sendo, portanto classificadas como proteínas centrais (*hub*) em redes de interação proteína-proteína (Wright e Dyson 1999; Dyson e Wright 2005; Hansen et al. 2006; Russell e Gibson 2008; Babu et al. 2011).

Adicionalmente, esta flexibilidade conformacional também ressalta outras características, tais como: a) uma variação de entropia (ΔS) e entalpia (ΔH) negativas; b) energia de ligação mais favorável devido à área de contato de uma IDP e/ou IDR ser maior do que de uma proteína tridimensionalmente estruturada; c) facilita diversas modificações pós-traducionais, o que permite a regulação de sua função e estabilidade no meio celular e d) propicia uma rápida associação para iniciar os eventos da cascata de sinalização celular (alta especificidade) e uma fácil dissociação ao término deste processo (baixa afinidade) (Wright e Dyson 1999; Dyson e Wright 2005; Hansen et al. 2006; Russell e Gibson 2008; Babu et al. 2011).

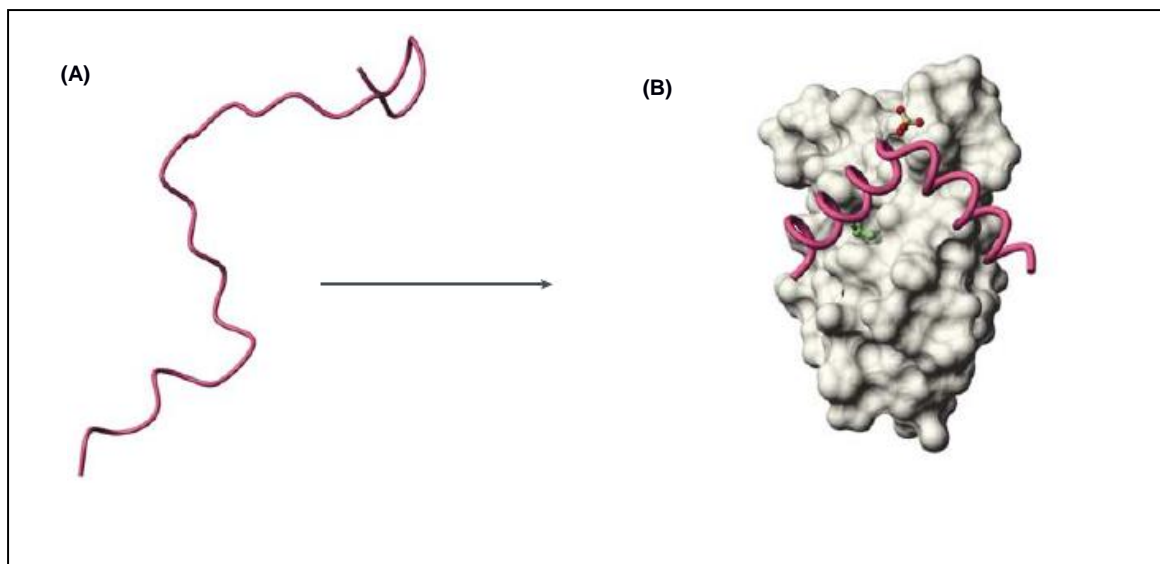


Figura 1.2: Exemplo de uma proteína totalmente desordenada, em seu estado nativo, adquirindo uma conformação tridimensional ao se ligar a seu alvo proteico. (A) Proteína conhecida como fator de transcrição, CREB, em seu estado nativo não apresenta nenhuma conformação tridimensional. **(B)** A proteína de Fator de Transcrição CREB adquire uma conformação tridimensional quando se liga ao seu alvo proteico. **Fonte:** Dyson & Wright (2005) - adaptada

1.2.3 Papel biológico

Assim como as IDPs e as IDRs estão relacionadas a diversas funções importantes para um funcionamento celular adequado, tais como mecanismos de controle e sinalização celular. Estas proteínas também podem estar associadas a várias doenças em seres humanos, seja pela alteração de sua disponibilidade ou de algum processo biológico através de interações proteína-proteína, seja pela interação parasito-hospedeiro (Torrieri 2010; Babu et al. 2011).

No primeiro caso, níveis inadequados de concentração de IDPs ou IDRs estão intimamente ligados a diversas patologias. Como exemplo temos vários tipos de cânceres e doenças neurodegenerativas que são mediados por alterações de expressão das proteínas estruturalmente desordenadas (Babu et al. 2011; Uversky 2013).

Outros estudos afirmam ainda, que a superexpressão ou mutações das IDPs tau, ataxina-1, α -sinucleína e huntingtina, aumentam o risco de formação de agregados fibrilares além de ocasionarem várias outras patologias humanas (Dunker et al. 2008a; Babu et al. 2011).

Entretanto, na relação parasito-hospedeiro, a plasticidade das IDPs e IDRs pode favorecer, por exemplo, a adesão, invasão e sobrevivência de parasitos no hospedeiro. O parasita *Plasmodium falciparum*, por exemplo, aproveita desta

vantagem para garantir sua sobrevivência dentro do organismo parasitado através da inibição de respostas efetivas de anticorpos com alta afinidade ou por facilitar a interação com moléculas do hospedeiro necessárias à ligação e invasão celular (Feng et al. 2006; Torrieri 2010).

1.2.4 Desenvolvimento de novas drogas

Segundo (Uversky et al. 2008), as IDPs e IDRs não só podem estar relacionadas às patogenias citadas acima, mas também podem auxiliar no tratamento de doenças, através do desenvolvimento de novos fármacos.

Um exemplo bem descrito está relacionado à trombina. Este polipeptídeo contém IDRs quando em seu estado ativo e é responsável pela conversão do fibrinogênio em fibrina no evento final da cascata de coagulação, além de influenciar na regulação de todo o processo de sinalização celular da via. Desta forma, a formação de trombos oclusivos está estritamente relacionada a essa proteína e devido a essas características, acredita-se que a inibição de sua atividade ou de seu receptor deva interromper a formação de trombos, evento que diminuiria os riscos de doenças cardiovasculares (Blaya et al. 1998).

Inibidores específicos para trombina são, portanto, alvos de vários estudos. Em especial a hirudina, um peptídeo com ação anticoagulante, encontrado naturalmente em glândulas salivares de sanguessugas da espécie *Hirudo medicinalis*, tem sido bastante utilizada em tratamentos clínicos como antitrombóticos. Evidências experimentais revelaram que a região C-terminal desta proteína é altamente desordenada, mas é capaz de adquirir uma conformação parcial ao se ligar à cadeia pesada da trombina. Esta interação é capaz de inibir a atividade da trombina e conseqüentemente diminui o risco de formação de trombos no organismo (Blaya et al. 1998; Uversky et al. 2008).

1.3 Predição computacional de IDPs e IDRs

1.3.1 Características gerais

Com o elevado aumento do número de estudos relacionados à identificação e caracterização funcional de IDPs e IDRs, houve simultaneamente o desenvolvimento

de inúmeros programas para predição dessas proteínas, fato que possibilitou um grande avanço para o campo (Tompa 2012).

Há apenas dez anos, somente o preditor PONDR[®] estava disponível, contudo, hoje há cerca de 50 programas disponíveis baseados em vários princípios diferentes de predição, como demonstrado por (He et al. 2009) e colaboradores (2009). Estes autores destacam vários programas que são capazes de identificar longos trechos de sequência de aminoácidos desordenados (maior que 40 resíduos consecutivos), como, por exemplo, o PONDR[®], FoldIndex, GlobPlot, Dis-EMBL, DISOPRED e DISOPRED2, DRIPPRED, IUPred, FoldUnfold, RONN, DISpro, DisPSSMP e DisPSSMP2, Spritz, PrDOS, dentre tantos outros. Entretanto, estes preditores diferem enormemente na abordagem utilizada, além do formato de apresentação e detalhamento dos resultados, o que limita a previsão de potenciais IDPs e IDRs por uma única abordagem (Tompa 2012).

Assim, diante deste fato, vários programas foram recentemente incrementados, passando a ter características de meta-preditor, como por exemplo, o PONDR-FIT. Além disso, de acordo com a avaliação experimental realizada pela *Critical Assessment of Structure Prediction (CASP)*, a precisão dos preditores atuais deve ser de aproximadamente 85%. Entretanto, mesmo com estas melhorias e as diretrizes avaliadas pela CASP, ainda não há um consenso descrito na literatura de uma metodologia computacional para identificação de potenciais IDPs e IDRs (Tompa 2012).

Outro programa bastante utilizado e que merece ser mencionado em termos do número de citações é o DISOPRED2. Este preditor realiza a identificação de IDPs utilizando a informação de alinhamentos múltiplos e seus dados de entrada são baseados em perfis de sequência gerados pelo PSI-BLAST (Altschul et al., 1990). A matriz, gerada por estes alinhamentos em última instância, quantifica o grau de similaridade em cada posição do alinhamento múltiplo viabilizando a identificação das regiões de alta variabilidade de sequência que auxiliam a predição da região desordenada (He et al. 2009).

Além dos programas mencionados, há preditores de desordem proteica que avaliam, características bioquímicas específicas, a probabilidade de uma proteína ser inteiramente estruturada ou apresentar regiões de desordem estrutural, enquanto outros são capazes de filtrar e identificar as regiões de baixa complexidade dentro da sequência e indicar regiões sem conformação estrutural

definida (He et al. 2009). As características utilizadas por cada preditor utilizado neste estudo será detalhado mais adiante.

1.4 Organismo modelo: *Cryptococcus* spp.

1.4.1 *Cryptococcus* spp e epidemiologia.

Os primeiros representantes do gênero *Cryptococcus* foram isolados em 1894 na Itália por Sanfelice e identificados como *Saccharomyces neoformans*. Entretanto, em 1901 Jean-Paul Vuillemin identificou que estes organismos não eram capazes de produzir ascósporos, característica que define o gênero *Saccharomyces*. Por esse motivo, o organismo foi renomeado como *Cryptococcus neoformans* (Mitchell e Perfect 1995; Srikanta et al. 2014).

Atualmente são conhecidas mais de 37 espécies de *Cryptococcus* distribuídas em todo o mundo. Contudo duas espécies, *C. neoformans* e *C. gattii*, que são comumente encontradas em fezes de pássaros e em árvores de eucalipto, respectivamente, apresentam grande importância médica, pois são responsáveis pelo alto índice de infecção e mortalidade entre os pacientes com imunidade comprometida (Li e Mody 2010).

Adicionalmente, apesar de ambas as espécies serem classificadas inicialmente como *C. neoformans* e *C. neoformans* var. *gattii*, a atual classificação indica que a variedade *gattii* é uma espécie distinta, sendo conhecida atualmente como *C. gattii*, enquanto que *C. neoformans* apresenta duas novas variedades, conhecidas como *C. neoformans* var. *grubii* e *C. neoformans* var. *neoformans*. Além disso, baseando-se nas reações de aglutinação capsular e em características moleculares, estas espécies estão divididas entre cinco sorotipos e nove genótipos, respectivamente (Figura 1.3) (Lin e Heitman 2006; Cogliati 2013).

O *C. neoformans* é classificado como um patógeno predominantemente oportunista, pois acomete principalmente pacientes imunocomprometidos. Já o *C. gattii* é classificado como um patógeno primário, uma vez que além de infectar pacientes imunocomprometidos, também acomete uma porcentagem estatisticamente significativa de indivíduos imunocompetentes, quando comparado às infecções causadas por *C. neoformans* (Lin e Heitman 2006; Ma e May 2009;

Santos et al. 2012; Chaturvedi e Wormley 2013; Chen et al. 2013; Cogliati 2013; Santos 2015).

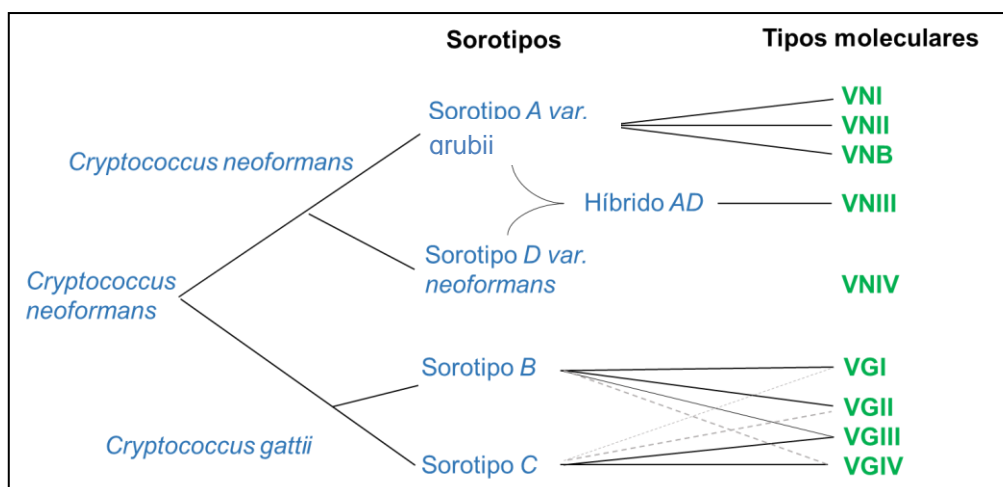


Figura 1.3: Classificação das espécies do complexo *C. neoformans*. A figura mostra a divergência das espécies do complexo *C. neoformans* e suas variedades, em relação ao seu ancestral comum, além da atual classificação destas espécies. **Fonte:** Lin & Heitman, 2006 – adaptado.

Esses patógenos são leveduras haploides que em seu estado leveduriforme corresponde à forma assexuada, com a sua reprodução ocorrendo por brotamento. Já na forma telemórfica, também conhecida como forma sexuada, a reprodução é regulada por um sistema de acasalamento bipolar denominado *mating-type* (MAT α e a), em que ocorre a produção de basidiósporos e, ocasionalmente, pseudo-hifas. Entretanto, em sua forma telemórfica, a espécie *C. neoformans* é denominada como *Filobasidiella neoformans* e a espécie *C. gattii* como *F. bacillispora* (Mitchell e Perfect 1995; Silva 2006; Ma e May 2009; Santos 2015).

Considerando a ocorrência de *Cryptococcus* spp em âmbito mundial, Cogliati (2013) verificou em sua revisão que as espécies *C. neoformans* e *C. gattii* são encontradas em vários países da Oceania, Ásia, África, Europa e nas Américas. O autor ainda afirma que 53% das 10548 amostras reportadas na literatura e isoladas nas Américas Central e do Sul foram encontradas no Brasil. Além disso, todos os tipos moleculares de ambas as espécies, exceto o VGIV (*C. gattii*), já foram encontrados no país, sendo portanto, a maioria dos isolados pertencentes aos tipos moleculares VNI, seguido por VGII, enquanto que VNII, VNIV, VGI, VGIII e o híbrido VNIII são encontrados em menor quantidade (Cogliati 2013).

Entretanto, levando-se em consideração o território brasileiro, um levantamento epidemiológico e ecológico dos sorotipos de *Cryptococcus* spp. foi realizado por Nishikawa e colaboradores em 2003. Nesse estudo, os autores

demonstraram que o sorotipo A, *C. neoformans* var. *neoformans*, prevalece em isolados clínicos e ambientais em três (Sul, Sudeste e Centro-Oeste), das cinco regiões de estudo (Norte, Sul, Nordeste, Sudeste e Centro-Oeste), enquanto que o sorotipo B, *C. gattii*, predomina nas outras duas regiões (Norte e Nordeste). Já os demais sorotipos, C, AD e D, também foram encontrados em algumas regiões do estudo, mas em menor quantidade, sendo: sorotipo C encontrado em uma amostra clínica da região Sudeste, o sorotipo D encontrado em uma amostra ambiental e uma clínica nas regiões Sudeste e Sul e por fim, o sorotipo AD foi encontrado em cinco amostras clínicas e uma ambiental, nas regiões Nordeste, Sudeste e Sul. Esta distribuição epidemiológica também está apresentada na Figura 1.4.

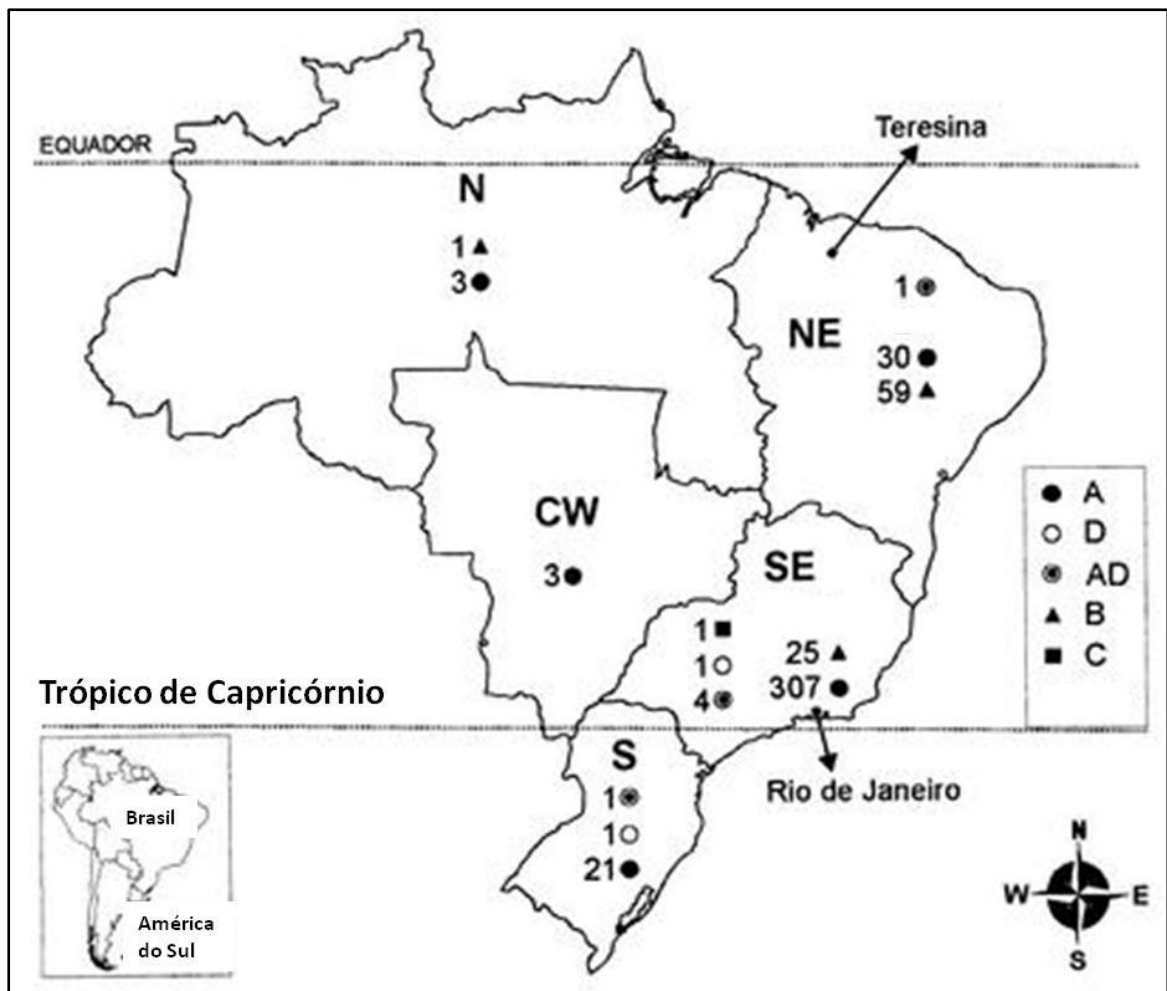


Figura 1.4: Distribuição dos sorotipos (A, B, C, D e AD) de *Cryptococcus* spp. pelas regiões geográficas do Brasil. Fonte: Nishikawa et al 2003 – adaptado.

1.4.2 Criptococose

Adquirida pela inalação de basidiósporos dessecados das espécies *C. neoformans* e *C. gattii*, a criptococose é uma micose sistêmica, predominantemente oportunista que pode provocar, em primeiro estágio, doenças respiratórias agudas ou crônicas. No entanto, com a disseminação sistêmica do fungo (segundo estágio), vários órgãos e tecidos podem ser afetados. O elevado tropismo pelo Sistema Nervoso Central (SNC) apresentado pela levedura também é uma característica que deve ser levada em consideração, uma vez que o quadro clínico do paciente pode evoluir para encefalite ou meningite criptocócica, o que pode levar o indivíduo ao óbito (Gullo et al. 2013).

Até o momento, existem três hipóteses que podem explicar o neurotropismo do fungo. A primeira proposta relata que os substratos neuronais presentes nos gânglios basais promovem o crescimento e a sobrevivência da levedura e desta forma, os espaços perivasculares podem servir como um nicho para *Cryptococcus* (Franco-Paredes et al. 2015) A segunda descreve que a existência de receptores neuronais específicos podem atrair *Cryptococcus* para o SNC (Lin e Heitman 2006). Por fim, a terceira hipótese, uma das mais difundidas, é que o fungo utiliza neurotransmissores como a dopamina que ajuda na síntese de melanina e por isso é atraído para o SNC (Lin e Heitman 2006; Byrnes et al. 2011).

Assim, considerando o processo de migração da levedura para o SNC, este pode acontecer por diferentes vias de disseminação hematogênica. Um destes mecanismos é conhecido como “Cavalo de Tróia”, que como o próprio nome sugere, ocorre quando células efectoras fagocíticas contendo *Cryptococcus* em seu interior, cruzam o endotélio dos vasos sanguíneos do cérebro. Outro mecanismo conhecido como “transferência lateral” refere-se à capacidade do *Cryptococcus* infectar outro macrófago sem que o fungo seja exposto ao ambiente extracelular. Por outro lado, a levedura pode cruzar a barreira hematoencefálica por meio da passagem direta entre as células endoteliais dos vasos sanguíneos, processo conhecido como “transcitose”. Durante este último processo, o fungo adquire uma morfologia diferenciada com formato ovóide e um broto na superfície celular, sugerindo que uma mudança morfológica possa ocorrer como parte desse processo (Figura 1.5) (Ma e May, 2009; Santos, 2015).

Entretanto, a evolução da doença e o quadro clínico apresentado pelo paciente estão intimamente relacionados à condição imunológica do indivíduo e à

virulência do fungo, uma vez que as espécies de *Cryptococcus* spp. além de serem geneticamente distintas, também diferem quanto à resposta ao tratamento antifúngico. Assim, estima-se que um milhão de casos de meningite criptocócica por ano esta associado aos pacientes portadores do Vírus da Imunodeficiência Humana (HIV), resultando em aproximadamente 625 mil mortes (Santos et al., 2012; Cogliati, 2013; Gullo et al., 2013; Santos, 2015).

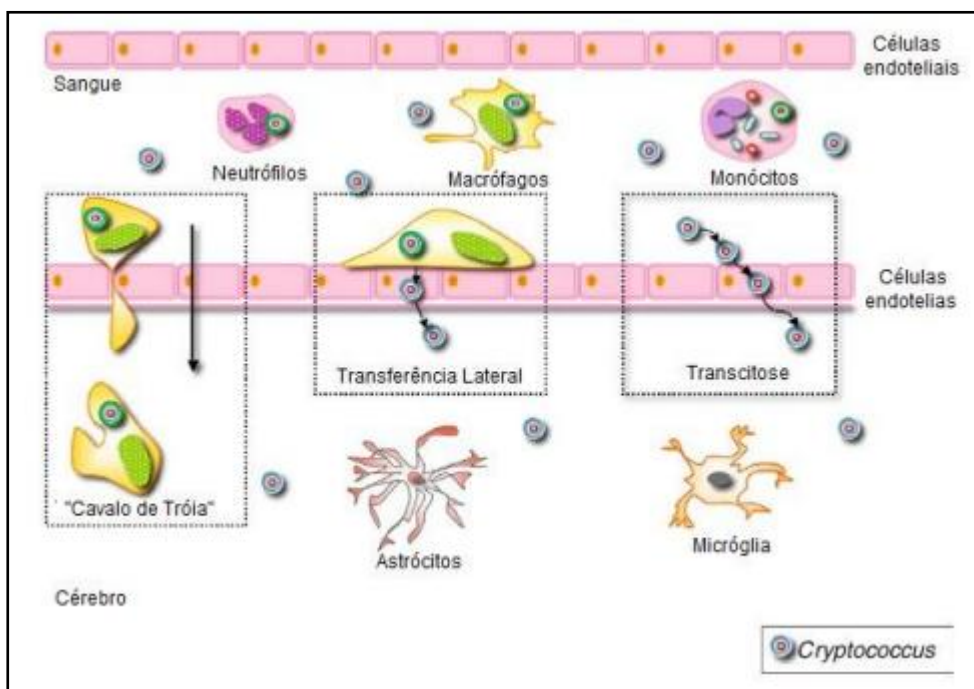


Figura 1.5: Esquematização dos processos de migração do *Cryptococcus* ssp. para o SNC, onde: (A) representa o processo conhecido como “Cavalo de Tróia”, (B) representa o processo conhecido como “Transferência Lateral” e (C) representa o processo de “Transcitose”. Fonte: Ma e May, 2009 – adaptado.

Embora nos últimos anos a distribuição em larga escala de antirretrovirais para os indivíduos soropositivos tenha influenciado a diminuição da incidência de criptococose nos países desenvolvidos, a doença ainda é um grande problema de saúde pública nos países em desenvolvimento, onde o acesso aos cuidados de saúde é limitado. Na África Sub-saariana, por exemplo, estima-se que a mortalidade por meningite criptocócica, entre os pacientes portadores do HIV, esteja entre 50 e 70% dos casos (CDC, 2014).

1.4.3 Fatores de virulência

A capacidade de sobrevivência e crescimento a temperaturas elevadas (37-40 °C), como em um hospedeiro mamífero, é uma característica que poucos fungos

possuem, uma vez que a maioria cresce a uma temperatura de 25-30 °C (Li e Mody 2010).

As espécies *C. neoformans* e *C. gattii* são fungos termotolerantes, pois apresentam características que proporcionam proteção ao fungo contra altas temperaturas, além da capacidade de invasão e sobrevivência no hospedeiro mamífero, permitindo o desenvolvimento de uma micose sistêmica (Li e Mody 2010).

Odom e colaboradores (1997) e Chen e colaboradores (2013), por exemplo, demonstraram que a produção de Calcinerina-A é essencial para o desenvolvimento de *C. neoformans* e *C. gattii* a temperaturas de 37°C. Estes estudos relatam que a inibição do gene CNA1, gene que codifica a Calcinerina-A, resulta em mutantes que não são capazes de crescer a temperatura corporal de 37°C.

Além disso, outros estudos demonstraram que a produção da cápsula polissacarídica é o principal fator de virulência de *Cryptococcus* (Srikanta et al. 2014). A cápsula está diretamente associada à parede celular da levedura, envolvendo-a completamente, e é basicamente formada por glucuronoxilomanana (GXM), galactoxilomanana (GalXM) e manoproteínas (Almeida et al. 2015).

Além da proteção contra o sistema imune do hospedeiro e ações oxidativas, esta estrutura também protege o organismo contra condições não favoráveis à sua sobrevivência. Tais condições, como por exemplo, alterações nas concentrações de gás carbônico e ferro, pH básico e presença de soro de mamífero, podem induzir a formação da cápsula e/ou o aumento de sua espessura (Mitchell e Perfect 1995; Ma e May 2009).

Ao ser fagocitado, o patógeno ativa uma resposta imunológica do hospedeiro. Entretanto, a interação do fungo com os fagócitos também pode aumentar a replicação intracelular do *Cryptococcus* spp. e lise celular. Essa resposta ao sistema imune envolve o acúmulo citoplasmático de vesículas contendo polissacarídeos capsulares e, posteriormente, o fungo libera essas vesículas ao redor dos fagossomos que são acumulados no citoplasma, promovendo a disfunção e lise dos macrófagos. O acúmulo de leveduras no interior dos macrófagos promove a ruptura das células do hospedeiro ou extrusão do fagossoma (Alvarez e Casadevall 2006; Santos 2015).

Segundo a revisão realizada por Almeida e colaboradores (2015), mutações nos genes CAP10, CAP59, CAP60, CAP64, CAS1 ou CAS2 conferem uma

diminuição na taxa de crescimento do mutante e na taxa de infecção em modelos murinos, uma vez que estes genes estão envolvidos na síntese do material capsular.

Outra estrutura, capaz de proteger a levedura contra a ação do sistema imune do hospedeiro é a melanina. Entretanto, diferentemente da cápsula, a melanina protege a célula contra agentes antifúngicos, como a Anfotericina B e o Fluconazol. Essa estrutura é formada a partir da polimerização oxidativa de catecolaminas, que por sua vez são sintetizadas através da ação catalítica de enzimas do grupo das fenoloxidasas, como a lacase (Almeida et al. 2015; Santos 2015).

Os genes LAC1 e LAC2 são responsáveis por codificar duas lacases envolvidas no processo de síntese da melanina e a simples deleção do gene LAC1 impede essa via e a síntese de melanina (Almeida et al. 2015; Santos 2015).

Por fim, enzimas extracelulares liberadas durante o processo de infecção também contribuem para a virulência do *Cryptococcus*, destruindo os tecidos, promovendo a sobrevivência do fungo e interferindo nas respostas imunes do hospedeiro (Almeida et al. 2015).

A urease, por exemplo, é comumente expressa em isolados de *C. neoformans*, pois é uma enzima capaz de catalisar a hidrólise da ureia em amoníaco e carbamato. Esta conversão é essencial para a sobrevivência da levedura em fezes de pássaros. Além disso, a urease desempenha um papel importante na disseminação da levedura para o SNC, pois promove o rompimento da barreira hematoencefálica (Almeida et al. 2015).

1.4.4 Resposta imune do hospedeiro

Ao ser inalado, o fungo ativa o sistema imune do hospedeiro, sendo as primeiras células envolvidas nesse processo os macrófagos alveolares. Estes, por sua vez, ativam outras células de defesa por meio da ação de citocinas e quimiocinas (Santos 2015).

A fagocitose é desencadeada pelo reconhecimento direto da levedura ou mediada por um receptor ativado pela via complemento ou pela síntese de anticorpos. Estruturas conservadas tais como os componentes da cápsula polissacarídica, principalmente o GXM, podem ser reconhecidas por receptores do tipo *Toll* (TLR-4), já as manoproteínas, que são expressas na superfície do fungo,

são reconhecidas pelos receptores de manose presentes nas células dendríticas (Voelz e May 2010; Santos 2015).

Ao invadir o SNC, as primeiras células a reconhecerem o patógeno são as micróglias, fagócitos cerebrais. Estas células respondem à infecção com a ativação do fator de transcrição NF- κ B seguido pela liberação de metaloproteínas de matriz, espécies reativas de oxigênio e citocinas pró-inflamatórias, tais como interleucina 1 α (IL1 α), fator de necrose tumoral α (TNF α) e proteína inflamatória de macrófagos 2 α (MIP-2 α) e Interferon γ (IFN- γ) (Orsi et al. 2009; Mendes et al. 2010; Santos 2015).

Além disso, Schoffelen e colaboradores (2013) realizaram um estudo comparativo entre *C. gattii* e dois sorotipos de *C. neoformans* quanto ao padrão de citocinas e a importância de alguns receptores do tipo *Toll*. Nesse estudo foi observado que *C. gattii* induz uma produção mais acentuada de citocinas pró-inflamatórias (IL-1 β , IL-6, IL-17, IL-22 e TNF- α) em comparação a *C. neoformans*. Adicionalmente, os autores também observaram que os receptores TLR-4 e TLR-9 participam da resposta do hospedeiro contra *C. gattii*.

A resposta imune exercida pelo organismo do hospedeiro frente à infecção por *Cryptococcus* também é dependente de outros mediadores, como por exemplo a produção de óxido nítrico e de espécies reativas de oxigênio por Leucócitos. Essa resposta representa um papel importante na defesa do hospedeiro a fim de se evitar a disseminação do patógeno pelo organismo (Fang 2004; Santos 2015)

2 JUSTIFICATIVA

Proteínas Intrinsecamente Desordenadas (IDPs) e Regiões Intrinsecamente Desordenadas (IDRs) representam uma classe de proteínas sem estrutura tridimensional bem definida em seu estado nativo, mas são funcionalmente ativas e tem ganhado grande destaque graças às suas características peculiares e ao fato de estarem relacionadas à várias doenças e à importantes funções celulares.

Um ponto que também merece destaque está relatado nos trabalhos de (Russell e Gibson 2008) e de (Cheng et al. 2006) e colaboradores (2006), os quais demonstraram que a presença de desordem estrutural em uma proteína pode ser caracterizado como um facilitador para o desenho de drogas devido à alta especificidade, uma vez que sua conformação se “adapta” ao seu alvo, e à baixa afinidade de interação com seu alvo, que são quesitos primordiais para atuação de um fármaco.

Contudo, a parasitologia, biologia molecular e computacional ainda possuem um grande desafio no campo da relação parasito/hospedeiro, no qual a análise da rede regulatória de eventos associados à essa interação, ainda possui muitas perguntas que somente poderão ser respondidas com a utilização de abordagens integradas de estudo. Assim um estudo pioneiro que avalia esta relação, realizado por Feng e colaboradores (2006), demonstrou a existência de IDPs e IDRs, em *P. falciparum* que estão associados à adesão e invasão celular do parasito em seu hospedeiro, o que possibilitou uma melhor compressão de alguns processos associados.

Diante deste fato, verificou-se a importância de um estudo mais profundo dessa intrigante classe de proteínas, que pode levar a um melhor entendimento dos mecanismos de interação parasito/hospedeiro. Além disso, a inexistência de estudos relatando a identificação e/ou caracterização funcional de IDPs e IDRs em *Cryptococcus* spp., relacionando-as também aos fatores de virulência em linhagens patogênicas e não patogênicas ao Homem, proporciona uma nova perspectiva de estudo e linha de atuação para o controle e/ou tratamento da criptococose.

3 OBJETIVOS

3.1 Objetivo Geral

Identificar o conteúdo de Proteínas Intrinsecamente Desordenadas (IDPs) e Regiões Intrinsecamente Desordenadas (IDRs) nos proteomas preditos de *Cryptococcus* spp. visando ao entendimento do papel biológico funcional dessa classe de proteínas.

3.2 Objetivos Específicos

- ✓ Reanotar automaticamente os genomas disponíveis visando à normalização da análise genômica comparativa;
- ✓ avaliar o estado da arte dos preditores computacionais de desordem estrutural, as estratégias e metodologias implementadas e a adequabilidade em termos de aplicação ao organismo de estudo;
- ✓ realizar a predição computacional e a anotação funcional das proteínas preditas contendo domínios não estruturados em *Cryptococcus* spp.;
- ✓ avaliar e comparar o conteúdo de IDPs e IDRs aos fatores de virulência entre as linhagens virulentas.

4 MATERIAIS E MÉTODOS

4.1 Scripts desenvolvidos

Todos os *scripts* desenvolvidos neste trabalho foram escritos em linguagem de programação *PERL* e/ou em *shell script*.

4.2 Genomas de *Cryptococcus* spp.

Os genomas de *Cryptococcus* spp. que foram utilizados no presente estudo foram obtidos a partir do banco de dados do *National Center for Biotechnology Information* (NCBI), o GenBank (em: <https://www.ncbi.nlm.nih.gov/>), e do *Broad Institute of MIT and Harvard* (em: <http://www.broadinstitute.org/>) em Fevereiro de 2014.

4.3 Anotação e Reanotação dos proteomas preditos de *Cryptococcus* spp.

As etapas analíticas que envolvem os processos de anotação e reanotação genômica utilizados neste trabalho estão descritas abaixo e também estão resumidas na Figura 4.1.

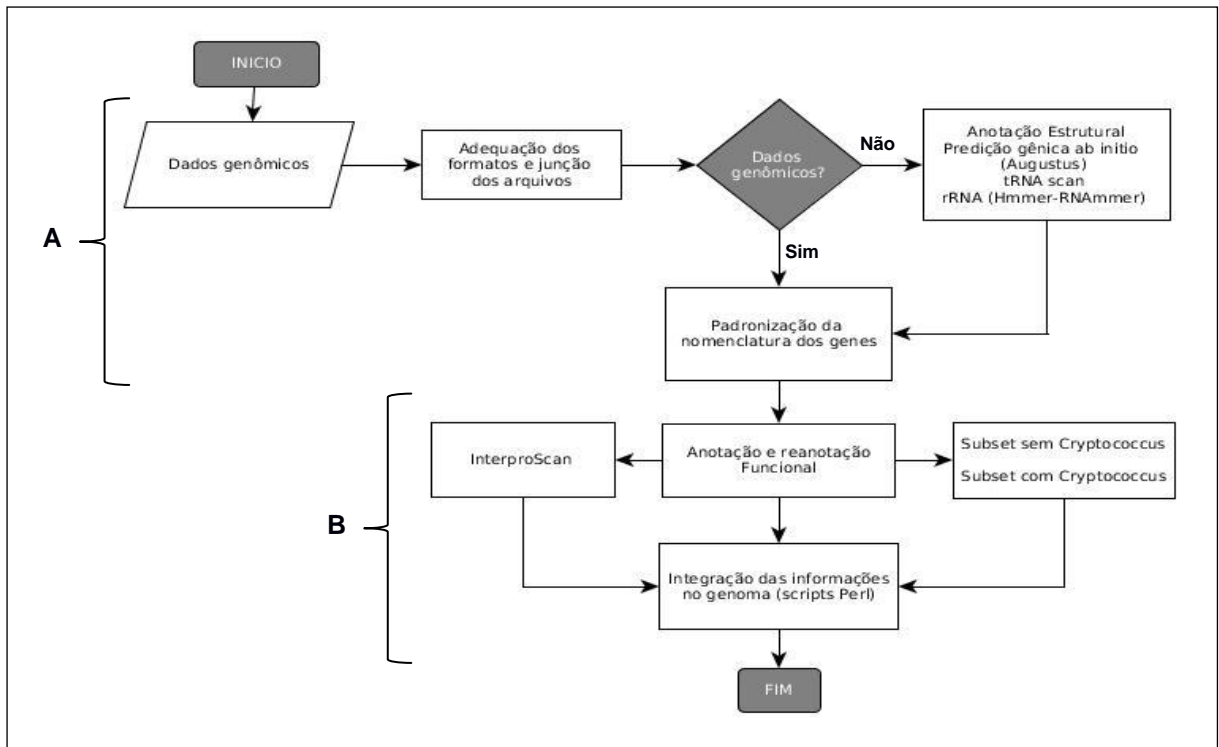


Figura 4.1: Fluxograma esquematizando as etapas analíticas do processo de anotação e reanotação genômica utilizado neste estudo, onde: **(A):** após a obtenção dos dados genômicos, foi realizada uma padronização das informações, posteriormente, foi realizada uma predição gênica: *ab initio* para os dados genômicos que não apresentavam informações gênicas e, por fim, a nomenclatura dos genes também foi padronizada; **(B):** Anotação e reanotação funcional, utilizando buscas por similaridade de sequências entre dois subset distintos (Com *Cryptococcus* e sem *Cryptococcus*, além da realizada pelo programa InterproScan. Após estas etapas, todas as informações foram integradas ao genoma, utilizando scripts em linguagem de programação PERL. **Fonte:** Acervo Pessoal.

4.3.1 Artemis

Como ferramenta de visualização de dados genômicos, foi utilizado o Artemis, na versão 16.0 (<http://www.sanger.ac.uk/Software/Artemis>), que é uma ferramenta escrita em JAVA, capaz de disponibilizar os resultados das análises de predição gênica e funcional em sua própria interface e em seus respectivos quadros de leitura (ORFs) (Rutherford et al. 2000).

4.3.2 Normalização e reanotação genômica

A avaliação dos dados genômicos disponíveis revelou duas situações distintas relacionadas ao status da anotação dos genomas selecionados. A primeira envolve a inexistência de qualquer informação relacionada aos genes preditos (ausência de predição gênica). A segunda situação envolve a existência da informação relacionada aos genes preditos, mas apresenta-se com versões de anotação funcional defasada. Conseqüentemente, estratégias distintas foram

empregadas para cada um desses casos. Essas estratégias serão descritas a seguir.

4.3.3 Dados genômicos sem predição gênica

De forma sucinta, para a predição *ab initio* foi utilizado o preditor gênico Augustus na versão 3.0.1 (<http://augustus.gobics.de/>). Adicionalmente, o Augustus já é treinado para os seguintes genomas de *Cryptococcus*: *C. gattii* e *C. neoformans* (Stanke e Morgenstern 2005).

Após essa etapa, foi realizada a padronização da nomenclatura dos genes e ajustes dos códons de início (*start codon*) e de parada (*stop codon*). Para a nomenclatura, adicionou-se a cada sequência gênica predita um nome único que foi associado ao termo anotador *systematic_id*. Esse identificador contém as iniciais do respectivo genoma, acrescido de seis dígitos numéricos, como por exemplo para o genoma de *C. bestiolae*, à primeira sequência gênica identificada foi atribuído como identificador cbest.000010, à segunda sequência o identificador cbest.000020 e assim por diante.

Para a predição de RNA transportador e RNA ribossomal, foram utilizados os softwares tRNAscan-SE versão 1.21 (<http://lowelab.ucsc.edu/tRNAscan-SE/>) e RNAmmer versão 1.2 (http://www.cbs.dtu.dk/cgi-bin/nph-sw_request?rnammer), respectivamente.

Além das etapas descritas, foram realizadas buscas por similaridade de sequências contra bancos de dados primários e secundários de domínio público, como por exemplo, o banco de dados (BD) não redundante do NCBI (NR) (<ftp://ftp.ncbi.nlm.nih.gov/blast/db/>), o BD de Domínios Conservados (CDD) (<http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>), o BD de Famílias Proteicas (Pfam) (<http://pfam.xfam.org/>), dentre outros. Para tanto, foram utilizados os algoritmos BLAST (*Basic Alignment Search Tool*) versão 2.2.26 e o InterProScan versão 5.8 (<https://code.google.com/p/interproscan/>) (Altschul et al. 1990; Jones et al. 2014).

Em especial, para as comparações feitas contra o NR optamos pela divisão do referido banco em dois grupos, o primeiro deles sem sequências de *Cryptococcus* spp. e o segundo grupo contendo somente as informações das espécies do gênero *Cryptococcus* presentes no NR obtido em 22/10/2014. Nas

buscas por similaridade de sequências contra esses bancos, utilizou-se como critério de corte um e-value de 1×10^{-06} .

Para todos esses processos da anotação estrutural e funcional, os resultados obtidos foram integrados ao genoma utilizando os *scripts* *inseri-anota-from-blast_nr_com-Crypto.pl*, *inseri-anota-from-blast_nr_sem-Crypto.pl* e *anota-two-blastsresults.pl* que foram desenvolvidos para as análises, obtendo-se ao final um único arquivo, para cada genoma analisado, contendo todas as informações.

4.3.4 Dados genômicos com predição gênica

Antes do processo de reanotação funcional, a nomenclatura dos genes também foi padronizada seguindo os mesmos critérios adotados na predição *ab initio* dos genomas que não apresentavam dados de predição gênica.

Após essa etapa, foi realizada a reanotação funcional dos genomas. O processo envolveu buscas por similaridade de sequências entre os genes e os bancos de dados proteicos, como descritos anteriormente.

Em seguida, para que as informações já presentes nesses arquivos não fossem perdidas, os termos anotadores presentes foram identificados com a terminação *_obsolete*. Após, as informações obtidas durante a etapa de reanotação foram integradas juntamente com as informações já contidas no respectivo genoma. Utilizou-se os mesmos *scripts* relacionados para a integração dos resultados do BLAST aos genomas sem predição gênica. Ao final, obtivemos um único arquivo, para cada genoma analisado, contendo todas as informações.

A tradução conceitual das sequências codificadoras de cada organismo foi realizada utilizando o programa Artemis. Esse programa permite selecionar todas as sequências e, a partir dessa seleção, é capaz de gerar um arquivo *fasta* contendo todas as proteínas com suas respectivas funções.

Essas informações foram posteriormente formatadas, utilizando o *script* *fasta2tab.pl* e integradas ao BD *idp_crypto* (vide o item 4.4.3).

4.4 Predição de desordem estrutural proteica

As etapas analíticas que envolvem os processos de predição de desordem estrutural proteica utilizados neste trabalho estão descritas abaixo e também estão resumidas na Figura 4.2.

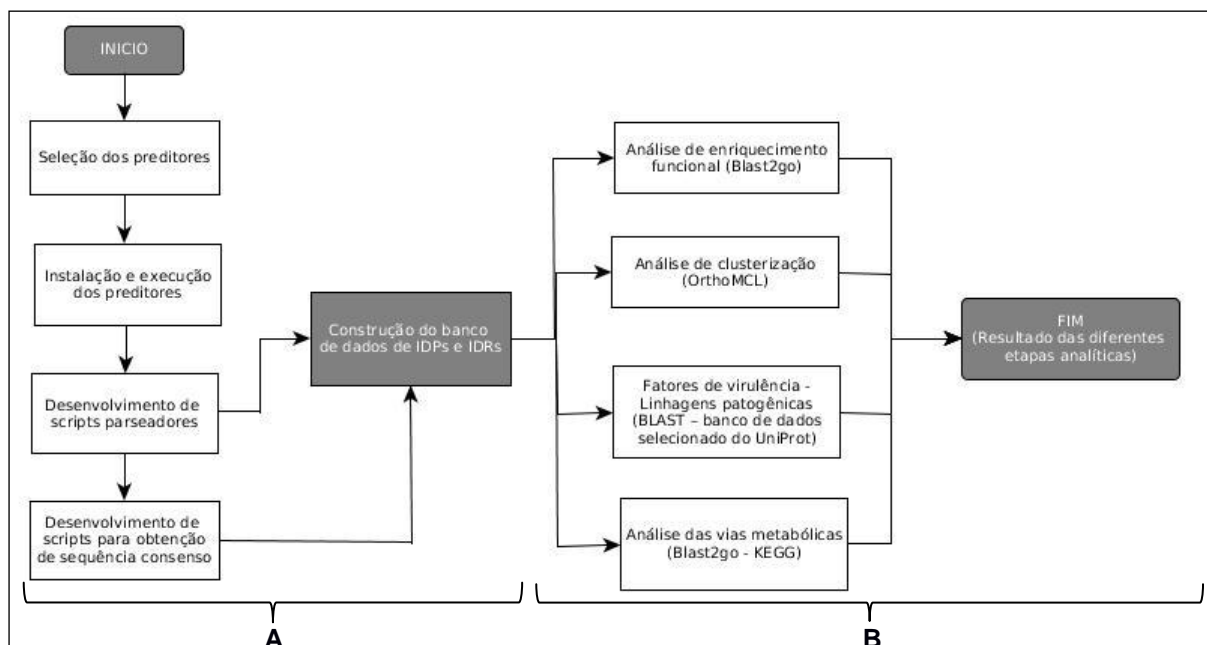


Figura 4.2: Fluxograma esquematizando as etapas analíticas do processo de predição de desordem estrutural proteica utilizado neste estudo, onde: **(A):** Seleção, instalação e execução dos preditores de desordem proteica e, posteriormente, foram utilizados *scripts* escritos em linguagem de programação *PERL* para parseamento das informações e obtenção das sequências consenso. Após esta etapa, as informações foram armazenadas em um banco de dados relacional; **(B):** Esta etapa refere-se às análises realizadas com as sequências consenso, como: enriquecimento funcional (utilizando o programa Blast2go), *clusterização* (utilizando o programa OrthoMCL), proteínas relacionadas aos fatores de virulência nas linhagens virulentas (utilizando busca por similaridade de sequências em um banco de dados selecionado a partir do UniProt) e por fim, análise das vias metabólicas (utilizando o programa Blast2go e a base de dados do KEGG). **Fonte:** Acervo Pessoal.

4.4.1 Seleção dos preditores computacionais de desordem estrutural proteica

Seguindo o delineamento estabelecido por Torrieri (2010), utilizamos para a seleção dos preditores de desordem estrutural e que foram empregados neste estudo, os seguintes critérios:

- Disponibilidade do artigo ou manual descrevendo a metodologia de predição empregada pelo programa;
- disponibilidade dos programas para *download* gratuito;
- disponibilidade de programas acessórios e/ou bancos anexos necessários à execução do preditor;
- disponibilidade de documentação suficiente para a correta instalação e execução do preditor;

- e) número de citações em trabalhos relacionados;
- f) implementação de diferentes definições de desordem estrutural.

4.4.2 Predição das IDPs e IDRs

Após a seleção do conjunto de preditores que foram utilizados neste estudo (vide item 5.3), as IDPs e IDRs foram identificadas utilizando o melhor conjunto de parâmetros relatado para cada programa.

Entretanto, devido aos diferentes formatos de saída gerados por cada um dos preditores, foram desenvolvidos *scripts* que geram formatos de saída padronizados, como mostrado na Figura 4.3. Esses *scripts* selecionam somente as sequências identificadas como desordenadas pelos preditores e que possuam, no mínimo, 40 resíduos de aminoácidos consecutivos.

Posteriormente, essas informações foram integradas ao BD *idp_crypto* (vide item abaixo).

protein_name	predictor	coord_init	coord_end	mean_score	sequence
cbest.000030	iupred	28	72	0.65	SHTVSTQNQWSPVSRAGELDGGDPSSSSPPLTKSTKRTRTGSRK
cbest.000070	iupred	252	500	0.74	KHRAPISFSSPTPKSTPVPTLAPPISRGSSKRPPPIPLRASSLKSFLFAQSASDNQNHITRSSSPTF
cbest.000070	iupred	609	649	0.65	PPFRVTRSESTPGLSTSGTLESMISDPDQTPSESEFVLHGE
cbest.000090	iupred	372	418	0.8	MTTAAAAYLTPDGGVENTESSLGGGGMGGNSTSGGDGSAAPSGSPPS
cbest.000100	iupred	64	156	0.71	QHYPANMSTAYGVPVDEEGENVPLALSSQTHHTPRISLRRESASSTEFRDQLEVEPFDEKGRDF
cbest.000150	iupred	71	176	0.7	NESIATSSCIGTATPAASEHPADRDNQANLIDLSSDLKTAAPRLVGRSSPQSNQSGPPPIPEHVDARSDDL
cbest.000170	iupred	458	563	0.71	SNTGGISAPASSAAAAPSSAAPFTNGSSPSSVSAPPDASSVITAPGSMAPSSAYSTPSDGGVAT
cbest.000200	iupred	197	344	0.83	KDKELELDMEDDESLLDSDEEREEEEEGGEREYVEDTDESVDGLEDYSGSEFDEFDSEEEES
cbest.000210	iupred	153	215	0.77	LFQMSGNEGMMGMRTGEVDETKEPTIEAQPTIIEEPSRRYDPNAVFTLDLNSDDSDDD
cbest.000230	iupred	1	294	0.84	MTDTGATPMPNPHPSHQPPPAVSRVYHRPSRNSAHYSAGTDVSLGMVSAASANTSSPRIF
cbest.000230	iupred	329	576	0.9	GMPQEWQKILDENGITRAEQEENPNEVLAVVQYFKNRDAQENQEVWQKMRNAGPANTPDS
cbest.000260	iupred	203	245	0.59	LHQTLQTNPLIPNLRKLAKEKETELTNLHRQSRNLAQREIKKV
cbest.000260	iupred	277	331	0.65	QRKAGLKAMEERENDTKRRLEEITQLKKEKEDRRSKETTLQERNKQLKMFQKKV
cbest.000260	iupred	350	622	0.78	RLRSSMYEGIVPSDVEDLKGRGGDTISQIESETYFPSSSRSTRNGRHSTAKMEEDDSLQIEMPSF
cbest.000270	iupred	158	244	0.74	GDDASVTKLEEEQGLTKNIATDEASAGEASQAAVGGGASNAAADTAEDEDAADTAEVDAA
cbest.000280	iupred	74	128	0.71	EKSSSELDLDFGGDDVESQSHEDKYENEDDLVDDQSSSIPTINGDQKEDARII
cbest.000280	iupred	191	262	0.63	SRNSYRNSTNSIKVLVHPNLRPRLKRTTSSPLPLATMSFKTPEGQSSFDGHDGHANANGNTSPV
cbest.000280	iupred	267	342	0.7	SSSSVGSKRSTIRSSRQLKGTGACTPTPTPINRIYTNQASSSKRHDQIQFQSSSENDIPDIHRSA
cbest.000280	iupred	392	650	0.79	GSSSGSFDFPNPNREIENSTSPSHLDSSPTDPNFPPLSQSTNHPHTHTHTQTQTQTQTQTQ
cbest.000280	iupred	830	991	0.74	VFEENSVDVFSEQPLEDQDQKQKDYEFGIKNTESNHQRFHTLNSQSSNTLVNRPPTKGYFSRE
cbest.000280	iupred	1062	1220	0.77	DGATKRDKLAVLMGQPGQNGARKSSQSAQSARSVGGRRNQNTSPTSPDNADDPVENGNTSSR
cbest.000340	iupred	1	243	0.82	MPSAHSHTSHKGGVKAAHAASPNSLNSQSPRLKLRNNSPSASTSTSSSKAIRSRPTHSHKILSI
cbest.000340	iupred	360	601	0.82	VLVQPRHERSRHRGDKGSRTAGSVSGSTVISGTDQAALLIDPDADDGEFDTEDSYWSGQSDI
cbest.000340	iupred	604	675	0.73	SIPGHPRYAARRAAEALCDPQDRETTPTDSDEAFSLEDMLETSLVLMHELEEHQNGNEHDGHDQD
cbest.000340	iupred	734	783	0.82	SPVLGPTSYQAKREKKKKPKATIRSSHHHHHHHQDHPHQIENGLANGM

Figura 4.3: Resultado gerado pelo *script iupred-parser-v3.pl* referente ao resultado do preditor IUPred. Esta figura mostra um exemplo do resultado padronizado que é gerado por todos os *scripts* desenvolvidos para tal finalidade. Na coluna *protein_name* encontram-se relacionados os nomes das proteínas, na coluna *predictor* está relacionado o nome do preditor. Nas colunas *coord_init* e *coord_end*, estão contidas as informações referentes às coordenadas de início e fim da sequência, preditas como desordenadas, respectivamente. Já coluna *mean_score*, está relacionada à informação do *score* médio (quando esse valor não é informado pelo preditor, atribui-se valor zero a este campo). Por fim, a coluna *sequence* contém a sequência de resíduos de aminoácidos preditos como desordenada. **Fonte:** Acervo Pessoal.

4.4.3 Banco de dados relacional de IDPs e IDRs em *Cryptococcus spp.*

Para viabilizar a análise dos dados relacionados às IDPs e IDRs identificadas, foi criado um BD relacional utilizando o sistema gerenciador de bancos de dados MySQL versão 14.14. Esse BD, denominado *idp_crypto*, contém as informações de todas as proteínas e suas respectivas funções preditas na tabela *proteome*. Além desta, as tabelas *predictions* e *consensus* contêm todas as informações relacionadas às predições de IDPs e IDRs e as sequências consensos geradas (vide item abaixo), respectivamente. O esquema relacional do banco está representado na Figura 4.4.

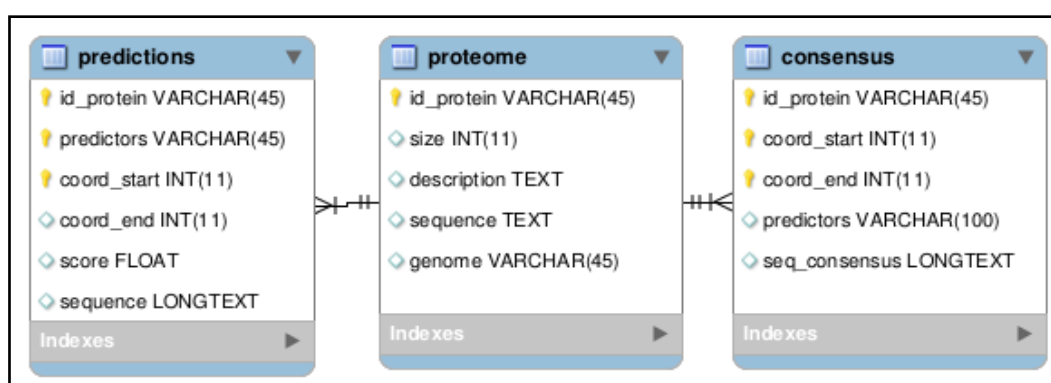


Figura 4.4: Esquema do Banco de Dados Relacional *idp_crypto*. As tabelas *predictions* e *consensus* estão relacionadas à tabela *proteome* através da chave estrangeira *id_protein* que também é a chave primária de cada tabela. Fonte: Acervo Pessoal.

4.5.2 Sequência consenso das IDPs e IDRs preditas

Para a remoção da redundância ocasionada pelas predições das IDPs e IDRs por diferentes preditores, foi necessário desenvolver os *scripts* *consenso-coord-idp-v4.pl* e *consenso-idp-coord--vs—proteome.pl*. O *script* *consenso-coord-idp-v4.pl* conecta ao BD gerado (*idp_crypto*), utilizando os parâmetros descritos na Tabela 4.1, e seleciona a tabela *predictions*. Após essa seleção, o *script* irá considerar uma coordenada consenso somente quando há uma sobreposição de, no mínimo, 10% da coordenada inicial seguinte em relação à coordenada final anteriormente formada (Figura 4.5) (Ruy 2011). As coordenadas que não atenderam ao critério de sobreposição mínimo (10%) foram consideradas como sendo outra região, ou seja, não fazem parte do consenso gerado.

Tabela 4.1: Parâmetros solicitados pelo *script consenso-coord-idp-v4.pl* para sua execução.

Parâmetro	Descrição
-d	Nome do banco de dados
-h	Servidor de conexão com o banco de dados
-u	Usuário cadastrado no banco de dados
-p	Senha do usuário para conexão no banco de dados
-o	Nome do arquivo de saída

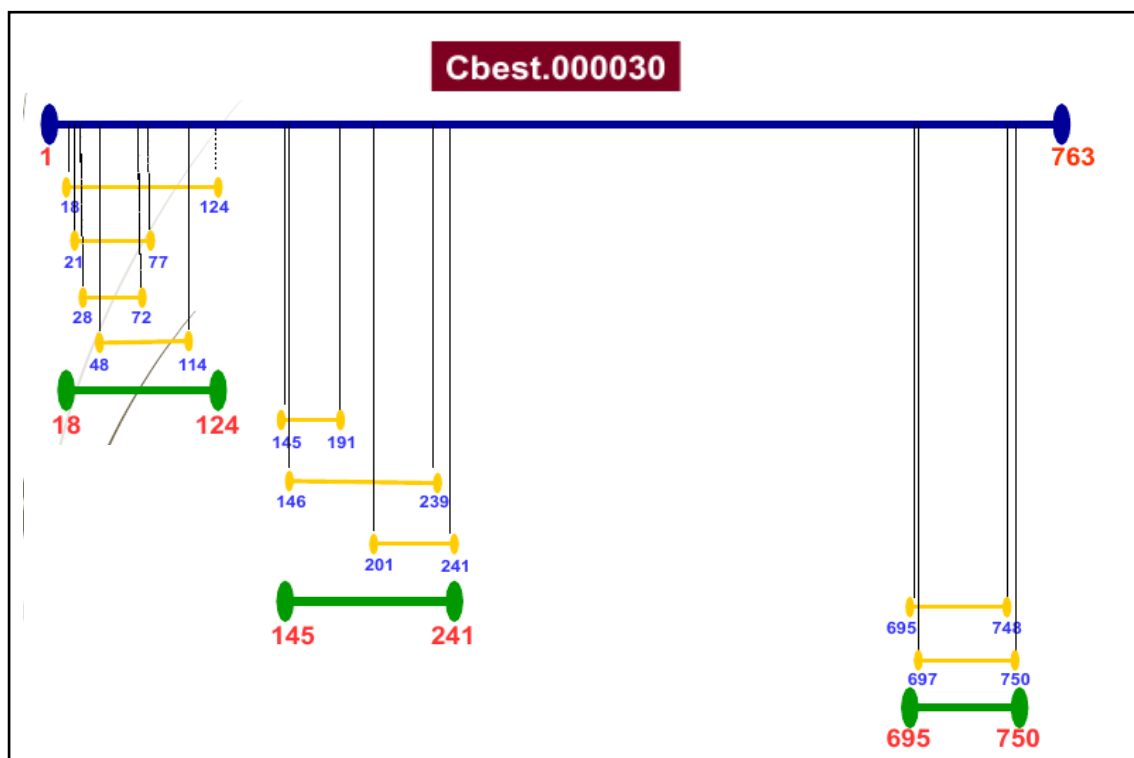


Figura 4.5: Diferentes regiões consenso formadas em uma mesma proteína e que atendiam ao critério de 10% de sobreposição. Em azul, está representado o tamanho total da proteína, em amarelo, as regiões preditas por diferentes preditores e, por fim, em verde as regiões consenso formadas. Figura fora de escala. Fonte: Acervo pessoal.

O arquivo gerado pelo *script consenso-coord-idp-v4.pl*, que contém as coordenadas consenso das predições, assim como o nome do diretório que contém todas as proteínas separadas em arquivos distintos e no formato Fasta, são utilizados como entrada para o *script consenso-idp-coord-vs—proteome.pl*. Assim, esse *script* compara cada arquivo com as coordenadas consenso e, se o nome da proteína em ambos os arquivos for idêntico, é selecionada a sequência de resíduos de aminoácidos correspondente às coordenadas.

Posteriormente, esses dados foram integrados à tabela *consensus* do BD *idp_crypto*.

4.4.4 Análise de Enriquecimento Funcional

Após a identificação do conteúdo de IDPs e IDRs nos proteomas preditos de *Cryptococcus* spp., foi realizada uma análise de enriquecimento funcional para cada proteoma. A abordagem utilizada envolve o banco de dados *Gene Ontology* (GO), o programa Blast2GO Pipeline Version 2.5 (B2G4Pipe) (Ashburner et al. 2000; The Gene Ontology Consortium 2015) (*The Gene Ontology Consortium*, 2015; Ashburner et al., 2000) e o pacote GO::TermFinder (<http://search.cpan.org/~sherlock/GO-TermFinder-0.86/lib/GO/TermFinder.pm>) (Conesa et al. 2005).

O teste estatístico aplicado para esta etapa foi o Teste Bilateral de Fischer, com Correção de Bonferroni e o critério de corte utilizado foi o p-valor ajustado menor que 0,05.

4.4.5 Mapeamento das vias metabólicas e associação aos fatores de virulência

Para a realização do mapeamento das IDPs e IDRs nas vias metabólicas, também utilizamos os resultados obtidos pelo Blast2GO para posterior mapeamento dessas proteínas em vias metabólicas presentes na Enciclopédia Kyoto de Genes de Genomas (KEGG) (Kanehisa et al. 2004).

Para a associação das proteínas identificadas como desordenadas, primeiramente foi realizada uma busca no BD do *Universal Protein Resource* (UniProt) com as palavras-chave: *Cryptococcus* e *virulence*. Os resultados obtidos através dessa busca formaram o BD utilizado posteriormente para a busca por similaridade de sequências realizada pelo algoritmo BLAST.

4.4.6 Análise de agrupamento

Levando ainda em consideração o objetivo de identificarmos o *core genome*, ou seja, o conjunto de genes compartilhados entre todos os organismos estudados e adicionalmente o conjunto de genes espécie-específico, realizamos uma análise de agrupamento utilizando o programa OrthoMCL versão 1.4 (Li et al. 2003). Os parâmetros utilizados para a execução do programa foram:

```
./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***

Primeiramente realizamos a análise com o proteoma de todos os organismos e posteriormente identificamos as IDPs e IDRs em cada situação.

5 RESULTADOS E DISCUSSÃO

5.1 Seleção dos genomas de *Cryptococcus* spp.

A grande quantidade de informações provenientes dos diferentes projetos finalizados ou em andamento relacionadas ao nosso organismo modelo, *Cryptococcus* spp., estava distribuída em diferentes BD de domínio público.

Assim, na etapa inicial deste trabalho foram selecionados 12 genomas de diferentes linhagens de *Cryptococcus* spp., sendo dois obtidos no BD do *Broad Institute* e 10 obtidos no BD do NCBI, *GenBank*. Adicionalmente, os dados utilizados integram duas espécies (*C. neoformans* e *C. gattii*) que correspondem a seis linhagens diferentes de importância médica (Tabela 5.1) (Li e Mody 2010).

Tabela 5.1: Lista de genomas de *Cryptococcus* spp. selecionados.

Organismo	Linhagem	Banco de Dados	GenBank Assembly	Virulento
<i>C. neoformans</i> var. <i>neoformans</i>	JEC21	NCBI	GCA_0000910	Sim
<i>C. neoformans</i> var. <i>grubii</i>	H99	NCBI	GCA_0001492	Sim
<i>C. neoformans</i> var. <i>neoformans</i>	B-3501A	NCBI	GCA_0001493	Sim
<i>C. gattii</i>	WM276	NCBI	GCA_0001859	Sim
<i>C. gattii</i>	CBS 7750	NCBI	GCA_0004995	Sim
<i>C. gattii</i>	R265	NCBI	GCA_0001494	Sim
<i>C. dejecticola</i>	CBS 10117	NCBI	GCA_0005125	Não
<i>C. bestiolae</i>	CBS 10118	NCBI	GCA_0005125	Não
<i>C. pinus</i>	CBS 10737	NCBI	GCA_0005126	Não
<i>C. flavescens</i>	NRRL Y-50378	NCBI	GCA_0004427	Não
<i>C. heveanensis</i>	BCC8398	Broad	ASQB0100000	Não
<i>C. heveanensis</i>	CBS569	Broad	ASQC0100000	Não

* Número de acesso pelo BD do *Broad Institute*

5.2 Anotação genômica

Rotineiramente, os dados disponíveis nos bancos de dados de domínio público apresentavam-se em diferentes formatos, qualidade e profundidade da informação biológica associada a cada sequência. Dentro desse contexto, usualmente algum tipo de reformatação e/ou reanotação se fez necessária como etapa preliminar das análises comparativas desejadas ou almejadas.

A análise preliminar dos dados selecionados para este trabalho evidenciou duas situações distintas, a existência de um primeiro grupo contendo quatro genomas com anotação estrutural e funcional e de um segundo grupo contendo oito genomas sem qualquer anotação genômica.

Com relação ao primeiro grupo (genomas com anotação estrutural e funcional), estava disponível o detalhamento de genes e cromossomos que integravam cada um dos genomas, o que corresponde a uma média 7036 genes. Por outro lado, com relação ao segundo grupo (genomas sem anotação estrutural e funcional) as únicas informações disponíveis estavam relacionadas aos arquivos contendo as sequências de nucleotídeos organizadas em *Scaffolds*, *Contigs* ou *Supercontigs*.

Na Tabela 5.2, apresentamos o detalhamento do *status* de anotação estrutural e funcional de cada um dos genomas analisados neste estudo. Assim, em relação aos pontos levantados acima e aos apresentados nessa mesma tabela, ficam evidentes: a) a necessidade de anotação dos genomas do segundo grupo; e b) a necessidade de reanotação dos genomas do primeiro grupo em função da atualização das versões dos bancos de dados.

Sendo assim, todos os genomas foram anotados ou reanotados. Logo, foram essas as informações que viabilizaram as análises preditivas e comparativas posteriores.

Tabela 5.2: Lista de genomas de *Cryptococcus* spp. com o número de genes encontrados pela predição *ab initio* e os provenientes da anotação original.

Organismo	Linhagem	Banco de Dados	Status de anotação*	Genes da anotação original	Genes Preditos pelo Augustus
<i>C. neoformans</i> var. <i>neoformans</i>	JEC21	NCBI	Cr	6475	-
<i>C. neoformans</i> var. <i>grubii</i>	H99	NCBI	Cr	7798	-
<i>C. neoformans</i> var. <i>neoformans</i>	B-3501A	NCBI	Cr	6578	-
<i>C. gattii</i>	WM276	NCBI	Cr	6575	-
<i>C. gattii</i>	CBS 7750	NCBI	S/C	-	7372
<i>C. gattii</i>	R265	NCBI	S/C	-	6209
<i>C. dejecticola</i>	CBS 10117	NCBI	S/C	-	7423
<i>C. bestiolae</i>	CBS 10118	NCBI	S/C	-	8128
<i>C. pinus</i>	CBS 10737	NCBI	S/C	-	5160
<i>C. flavescens</i>	NRRL Y-50378	NCBI	S/C	-	8714
<i>C. heveanensis</i>	BCC8398	Broad	C/Sc	-	7379
<i>C. heveanensis</i>	CBS569	Broad	C/Sc	-	7295

* Cr: Cromossomo; S/C: *Scaffold* ou *Contig*; C/Sc: *Contigs* ou *Supercontigs*

5.2.1 Anotação estrutural

Com relação aos oito genomas organizados em *Scaffolds*, *Contigs* ou *Supercontigs* (*C. gattii* CBS 7750, *C. gattii* R265, *C. dejecticola* CBS1011, *C. bestiolae* CBS10118, *C. pinus* CBS10737, *C. flavescens* NRRL Y-50378, *C. heveanensis* BCC8398, *C. heveanensis* CBS569), foi realizada a predição gênica *ab initio*. Para tanto, utilizamos o programa Augustus (Stanke e Morgenstern 2005). Este software baseia-se em modelos ocultos de cadeia de Markov, além de empregar um modelo probabilístico relacionado à sequências gênicas para genomas de eucariotos. Adicionalmente, este programa emprega como modelo de treinamento os genomas de *C. gattii* e *C. neoformans*. A Tabela 5.2 sumariza o número de genes preditos para cada um dos genomas analisados.

Como resultado dessa etapa, este estudo gerou a anotação estrutural de oito genomas, totalizando 57.680 genes (o que corresponde em média 7210 genes) preditos que posteriormente foram anotados funcionalmente e analisados em termos do seu conteúdo de desordem estrutural. Complementarmente às análises de desordem proteica, esse resultado compõe um outro aspecto da contribuição do presente estudo.

5.2.2 Anotação funcional

Todos os 12 genomas analisados passaram pela etapa de anotação funcional descrita a seguir.

O processo analítico empregado envolveu a comparação dos genes preditos em todos os genomas (85.106 genes) contra diferentes bancos de dados de domínio público, incluindo: Pfam, CDD, Interproscan e dois *subsets* diferentes do BD NR do NCBI. O primeiro *subset* do BD NR utilizado integrava todas as sequências de *Cryptococcus* spp. depositadas até 22/10/2014 e o segundo *subset* integrava todas as demais sequências depositadas até a mesma data.

Na Tabela 5.3 podemos observar que o número de proteínas funcionalmente anotadas com o BD NR sem *Cryptococcus* está em torno de 15% nos genomas analisados, sendo que o conteúdo de proteínas hipotéticas está em torno de 85%. Essa análise evita a propagação de erros inerentes ao processo de anotação automática, ou seja, diminui consideravelmente a inferência equivocada de função atribuída por buscas de similaridade de sequência (Figura 5.1).

Através da utilização dos bancos mencionados acima e dos *subsets* do BD NR foi possível identificar o grau de conservação das sequências entre os genomas de *Cryptococcus*. Assim, a análise de agrupamento do produto proteico, codificado pelos genes dos 12 genomas, revelou um *core genome* contendo 18900 genes. O detalhamento dos genes espécie-específicos e compartilhados entre as espécies e linhagens pode ser visualizado na Figura 1.5/Figura 5.2.

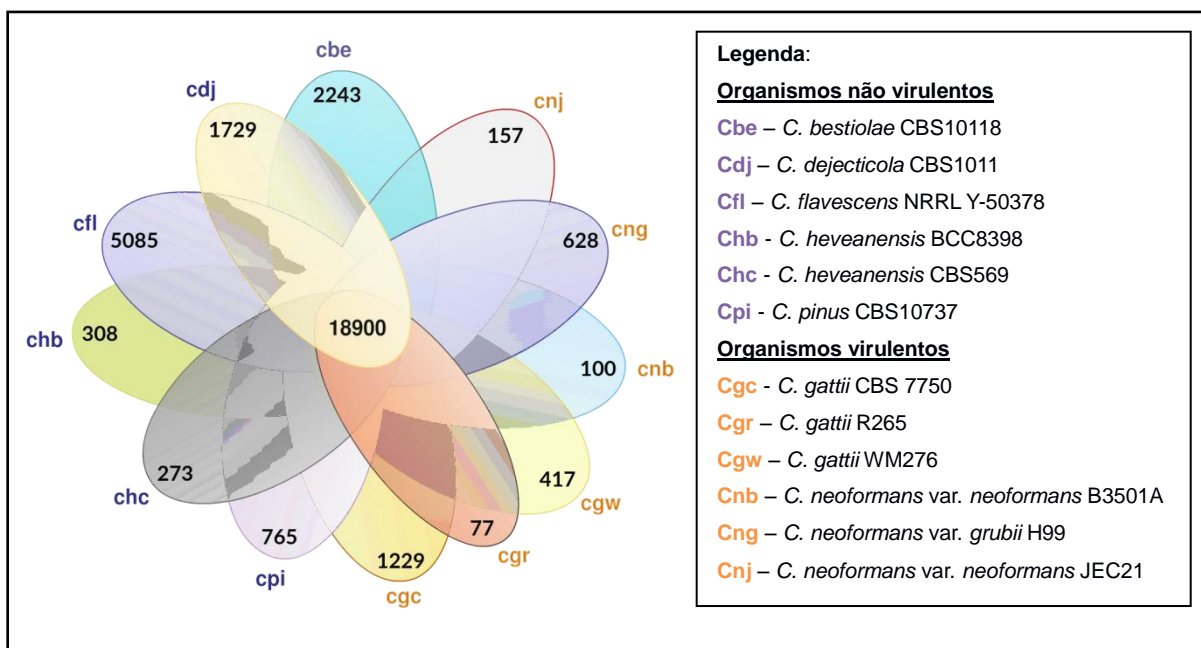


Figura 5.2: Diagrama de venn referente à análise de clusterização. Neste diagrama podemos observar o número de genes relacionados ao *core genome*, além do número de genes relacionados para cada espécie. **Fonte:** Acervo Pessoal.

5.3 Seleção dos preditores de IDPs e IDRs

De acordo com os critérios estabelecidos no item 4.4.1 deste documento, foram selecionados inicialmente nove preditores de IDPs e IDRs, são eles: DisEMBL (Linding et al. 2003a), GlobPlot (Linding et al. 2003b), IUPred (Dosztányi et al. 2005), ANCHOR (Dosztányi et al. 2009), DynaMine (Cilia et al. 2013), DISOPRED3 (Jones e Cozzetto 2015), VSL2 (Peng et al. 2006), ESPRITZ (Walsh et al. 2012) e Dis-PRO (Cheng et al. 2005) (Tabela 5.4).

Tabela 5.4: Relação de preditores selecionados a partir dos critérios estabelecidos

Preditor	Tipo de abordagem analítica empregada	Observação	Número de Citações
DisEMBL	Três metodologias implementadas na forma de redes neurais.	Instalado	727
GlobPlot	Determina a propensão de cada resíduo estar numa região desordenada (random-coil) ou em uma estrutura secundária, ambos segundo o DSSP. Esta propensão é definida pela fórmula $P=RC-SS$.	Instalado	654
IUPred	Sequência de aminoácidos com base no teor de energia estimada entre os resíduos	Instalado	696
ANCHOR	Identifica segmentos que desordenados que estão em regiões de ligação. A metodologia empregada baseia-se na abordagem de estimativa de energia entre os resíduos.	Instalado	139
DynaMine	Utiliza regressão linear para prever a dinâmica da cadeia principal da proteína.	Instalado, mas não utilizado	16
DISOPRED3	Predição de sítios de ligação de proteínas no interior das regiões desordenadas usando SVM, que analisa padrões de conservação da sequência evolutiva, informações sobre a posição e composição de aminoácidos das potenciais regiões desordenadas.	Instalado, mas não utilizado	365
VSL2	1º) analisa a sequência de aminoácidos. 2º) informações sobre a estrutura secundária 3º) informações sobre a variabilidade sequências (derivadas de alinhamentos).	Instalado, mas não utilizado	353
ESPRITZ	É baseada em um método de aprendizado de máquina que não necessita de janelas deslizantes ou quaisquer fontes complexas de informações (redes neurais recursivas e bi-direcionais)	Não foi possível realizar a instalação	55
Dis-Pro	Baseia em métodos de aprendizado de máquina e informação relacionadas à evolução. Além disso, prediz estruturas secundárias de acordo com a acessibilidade relativa ao solvente.	Instalado, mas apresentou problemas durante a execução	135

A análise individual de cada programa revelou alguns problemas. Particularmente, os preditores VSL2 (Peng et al. 2006) e o Dis-PRO (Cheng et al. 2005) não foram implementados nas análises, pois erros no código fonte associados à aquisição dos parâmetros fornecidos pelo usuário (VSL2) ou à execução de rotinas ligadas a programas secundários e/ou dependências (Dis-Pro) impediram a execução dos mesmos.

Adicionalmente, o programa ESPRITZ (Walsh et al. 2012) apresentou erros de compilação e alto tempo de processamento, que mesmo após o contato com os desenvolvedores não puderam ser solucionados a tempo de viabilizar seu emprego.

Similarmente, apesar de ter sido instalado, o DISOPRED3 (Jones e Cozzetto 2015) foi descartado, uma vez que apresentava um tempo de processamento demasiadamente grande (aproximadamente dois meses por genoma em um servidor onde foram disponibilizados 20 núcleos de processamento).

Por fim, resolvemos remover das análises os resultados gerados pela abordagem *Coils* implementada no preditor DisEMBL, uma vez que foi observada uma discrepância (aproximadamente 77-90% do conteúdo proteico dos organismos, conforme mostrado na Tabela 5.5 e na Figura 5.3 quanto ao número de proteínas identificadas por esta abordagem, em relação aos demais preditores. Confirmamos a constatação feita anteriormente por Torrieri (2010), com dados de Tripanosomatídeos, que evidenciam que essa abordagem apresenta uma alta taxa de Falso-positivos. Resultado similar foi observado para o programa DynaMine o que também nos levou a remover os resultados gerados por essa abordagem preditiva das etapas analíticas posteriores.

Tabela 5.5: Resultado do preditor DynaMine e da abordagem COILS implementada pelo DisEMBL

Organismo	Linhagem	Número de Genes	DynaMine (%)	COILS (%)
<i>C. neoformans</i> var. <i>neoformans</i>	JEC21	6475	4642 (71,7)	6419 (81,9)
<i>C. neoformans</i> var. <i>grubii</i>	H99	7798	5535 (71,0)	6471 (77,6)
<i>C. neoformans</i> var. <i>neoformans</i>	B-3501A	6578	3784 (57,3)	5571 (84,3)
<i>C. gattii</i>	WM276	6575	4679 (57,5)	5447 (82,8)
<i>C. gattii</i>	CBS 7750	7372	5139 (69,7)	6072 (82,4)
<i>C. gattii</i>	R265	6209	4552 (73,3)	5300 (85,4)
<i>C. dejecticola</i>	CBS 10117	7423	5771 (77,7)	6605 (89,0)
<i>C. bestiolae</i>	CBS 10118	8128	6118 (75,3)	7156 (88,0)
<i>C. pinus</i>	CBS 10737	5160	3933 (76,2)	4507 (87,3)
<i>C. flavescens</i>	NRRL Y-50378	8714	6491 (74,5)	7281 (83,6)
<i>C. heveanensis</i>	BCC8398	7379	5975 (81,0)	6630 (89,8)
<i>C. heveanensis</i>	CBS569	7295	5923 (81,2)	6549 (89,8)

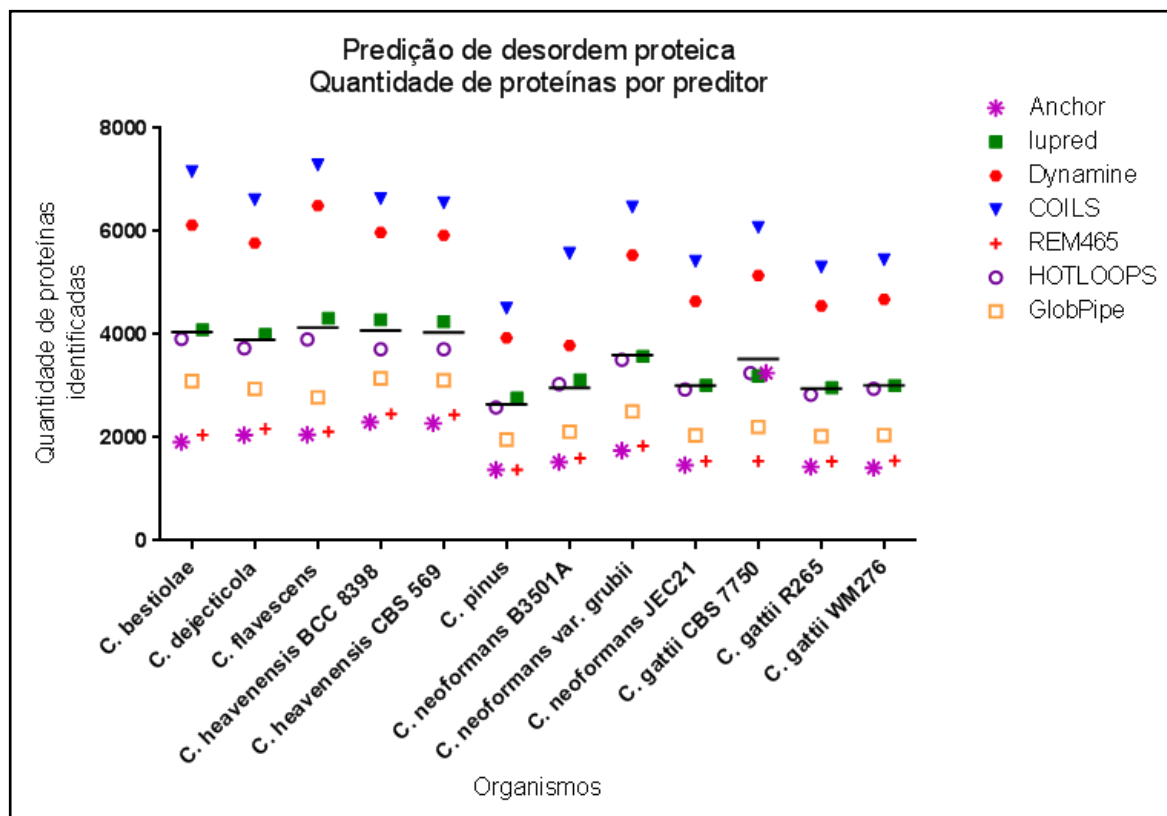


Figura 5.3: Gráfico referente à predição de desordem proteica por preditor e organismo.

Assim, os programas DisEMBL, GlobPlot, IUPred e ANCHOR, foram implementados neste estudo. Tais programas, exceto o ANCHOR, também foram implementados no estudo de Ruy (2011) e Torrieri (2010) e foram indicados, por esses autores, como a melhor combinação de preditores para a identificação de desordem estrutural.

Portanto, a seguir apresentamos detalhadamente estes programas e suas abordagens utilizadas para identificação de IDPs e IDRs. Adicionalmente, como parte complementar as análises realizadas por estes preditores, descrevemos também, as informações que foram extraídas dos resultados gerados por cada um deles.

5.3.1 DisEMBL

O DisEMBL utiliza três metodologias diferentes de predição que são implementadas na forma de redes neurais, ou seja, o algoritmo foi treinado com um conjunto de proteínas com características conhecidas (aprendizado supervisionado). Complementarmente, calculados por funções sigmóides, são atribuídos valores associados a cada característica de composição da sequência. Após este

treinamento, o DisEMBL é capaz de identificar tais características de acordo com os melhores critérios obtidos (Linding et al. 2003a; Torrieri 2010).

Duas das metodologias empregadas por este preditor (*Loops/coils* e *Hotloops*) baseiam-se, primeiramente, na identificação de estruturas secundárias de acordo com os critérios estabelecidos pelo Dicionário de Estruturas Secundárias de Proteínas (DSSP) (Linding et al. 2003a; Torrieri 2010).

A primeira abordagem, *Loops/coils*, identifica os resíduos presentes em cada tipo de estrutura secundária. Contudo, os resíduos identificados nas regiões de Folha- β e α -hélice são identificados como ordenados, assim, todos os resíduos que se localizam nas demais estruturas (*random coils*, *loops* e *turns*) são considerados como pertencentes a uma região estruturalmente desordenada (Linding et al. 2003a; Torrieri 2010). Entretanto, como mencionamos anteriormente, essa análise preditiva foi removida uma vez que gerava uma alta taxa de Falso-positivos.

Já a segunda abordagem, *Hotloops*, como descrito por Linding e colaboradores (2003), é um refinamento da primeira metodologia. Esta análise identifica regiões de *loops* com alto grau de mobilidade que é determinado pela flutuação dos átomos de Carbono α ao redor de suas posições médias e, assim, informações sobre a dinâmica da proteína são identificadas. Este tipo de análise também é conhecido como fator de temperatura do Carbono α ou Fator B, e possui uma maior sensibilidade se comparado com a primeira abordagem descrita (Linding et al. 2003b; Torrieri 2010).

Por fim, a última metodologia empregada por este programa, *REMARK465*, identifica regiões que apresentam falta de coordenadas espaciais (x, y, z) dentro de arquivos do BD do *Protein Data Bank* (PDB). Esta abordagem utiliza dados armazenados pelo treinamento do algoritmo com um conjunto de informações do BD PDB e as compara com as características encontradas na análise em questão.

Para a execução do DisEMBL utilizou-se os parâmetros estabelecidos pelos autores uma vez que os mesmos foram descritos como sendo os de melhor desempenho. Como arquivo de entrada, este programa utiliza dados de sequência no formato Fasta, contendo todas as proteínas que se deseja analisar. No caso deste estudo, o arquivo de entrada foi o proteoma predito para cada um dos organismos selecionados (Linding et al. 2003a). A linha de comando utilizada para tanto foi:

```
./DisEMBL-SeqIO.py 8 8 4 1.2 1.4 1.2 <arquivo-multi-fasta> > <nome-do-arquivo-de-saída>
```

Legenda dos parâmetros:

- ✓ *Savitzky-Golay smoothing frame:* 8
- ✓ *Minimum peak width:* 8
- ✓ *Maximum join distance:* 4
- ✓ *Coils threshold:* 1.2
- ✓ *Hot loops threshold:* 1.4
- ✓ *Remark465 threshold:* 1.2

Este preditor gera como resultado um arquivo no formato Fasta. Neste arquivo, no campo destinado às informações adicionais sobre a sequência, fica explícita a identificação de cada abordagem utilizada (*Coils*, *Hotloops*, *REMARK465*) e as respectivas coordenadas das regiões desordenadas identificadas.

Com o objetivo de automatizar a extração de informações dos resultados gerados e posterior carga no banco de dados MySQL (*idp_crypto*) desenvolvido para este estudo, o *script disembl-parser-v5.pl* foi desenvolvido. Este *script* retira todas as informações relacionadas à identificação das abordagens utilizadas, assim como as coordenadas das regiões desordenadas e suas respectivas sequências de aminoácidos, viabilizando a carga no banco de dados de maneira correta.

5.3.2 GlobPlot

Este programa identifica regiões de desordem estrutural, baseando-se na propensão de cada resíduo estar numa região desordenada (*random-coil*) ou em uma região de estrutura secundária. Para tanto, o GlobPlot compara os valores de propensão (RC e SS), obtidos a partir da análise de sequências do BD do *Structural Classification of Proteins* (SCOP), com os valores de propensão obtidos nas sequências que serão analisadas. Esta propensão é definida pela fórmula:

$$P=RC-SS$$

onde P é a propensão resultante para cada resíduo, RC é a propensão do resíduo estar em uma região de *random-coil* e SS é a propensão do resíduo estar em uma região de estrutura secundária (Linding et al. 2003b).

Para a execução deste preditor, assim como para o DisEMBL, é necessário um arquivo de entrada, no formato Fasta, contendo todas as sequências a serem analisadas. Além disso, também foram utilizados os parâmetros de melhor desempenho estabelecidos pelos autores, conforme a seguinte linha de comando:

```
./GlobPipe-updated2-v2.py 10 15 74 4 5 <arquivo-de-entrada-em-formato-fasta>><arquivo-de-saída>
```

Legenda dos parâmetros:

- ✓ Savitzky-Golay frame (*Smooth_frame*): 10
- ✓ Maximum linker length (*join distance*): 15
- ✓ Minimum domain length (*peak width*): 74
- ✓ Maximum join distance: 4
- ✓ Minimum peak width: 5

O arquivo de saída gerado pelo GlobPlot é similar ao arquivo gerado pelo DisEMBL. Entretanto, este preditor, além de informar as coordenadas das regiões identificadas como desordenadas, também informa as coordenadas das regiões globulares (Linding et al. 2003b).

Similarmente o *script disembl-parser-v5.pl* foi desenvolvido com o objetivo de selecionar todas as informações referentes à identificação de regiões desordenadas e excluir todas as informações relacionadas às regiões globulares. Posteriormente, tais informações foram armazenadas no BD *idp_crypto*.

5.3.3 IUPred

Diferentemente dos preditores citados acima, o IUPred identifica regiões desordenadas baseando-se na energia livre de Gibbs (ΔG) proveniente da ligação peptídica de dados conhecidos. Esta energia é calculada a partir da diferença entre ΔH e ΔS . A partir desta análise, o IUPred realiza um cálculo estatístico que estima o potencial dos polipeptídios formarem interações estáveis. Dentro desse contexto, os valores referentes a este potencial formam uma matriz de tamanho 20 x 20, que posteriormente é comparada com os dados a serem analisados (Dosztányi et al. 2005).

Segundo Dosztányi e colaboradores (2005), este tipo de abordagem é possível devido ao fato de IDPs e IDRs apresentarem uma grande quantidade de resíduos que não conseguem formar interações inter-resíduos suficientes e, assim, geram um perfil de energia livre diferente em relação às regiões globulares.

Para a execução deste preditor foi necessário separar o proteoma predito de cada organismo em arquivos individuais, além da opção *long*, indicada para identificação de grandes regiões de desordem (> 30 resíduos consecutivos de aminoácidos). Esta etapa de separação do proteoma foi necessária, pois este programa solicita um arquivo de entrada por vez e no formato Fasta.

Considerando esses pontos, desenvolvemos um *script* em *shell* para a execução das análises em todas as sequências, conforme descrito abaixo.

```
for i in `ls <arquivo>`; do iupred<caminho-do-diretório-dos-arquivos> $ilong ><arquivo-de-saída> ; done &
```

Como resultado, o IUPred gera arquivos individuais para cada sequência analisada e atribui aos resíduos um valor de *score*, resultante da análise estatística realizada, que variam entre 0 e 1. Segundo Dosztányi e colaboradores (2005), resíduos cujo valor de *score* esteja acima de 0,5 podem ser considerados desordenados.

Logo, considerando também a necessidade de extração de informações úteis contidas nos relatórios gerados pelo programa, desenvolvemos o *script* *iupred-parser-v3.pl*. Este *script* é capaz de abrir todos os arquivos gerados pelo IUPred, selecionar e extrair todas as informações referentes à identificação de regiões desordenadas, além de excluir todas as informações relacionadas às regiões globulares. Adicionalmente, assim como para os *scripts* que realizam o tratamento dos resultados descritos anteriormente, tais informações são armazenadas em um único arquivo que pode ser utilizado para carregar as informações geradas no BD *idp_crypto*.

5.3.4 ANCHOR

Este programa baseia-se nas propriedades físicas das regiões de ligação entre resíduos. Cálculos de energia de ligação são estimados pelo IUPred e segundo Dosztányi e colaboradores (2009), este cálculo também pode identificar regiões que são propensas a ganhar energia ao interagir com proteínas globulares. Assim, o ANCHOR combina a tendência geral do resíduo ser identificado como desordenado com a influência do ambiente em sua estrutura.

Para a execução deste preditor também foi necessário separar o proteoma predito de cada organismo em arquivos individuais contendo uma única sequência, uma vez que o programa solicita um arquivo de entrada por vez e no formato Fasta. Portanto, foi utilizado um *script* em *shell* para a execução das análises de todas as sequências, conforme descrito abaixo.

```
for i in `ls <arquivo>`; do anchor $i -d <caminho-do-diretório-dos-arquivos>><arquivo-de-saída> ; done &
```

Assim como o IUPred, o ANCHOR gera arquivos individuais para cada arquivo de sequência analisado e atribui aos resíduos um valor de *score* resultante da análise estatística realizada, que variam entre 0 e 1. Segundo Dosztányi e colaboradores (2005), resíduos cujo valor de *score* esteja acima de 0,5 podem ser considerados desordenados.

Similarmente à estratégia utilizada na análise dos dados gerados pelo IUPred, o *script iupred-parser-v3.pl* foi desenvolvido para extração dos resultados gerados e posterior carga de informações no BD *idp_crypto*.

5.4 Desordem proteica

Segundo Dunker e colaboradores (2002), a disponibilidade de informações genômicas tem demonstrado que a ocorrência de regiões desordenadas com tamanho significativo (maior ou igual a 40 resíduos de aminoácidos consecutivos) é surpreendentemente comum em proteínas funcionais e nos mais diversos organismos. Os autores também afirmam que a desordem proteica parece ser mais abundante em eucariotos do que em procariotos, ou seja, o número de IDPs e IDRs aumenta de acordo com a complexidade do organismo.

Corroborando esse relato, em nossas análises demonstramos que mesmo utilizando metodologias diferentes para identificação de desordem proteica, há um elevado conteúdo de proteínas desordenadas com pelo menos 40 resíduos de aminoácidos consecutivos. Além disso, observamos que os preditores selecionados para este estudo identificaram um conteúdo de 25-55% de proteínas com algum grau de desordem em relação ao proteoma predito de *Cryptococcus* spp.

Ao dividirmos os 12 organismos em dois grupos (não virulentos e virulentos) e considerarmos o grupo de organismos não virulentos, a menor porcentagem média

está relacionada ao conteúdo identificado pelo preditor ANCHOR, com 27,2% e a maior pelo preditor IUPred, com 53,9%. Já considerando o grupo de organismos virulentos a menor porcentagem média está relacionada ao conteúdo identificado pelo preditor REM465, com 23,4% e a maior pelo preditor IUPred, com 46,0% (Figura 5.4). Entretanto, o teste estatístico ANOVA apresentou um valor de $p \geq 0,05$, indicando não haver diferença estatística significativa.

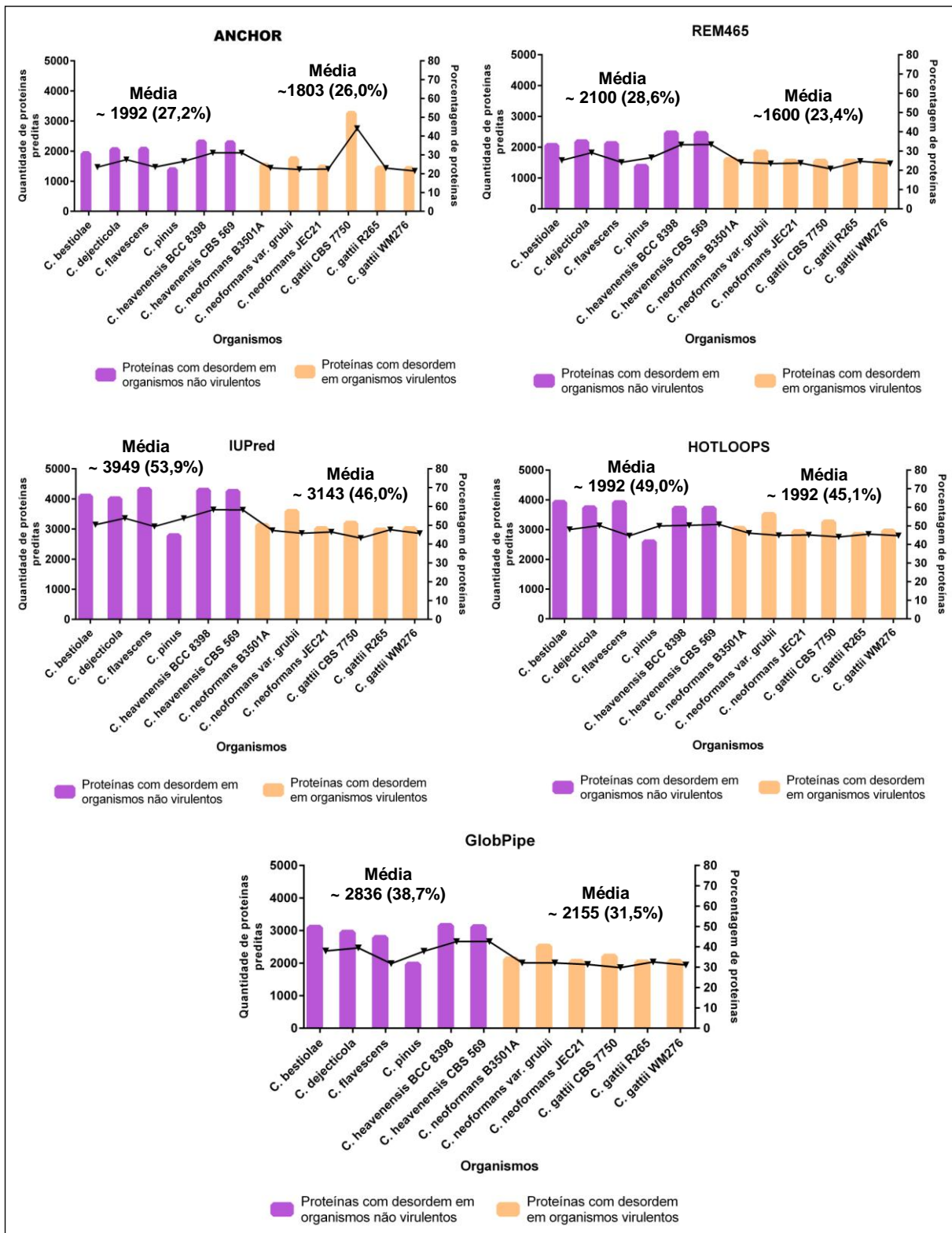


Figura 5.4: Gráficos relacionados ao conteúdo de desordem proteica identificado por cada preditor e em cada organismo. Nesses gráficos, as barras representam o conteúdo de proteínas, com pelo menos uma IDR, identificadas por cada preditor, sendo que as barras em lilás referem-se ao grupo de organismos não virulentos e as barras em laranja ao grupo de organismos virulentos. Esses dados estão relacionados ao eixo y principal (à esquerda). A média e a porcentagem média de proteínas identificadas para cada grupo também estão explicitados em cada situação. A linha, por sua vez, representa a porcentagem de proteínas identificadas pelos programas em relação ao conteúdo analisado, ou seja, ao proteoma de cada organismo e estão relacionadas ao eixo y secundário (à direita).

Após a realização das predições e da carga de seus resultados no BD *idp_crypto*, desenvolvemos o *script consenso-coord-idp-v4.pl* com o objetivo de encontrar o consenso das predições obtidas. Resumidamente, o *script* é capaz de realizar uma busca na tabela *predictions*, onde seleciona todas as informações referentes às proteínas e às coordenadas de início e fim de trechos identificados, pelos preditores selecionados, como desordenados. Após essa seleção, como já mencionado neste documento, o *script* realiza a comparação das coordenadas de cada proteína e, de acordo com o critério de 10% de sobreposição adotado, uma sequência consenso é formada. Caso esse critério não tenha sido satisfeito, as regiões foram consideradas como IDRs individuais. Após a execução desse *script*, o *script consenso-idp-coord--vs—proteome.pl* foi executado e posteriormente os dados gerados foram armazenados na tabela *consensus* do BD *idp_crypto*.

Esse critério de sobreposição mínima de 10% (relatado no item 4.5.2 de Materiais e Métodos) foi o critério adotado para geração das predições consenso. Critérios similares já foram descritos em Ruy (2011) e Torrieri (2010) que obtiveram bons resultados.

Após a etapa de geração do consenso das predições de IDPs e IDRs várias consultas utilizando a linguagem SQL foram realizadas no BD *idp_crypto*. Como resultado dessas buscas, observamos um conteúdo aproximado entre 60 e 70% de proteínas desordenadas contendo um domínio desordenado de pelo menos 40 resíduos de aminoácidos (Figura 5.5).

Resultados similares foram descritos em Tripanosomatídeos por Ruy (2011) que relatam a existência de um conteúdo médio que varia de 55 a 71% de proteínas preditas como desordenadas e utilizando os mesmos critérios de corte. Complementarmente e aplicando os mesmos critérios de classificação, Dunker e colaboradores (2002), utilizando somente as predições realizadas pelo PONDR, relatam um conteúdo de 35% de IDPs e IDRs para os cromossomos II e III do protozoário *P. falciparum* e de 40% para o proteoma da levedura *S. cerevisiae*.

O resultado pormenorizado para cada um dos genomas também pode ser avaliado na Figura 5.5 que mostra o conteúdo total de proteínas (em verde) e a porcentagem de proteínas totalmente e parcialmente desordenadas, IDPs e IDRs, respectivamente. Os valores obtidos por esse gráfico estão relacionados ao conteúdo proteico de cada organismo.

Adicionalmente, a Figura 5.5 também mostra o conteúdo relacionado às IDPs e IDRs preditas para cada organismo analisado. Vale ressaltar que o conteúdo de IDPs (proteínas totalmente desordenadas) varia entre 0,8 e 2,1% e, por outro lado, aproximadamente 65 a 70% dos produtos proteicos codificados pelos genes dos organismos analisados contêm pelo menos um domínio desordenado com 40 resíduos de aminoácidos consecutivos.

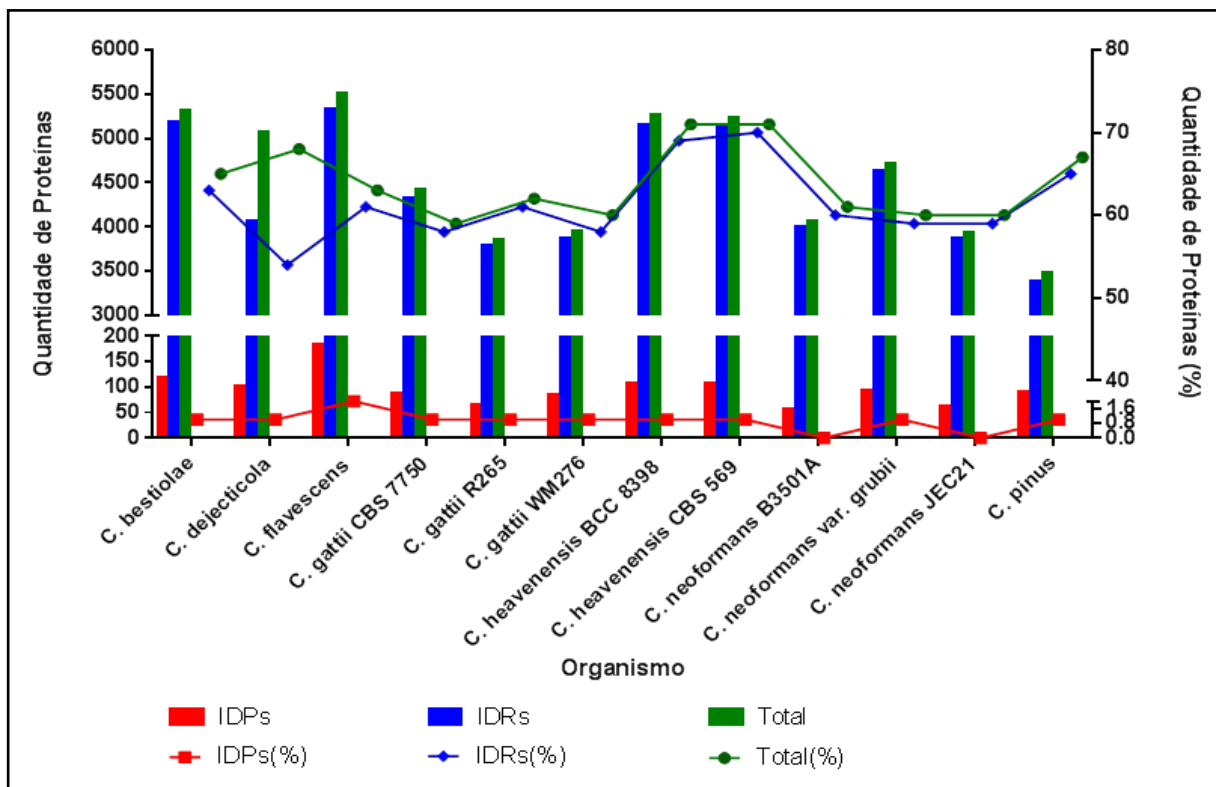


Figura 5.5: Gráfico relacionado às IDPs e IDRs obtidas pelo consenso das análises preditivas. Nesse gráfico, as barras representam o total de proteínas (em verde) obtidas a partir do consenso das análises preditivas. Já as barras em vermelho representam o número de IDPs identificado, enquanto as barras em verde representam o número de proteínas que apresentam, pelo menos, um domínio desordenado de acordo com o critério estabelecido. Esses dados estão relacionados ao eixo y principal (à esquerda). Já as linhas, por sua vez, representam a porcentagem desses dados em relação ao conteúdo analisado, ou seja, ao proteoma predito. Esses dados estão relacionados ao eixo y secundário (à direita).

Levando-se em consideração a análise de agrupamento, identificamos que dos 18.900 proteínas presentes no *core genome*, 11.409 proteínas apresentavam pelo menos uma IDR, correspondendo a 60,4% das proteínas compartilhadas pelos 12 organismos, conforme mostrado na Figura 5.6.

Além disso, também podemos observar na Figura 5.6 a quantidade de proteínas linhagem-específicas que possuem pelo menos uma IDR. Neste contexto, considerando o grupo de organismos virulentos e de não virulentos, podemos observar que o segundo grupo apresenta uma maior quantidade de proteínas espécie-específicas.

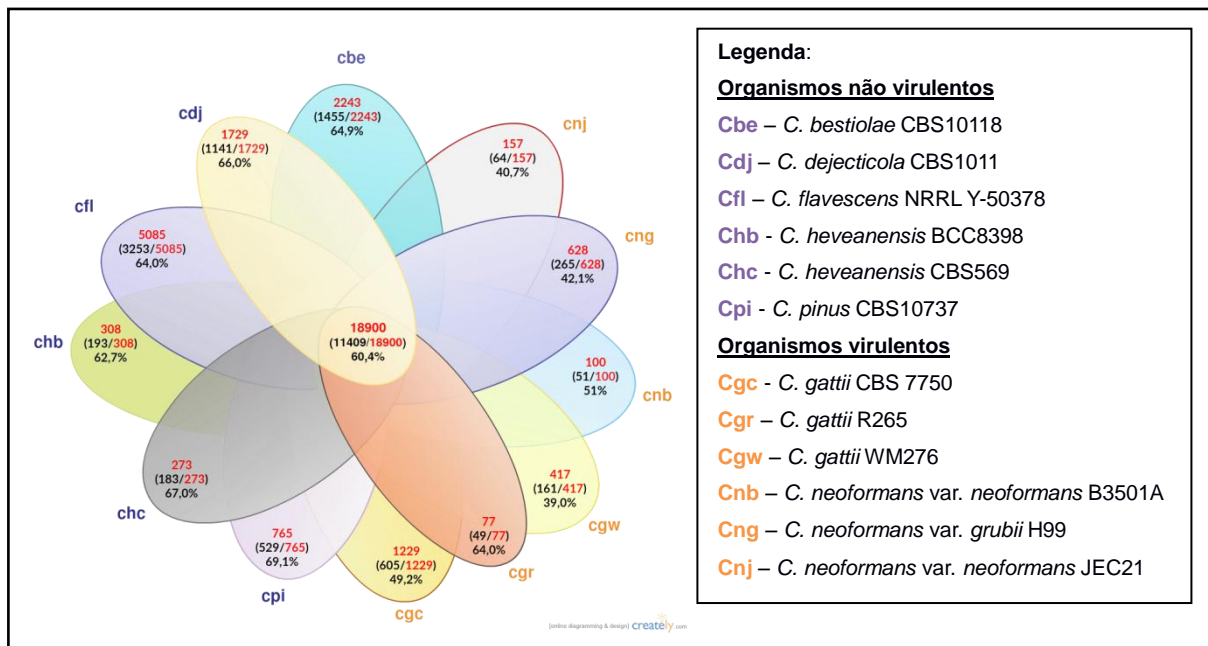


Figura 5.6: Diagrama de venn referente à análise de clusterização. Neste diagrama podemos observar o número de genes relacionados ao *core genome*, além do número de genes relatados para cada espécie. Em vermelho está destacado o número total de proteínas agrupadas em cada situação e em preto o número de proteínas identificadas com pelo menos uma IDR e porcentagem correspondente. **Fonte:** Acervo Pessoal.

Sendo assim, considerando o grupo de organismos não virulentos e as proteínas linhagem-específicas, podemos observar que o *C. pinus* apresentou maior proporção de proteínas com pelo menos uma IDR (69,1%), enquanto que o *C. heveanensis*, linhagem BCC8398, apresentou menor proporção (62,7%). Já considerando o grupo de organismos virulentos, observamos que o *C. gattii*, linhagem R265, apresentou maior proporção (64,0%) e o *C. gattii*, linhagem WM276 apresentou a menor proporção (39%).

5.4.1 Análise de enriquecimento funcional

Após a análise referente ao conteúdo relacionado às IDPs e IDRs preditas para cada organismo deste estudo, selecionamos os dados referentes aos organismos analisados.

Diante desses dados, foi realizada uma análise de enriquecimento funcional entre os termos anotadores das sequências relacionadas ao GO para cada um dos organismos. Essa metodologia, como mencionado na seção 4.4.4 de Materiais e Métodos, utiliza um teste estatístico, conhecido como Teste Bilateral de Fischer com Correção de Bonferroni. Além disso, foi utilizado um critério de corte de p-valor ajustado menor que 0,05 (Korpelainen et al. 2015).

Essa abordagem permitiu a identificação e um melhor entendimento sobre as funções mais presentes dentro do conjunto de proteínas com desordem proteica, quando esse é comparado aos dados de referência, que nesse caso são as proteínas presentes no proteoma predito de cada organismo (Korpelainen et al., 2015).

Assim como já relatado por vários autores, as IDPs e IDRs apresentam um papel funcional importante em vários processos biológicos, como a regulação da transcrição, tradução e transdução de sinais (Wright e Dyson 1999; Brown et al. 2002; Radivojac et al. 2004).

As Figura A. 1 a Figura A. 6, presentes no Apêndice A deste documento, referem-se à categoria do GO conhecida como “Processo biológico”. Analisando esses gráficos, observamos que todos os termos, identificados para essa categoria, são enriquecidos com proteínas que contêm pelo menos 40 resíduos de aminoácidos consecutivos preditos como desordenados para os organismos classificados como virulentos. Podemos destacar que os termos relacionados à regulação da transcrição e fosforilação de proteínas, são altamente enriquecidos e estão presentes nos seis organismos analisados.

Além disso, as Figura B. 1 a Figura B. 6, presentes no Apêndice B, também estão relacionados à categoria GO “Processo biológico”. Entretanto, esses gráficos referem-se aos organismos analisados que são identificados como não virulentos. Analisando esses resultados, podemos perceber que, assim como para o grupo dos organismos não virulentos, os termos relacionados à transcrição e fosforilação de proteínas estão entre os mais enriquecidos do grupo.

5.4.2 IDPs e IDRs relacionadas aos fatores de virulência de *Cryptococcus* spp.

A busca no BD *UniProt* pelas palavras-chave *Cryptococcus* e *virulence*, resultou em 7.608 sequências, com as quais foi criado um BD de fatores de virulência de *Cryptococcus*. Esse resultado mostra que, além das proteínas diretamente relacionadas aos principais fatores de virulência, outras proteínas relacionadas à virulência também compõem esse BD e que será disponibilizado no Material Suplementar do artigo a ser submetido.

A identificação do conteúdo de desordem proteica nos proteomas de *C. gattii* e *C. neoformans* também possibilitou a identificação, através de buscas por similaridade de sequências, de proteínas que estão relacionadas aos fatores de

virulência presentes no BD mencionado acima. Do resultado obtido, foram consideradas somente as proteínas com no mínimo 80% de cobertura e 80% de similaridade.

Posteriormente, em uma análise detalhada desses dados, observamos que dentro do conteúdo de IDPs e IDRs identificado por este estudo, não há proteínas relacionadas à termotolerância (gene CNA1). Entretanto, foi encontrada para cada organismo, pelo menos uma proteína relacionada à Família CAP (Tabela 7), que por sua vez, está diretamente relacionada à formação da cápsula polissacarídica.

Outras proteínas que também são capazes de regular a síntese da cápsula em *Cryptococcus* spp., foram descritas por Jung e colaboradores (2006). Esses autores identificaram em *C. neoformans* que a sensibilidade da levedura frente à quantidade de ferro no ambiente é mediada por proteínas que apresentam motivos conservados ricos em Cisteína e/ou proteínas classificadas como *Zinc finger*. Em nossas análises também observamos para as espécies de *C. gattii* e *C. neoformans*, que há pelo menos duas proteínas funcionalmente identificadas como *Zinc finger* e que apresentam no mínimo um domínio funcional com 40 resíduos de aminoácidos consecutivos estruturalmente desordenados. Esse resultado também está descrito na Tabela 5.6.

Já para a família LAC, somente foram encontradas duas proteínas, sendo uma relacionada à espécie *C. neoformans* var. *grubii* H99 e outra relacionada à espécie *C. gattii*, linhagem R265. Como já mencionado neste documento, proteínas codificadas por este gene possuem uma relação direta com a síntese de melanina, que por sua vez, confere proteção contra estresses oxidativos, provenientes da ação de antifúngicos como a Anfotericina B e o Fluconazol, como de células fagocitárias (Almeida et al., 2015).

Tabela 5.6: Número de IDPs e IDRs preditas para as espécies *C. neoformans* e *C. gattii* identificadas entre os fatores de virulência

Organismo	Linhagem	Família CAP	Família LAC	Transportadores ABC	Zinc Finger
<i>C. neoformans</i> var. <i>neoformans</i>	JEC21	1	-	20	3
<i>C. neoformans</i> var. <i>grubii</i>	H99	1	1	21	5
<i>C. neoformans</i> var. <i>neoformans</i>	B-3501 ^a	1	-	21	3
<i>C. gattii</i>	WM276	1	-	16	4
<i>C. gattii</i>	CBS 7750	2	-	14	2
<i>C. gattii</i>	R265	1	1	17	2

Adicionalmente, também é interessante ressaltar que há um número considerável de proteínas (preditas com no mínimo 40 resíduos de aminoácidos consecutivos), que são funcionalmente identificadas como Transportadores do tipo *ATP Binding Cassette* (ABC) e estão presentes em todas as linhagens analisadas (Tabela 5.6). Esses Transportadores ABC são proteínas transmembrana envolvidas na proteção das células contra compostos tóxicos, e estão relacionadas a uma forma de defesa contra substâncias antimicrobianas produzidas pelas células hospedeiras fagocíticas e pelos antifúngicos. Goulart e colaboradores (2010) observaram em *C. neoformans*, que ao ser fagocitada, a levedura é capaz de expressar a proteína *adrenoleukodystrophy*, que é funcionalmente identificada como um Transportador ABC.

De forma complementar, na análise das vias metabólicas relacionadas ao grupo de organismos virulentos, identificamos que há um número considerável de IDPs/IDRs que além de estarem atuando diretamente ou indiretamente como um fator de virulência, também estão atuando em diversas vias metabólicas, como pode ser visualizado nas Tabela C. 1 a Tabela C. 6, presentes no Apêndice C deste documento. Além disso, podemos destacar que dentre as vias identificadas, 33 são comuns a todas as linhagens desse grupo de organismos (Tabela 5.7) e pelo menos uma proteína de cada um está atuando como uma enzima nessas vias metabólicas.

Tabela 5.7: Relação de vias metabólicas comum a todos os organismos virulentos

Via metabólica	Mapa KEGG
<i>Alanine, aspartate and glutamate metabolism</i>	map00250
<i>Amino sugar and nucleotide sugar metabolism</i>	map00520
<i>Aminoacyl-tRNA biosynthesis</i>	map00970
<i>Aminobenzoate degradation</i>	map00627
<i>Arginine biosynthesis</i>	map00220
<i>Biosynthesis of antibiotics</i>	map01130
<i>Butanoate metabolism</i>	map00650
<i>C5-Branched dibasic acid metabolism</i>	map00660
<i>Carbon fixation pathways in prokaryotes</i>	map00720
<i>Citrate cycle (TCA cycle)</i>	map00020
<i>Folate biosynthesis</i>	map00790
<i>Fructose and mannose metabolism</i>	map00051
<i>Glycine, serine and threonine metabolism</i>	map00260
<i>Glycolysis / Gluconeogenesis</i>	map00010
<i>Inositol phosphate metabolism</i>	map00562
<i>Lysine degradation</i>	map00310
<i>Methane metabolism</i>	map00680
<i>mTOR signaling pathway</i>	map04150
<i>N-Glycan biosynthesis</i>	map00510
<i>Nitrogen metabolism</i>	map00910
<i>One carbon pool by folate</i>	map00670
<i>Pantothenate and CoA biosynthesis</i>	map00770
<i>Pentose phosphate pathway</i>	map00030
<i>Phosphatidylinositol signaling system</i>	map04070
<i>Propanoate metabolism</i>	map00640
<i>Purine metabolism</i>	map00230
<i>Pyrimidine metabolism</i>	map00240
<i>Pyruvate metabolism</i>	map00620
<i>Riboflavin metabolism</i>	map00740
<i>Starch and sucrose metabolism</i>	map00500
<i>Sulfur metabolism</i>	map00920
<i>T cell receptor signaling pathway</i>	map04660
<i>Valine, leucine and isoleucine biosynthesis</i>	map00290

Considerando este conjunto de vias metabólicas, destacamos a via de biossíntese de N-glycanos (do inglês *N-Glycan biosynthesis*). Esses compostos são de extrema importância para a virulência do *Cryptococcus* spp., uma vez que participam da síntese de GXM, um importante componente para a formação capsular, como relatado no item 1.4.3 da Introdução deste documento. Como exemplo, selecionamos a via relacionada à espécie *C. neoformans* var. *grubii*, linhagem H99 que apresenta duas proteínas que atuam como quinase (*EC number*: 2.7.1.108) nessa via, que está destacado em rosa (Figura 5.7).

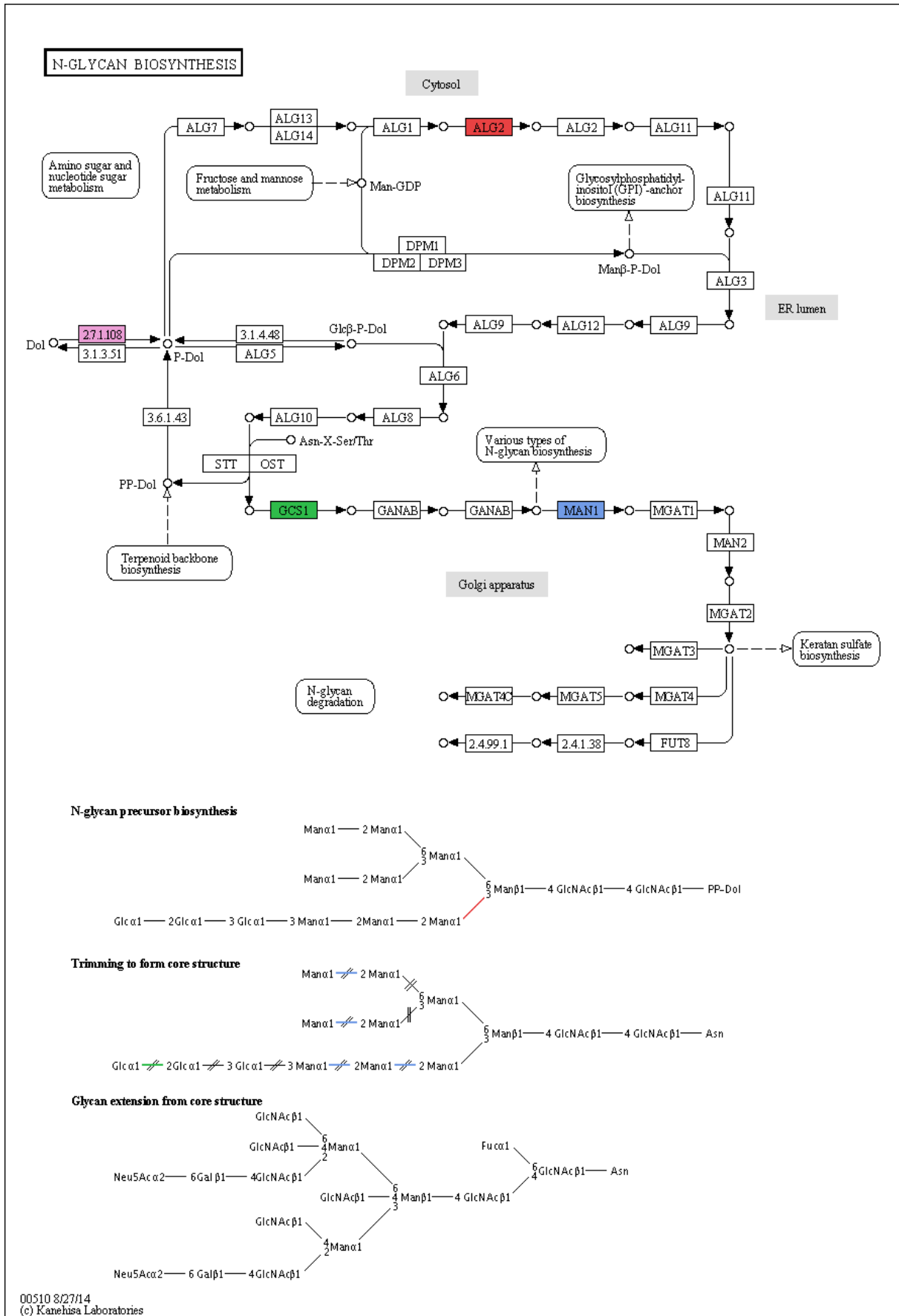


Figura 5.7: Via de Biossíntese de N-glicanos. Nessa via metabólica, foram identificadas pelo mapeamento das proteínas estruturalmente desordenadas preditas neste trabalho, duas proteínas que atuam como quinases (EC number: :2.7.1.108), destacada em rosa. **Fonte:** Acervo Pessoal.

6 CONCLUSÃO

Os resultados obtidos ao final deste estudo nos permitiram concluir que:

- a) O banco de dados relacional desenvolvido facilitou a identificação dos dados obtidos pelas diferentes análises preditivas, permitindo buscas específicas e de forma acurada para a geração de sequências consenso e dos dados obtidos após essa análise;
- b) as espécies de *Cryptococcus* spp apresentam um conteúdo de proteínas estruturalmente desordenadas entre 60 e 70% em relação ao proteoma predito. Adicionalmente, o conteúdo de IDPs, como já era esperado, é menor quando comparado ao conteúdo de proteínas com pelo menos uma IDR;
- c) entre as espécies virulentas de *Cryptococcus* analisadas no presente estudo, há uma grande quantidade de IDPs e IDRs que têm papel funcional vital em vários processos biológicos.
- d) a análise de IDPs e IDRs com os fatores de virulência permitiu identificar um conteúdo significativo dessas proteínas que estão relacionadas aos principais fatores de virulência;
- e) IDPs e IDRs foram identificadas atuando como enzimas em vias metabólicas importantes ligadas aos fatores de virulência do organismo.

7 PERSPECTIVAS

A análise de desempenho para esse conjunto de preditores é uma etapa importante na comparação dos resultados obtidos por cada programa, uma vez que utilizaram diferentes metodologias para identificação de IDPs e IDRs. Além disso, esse tipo de análise também permite identificar a melhor combinação de programas a ser utilizado para identificação dessas proteínas. Por esse motivo, essa análise começou a ser realizada neste estudo, mas devido ao tempo para finalização deste trabalho, essa análise não pôde ser concluída em tempo hábil. Entretanto, ela não deixará de ser realizada, uma vez que o conjunto de proteínas e suas regiões, identificadas experimentalmente como desordenadas (presentes no BD do PDB e do DisProt), já foi selecionado.

Adicionalmente, diante de todas as análises que foram realizadas para este trabalho, a construção de um *pipeline* também é desejada, uma vez que ele permitirá a execução automática de todos os processos realizados individualmente.

Outro ponto que poderá ser explorado através de um estudo mais detalhado dos dados é a associação do tamanho da desordem proteica em proteínas não identificadas como IDPs, ou seja, totalmente desordenadas, e assim um perfil desse conteúdo poderá ser melhor analisado.

Além disso, com base nos objetivos traçados para este estudo, conseguimos identificar uma relação entre IDPs e/ou IDRs com as proteínas que atuam nos fatores de virulência de *Cryptococcus* spp. Assim, tais proteínas poderão ser alvos de estudos para o desenvolvimento de novas drogas antifúngicas ou para a identificação de resistência dos organismos aos azóis (antifúngicos) conhecidos.

8 REFERÊNCIAS BIBLIOGRÁFICAS

- Almeida F, Wolf JM, Casadevall A. Virulence-Associated Enzymes of *Cryptococcus neoformans*. *Eukaryot Cell*. 1 de dezembro de 2015;14(12):1173–85.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 5 de outubro de 1990;215(3):403–10.
- Alvarez M, Casadevall A. Phagosome extrusion and host-cell survival after *Cryptococcus neoformans* phagocytosis by macrophages. *Curr Biol CB*. 7 de novembro de 2006;16(21):2161–5.
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology: tool for the unification of biology. *Nat Genet*. Maio de 2000;25(1):25–9.
- Babu MM, van der Lee R, de Groot NS, Gsponer J. Intrinsically disordered proteins: regulation and disease. *Curr Opin Struct Biol*. junho de 2011;21(3):432–40.
- Bernstein FC, Koetzle TF, Williams GJ, Meyer EF, Brice MD, Rodgers JR, et al. The Protein Data Bank: a computer-based archival file for macromolecular structures. *J Mol Biol*. 25 de maio de 1977;112(3):535–42.
- Blaya C, Prati C, Boneti C, Bonamigo DR, Krumenauer RCP, Rosito GA. Análise da utilização dos novos inibidores da trombina na prática médica. *Arq Bras Cardiol*. agosto de 1998;71(2):163–7.
- Brown CJ, Takayama S, Campen AM, Vise P, Marshall TW, Oldfield CJ, et al. Evolutionary Rate Heterogeneity in Proteins with Long Disordered Regions. *J Mol Evol*. julho de 2002;55(1):104–10.
- Byrnes EJ, Bartlett KH, Perfect JR, Heitman J. *Cryptococcus gattii*: an emerging fungal pathogen infecting humans and animals. *Microbes Infect Inst Pasteur*. outubro de 2011;13(11):895–907.
- Chaturvedi AK, Wormley FL. *Cryptococcus* antigens and immune responses: implications for a vaccine. *Expert Rev Vaccines*. novembro de 2013;12(11):1261–72.
- Chen Y-L, Lehman VN, Lewit Y, Averette AF, Heitman J. Calcineurin Governs Thermotolerance and Virulence of *Cryptococcus gattii*. *G3 Genes Genomes Genetics*. 1 de março de 2013;3(3):527–39.
- Cheng J, Sweredoski MJ, Baldi P. Accurate Prediction of Protein Disordered Regions by Mining Protein Structure Data. *Data Min Knowl Discov*. 14 de julho de 2005;11(3):213–22.
- Cheng Y, LeGall T, Oldfield CJ, Mueller JP, Van Y-YJ, Romero P, et al. Rational drug design via intrinsically disordered protein. *Trends Biotechnol*. outubro de 2006;24(10):435–42.

Cilia E, Pancsa R, Tompa P, Lenaerts T, Vranken WF. From protein sequence to dynamics and disorder with DynaMine. *Nat Commun* [Internet]. 14 de novembro de 2013;4. Recuperado de: <http://www.nature.com/ncomms/2013/131114/ncomms3741/full/ncomms3741.html?message-global=remove>

Cogliati M. Global Molecular Epidemiology of *Cryptococcus neoformans* and *Cryptococcus gattii*: An Atlas of the Molecular Types. *Scientifica* [Internet]. 2013;2013. Recuperado de: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3820360/>

Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*. 15 de setembro de 2005;21(18):3674–6.

Dosztányi Z, Csizmok V, Tompa P, Simon I. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics*. 15 de agosto de 2005;21(16):3433–4.

Dosztányi Z, Mészáros B, Simon I. ANCHOR: web server for predicting protein binding regions in disordered proteins. *Bioinforma Oxf Engl*. 15 de outubro de 2009;25(20):2745–6.

Dunker AK, Babu MM, Barbar E, Blackledge M, Bondos SE, Dosztányi Z, et al. What's in a name? Why these proteins are intrinsically disordered. *Intrinsically Disord Proteins*. 1 de janeiro de 2013;1(1):e24157.

Dunker AK, Brown CJ, Lawson JD, Iakoucheva LM, Obradović Z. Intrinsic disorder and protein function. *Biochemistry (Mosc)*. 28 de maio de 2002;41(21):6573–82.

Dunker AK, Lawson JD, Brown CJ, Williams RM, Romero P, Oh JS, et al. Intrinsically disordered protein. *J Mol Graph Model*. 2001;19(1):26–59.

Dunker AK, Oldfield CJ, Meng J, Romero P, Yang JY, Chen JW, et al. The unfoldomics decade: an update on intrinsically disordered proteins. *BMC Genomics*. 16 de setembro de 2008a;9(Suppl 2):S1.

Dunker AK, Oldfield CJ, Meng J, Romero P, Yang JY, Chen JW, et al. The unfoldomics decade: an update on intrinsically disordered proteins. *BMC Genomics*. 16 de setembro de 2008b;9(Suppl 2):S1.

Dyson HJ, Wright PE. Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol*. março de 2005;6(3):197–208.

Fang FC. Antimicrobial reactive oxygen and nitrogen species: concepts and controversies. *Nat Rev Microbiol*. outubro de 2004;2(10):820–32.

Feng Z-P, Zhang X, Han P, Arora N, Anders RF, Norton RS. Abundance of intrinsically unstructured proteins in *P. falciparum* and other apicomplexan parasite proteomes. *Mol Biochem Parasitol*. dezembro de 2006;150(2):256–67.

Fischer E. Einfluss der Configuration auf die Wirkung der Enzyme. *Berichte Dtsch Chem Ges*. 1 de outubro de 1894;27(3):2985–93.

Franco-Paredes C, Womack T, Bohlmeier T, Sellers B, Hays A, Patel K, et al. Management of *Cryptococcus gattii* meningoencephalitis. *Lancet Infect Dis*. março de 2015;15(3):348–55.

Goulart L, Silva LKRE, Chiapello L, Silveira C, Crestani J, Masih D, et al. *Cryptococcus neoformans* and *Cryptococcus gattii* genes preferentially expressed during rat macrophage infection. *Med Mycol*. 1 de novembro de 2010;48(7):932–41.

Gullo FP, Rossi SA, Sardi J de CO, Teodoro VLI, Mendes-Giannini MJS, Fusco-Almeida AM. Cryptococcosis: epidemiology, fungal resistance, and new alternatives for treatment. *Eur J Clin Microbiol Infect Dis Off Publ Eur Soc Clin Microbiol*. novembro de 2013;32(11):1377–91.

Habchi J, Tompa P, Longhi S, Uversky VN. Introducing Protein Intrinsic Disorder. *Chem Rev*. 9 de julho de 2014;114(13):6561–88.

Hansen JC, Lu X, Ross ED, Woody RW. Intrinsic Protein Disorder, Amino Acid Composition, and Histone Terminal Domains. *J Biol Chem*. 27 de janeiro de 2006;281(4):1853–6.

He B, Wang K, Liu Y, Xue B, Uversky VN, Dunker AK. Predicting intrinsic disorder in proteins: an overview. *Cell Res*. agosto de 2009;19(8):929–49.

Huang T, He Z-S, Cui W-R, Cai Y-D, Shi X-H, Hu L-L, et al. A sequence-based approach for predicting protein disordered regions. *Protein Pept Lett*. março de 2013;20(3):243–8.

Jones DT, Cozzetto D. DISOPRED3: precise disordered region predictions with annotated protein-binding activity. *Bioinformatics*. 15 de março de 2015;31(6):857–63.

Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, et al. InterProScan 5: genome-scale protein function classification. *Bioinformatics*. 1 de maio de 2014;30(9):1236–40.

Jung WH, Sham A, White R, Kronstad JW. Iron Regulation of the Major Virulence Factors in the AIDS-Associated Pathogen *Cryptococcus neoformans*. *PLoS Biol*. 21 de novembro de 2006;4(12):e410.

Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M. The KEGG resource for deciphering the genome. *Nucleic Acids Res*. 1 de janeiro de 2004;32(Database issue):D277–80.

Karush F. Heterogeneity of the Binding Sites of Bovine Serum Albumin¹. *J Am Chem Soc*. 1 de junho de 1950;72(6):2705–13.

Korpelainen E, Tuimala J, Somervuo P, Huss M, Wong G. *RNA-seq Data Analysis A Practical Approach*. CRC Press; 2015. (CHAPMAN & HALL/CRC Mathematical and Computational Biology Series).

Koshland DE. Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proc Natl Acad Sci U S A*. fevereiro de 1958;44(2):98–104.

van der Lee R, Buljan M, Lang B, Weatheritt RJ, Daughdrill GW, Dunker AK, et al. Classification of Intrinsically Disordered Regions and Proteins. *Chem Rev.* 9 de julho de 2014;114(13):6589–631.

Li SS, Mody CH. *Cryptococcus*. *Proc Am Thorac Soc.* 15 de maio de 2010;7(3):186–96.

Lin X, Heitman J. The biology of the *Cryptococcus neoformans* species complex. *Annu Rev Microbiol.* 2006;60:69–105.

Linding R, Jensen LJ, Diella F, Bork P, Gibson TJ, Russell RB. Protein Disorder Prediction. *Structure.* 11 de janeiro de 2003a;11(11):1453–9.

Linding R, Russell RB, Neduva V, Gibson TJ. GlobPlot: exploring protein sequences for globularity and disorder. *Nucleic Acids Res.* 1 de julho de 2003b;31(13):3701–8.

Ma H, May RC. Virulence in *Cryptococcus* Species. In: Laskin AI, Sariaslani S, Gadd GM, organizadores. *Adv Appl Microbiol* [Internet]. Burlington: Elsevier Inc.; 2009. p. 131–90. (Academic Press, 2009; vol. 67). Recuperado de: <http://www.sciencedirect.com/science/article/pii/S0065216408010058>

Mendes FES, Oliveira LVN, Faria ES, Alvarenga DG, Pinto MR, Taborda CP, et al. Correlation of the in vitro antifungal drug susceptibility with the in vivo activity of fluconazole in a murine model of cerebral infection caused by *Cryptococcus gattii*. *Eur J Clin Microbiol Infect Dis.* 28 de agosto de 2010;29(12):1525–32.

Mitchell TG, Perfect JR. Cryptococcosis in the era of AIDS--100 years after the discovery of *Cryptococcus neoformans*. *Clin Microbiol Rev.* outubro de 1995;8(4):515–48.

Nishikawa MM, Lazera MS, Barbosa GG, Trilles L, Balassiano BR, Macedo RCL, et al. Serotyping of 467 *Cryptococcus neoformans* Isolates from Clinical and Environmental Sources in Brazil: Analysis of Host and Regional Patterns. *J Clin Microbiol.* janeiro de 2003;41(1):73–7.

Odom A, Muir S, Lim E, Toffaletti DL, Perfect J, Heitman J. Calcineurin is required for virulence of *Cryptococcus neoformans*. *EMBO J.* 15 de maio de 1997;16(10):2576–89.

Oldfield CJ, Dunker AK. Intrinsically Disordered Proteins and Intrinsically Disordered Protein Regions. *Annu Rev Biochem.* 2014;83(1):553–84.

Orsi CF, Colombari B, Ardizzoni A, Peppoloni S, Neglia R, Posteraro B, et al. The ABC transporter-encoding gene *AFR1* affects the resistance of *Cryptococcus neoformans* to microglia-mediated antifungal activity by delaying phagosomal maturation. *FEMS Yeast Res.* março de 2009;9(2):301–10.

Peng K, Radivojac P, Vucetic S, Dunker AK, Obradovic Z. Length-dependent prediction of protein intrinsic disorder. *BMC Bioinformatics.* 2006;7:208.

Radivojac P, Obradovic Z, Smith DK, Zhu G, Vucetic S, Brown CJ, et al. Protein flexibility and intrinsic disorder. *Protein Sci Publ Protein Soc.* janeiro de 2004;13(1):71–80.

Russell RB, Gibson TJ. A careful disorderliness in the proteome: Sites for interaction and targets for future therapies. *FEBS Lett.* 9 de abril de 2008;582(8):1271–5.

Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, et al. Artemis: sequence visualization and annotation. *Bioinforma Oxf Engl.* outubro de 2000;16(10):944–5.

Ruy P de C. Análise computacional baseada no desenvolvimento de um pipeline de técnicas ab initio para predição de desordem estrutural protéica em genomas de tripanosomatídeo [Internet]. [Belo Horizonte - Minas Gerais]: Centro de pesquisas René Rachou -Fundacao Oswaldo Cruz; 2011. Recuperado de: http://www.cpqrr.fiocruz.br/texto-completo/D_64.pdf

Santos JRA. Influência da susceptibilidade ao fluconazol na progressão da criptococose experimental causada por *Cryptococcus gattii*. [Belo Horizonte - Minas Gerais]: Universidade Federal de Minas Gerais - UFMG; 2015.

Santos JRA, Gouveia LF, Taylor ELS, Resende-Stoianoff MA, Pianetti GA, César IC, et al. Dynamic Interaction between Fluconazole and Amphotericin B against *Cryptococcus gattii*. *Antimicrob Agents Chemother.* maio de 2012;56(5):2553–8.

Schoffelen T, Illnait-Zaragozi M-T, Joosten LAB, Netea MG, Boekhout T, Meis JF, et al. *Cryptococcus gattii* Induces a Cytokine Pattern That Is Distinct from Other Cryptococcal Species. *PLoS ONE* [Internet]. 31 de janeiro de 2013; 8(1). Recuperado de: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3561320/>

Silva PR da. SUSCEPTIBILIDADE A ANTIFÚNGICOS E RESPOSTA IMUNOLÓGICA IN VITRO ÀS VARIEDADES DO *Cryptococcus neoformans*. [Internet] [Tese]. [Uberaba - MG]: Universidade Federal do Triângulo Mineiro; 2006. Recuperado de: http://www.uftm.edu.br/pgcs/imagem/Tese_PauloRobertoSilvaME.pdf

Srikanta D, Santiago-Tirado FH, Doering TL. *Cryptococcus neoformans*: historical curiosity to modern pathogen. *Yeast Chichester Engl.* fevereiro de 2014;31(2):47–60.

Stanke M, Morgenstern B. AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Res.* 1 de julho de 2005;33(Web Server issue):W465-467.

The Gene Ontology Consortium. Gene Ontology Consortium: going forward. *Nucleic Acids Res.* 28 de janeiro de 2015;43(D1):D1049–56.

Tompa P. Structure and Function of Intrinsically Disordered Proteins. Vol. 1. Chapman & Hall/CRC Taylor & Francis Group; 2010.

Tompa P. Intrinsically disordered proteins: a 10-year recap. *Trends Biochem Sci.* dezembro de 2012;37(12):509–16.

Torrieri R. Identificação e caracterização computacional de proteínas do tipo IUP no proteoma predito de *Schistosoma mansoni* [Internet] [Dissertação de Mestrado]. [Belo Horizonte - Minas Gerais]: Centro de Pesquisas René Rachou - Fundação Oswaldo Cruz; 2010. Recuperado de: http://www.cpqrr.fiocruz.br/texto-completo/D_49.pdf

Uversky VN. A decade and a half of protein intrinsic disorder: Biology still waits for physics. *Protein Sci Publ Protein Soc.* junho de 2013;22(6):693–724.

Uversky VN, Oldfield CJ, Dunker AK. Intrinsically disordered proteins in human diseases: introducing the D2 concept. *Annu Rev Biophys.* 2008;37:215–46.

Voelz K, May RC. Cryptococcal Interactions with the Host Immune System. *Eukaryot Cell.* junho de 2010;9(6):835–46.

Walsh I, Martin AJM, Di Domenico T, Tosatto SCE. ESpritz: accurate and fast prediction of protein disorder. *Bioinforma Oxf Engl.* 15 de fevereiro de 2012;28(4):503–9.

Wright PE, Dyson HJ. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J Mol Biol.* 22 de outubro de 1999;293(2):321–31.

APENDICE A – GRÁFICOS RELACIONADOS À ANÁLISE DE ENRIQUECIMENTO ENTRE AS ESPÉCIES/LINHAGENS VIRULENTAS

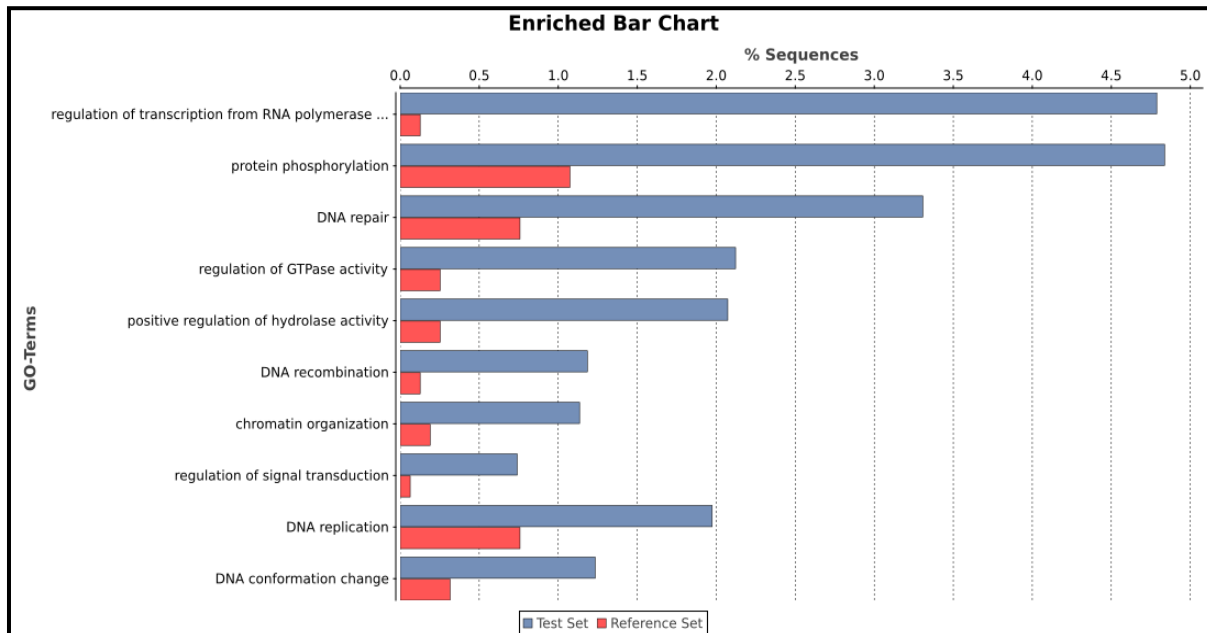


Figura A. 1: Análise de enriquecimento de desordem proteica em *C. gattii*, linhagem WM276. Este gráfico foi obtido a partir da análise de enriquecimento referente ao conteúdo total de desordem proteica dentro do proteoma de *C. gattii* linhagem WM276. Em azul, está representado o conteúdo de proteínas desordenadas contidas nos termos anotadores. Em vermelho, está representado o conteúdo de proteínas globulares presentes no proteoma.

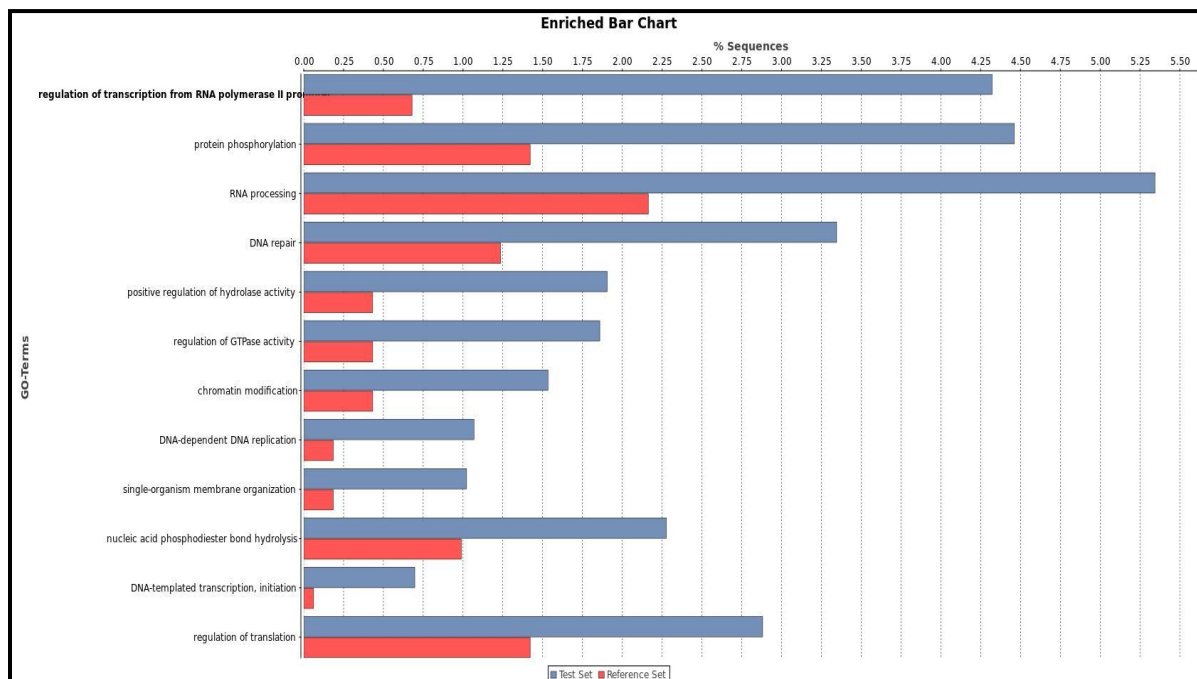


Figura A. 2: Gráfico obtido a partir da análise de enriquecimento referente ao conteúdo total de desordem proteica dentro do proteoma de *C. gattii*, linhagem R265. Em azul, está representado o conteúdo de proteínas desordenadas contidas nos termos anotadores. Em vermelho, está representado o conteúdo de proteínas globulares presentes no proteoma.

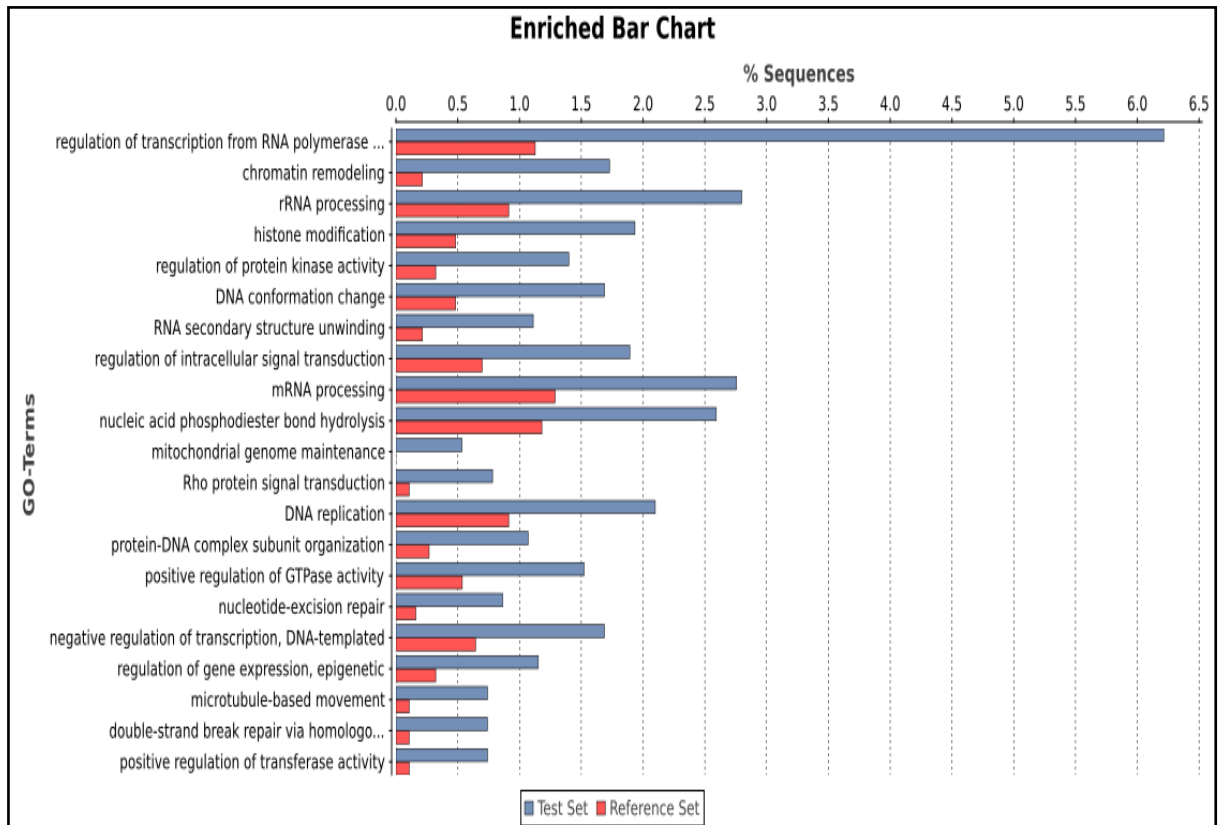


Figura A. 3: Gráfico referente à Análise de enriquecimento de desordem proteica em *C. neoformans*, linhagem B3501A. Este gráfico foi obtido a partir da análise de enriquecimento referente ao conteúdo total de desordem proteica dentro do proteoma de *C. neoformans*, linhagem B3501A. Em azul, está representado o conteúdo de proteínas desordenadas contidas nos termos anotadores. Em vermelho, está representado o conteúdo de proteínas globulares presentes no proteoma

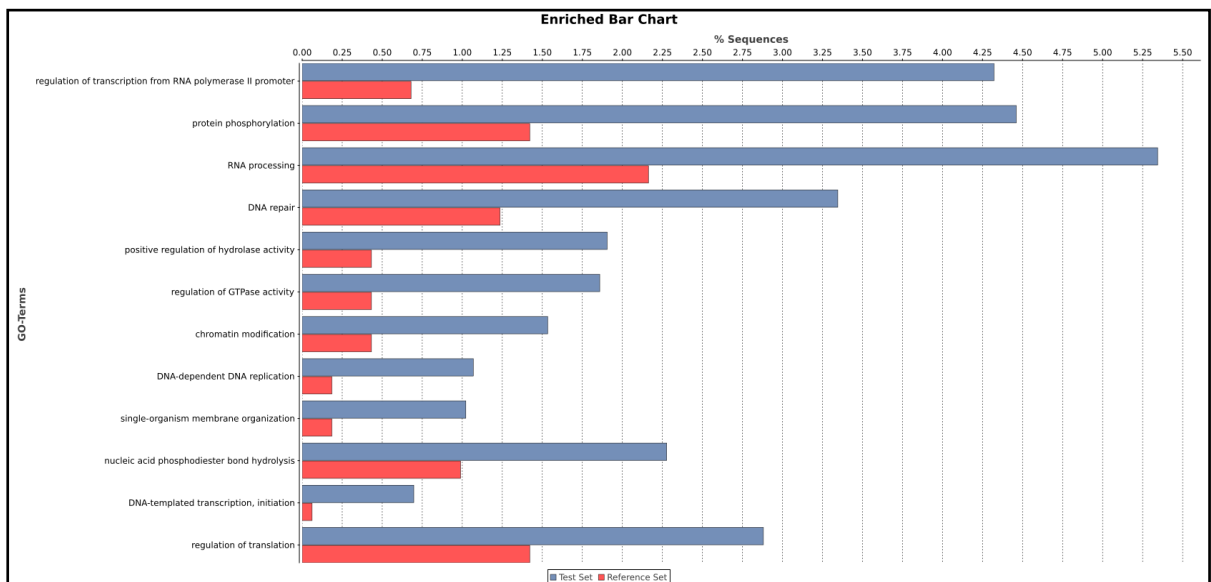


Figura A. 4: Gráfico referente à análise de enriquecimento de desordem proteica em *C. gattii*, linhagem CBS7750. Este gráfico obtido a partir da análise de enriquecimento referente ao conteúdo total de desordem proteica dentro do proteoma predito de *C. gattii* linhagem CBS7750. Em azul, está representado o conteúdo de proteínas desordenadas contidas nos termos anotadores. Em vermelho, está representado o conteúdo de proteínas globulares presentes no proteoma.

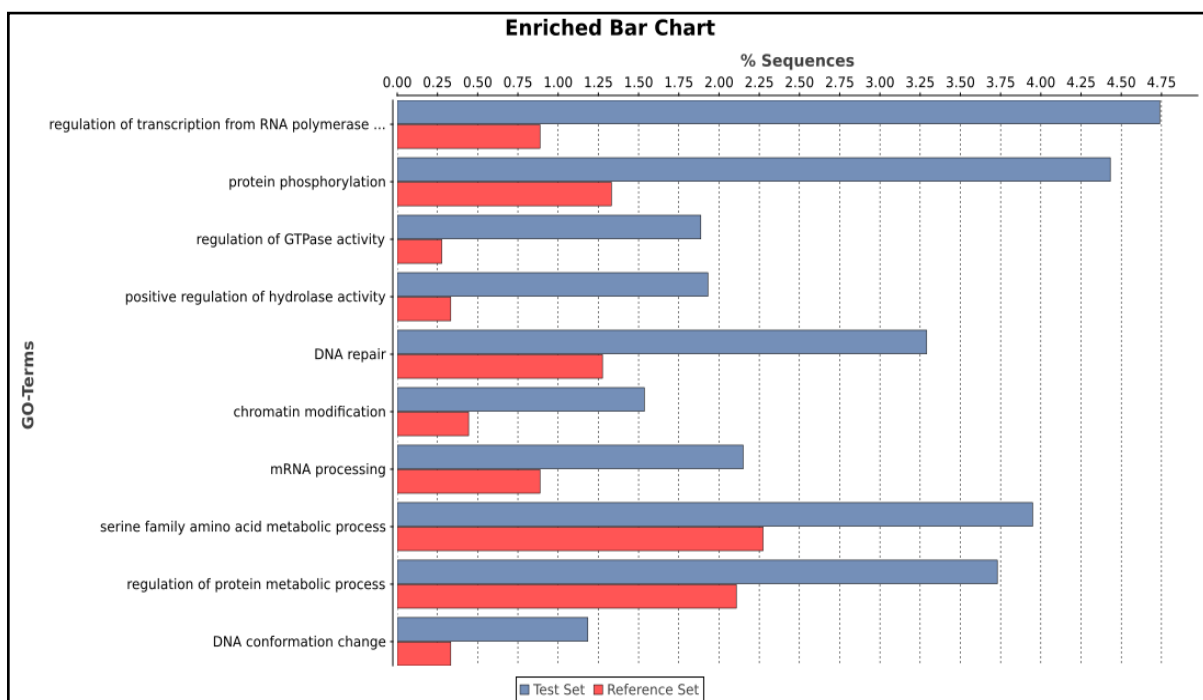


Figura A. 5: Gráfico referente à Análise de enriquecimento de desordem proteica em *C. neoformans*, linhagem JEC21. Este gráfico foi obtido a partir da análise de enriquecimento referente ao conteúdo total de desordem proteica dentro do proteoma de *C. neoformans*, linhagem JEC21. Em azul, está representado o conteúdo de proteínas desordenadas contidas nos termos anotadores. Em vermelho, está representado o conteúdo de proteínas globulares presentes no proteoma.

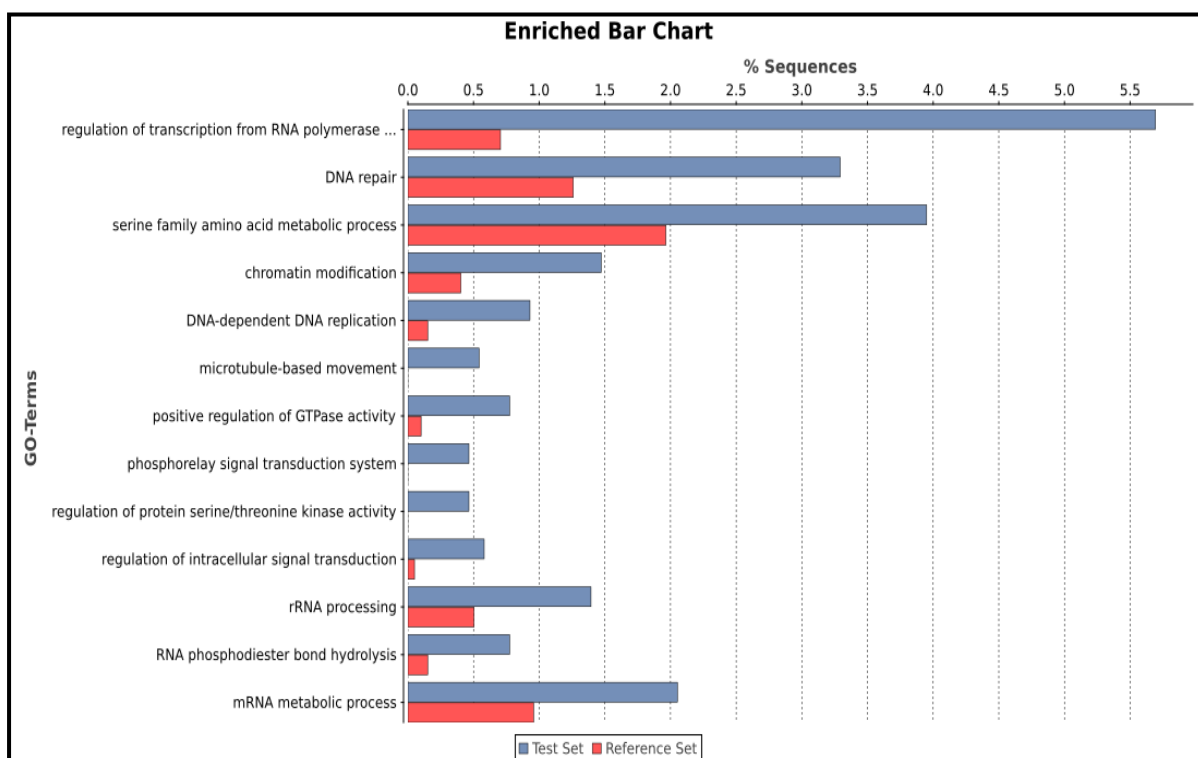


Figura A. 6: Gráfico referente à Análise de enriquecimento de desordem proteica em *C. neoformans* var. *grubii*, linhagem H99. Este gráfico foi obtido a partir da análise de enriquecimento referente ao conteúdo total de desordem proteica dentro do proteoma de *C. neoformans* var. *grubii*, linhagem H99. Em azul, está representado o conteúdo de proteínas desordenadas contidas nos termos anotadores. Em vermelho, está representado o conteúdo de proteínas globulares presentes no proteoma.

APENDICE B– GRÁFICOS RELACIONADOS À ANÁLISE DE ENRIQUECIMENTO ENTRE AS ESPÉCIES NÃO VIRULENTAS

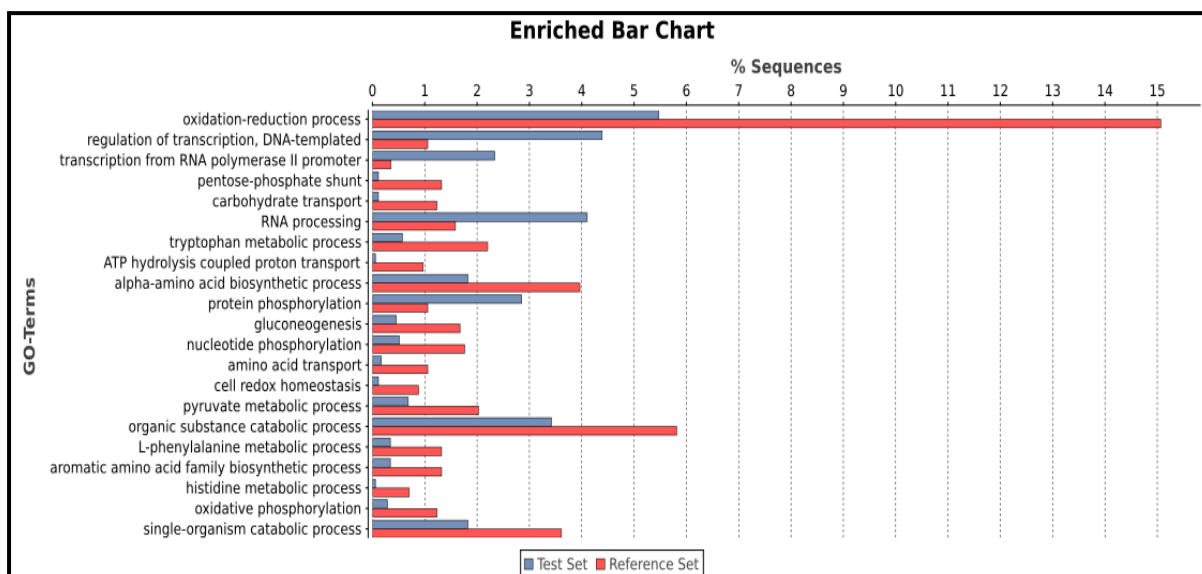


Figura B. 1: Gráfico referente à Análise de enriquecimento de desordem proteica em *C. dejecticola*, linhagem CBS10117. Este gráfico foi obtido a partir da análise de enriquecimento referente ao conteúdo total de desordem proteica dentro do proteoma de *C. dejecticola*, linhagem CBS10117. Em azul, está representado o conteúdo de proteínas desordenadas contidas nos termos anotadores. Em vermelho, está representado o conteúdo de proteínas globulares presentes no proteoma.

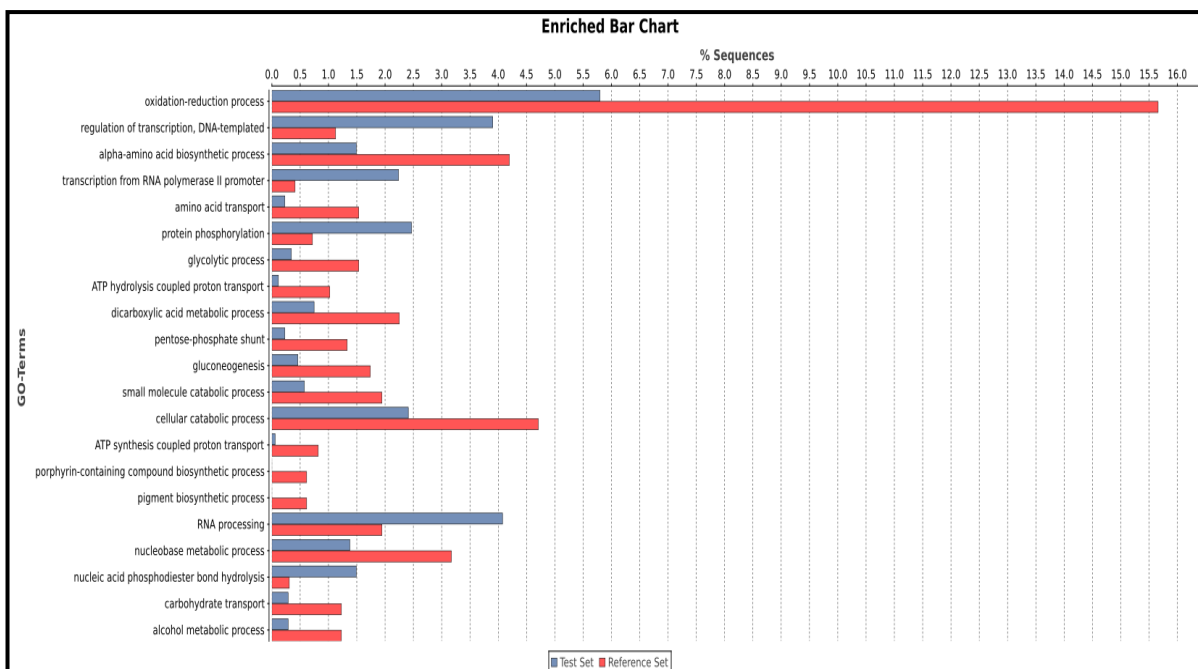


Figura B. 2: Gráfico referente à Análise de enriquecimento de desordem proteica em *C. bestiolarum*, linhagem CBS10118. Este gráfico foi obtido a partir da análise de enriquecimento referente ao conteúdo total de desordem proteica dentro do proteoma de *C. bestiolarum*, linhagem CBS10118. Em azul, está representado o conteúdo de proteínas desordenadas contidas nos termos anotadores. Em vermelho, está representado o conteúdo de proteínas globulares presentes no proteoma.

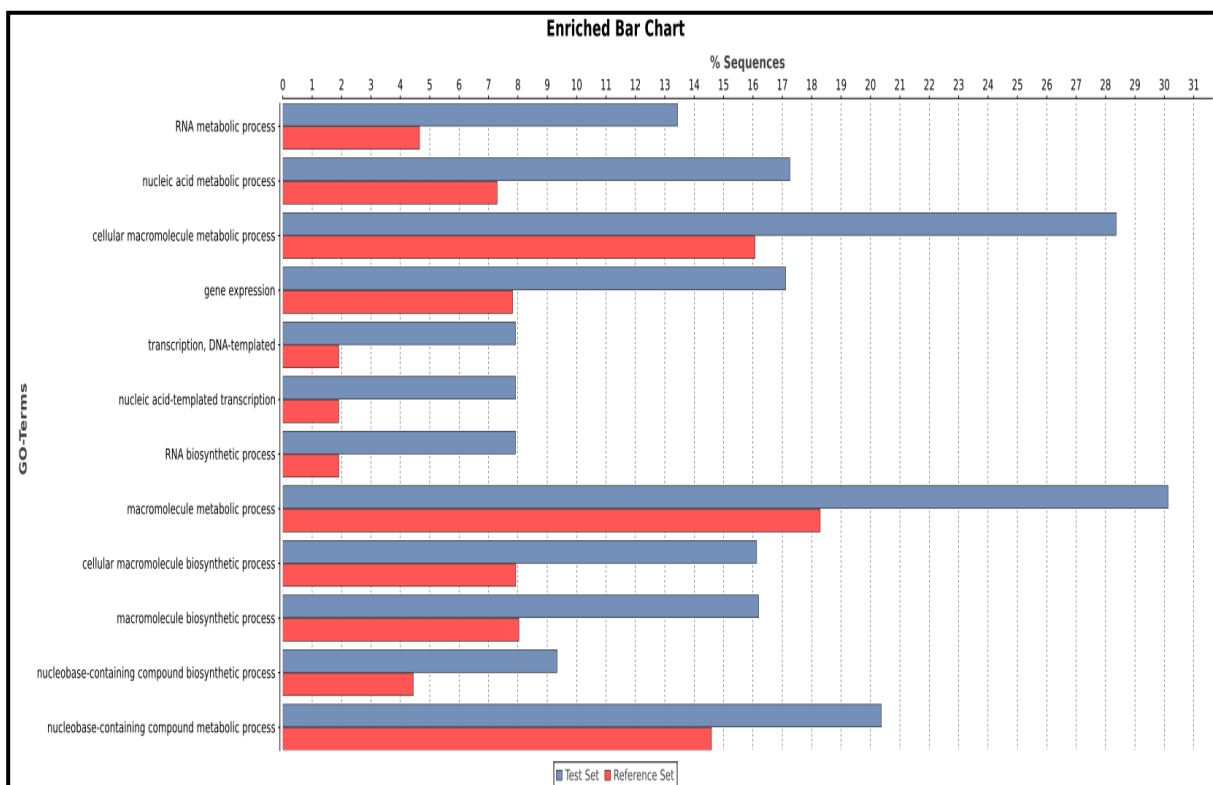


Figura B. 3: Gráfico referente à análise de enriquecimento de desordem proteica em *C. heveanensis*, linhagem CBS569. Este gráfico foi obtido a partir da análise de enriquecimento referente ao conteúdo total de desordem proteica dentro do proteoma de *C. heveanensis*, linhagem CBS569. Em azul, está representado o conteúdo de proteínas desordenadas contidas nos termos anotadores. Em vermelho, está representado o conteúdo de proteínas globulares presentes no proteoma.

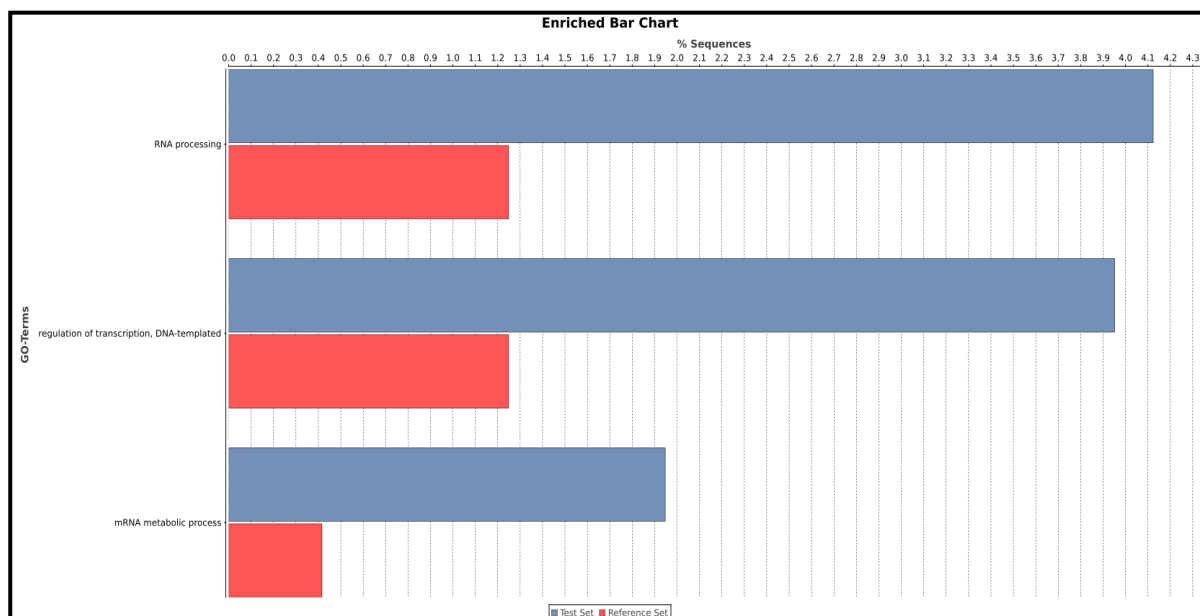


Figura B. 4: Gráfico referente à análise de enriquecimento de desordem proteica em *C. flavecens*, linhagem NRRL Y-50378. Este gráfico foi obtido a partir da análise de enriquecimento referente ao conteúdo total de desordem proteica dentro do proteoma de *C. flavecens*, linhagem NRRL Y-50378. Em azul, está representado o conteúdo de proteínas desordenadas contidas nos termos anotadores. Em vermelho, está representado o conteúdo de proteínas globulares presentes no proteoma.

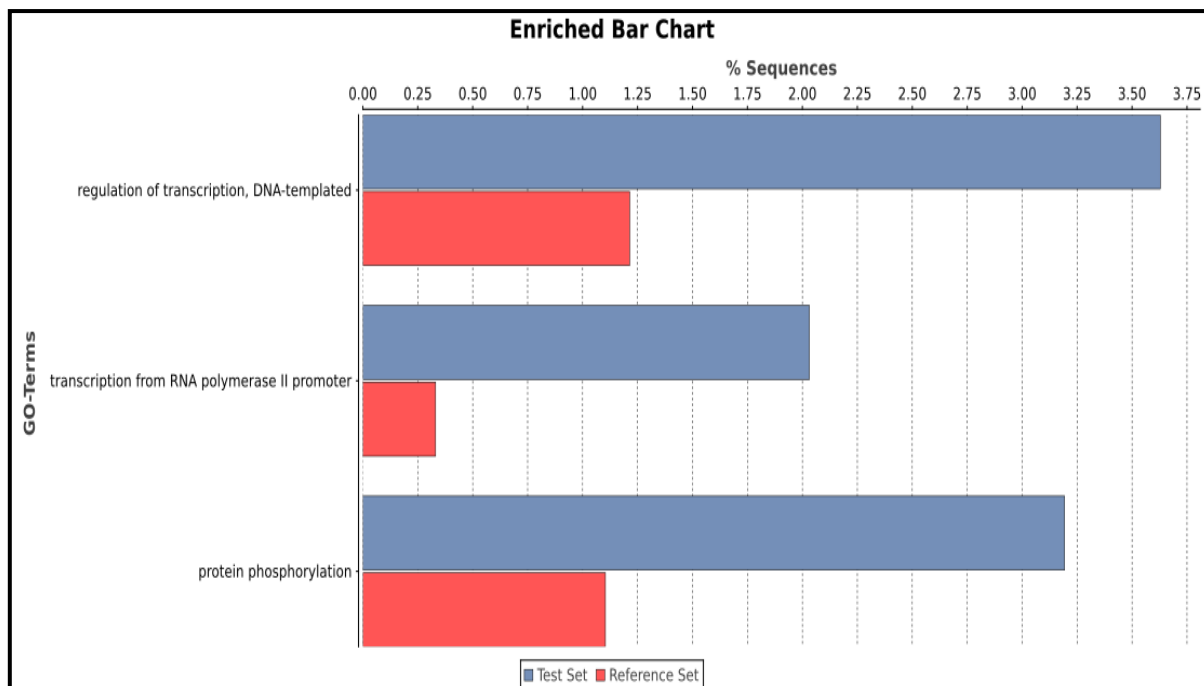


Figura B. 5: Gráfico referente à análise de enriquecimento de desordem proteica em *C. pinus*, linhagem CBS10737. Este gráfico foi obtido a partir da análise de enriquecimento referente ao conteúdo total de desordem proteica dentro do proteoma de *C. pinus*, linhagem CBS10737. Em azul, está representado o conteúdo de proteínas desordenadas contidas nos termos anotadores. Em vermelho, está representado o conteúdo de proteínas globulares presentes no proteoma.

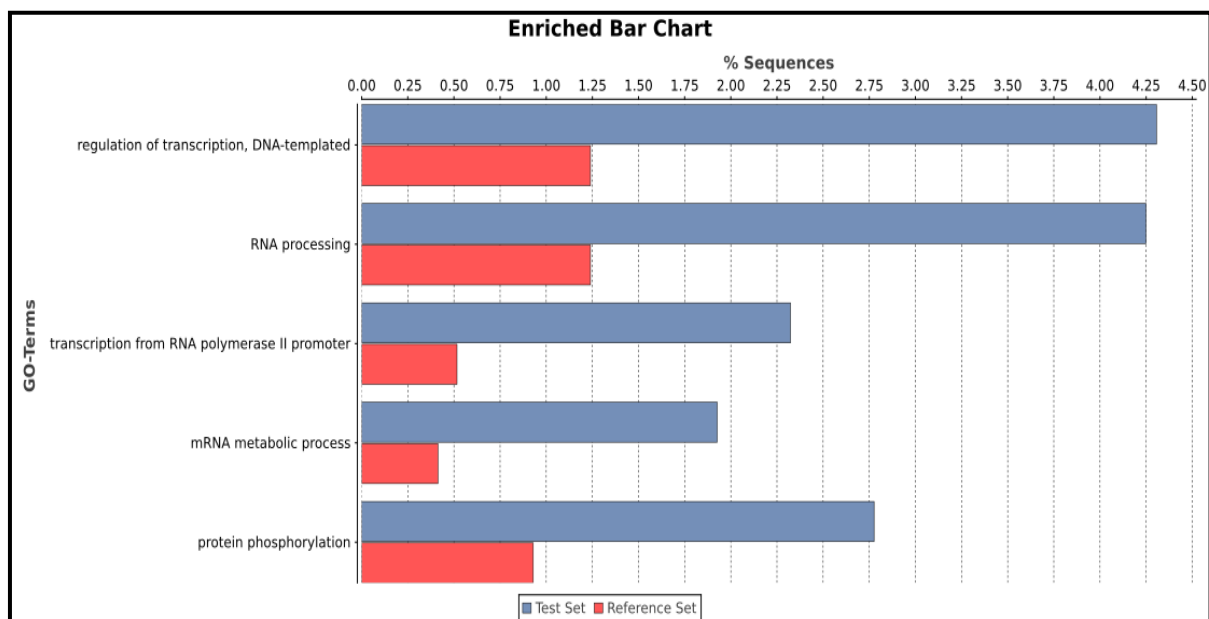


Figura B. 6: Gráfico referente à análise de enriquecimento de desordem proteica em *C. heveanensis*, linhagem BCC8938. Este gráfico foi obtido a partir da análise de enriquecimento referente ao conteúdo total de desordem proteica dentro do proteoma de *C. heveanensis*, linhagem BCC8938. Em azul, está representado o conteúdo de proteínas desordenadas contidas nos termos anotadores. Em vermelho, está representado o conteúdo de proteínas globulares presentes no proteoma.

APENDICE C – TABELAS RELACIONADAS AO MAPEAMENTO DAS IDPS E IDRS NAS VIAS METABÓLICAS ENTRE AS ESPÉCIES/LINHAGENS VIRULENTAS

Tabela C. 1: Vias metabólicas mapeadas em *C. gattii* CBS 7750. Todas as proteínas mapeadas referem-se às IDPs e IDRs identificadas neste estudo.

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Alanine, aspartate and glutamate metabolism</i>	7	<i>Transaminase</i>	ec:2.6.1.19	1	Cgcbs.045110	map00250
<i>Alanine, aspartate and glutamate metabolism</i>	7	<i>Synthase</i>	ec:6.3.4.5	1	Cgcbs.026530	map00250
<i>Alanine, aspartate and glutamate metabolism</i>	7	<i>synthase (NADH)</i>	ec:1.4.1.14	1	Cgcbs.067220	map00250
<i>Alanine, aspartate and glutamate metabolism</i>	7	<i>synthase (NADPH)</i>	ec:1.4.1.13	1	Cgcbs.067220	map00250
<i>Alanine, aspartate and glutamate metabolism</i>	7	<i>Carbamoyltransferase</i>	ec:2.1.3.2	1	Cgcbs.015160	map00250
<i>Alanine, aspartate and glutamate metabolism</i>	7	<i>synthase (glutamine-hydrolysing)</i>	ec:6.3.5.4	1	Cgcbs.047590	map00250
<i>Alanine, aspartate and glutamate metabolism</i>	7	<i>Decarboxylase</i>	ec:4.1.1.15	1	Cgcbs.069340	map00250
<i>Alanine, aspartate and glutamate metabolism</i>	7	<i>phosphoribosylidiphosphate 5-amidotransferase</i>	ec:2.4.2.14	1	Cgcbs.041260	map00250
<i>Amino sugar and nucleotide sugar metabolism</i>	12	<i>Isomerase</i>	ec:5.3.1.9	1	Cgcbs.065590	map00520
<i>Amino sugar and nucleotide sugar metabolism</i>	12	<i>Decarboxylase</i>	ec:4.1.1.35	1	Cgcbs.026930	map00520
<i>Amino sugar and nucleotide sugar metabolism</i>	12	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	2	Cgcbs.001910, Cgcbs.051600	map00520
<i>Amino sugar and nucleotide sugar metabolism</i>	12	<i>Synthase</i>	ec:1.1.1.271	1	Cgcbs.013800	map00520
<i>Amino sugar and nucleotide sugar metabolism</i>	12	<i>Synthase</i>	ec:2.4.1.16	5	Cgcbs.015920, Cgcbs.063260, Cgcbs.043810, Cgcbs.026880, Cgcbs.058710	map00520
<i>Amino sugar and nucleotide sugar metabolism</i>	12	<i>Chitodextrinase</i>	ec:3.2.1.14	2	Cgcbs.033050, Cgcbs.037670	map00520
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>synthase (glutamine-hydrolysing)</i>	ec:6.3.5.7	1	Cgcbs.070670	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.18	1	Cgcbs.005380	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.19	1	Cgcbs.025390	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.14	1	Cgcbs.019110	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.16	1	Cgcbs.071960	map00970

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.10	1	Cgcb.009860	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.12	2	Cgcb.024220, Cgcb.068700	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.6	1	Cgcb.033860	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.7	1	Cgcb.021210	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.9	1	Cgcb.001050	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.2	1	Cgcb.011880	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.3	1	Cgcb.072370	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.5	1	Cgcb.003670	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.1	2	Cgcb.069730, Cgcb.004660	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.20	1	Cgcb.066590	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.22	2	Cgcb.045580, Cgcb.009480	map00970
<i>Aminobenzoate degradation</i>	2	<i>Phosphatase</i>	ec:3.1.3.2	2	Cgcb.062550, Cgcb.023860	map00627
<i>Arginine and proline metabolism</i>	3	<i>Oxidase</i>	ec:1.4.3.3	2	Cgcb.051220, Cgcb.036410	map00330
<i>Arginine and proline metabolism</i>	3	<i>5-kinase</i>	ec:2.7.2.11	1	Cgcb.052600	map00330
<i>Arginine biosynthesis</i>	3	<i>Synthase</i>	ec:6.3.4.5	1	Cgcb.026530	map00220
<i>Arginine biosynthesis</i>	3	<i>Kinase</i>	ec:2.7.2.8	1	Cgcb.020820	map00220
<i>Arginine biosynthesis</i>	3	<i>Reductase</i>	ec:1.2.1.38	1	Cgcb.020820	map00220
<i>Arginine biosynthesis</i>	3	<i>N-acetyltransferase</i>	ec:2.3.1.1	1	Cgcb.040960	map00220
<i>Benzoate degradation</i>	1	<i>Dehydrogenase</i>	ec:1.1.1.35	1	Cgcb.032400	map00362
<i>beta-Alanine metabolism</i>	3	<i>Transaminase</i>	ec:2.6.1.19	1	Cgcb.045110	map00410
<i>beta-Alanine metabolism</i>	3	<i>Decarboxylase</i>	ec:4.1.1.15	1	Cgcb.069340	map00410
<i>beta-Alanine metabolism</i>	3	<i>Oxidase</i>	ec:1.4.3.21	1	Cgcb.053910	map00410
<i>Biosynthesis of ansamycins</i>	1	<i>glycolaldehydetransferase</i>	ec:2.2.1.1	1	Cgcb.057400	map01051
<i>Biosynthesis of antibiotics</i>	50	<i>Dehydrogenase</i>	ec:1.8.1.4	1	Cgcb.068190	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>Synthase</i>	ec:2.3.3.14	1	Cgcb.006470	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>Succinyltransferase</i>	ec:2.3.1.61	1	Cgcb.000020	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>Kinase</i>	ec:2.7.4.3	1	Cgcb.071770	map01130

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Biosynthesis of antibiotics</i>	50	<i>Isomerase</i>	ec:5.3.1.9	1	Cgcbs.065590	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>carboxykinase (ATP)</i>	ec:4.1.1.49	1	Cgcbs.033430	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>Synthase</i>	ec:6.3.4.5	1	Cgcbs.026530	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>O-methyltransferase</i>	ec:2.1.1.100	1	Cgcbs.048460	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>Decarboxylase</i>	ec:4.1.1.33	1	Cgcbs.053790	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>Oxidase</i>	ec:1.4.3.3	2	Cgcbs.051220, Cgcbs.036410	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>Kinase</i>	ec:2.7.1.71	1	Cgcbs.001160	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>Hydratase</i>	ec:4.2.1.3	2	Cgcbs.057480, Cgcbs.007990	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>Dehydrogenase</i>	ec:1.1.1.95	1	Cgcbs.013380	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>dihydroxyacetonetransferase</i>	ec:2.2.1.2	1	Cgcbs.064370	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>Dehydratase</i>	ec:4.2.1.10	1	Cgcbs.001160	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>glycolaldehydetransferase</i>	ec:2.2.1.1	1	Cgcbs.057400	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>Synthase</i>	ec:2.2.1.6	2	Cgcbs.003380, Cgcbs.059430	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>Equilase</i>	ec:1.11.1.6	2	Cgcbs.054930, Cgcbs.055290	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>Synthase</i>	ec:2.5.1.54	1	Cgcbs.040640	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>Kinase</i>	ec:2.7.2.8	1	Cgcbs.020820	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>Kinase</i>	ec:2.7.2.4	1	Cgcbs.031820	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>Kinase</i>	ec:2.7.2.3	1	Cgcbs.026530	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>Reductase</i>	ec:1.2.1.38	1	Cgcbs.020820	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>5-kinase</i>	ec:2.7.2.11	1	Cgcbs.052600	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>Deaminase</i>	ec:3.5.4.6	2	Cgcbs.040010, Cgcbs.040020	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>Synthase</i>	ec:6.3.2.6	1	Cgcbs.004660	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>synthase (NADH)</i>	ec:1.4.1.14	1	Cgcbs.067220	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>synthase (NADPH)</i>	ec:1.4.1.13	1	Cgcbs.067220	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>ammonia-lyase</i>	ec:4.3.1.19	1	Cgcbs.037080	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	2	Cgcbs.001910, Cgcbs.051600	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>reductoisomerase (NADP+)</i>	ec:1.1.1.86	1	Cgcbs.021180	map01130

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Biosynthesis of antibiotics</i>	50	<i>Synthase</i>	ec:4.2.3.4	1	Cgcbs.001160	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>Phosphohexokinase</i>	ec:2.7.1.11	1	Cgcbs.011080	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>N-acetyltransferase</i>	ec:2.3.1.1	1	Cgcbs.040960	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>ligase (ADP-forming)</i>	ec:6.2.1.5	1	Cgcbs.026350	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>Synthase</i>	ec:5.5.1.4	1	Cgcbs.049050	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>1-carboxyvinyltransferase</i>	ec:2.5.1.19	1	Cgcbs.001160	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>ligase (GDP-forming)</i>	ec:6.2.1.4	1	Cgcbs.026350	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>Ligase</i>	ec:6.2.1.1	1	Cgcbs.013090	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>Dehydrogenase</i>	ec:1.1.1.35	1	Cgcbs.032400	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>reductase (NADPH)</i>	ec:1.1.1.34	1	Cgcbs.023710	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>Acetyltransferase</i>	ec:2.3.1.12	1	Cgcbs.007790	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>dehydrogenase (NADP+-dependent, decarboxylating)</i>	ec:1.1.1.44	1	Cgcbs.049590	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>dehydrogenase (NADP+)</i>	ec:1.1.1.42	1	Cgcbs.065540	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>dehydrogenase (NAD+)</i>	ec:1.1.1.41	2	Cgcbs.066390, Cgcbs.016610	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>O-acetyltransferase</i>	ec:2.3.1.31	2	Cgcbs.044090, Cgcbs.053820	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>Dehydrogenase</i>	ec:1.1.1.25	1	Cgcbs.001160	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>citrate synthase</i>	ec:2.3.3.8	1	Cgcbs.026350	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>phosphoribosylidiphosphate 5-amidotransferase</i>	ec:2.4.2.14	1	Cgcbs.041260	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>(Si)-synthase</i>	ec:2.3.3.1	1	Cgcbs.005630	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>dehydrogenase (succinyl-transferring)</i>	ec:1.2.4.2	1	Cgcbs.000910	map01130
<i>Biosynthesis of antibiotics</i>	50	<i>Kinase</i>	ec:2.7.4.10	1	Cgcbs.071770	map01130
<i>Biosynthesis of unsaturated fatty acids</i>	1	<i>9-desaturase</i>	ec:1.14.19.1	1	Cgcbs.010950	map01040
<i>Biotin metabolism</i>	1	<i>Synthase</i>	ec:2.8.1.6	1	Cgcbs.022710	map00780
<i>Butanoate metabolism</i>	7	<i>Transaminase</i>	ec:2.6.1.19	1	Cgcbs.045110	map00650
<i>Butanoate metabolism</i>	7	<i>Synthase</i>	ec:2.2.1.6	2	Cgcbs.003380, Cgcbs.059430	map00650
<i>Butanoate metabolism</i>	7	<i>Dehydrogenase</i>	ec:1.1.1.35	1	Cgcbs.032400	map00650
<i>Butanoate metabolism</i>	7	<i>Decarboxylase</i>	ec:4.1.1.15	1	Cgcbs.069340	map00650

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Butanoate metabolism</i>	7	<i>Ligase</i>	ec:6.2.1.16	1	Cgcbs.061300	map00650
<i>Butanoate metabolism</i>	7	<i>CoA-transferase</i>	ec:2.8.3.5	1	Cgcbs.054730	map00650
<i>Butirosin and neomycin biosynthesis</i>	2	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	2	Cgcbs.001910, Cgcbs.051600	map00524
<i>C5-Branched dibasic acid metabolism</i>	4	<i>Synthase</i>	ec:2.2.1.6	2	Cgcbs.003380, Cgcbs.059430	map00660
<i>C5-Branched dibasic acid metabolism</i>	4	<i>Dehydrogenase</i>	ec:1.1.1.85	1	Cgcbs.016960	map00660
<i>C5-Branched dibasic acid metabolism</i>	4	<i>ligase (ADP-forming)</i>	ec:6.2.1.5	1	Cgcbs.026350	map00660
<i>Caprolactam degradation</i>	1	<i>Dehydrogenase</i>	ec:1.1.1.35	1	Cgcbs.032400	map00930
<i>Carbapenem biosynthesis</i>	1	<i>5-kinase</i>	ec:2.7.2.11	1	Cgcbs.052600	map00332
<i>Carbon fixation in photosynthetic organisms</i>	5	<i>carboxykinase (ATP)</i>	ec:4.1.1.49	1	Cgcbs.033430	map00710
<i>Carbon fixation in photosynthetic organisms</i>	5	<i>glycolaldehydetransferase</i>	ec:2.2.1.1	1	Cgcbs.057400	map00710
<i>Carbon fixation in photosynthetic organisms</i>	5	<i>Kinase</i>	ec:2.7.2.3	1	Cgcbs.026530	map00710
<i>Carbon fixation in photosynthetic organisms</i>	5	<i>dehydrogenase (decarboxylating)</i>	ec:1.1.1.39	2	Cgcbs.032190, Cgcbs.022490	map00710
<i>Carbon fixation pathways in prokaryotes</i>	8	<i>reductase [NAD(P)H]</i>	ec:1.5.1.20	2	Cgcbs.053580, Cgcbs.013520	map00720
<i>Carbon fixation pathways in prokaryotes</i>	8	<i>Hydratase</i>	ec:4.2.1.3	2	Cgcbs.057480, Cgcbs.007990	map00720
<i>Carbon fixation pathways in prokaryotes</i>	8	<i>ligase (ADP-forming)</i>	ec:6.2.1.5	1	Cgcbs.026350	map00720
<i>Carbon fixation pathways in prokaryotes</i>	8	<i>Ligase</i>	ec:6.2.1.1	1	Cgcbs.013090	map00720
<i>Carbon fixation pathways in prokaryotes</i>	8	<i>Dehydrogenase</i>	ec:1.1.1.35	1	Cgcbs.032400	map00720
<i>Carbon fixation pathways in prokaryotes</i>	8	<i>dehydrogenase (NADP+)</i>	ec:1.1.1.42	1	Cgcbs.065540	map00720
<i>Carbon fixation pathways in prokaryotes</i>	8	<i>citrate synthase</i>	ec:2.3.3.8	1	Cgcbs.026350	map00720
<i>Citrate cycle (TCA cycle)</i>	12	<i>Dehydrogenase</i>	ec:1.8.1.4	1	Cgcbs.068190	map00020
<i>Citrate cycle (TCA cycle)</i>	12	<i>Succinyltransferase</i>	ec:2.3.1.61	1	Cgcbs.000020	map00020
<i>Citrate cycle (TCA cycle)</i>	12	<i>carboxykinase (ATP)</i>	ec:4.1.1.49	1	Cgcbs.033430	map00020
<i>Citrate cycle (TCA cycle)</i>	12	<i>Hydratase</i>	ec:4.2.1.3	2	Cgcbs.057480, Cgcbs.007990	map00020
<i>Citrate cycle (TCA cycle)</i>	12	<i>ligase (ADP-forming)</i>	ec:6.2.1.5	1	Cgcbs.026350	map00020
<i>Citrate cycle (TCA cycle)</i>	12	<i>ligase (GDP-forming)</i>	ec:6.2.1.4	1	Cgcbs.026350	map00020
<i>Citrate cycle (TCA cycle)</i>	12	<i>Acetyltransferase</i>	ec:2.3.1.12	1	Cgcbs.007790	map00020

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Citrate cycle (TCA cycle)</i>	12	<i>dehydrogenase (NADP+)</i>	ec:1.1.1.42	1	Cgcbs.065540	map00020
<i>Citrate cycle (TCA cycle)</i>	12	<i>dehydrogenase (NAD+)</i>	ec:1.1.1.41	2	Cgcbs.066390, Cgcbs.016610	map00020
<i>Citrate cycle (TCA cycle)</i>	12	<i>citrate synthase</i>	ec:2.3.3.8	1	Cgcbs.026350	map00020
<i>Citrate cycle (TCA cycle)</i>	12	<i>(Si)-synthase</i>	ec:2.3.3.1	1	Cgcbs.005630	map00020
<i>Citrate cycle (TCA cycle)</i>	12	<i>dehydrogenase (succinyl-transferring)</i>	ec:1.2.4.2	1	Cgcbs.000910	map00020
<i>Cyanoamino acid metabolism</i>	3	<i>glutamyl transpeptidase</i>	ec:2.3.2.2	3	Cgcbs.041630, Cgcbs.006850, Cgcbs.028980	map00460
<i>Cysteine and methionine metabolism</i>	6	<i>S-methyltransferase</i>	ec:2.1.1.10	1	Cgcbs.066860	map00270
<i>Cysteine and methionine metabolism</i>	6	<i>Kinase</i>	ec:2.7.2.4	1	Cgcbs.031820	map00270
<i>Cysteine and methionine metabolism</i>	6	<i>Ligase</i>	ec:6.3.2.2	1	Cgcbs.060730	map00270
<i>Cysteine and methionine metabolism</i>	6	<i>dioxygenase [iron(II)-requiring]</i>	ec:1.13.11.54	1	Cgcbs.067380	map00270
<i>Cysteine and methionine metabolism</i>	6	<i>O-acetyltransferase</i>	ec:2.3.1.31	2	Cgcbs.044090, Cgcbs.053820	map00270
<i>D-Arginine and D-ornithine metabolism</i>	2	<i>Oxidase</i>	ec:1.4.3.3	2	Cgcbs.051220, Cgcbs.036410	map00472
<i>Drug metabolism - other enzymes</i>	1	<i>Dehydrogenase</i>	ec:1.1.1.205	1	Cgcbs.017070	map00983
<i>Ether lipid metabolism</i>	3	<i>D</i>	ec:3.1.4.4	2	Cgcbs.041170, Cgcbs.006650	map00565
<i>Ether lipid metabolism</i>	3	<i>Esterase</i>	ec:3.1.1.47	1	Cgcbs.025010	map00565
<i>Fatty acid biosynthesis</i>	1	<i>Synthase</i>	ec:2.3.1.85	1	Cgcbs.061890	map00061
<i>Fatty acid degradation</i>	1	<i>Dehydrogenase</i>	ec:1.1.1.35	1	Cgcbs.032400	map00071
<i>Fatty acid elongation</i>	1	<i>Dehydrogenase</i>	ec:1.1.1.35	1	Cgcbs.032400	map00062
<i>Folate biosynthesis</i>	1	<i>Synthase</i>	ec:6.3.2.17	1	Cgcbs.050070	map00790
<i>Fructose and mannose metabolism</i>	6	<i>phosphofructokinase 2</i>	ec:2.7.1.105	2	Cgcbs.015190, Cgcbs.033380	map00051
<i>Fructose and mannose metabolism</i>	6	<i>2-phosphatase</i>	ec:3.1.3.46	2	Cgcbs.015190, Cgcbs.033380	map00051
<i>Fructose and mannose metabolism</i>	6	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	2	Cgcbs.001910, Cgcbs.051600	map00051
<i>Fructose and mannose metabolism</i>	6	<i>Phosphohexokinase</i>	ec:2.7.1.11	1	Cgcbs.011080	map00051
<i>Fructose and mannose metabolism</i>	6	<i>Synthase</i>	ec:1.1.1.271	1	Cgcbs.013800	map00051
<i>Galactose metabolism</i>	4	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	2	Cgcbs.001910, Cgcbs.051600	map00052
<i>Galactose metabolism</i>	4	<i>Phosphohexokinase</i>	ec:2.7.1.11	1	Cgcbs.011080	map00052
<i>Galactose metabolism</i>	4	<i>Maltase</i>	ec:3.2.1.20	1	Cgcbs.051720	map00052

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Geraniol degradation</i>	1	<i>Dehydrogenase</i>	ec:1.1.1.35	1	Cgcbs.032400	map00281
<i>Glutathione metabolism</i>	8	<i>Reductase</i>	ec:1.17.4.1	1	Cgcbs.063140	map00480
<i>Glutathione metabolism</i>	8	<i>Ligase</i>	ec:6.3.2.2	1	Cgcbs.060730	map00480
<i>Glutathione metabolism</i>	8	<i>dehydrogenase (NADP+-dependent, decarboxylating)</i>	ec:1.1.1.44	1	Cgcbs.049590	map00480
<i>Glutathione metabolism</i>	8	<i>dehydrogenase (NADP+)</i>	ec:1.1.1.42	1	Cgcbs.065540	map00480
<i>Glutathione metabolism</i>	8	<i>glutamyl transpeptidase</i>	ec:2.3.2.2	3	Cgcbs.041630, Cgcbs.006850, Cgcbs.028980	map00480
<i>Glutathione metabolism</i>	8	<i>thioredoxin peroxidase</i>	ec:1.11.1.15	1	Cgcbs.047240	map00480
<i>Glycerolipid metabolism</i>	5	<i>kinase (ATP)</i>	ec:2.7.1.107	2	Cgcbs.008300, Cgcbs.018500	map00561
<i>Glycerolipid metabolism</i>	5	<i>Kinase</i>	ec:2.7.1.30	2	Cgcbs.008190, Cgcbs.008180	map00561
<i>Glycerolipid metabolism</i>	5	<i>O-acyltransferase</i>	ec:2.3.1.51	1	Cgcbs.007640	map00561
<i>Glycerophospholipid metabolism</i>	12	<i>kinase (ATP)</i>	ec:2.7.1.107	2	Cgcbs.008300, Cgcbs.018500	map00564
<i>Glycerophospholipid metabolism</i>	12	<i>Cytidylyltransferase</i>	ec:2.7.7.41	1	Cgcbs.006240	map00564
<i>Glycerophospholipid metabolism</i>	12	<i>N-methyltransferase</i>	ec:2.1.1.17	2	Cgcbs.028830, Cgcbs.028850	map00564
<i>Glycerophospholipid metabolism</i>	12	<i>Decarboxylase</i>	ec:4.1.1.65	2	Cgcbs.026820, Cgcbs.066510	map00564
<i>Glycerophospholipid metabolism</i>	12	<i>D</i>	ec:3.1.4.4	2	Cgcbs.041170, Cgcbs.006650	map00564
<i>Glycerophospholipid metabolism</i>	12	<i>Dehydrogenase</i>	ec:1.1.5.3	1	Cgcbs.040830	map00564
<i>Glycerophospholipid metabolism</i>	12	<i>lecithinase B</i>	ec:3.1.1.5	1	Cgcbs.009790	map00564
<i>Glycerophospholipid metabolism</i>	12	<i>O-acyltransferase</i>	ec:2.3.1.51	1	Cgcbs.007640	map00564
<i>Glycine, serine and threonine metabolism</i>	11	<i>Dehydrogenase</i>	ec:1.8.1.4	1	Cgcbs.068190	map00260
<i>Glycine, serine and threonine metabolism</i>	11	<i>S-aminomethyldihydrolipoylprotein: (6S)-tetrahydrofolate aminomethyltransferase (ammonia-forming)</i>	ec:2.1.2.10	1	Cgcbs.040880	map00260
<i>Glycine, serine and threonine metabolism</i>	11	<i>Oxidase</i>	ec:1.4.3.3	2	Cgcbs.051220, Cgcbs.036410	map00260
<i>Glycine, serine and threonine metabolism</i>	11	<i>dehydrogenase (aminomethyl-transferring)</i>	ec:1.4.4.2	1	Cgcbs.048420	map00260
<i>Glycine, serine and threonine metabolism</i>	11	<i>Dehydrogenase</i>	ec:1.1.1.95	1	Cgcbs.013380	map00260
<i>Glycine, serine and threonine metabolism</i>	11	<i>Kinase</i>	ec:2.7.2.4	1	Cgcbs.031820	map00260
<i>Glycine, serine and threonine metabolism</i>	11	<i>ammonia-lyase</i>	ec:4.3.1.19	1	Cgcbs.037080	map00260

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Glycine, serine and threonine metabolism</i>	11	<i>Oxidase</i>	ec:1.5.3.1	1	Cgcb.042840	map00260
<i>Glycine, serine and threonine metabolism</i>	11	<i>Dehydrogenase</i>	ec:1.1.99.1	1	Cgcb.069310	map00260
<i>Glycine, serine and threonine metabolism</i>	11	<i>Oxidase</i>	ec:1.4.3.21	1	Cgcb.053910	map00260
<i>Glycolysis / Gluconeogenesis</i>	9	<i>Dehydrogenase</i>	ec:1.8.1.4	1	Cgcb.068190	map00010
<i>Glycolysis / Gluconeogenesis</i>	9	<i>Isomerase</i>	ec:5.3.1.9	1	Cgcb.065590	map00010
<i>Glycolysis / Gluconeogenesis</i>	9	<i>carboxykinase (ATP)</i>	ec:4.1.1.49	1	Cgcb.033430	map00010
<i>Glycolysis / Gluconeogenesis</i>	9	<i>Kinase</i>	ec:2.7.2.3	1	Cgcb.026530	map00010
<i>Glycolysis / Gluconeogenesis</i>	9	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	2	Cgcb.001910, Cgcb.051600	map00010
<i>Glycolysis / Gluconeogenesis</i>	9	<i>Phosphohexokinase</i>	ec:2.7.1.11	1	Cgcb.011080	map00010
<i>Glycolysis / Gluconeogenesis</i>	9	<i>Ligase</i>	ec:6.2.1.1	1	Cgcb.013090	map00010
<i>Glycolysis / Gluconeogenesis</i>	9	<i>Acetyltransferase</i>	ec:2.3.1.12	1	Cgcb.007790	map00010
<i>Glyoxylate and dicarboxylate metabolism</i>	5	<i>Hydratase</i>	ec:4.2.1.3	2	Cgcb.057480, Cgcb.007990	map00630
<i>Glyoxylate and dicarboxylate metabolism</i>	5	<i>Equilase</i>	ec:1.11.1.6	2	Cgcb.054930, Cgcb.055290	map00630
<i>Glyoxylate and dicarboxylate metabolism</i>	5	<i>(Si)-synthase</i>	ec:2.3.3.1	1	Cgcb.005630	map00630
<i>Histidine metabolism</i>	2	<i>histidinol phosphate phosphatase</i>	ec:3.1.3.15	1	Cgcb.041360	map00340
<i>Histidine metabolism</i>	2	<i>Diphosphatase</i>	ec:3.6.1.31	1	Cgcb.068500	map00340
<i>Histidine metabolism</i>	2	<i>Dehydrogenase</i>	ec:1.1.1.23	1	Cgcb.068500	map00340
<i>Histidine metabolism</i>	2	<i>Cyclohydrolase</i>	ec:3.5.4.19	1	Cgcb.068500	map00340
<i>Inositol phosphate metabolism</i>	4	<i>3-kinase</i>	ec:2.7.1.127	1	Cgcb.041740	map00562
<i>Inositol phosphate metabolism</i>	4	<i>4-kinase</i>	ec:2.7.1.67	1	Cgcb.024510	map00562
<i>Inositol phosphate metabolism</i>	4	<i>Synthase</i>	ec:5.5.1.4	1	Cgcb.049050	map00562
<i>Inositol phosphate metabolism</i>	4	<i>3-kinase</i>	ec:2.7.1.137	1	Cgcb.017870	map00562
<i>Isoquinoline alkaloid biosynthesis</i>	1	<i>Oxidase</i>	ec:1.4.3.21	1	Cgcb.053910	map00950
<i>Lipoic acid metabolism</i>	1	<i>Synthase</i>	ec:2.8.1.8	1	Cgcb.053070	map00785
<i>Lysine biosynthesis</i>	2	<i>Synthase</i>	ec:2.3.3.14	1	Cgcb.006470	map00300
<i>Lysine biosynthesis</i>	2	<i>Kinase</i>	ec:2.7.2.4	1	Cgcb.031820	map00300
<i>Lysine degradation</i>	5	<i>Succinyltransferase</i>	ec:2.3.1.61	1	Cgcb.000020	map00310

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Lysine degradation</i>	5	<i>N-methyltransferase</i>	ec:2.1.1.43	2	Cgcbs.036140, Cgcbs.009140	map00310
<i>Lysine degradation</i>	5	<i>Dehydrogenase</i>	ec:1.1.1.35	1	Cgcbs.032400	map00310
<i>Lysine degradation</i>	5	<i>dehydrogenase (succinyl-transferring)</i>	ec:1.2.4.2	1	Cgcbs.000910	map00310
<i>Methane metabolism</i>	3	<i>Dehydrogenase</i>	ec:1.1.1.95	1	Cgcbs.013380	map00680
<i>Methane metabolism</i>	3	<i>Phosphohexokinase</i>	ec:2.7.1.11	1	Cgcbs.011080	map00680
<i>Methane metabolism</i>	3	<i>Ligase</i>	ec:6.2.1.1	1	Cgcbs.013090	map00680
<i>Monobactam biosynthesis</i>	1	<i>Kinase</i>	ec:2.7.2.4	1	Cgcbs.031820	map00261
<i>mTOR signaling pathway</i>	4	<i>protein kinase</i>	ec:2.7.11.24	4	Cgcbs.046830, Cgcbs.030160, Cgcbs.054430, Cgcbs.032670	map04150
<i>N-Glycan biosynthesis</i>	6	<i>Kinase</i>	ec:2.7.1.108	2	Cgcbs.063780, Cgcbs.063770	map00510
<i>N-Glycan biosynthesis</i>	6	<i>1,2-alpha-mannosidase</i>	ec:3.2.1.113	2	Cgcbs.034960, Cgcbs.027700	map00510
<i>N-Glycan biosynthesis</i>	6	<i>N-acetylglucosaminyltransferase</i>	ec:2.4.1.141	1	Cgcbs.063990	map00510
<i>N-Glycan biosynthesis</i>	6	<i>Glucosidase</i>	ec:3.2.1.106	1	Cgcbs.061640	map00510
<i>Nicotinate and nicotinamide metabolism</i>	6	<i>Kinase</i>	ec:2.7.1.23	5	Cgcbs.032320, Cgcbs.008300, Cgcbs.032310, Cgcbs.018500, Cgcbs.032340	map00760
<i>Nicotinate and nicotinamide metabolism</i>	6	<i>diphosphorylase (carboxylating)</i>	ec:2.4.2.19	1	Cgcbs.040460	map00760
<i>Nitrogen metabolism</i>	1	<i>synthase (NADH)</i>	ec:1.4.1.14	1	Cgcbs.067220	map00910
<i>Nitrogen metabolism</i>	1	<i>synthase (NADPH)</i>	ec:1.4.1.13	1	Cgcbs.067220	map00910
<i>One carbon pool by folate</i>	3	<i>S-aminomethyldihydrolipoylprotein: (6S)-tetrahydrofolate aminomethyltransferase (ammonia-forming)</i>	ec:2.1.2.10	1	Cgcbs.040880	map00670
<i>One carbon pool by folate</i>	3	<i>reductase [NAD(P)H]</i>	ec:1.5.1.20	2	Cgcbs.053580, Cgcbs.013520	map00670
<i>Other glycan degradation</i>	1	<i>psychosine hydrolase</i>	ec:3.2.1.45	1	Cgcbs.015120	map00511
<i>Other types of O-glycan biosynthesis</i>	1	<i>Mannosyltransferase</i>	ec:2.4.1.109	1	Cgcbs.068530	map00514
<i>Pantothenate and CoA biosynthesis</i>	7	<i>hydroxymethyltransferase</i>	ec:2.1.2.11	1	Cgcbs.011980	map00770
<i>Pantothenate and CoA biosynthesis</i>	7	<i>Synthase</i>	ec:2.2.1.6	2	Cgcbs.003380, Cgcbs.059430	map00770
<i>Pantothenate and CoA biosynthesis</i>	7	<i>Kinase</i>	ec:2.7.1.33	1	Cgcbs.041400	map00770
<i>Pantothenate and CoA biosynthesis</i>	7	<i>reductoisomerase (NADP+)</i>	ec:1.1.1.86	1	Cgcbs.021180	map00770
<i>Pantothenate and CoA biosynthesis</i>	7	<i>Synthase</i>	ec:2.7.8.7	2	Cgcbs.061890, Cgcbs.033980	map00770

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Penicillin and cephalosporin biosynthesis</i>	2	<i>Oxidase</i>	ec:1.4.3.3	2	Cgcbs.051220, Cgcbs.036410	map00311
<i>Pentose phosphate pathway</i>	5	<i>Isomerase</i>	ec:5.3.1.9	1	Cgcbs.065590	map00030
<i>Pentose phosphate pathway</i>	5	<i>dihydroxyacetonetransferase</i>	ec:2.2.1.2	1	Cgcbs.064370	map00030
<i>Pentose phosphate pathway</i>	5	<i>glycolaldehydetransferase</i>	ec:2.2.1.1	1	Cgcbs.057400	map00030
<i>Pentose phosphate pathway</i>	5	<i>Phosphohexokinase</i>	ec:2.7.1.11	1	Cgcbs.011080	map00030
<i>Pentose phosphate pathway</i>	5	<i>dehydrogenase (NADP+-dependent, decarboxylating)</i>	ec:1.1.1.44	1	Cgcbs.049590	map00030
<i>Phenylalanine metabolism</i>	1	<i>Oxidase</i>	ec:1.4.3.21	1	Cgcbs.053910	map00360
<i>Phenylalanine, tyrosine and tryptophan biosynthesis</i>	2	<i>Kinase</i>	ec:2.7.1.71	1	Cgcbs.001160	map00400
<i>Phenylalanine, tyrosine and tryptophan biosynthesis</i>	2	<i>Dehydratase</i>	ec:4.2.1.10	1	Cgcbs.001160	map00400
<i>Phenylalanine, tyrosine and tryptophan biosynthesis</i>	2	<i>Synthase</i>	ec:2.5.1.54	1	Cgcbs.040640	map00400
<i>Phenylalanine, tyrosine and tryptophan biosynthesis</i>	2	<i>Synthase</i>	ec:4.2.3.4	1	Cgcbs.001160	map00400
<i>Phenylalanine, tyrosine and tryptophan biosynthesis</i>	2	<i>1-carboxyvinyltransferase</i>	ec:2.5.1.19	1	Cgcbs.001160	map00400
<i>Phenylalanine, tyrosine and tryptophan biosynthesis</i>	2	<i>Dehydrogenase</i>	ec:1.1.1.25	1	Cgcbs.001160	map00400
<i>Phenylpropanoid biosynthesis</i>	1	<i>Lactoperoxidase</i>	ec:1.11.1.7	1	Cgcbs.047240	map00940
<i>Phosphatidylinositol signaling system</i>	6	<i>kinase (ATP)</i>	ec:2.7.1.107	2	Cgcbs.008300, Cgcbs.018500	map04070
<i>Phosphatidylinositol signaling system</i>	6	<i>3-kinase</i>	ec:2.7.1.127	1	Cgcbs.041740	map04070
<i>Phosphatidylinositol signaling system</i>	6	<i>4-kinase</i>	ec:2.7.1.67	1	Cgcbs.024510	map04070
<i>Phosphatidylinositol signaling system</i>	6	<i>Cytidylyltransferase</i>	ec:2.7.7.41	1	Cgcbs.006240	map04070
<i>Phosphatidylinositol signaling system</i>	6	<i>3-kinase</i>	ec:2.7.1.137	1	Cgcbs.017870	map04070
<i>Porphyrin and chlorophyll metabolism</i>	3	<i>Synthase</i>	ec:4.4.1.17	1	Cgcbs.019450	map00860
<i>Porphyrin and chlorophyll metabolism</i>	3	<i>Dehydrogenase</i>	ec:1.3.1.76	1	Cgcbs.055680	map00860
<i>Porphyrin and chlorophyll metabolism</i>	3	<i>Oxidase</i>	ec:1.3.3.3	1	Cgcbs.046050	map00860
<i>Primary bile acid biosynthesis</i>	1	<i>Dehydrogenase</i>	ec:1.1.1.35	1	Cgcbs.032400	map00120
<i>Propanoate metabolism</i>	4	<i>Dehydrogenase</i>	ec:1.8.1.4	1	Cgcbs.068190	map00640
<i>Propanoate metabolism</i>	4	<i>Transaminase</i>	ec:2.6.1.19	1	Cgcbs.045110	map00640
<i>Propanoate metabolism</i>	4	<i>ligase (ADP-forming)</i>	ec:6.2.1.5	1	Cgcbs.026350	map00640

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Propanoate metabolism</i>	4	<i>ligase (GDP-forming)</i>	ec:6.2.1.4	1	Cgcbs.026350	map00640
<i>Propanoate metabolism</i>	4	<i>Ligase</i>	ec:6.2.1.1	1	Cgcbs.013090	map00640
<i>Purine metabolism</i>	52	<i>Cyclase</i>	ec:4.6.1.1	1	Cgcbs.028060	map00230
<i>Purine metabolism</i>	52	<i>Kinase</i>	ec:2.7.4.3	1	Cgcbs.071770	map00230
<i>Purine metabolism</i>	52	<i>Adenylpyrophosphatase</i>	ec:3.6.1.3	13	Cgcbs.023720, Cgcbs.032380, Cgcbs.051730, Cgcbs.014780, Cgcbs.071630, Cgcbs.013610, Cgcbs.010710, Cgcbs.024890, Cgcbs.061450, Cgcbs.041550, Cgcbs.005100, Cgcbs.009440, Cgcbs.007770	map00230
<i>Purine metabolism</i>	52	<i>Adenyltransferase</i>	ec:2.7.7.53	1	Cgcbs.068970	map00230
<i>Purine metabolism</i>	52	<i>Allantoicase</i>	ec:3.5.3.4	1	Cgcbs.007650	map00230
<i>Purine metabolism</i>	52	<i>Phosphodiesterase</i>	ec:3.1.4.17	1	Cgcbs.000240	map00230
<i>Purine metabolism</i>	52	<i>Reductase</i>	ec:1.17.4.1	1	Cgcbs.063140	map00230
<i>Purine metabolism</i>	52	<i>Deaminase</i>	ec:3.5.4.6	2	Cgcbs.040010, Cgcbs.040020	map00230
<i>Purine metabolism</i>	52	<i>Synthase</i>	ec:6.3.2.6	1	Cgcbs.004660	map00230
<i>Purine metabolism</i>	52	<i>Dehydrogenase</i>	ec:1.1.1.205	1	Cgcbs.017070	map00230
<i>Purine metabolism</i>	52	<i>Kinase</i>	ec:2.7.1.25	1	Cgcbs.063090	map00230
<i>Purine metabolism</i>	52	<i>RNA polymerase</i>	ec:2.7.7.6	15	Cgcbs.013160, Cgcbs.037150, Cgcbs.022810, Cgcbs.072030, Cgcbs.018710, Cgcbs.031620, Cgcbs.068170, Cgcbs.017530, Cgcbs.032630, Cgcbs.030170, Cgcbs.008470, Cgcbs.008580, Cgcbs.062620, Cgcbs.061100, Cgcbs.022440, Cgcbs.004260, Cgcbs.047940, Cgcbs.037070, Cgcbs.031370, Cgcbs.053230, Cgcbs.053890, Cgcbs.062700, Cgcbs.057010, Cgcbs.073160, Cgcbs.065970, Cgcbs.022870, Cgcbs.037970	map00230
<i>Purine metabolism</i>	52	<i>DNA polymerase</i>	ec:2.7.7.7	12		map00230
<i>Purine metabolism</i>	52	<i>phosphoribosylidiphosphate 5-amidotransferase</i>	ec:2.4.2.14	1	Cgcbs.041260	map00230
<i>Purine metabolism</i>	52	<i>Kinase</i>	ec:2.7.4.10	1	Cgcbs.071770	map00230
<i>Pyrimidine metabolism</i>	34	<i>carbamoylaspartic dehydrase</i>	ec:3.5.2.3	1	Cgcbs.013880	map00240
<i>Pyrimidine metabolism</i>	34	<i>synthase (glutamine hydrolysing)</i>	ec:6.3.4.2	1	Cgcbs.043620	map00240
<i>Pyrimidine metabolism</i>	34	<i>Reductase</i>	ec:1.17.4.1	1	Cgcbs.063140	map00240
<i>Pyrimidine metabolism</i>	34	<i>Carbamoyltransferase</i>	ec:2.1.3.2	1	Cgcbs.015160	map00240

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Pyrimidine metabolism</i>	34	<i>Synthase</i>	<i>ec:4.2.1.70</i>	1	Cgcbs.048120	map00240
<i>Pyrimidine metabolism</i>	34	<i>Diphosphatase</i>	<i>ec:3.6.1.23</i>	1	Cgcbs.067680	map00240
<i>Pyrimidine metabolism</i>	34	<i>RNA polymerase</i>	<i>ec:2.7.7.6</i>	15	Cgcbs.013160, Cgcbs.037150, Cgcbs.022810, Cgcbs.072030, Cgcbs.018710, Cgcbs.031620, Cgcbs.068170, Cgcbs.017530, Cgcbs.032630, Cgcbs.030170, Cgcbs.008470, Cgcbs.008580, Cgcbs.062620, Cgcbs.061100, Cgcbs.022440, Cgcbs.004260, Cgcbs.047940, Cgcbs.037070, Cgcbs.031370, Cgcbs.053230, Cgcbs.053890, Cgcbs.062700, Cgcbs.057010, Cgcbs.073160, Cgcbs.065970, Cgcbs.022870, Cgcbs.037970	map00240
<i>Pyrimidine metabolism</i>	34	<i>DNA polymerase</i>	<i>ec:2.7.7.7</i>	12	Cgcbs.018730	map00240
<i>Pyrimidine metabolism</i>	34	<i>Kinase</i>	<i>ec:2.7.4.14</i>	1	Cgcbs.018730	map00240
<i>Pyruvate metabolism</i>	9	<i>Dehydrogenase</i>	<i>ec:1.8.1.4</i>	1	Cgcbs.068190	map00620
<i>Pyruvate metabolism</i>	9	<i>Synthase</i>	<i>ec:2.3.3.14</i>	1	Cgcbs.006470	map00620
<i>Pyruvate metabolism</i>	9	<i>Synthase</i>	<i>ec:2.3.3.13</i>	1	Cgcbs.051530	map00620
<i>Pyruvate metabolism</i>	9	<i>carboxykinase (ATP)</i>	<i>ec:4.1.1.49</i>	1	Cgcbs.033430	map00620
<i>Pyruvate metabolism</i>	9	<i>Hydrolase</i>	<i>ec:3.1.2.6</i>	1	Cgcbs.007880	map00620
<i>Pyruvate metabolism</i>	9	<i>Ligase</i>	<i>ec:6.2.1.1</i>	1	Cgcbs.013090	map00620
<i>Pyruvate metabolism</i>	9	<i>dehydrogenase (decarboxylating)</i>	<i>ec:1.1.1.39</i>	2	Cgcbs.032190, Cgcbs.022490	map00620
<i>Pyruvate metabolism</i>	9	<i>Acetyltransferase</i>	<i>ec:2.3.1.12</i>	1	Cgcbs.007790	map00620
<i>Riboflavin metabolism</i>	4	<i>Synthase</i>	<i>ec:2.5.1.9</i>	1	Cgcbs.043070	map00740
<i>Riboflavin metabolism</i>	4	<i>cyclohydrolase II</i>	<i>ec:3.5.4.25</i>	2	Cgcbs.055880, Cgcbs.015620	map00740
<i>Riboflavin metabolism</i>	4	<i>reductase (NADPH)</i>	<i>ec:1.5.1.30</i>	1	Cgcbs.021370	map00740
<i>Selenocompound metabolism</i>	1	<i>Ligase</i>	<i>ec:6.1.1.10</i>	1	Cgcbs.009860	map00450
<i>Sphingolipid metabolism</i>	2	<i>psychosine hydrolase</i>	<i>ec:3.2.1.45</i>	1	Cgcbs.015120	map00600
<i>Sphingolipid metabolism</i>	2	<i>Sulfatase</i>	<i>ec:3.1.6.1</i>	1	Cgcbs.049520	map00600
<i>Starch and sucrose metabolism</i>	12	<i>endo-1,4-beta-D-glucanase</i>	<i>ec:3.2.1.4</i>	1	Cgcbs.052320	map00500
<i>Starch and sucrose metabolism</i>	12	<i>Glycogenase</i>	<i>ec:3.2.1.1</i>	1	Cgcbs.035470	map00500
<i>Starch and sucrose metabolism</i>	12	<i>Isomerase</i>	<i>ec:5.3.1.9</i>	1	Cgcbs.065590	map00500
<i>Starch and sucrose metabolism</i>	12	<i>Decarboxylase</i>	<i>ec:4.1.1.35</i>	1	Cgcbs.026930	map00500
<i>Starch and sucrose metabolism</i>	12	<i>hexokinase type IV glucokinase</i>	<i>ec:2.7.1.1</i>	2	Cgcbs.001910, Cgcbs.051600	map00500

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Starch and sucrose metabolism</i>	12	<i>Maltase</i>	ec:3.2.1.20	1	Cgcb.051720	map00500
<i>Starch and sucrose metabolism</i>	12	<i>Synthase</i>	ec:2.4.1.34	1	Cgcb.058520	map00500
<i>Starch and sucrose metabolism</i>	12	<i>Trehalase</i>	ec:3.2.1.28	1	Cgcb.002100	map00500
<i>Starch and sucrose metabolism</i>	12	<i>amyl-1,6-glucosidase</i>	ec:3.2.1.33	1	Cgcb.069560	map00500
<i>Starch and sucrose metabolism</i>	12	<i>synthase (UDP-forming)</i>	ec:2.4.1.15	1	Cgcb.067530	map00500
<i>Starch and sucrose metabolism</i>	12	<i>Synthase</i>	ec:2.4.1.11	1	Cgcb.011750	map00500
<i>Starch and sucrose metabolism</i>	12	<i>disproportionating enzyme</i>	ec:2.4.1.25	1	Cgcb.069560	map00500
<i>Steroid hormone biosynthesis</i>	1	<i>Sulfatase</i>	ec:3.1.6.1	1	Cgcb.049520	map00140
<i>Streptomycin biosynthesis</i>	3	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	2	Cgcb.001910, Cgcb.051600	map00521
<i>Streptomycin biosynthesis</i>	3	<i>Synthase</i>	ec:5.5.1.4	1	Cgcb.049050	map00521
<i>Sulfur metabolism</i>	2	<i>Kinase</i>	ec:2.7.1.25	1	Cgcb.063090	map00920
<i>Sulfur metabolism</i>	2	<i>reductase (thioredoxin)</i>	ec:1.8.4.8	1	Cgcb.065790	map00920
<i>Synthesis and degradation of ketone bodies</i>	1	<i>CoA-transferase</i>	ec:2.8.3.5	1	Cgcb.054730	map00072
<i>T cell receptor signaling pathway</i>	5	<i>Phosphatase</i>	ec:3.1.3.16	5	Cgcb.066550, Cgcb.034450, Cgcb.007120, Cgcb.008770, Cgcb.034540	map04660
<i>Taurine and hypotaurine metabolism</i>	4	<i>glutamyl transpeptidase</i>	ec:2.3.2.2	3	Cgcb.041630, Cgcb.006850, Cgcb.028980	map00430
<i>Taurine and hypotaurine metabolism</i>	4	<i>Decarboxylase</i>	ec:4.1.1.15	1	Cgcb.069340	map00430
<i>Terpenoid backbone biosynthesis</i>	3	<i>O-methyltransferase</i>	ec:2.1.1.100	1	Cgcb.048460	map00900
<i>Terpenoid backbone biosynthesis</i>	3	<i>Decarboxylase</i>	ec:4.1.1.33	1	Cgcb.053790	map00900
<i>Terpenoid backbone biosynthesis</i>	3	<i>reductase (NADPH)</i>	ec:1.1.1.34	1	Cgcb.023710	map00900
<i>Thiamine metabolism</i>	1	<i>Diphosphokinase</i>	ec:2.7.6.2	1	Cgcb.030920	map00730
<i>Toluene degradation</i>	1	<i>Dehydrogenase</i>	ec:1.1.1.35	1	Cgcb.032400	map00623
<i>Tropane, piperidine and pyridine alkaloid biosynthesis</i>	1	<i>Oxidase</i>	ec:1.4.3.21	1	Cgcb.053910	map00960
<i>Tryptophan metabolism</i>	5	<i>3-monooxygenase</i>	ec:1.14.13.9	1	Cgcb.040610	map00380
<i>Tryptophan metabolism</i>	5	<i>Equilase</i>	ec:1.11.1.6	2	Cgcb.054930, Cgcb.055290	map00380
<i>Tryptophan metabolism</i>	5	<i>Dehydrogenase</i>	ec:1.1.1.35	1	Cgcb.032400	map00380
<i>Tryptophan metabolism</i>	5	<i>dehydrogenase (succinyl-transferring)</i>	ec:1.2.4.2	1	Cgcb.000910	map00380
<i>Tyrosine metabolism</i>	1	<i>Oxidase</i>	ec:1.4.3.21	1	Cgcb.053910	map00350

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Valine, leucine and isoleucine biosynthesis</i>	6	<i>Synthase</i>	ec:2.3.3.13	1	Cgcbs.051530	map00290
<i>Valine, leucine and isoleucine biosynthesis</i>	6	<i>Synthase</i>	ec:2.2.1.6	2	Cgcbs.003380, Cgcbs.059430	map00290
<i>Valine, leucine and isoleucine biosynthesis</i>	6	<i>ammonia-lyase</i>	ec:4.3.1.19	1	Cgcbs.037080	map00290
<i>Valine, leucine and isoleucine biosynthesis</i>	6	<i>reductoisomerase (NADP+)</i>	ec:1.1.1.86	1	Cgcbs.021180	map00290
<i>Valine, leucine and isoleucine biosynthesis</i>	6	<i>Dehydrogenase</i>	ec:1.1.1.85	1	Cgcbs.016960	map00290
<i>Valine, leucine and isoleucine degradation</i>	5	<i>Dehydrogenase</i>	ec:1.8.1.4	1	Cgcbs.068190	map00280
<i>Valine, leucine and isoleucine degradation</i>	5	<i>Dehydrogenase</i>	ec:1.1.1.35	1	Cgcbs.032400	map00280
<i>Valine, leucine and isoleucine degradation</i>	5	<i>Dehydrogenase</i>	ec:1.1.1.31	1	Cgcbs.049590	map00280
<i>Valine, leucine and isoleucine degradation</i>	5	<i>Ligase</i>	ec:6.2.1.16	1	Cgcbs.061300	map00280
<i>Valine, leucine and isoleucine degradation</i>	5	<i>CoA-transferase</i>	ec:2.8.3.5	1	Cgcbs.054730	map00280
<i>Various types of N-glycan biosynthesis</i>	3	<i>1,2-alpha-mannosidase</i>	ec:3.2.1.113	2	Cgcbs.034960, Cgcbs.027700	map00513
<i>Various types of N-glycan biosynthesis</i>	3	<i>N-acetylglucosaminyltransferase</i>	ec:2.4.1.141	1	Cgcbs.063990	map00513
<i>Vitamin B6 metabolism</i>	1	<i>Kinase</i>	ec:2.7.1.35	1	Cgcbs.053000	map00750

Tabela C. 2: Vias metabólicas mapeadas em *C. gattii*, linhagem R265. Todas as proteínas mapeadas referem-se às IDPs e IDRs identificadas neste estudo.

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Alanine, aspartate and glutamate metabolism</i>	7	<i>Transaminase</i>	ec:2.6.1.19	1	Cgr26.031970	map00250
<i>Alanine, aspartate and glutamate metabolism</i>	7	<i>Synthase</i>	ec:6.3.4.5	1	Cgr26.004250	map00250
<i>Alanine, aspartate and glutamate metabolism</i>	7	<i>synthase (NADH)</i>	ec:1.4.1.14	1	Cgr26.012120	map00250
<i>Alanine, aspartate and glutamate metabolism</i>	7	<i>synthase (NADPH)</i>	ec:1.4.1.13	1	Cgr26.012120	map00250
<i>Alanine, aspartate and glutamate metabolism</i>	7	<i>carbamoyltransferase</i>	ec:2.1.3.2	1	Cgr26.059420	map00250
<i>Alanine, aspartate and glutamate metabolism</i>	7	<i>synthase (glutamine-hydrolysing)</i>	ec:6.3.5.4	1	Cgr26.027280	map00250
<i>Alanine, aspartate and glutamate metabolism</i>	7	<i>Decarboxylase</i>	ec:4.1.1.15	1	Cgr26.022420	map00250
<i>Alanine, aspartate and glutamate metabolism</i>	7	<i>phosphoribosyl-diphosphate amidotransferase</i>	5- ec:2.4.2.14	1	Cgr26.054870	map00250
<i>Amino sugar and nucleotide sugar metabolism</i>	12	<i>Isomerase</i>	ec:5.3.1.9	1	Cgr26.014370	map00520
<i>Amino sugar and nucleotide sugar metabolism</i>	12	<i>Decarboxylase</i>	ec:4.1.1.35	1	Cgr26.004580	map00520
<i>Amino sugar and nucleotide sugar metabolism</i>	12	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	1	Cgr26.020830	map00520
<i>Amino sugar and nucleotide sugar metabolism</i>	12	<i>Synthase</i>	ec:1.1.1.271	1	Cgr26.058270	map00520
<i>Amino sugar and nucleotide sugar metabolism</i>	12	<i>Synthase</i>	ec:2.4.1.16	6	Cgr26.006640, Cgr26.050160, Cgr26.060070, Cgr26.056990, Cgr26.049640, Cgr26.004540	map00520

Via metabolic	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Amino sugar and nucleotide sugar metabolism</i>	12	<i>Chitodextrinase</i>	ec:3.2.1.14	2	Cgr26.018110, Cgr26.025510	map00520
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>synthase (glutamine-hydrolysing)</i>	ec:6.3.5.7	1	Cgr26.029800	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.18	1	Cgr26.046810	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.19	1	Cgr26.003320	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.14	1	Cgr26.007360	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.16	1	Cgr26.052220	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.10	1	Cgr26.040070	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.12	2	Cgr26.034260, Cgr26.002400	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.6	1	Cgr26.018790	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.7	1	Cgr26.009120	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.9	1	Cgr26.043070	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.2	1	Cgr26.041800	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.3	1	Cgr26.052620	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.5	1	Cgr26.045370	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.1	2	Cgr26.046250, Cgr26.022750	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.20	1	Cgr26.011580	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	19	<i>Ligase</i>	ec:6.1.1.22	2	Cgr26.032410, Cgr26.039730	map00970
<i>Aminobenzoate degradation</i>	2	<i>Phosphatase</i>	ec:3.1.3.2	2	Cgr26.049090, Cgr26.002120	map00627
<i>Arginine and proline metabolism</i>	3	<i>Oxidase</i>	ec:1.4.3.3	2	Cgr26.024470, Cgr26.020460	map00330
<i>Arginine and proline metabolism</i>	3	<i>5-kinase</i>	ec:2.7.2.11	1	Cgr26.021670	map00330
<i>Arginine biosynthesis</i>	3	<i>Synthase</i>	ec:6.3.4.5	1	Cgr26.004250	map00220
<i>Arginine biosynthesis</i>	3	<i>Kinase</i>	ec:2.7.2.8	1	Cgr26.008830	map00220
<i>Arginine biosynthesis</i>	3	<i>Reductase</i>	ec:1.2.1.38	1	Cgr26.008830	map00220
<i>Arginine biosynthesis</i>	3	<i>N-acetyltransferase</i>	ec:2.3.1.1	1	Cgr26.054610	map00220
<i>Benzoate degradation</i>	1	<i>Dehydrogenase</i>	ec:1.1.1.35	1	Cgr26.017520	map00362

Via metabolic	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>beta-Alanine metabolism</i>	3	<i>Transaminase</i>	ec:2.6.1.19	1	Cgr26.031970	map00410
<i>beta-Alanine metabolism</i>	3	<i>Decarboxylase</i>	ec:4.1.1.15	1	Cgr26.022420	map00410
<i>beta-Alanine metabolism</i>	3	<i>Oxidase</i>	ec:1.4.3.21	1	Cgr26.035190	map00410
<i>Biosynthesis of ansamycins</i>	1	<i>glycolaldehydetransferase</i>	ec:2.2.1.1	1	Cgr26.001390	map01051
<i>Biosynthesis of antibiotics</i>	47	<i>Dehydrogenase</i>	ec:1.8.1.4	1	Cgr26.033850	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>Synthase</i>	ec:2.3.3.14	1	Cgr26.037080	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>succinyltransferase</i>	ec:2.3.1.61	1	Cgr26.042200	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>Kinase</i>	ec:2.7.4.3	1	Cgr26.030800	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>Isomerase</i>	ec:5.3.1.9	1	Cgr26.014370	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>carboxykinase (ATP)</i>	ec:4.1.1.49	1	Cgr26.018450	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>Synthase</i>	ec:6.3.4.5	1	Cgr26.004250	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>Decarboxylase</i>	ec:4.1.1.33	1	Cgr26.035080	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>Oxidase</i>	ec:1.4.3.3	2	Cgr26.024470, Cgr26.020460	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>Kinase</i>	ec:2.7.1.71	1	Cgr26.043160	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>Hydratase</i>	ec:4.2.1.3	2	Cgr26.001430, Cgr26.038430	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>Dehydrogenase</i>	ec:1.1.1.95	1	Cgr26.057930	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>dihydroxyacetone transferase</i>	ec:2.2.1.2	1	Cgr26.013380	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>glycolaldehydetransferase</i>	ec:2.2.1.1	1	Cgr26.001390	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>Dehydratase</i>	ec:4.2.1.10	1	Cgr26.043160	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>Equilase</i>	ec:1.11.1.6	2	Cgr26.036390, Cgr26.036100	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>Synthase</i>	ec:2.2.1.6	2	Cgr26.045080, Cgr26.050780	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>Synthase</i>	ec:2.5.1.54	1	Cgr26.054350	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>Kinase</i>	ec:2.7.2.8	1	Cgr26.008830	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>Kinase</i>	ec:2.7.2.4	1	Cgr26.017030	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>Kinase</i>	ec:2.7.2.3	1	Cgr26.004250	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>Reductase</i>	ec:1.2.1.38	1	Cgr26.008830	map01130

Via metabolic	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Biosynthesis of antibiotics</i>	47	<i>5-kinase</i>	ec:2.7.2.11	1	Cgr26.021670	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>Deaminase</i>	ec:3.5.4.6	2	Cgr26.053810, Cgr26.053800	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>Synthase</i>	ec:6.3.2.6	1	Cgr26.046250	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>synthase (NADH)</i>	ec:1.4.1.14	1	Cgr26.012120	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>synthase (NADPH)</i>	ec:1.4.1.13	1	Cgr26.012120	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	1	Cgr26.020830	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>reductoisomerase (NADP+)</i>	ec:1.1.1.86	1	Cgr26.009090	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>Synthase</i>	ec:4.2.3.4	1	Cgr26.043160	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>phosphohexokinase</i>	ec:2.7.1.11	1	Cgr26.041180	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>N-acetyltransferase</i>	ec:2.3.1.1	1	Cgr26.054610	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>ligase (ADP-forming)</i>	ec:6.2.1.5	1	Cgr26.004090	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>Synthase</i>	ec:5.5.1.4	1	Cgr26.028540	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>ligase (GDP-forming)</i>	ec:6.2.1.4	1	Cgr26.004090	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>1-carboxyvinyltransferase</i>	ec:2.5.1.19	1	Cgr26.043160	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>Ligase</i>	ec:6.2.1.1	1	Cgr26.057710	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>Dehydrogenase</i>	ec:1.1.1.35	1	Cgr26.017520	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>reductase (NADPH)</i>	ec:1.1.1.34	1	Cgr26.011130	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>Synthase</i>	ec:2.5.1.21	1	Cgr26.055880	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>acetyltransferase</i>	ec:2.3.1.12	1	Cgr26.038250	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>dehydrogenase (NADP+-dependent, decarboxylating)</i>	ec:1.1.1.44	1	Cgr26.028980	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>dehydrogenase (NADP+)</i>	ec:1.1.1.42	1	Cgr26.014330	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>dehydrogenase (NAD+)</i>	ec:1.1.1.41	2	Cgr26.011400, Cgr26.060660	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>O-acetyltransferase</i>	ec:2.3.1.31	2	Cgr26.031080, Cgr26.035110	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>Dehydrogenase</i>	ec:1.1.1.25	1	Cgr26.043160	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>citrate synthase</i>	ec:2.3.3.8	1	Cgr26.004090	map01130

Via metabolic	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Biosynthesis of antibiotics</i>	47	<i>phosphoribosyl-diphosphate 5-amidotransferase</i>	ec:2.4.2.14	1	Cgr26.054870	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>dehydrogenase (succinyl-transferring)</i>	ec:1.2.4.2	1	Cgr26.042940	map01130
<i>Biosynthesis of antibiotics</i>	47	<i>Kinase</i>	ec:2.7.4.10	1	Cgr26.030800	map01130
<i>Biosynthesis of unsaturated fatty acids</i>	1	<i>9-desaturase</i>	ec:1.14.19.1	1	Cgr26.041040	map01040
<i>Biotin metabolism</i>	2	<i>Synthase</i>	ec:2.8.1.6	1	Cgr26.010380	map00780
<i>Biotin metabolism</i>	2	<i>Ligase</i>	ec:6.3.4.15	1	Cgr26.028290	map00780
<i>Butanoate metabolism</i>	7	<i>Transaminase</i>	ec:2.6.1.19	1	Cgr26.031970	map00650
<i>Butanoate metabolism</i>	7	<i>Synthase</i>	ec:2.2.1.6	2	Cgr26.045080, Cgr26.050780	map00650
<i>Butanoate metabolism</i>	7	<i>Dehydrogenase</i>	ec:1.1.1.35	1	Cgr26.017520	map00650
<i>Butanoate metabolism</i>	7	<i>Decarboxylase</i>	ec:4.1.1.15	1	Cgr26.022420	map00650
<i>Butanoate metabolism</i>	7	<i>Ligase</i>	ec:6.2.1.16	1	Cgr26.048040	map00650
<i>Butanoate metabolism</i>	7	<i>CoA-transferase</i>	ec:2.8.3.5	1	Cgr26.035940	map00650
<i>Butirosin and neomycin biosynthesis</i>	1	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	1	Cgr26.020830	map00524
<i>C5-Branched dibasic acid metabolism</i>	4	<i>Synthase</i>	ec:2.2.1.6	2	Cgr26.045080, Cgr26.050780	map00660
<i>C5-Branched dibasic acid metabolism</i>	4	<i>Dehydrogenase</i>	ec:1.1.1.85	1	Cgr26.060940	map00660
<i>C5-Branched dibasic acid metabolism</i>	4	<i>ligase (ADP-forming)</i>	ec:6.2.1.5	1	Cgr26.004090	map00660
<i>Caprolactam degradation</i>	1	<i>Dehydrogenase</i>	ec:1.1.1.35	1	Cgr26.017520	map00930
<i>Carbapenem biosynthesis</i>	1	<i>5-kinase</i>	ec:2.7.2.11	1	Cgr26.021670	map00332
<i>Carbon fixation in photosynthetic organisms</i>	5	<i>carboxykinase (ATP)</i>	ec:4.1.1.49	1	Cgr26.018450	map00710
<i>Carbon fixation in photosynthetic organisms</i>	5	<i>glycolaldehydetransferase</i>	ec:2.2.1.1	1	Cgr26.001390	map00710
<i>Carbon fixation in photosynthetic organisms</i>	5	<i>Kinase</i>	ec:2.7.2.3	1	Cgr26.004250	map00710
<i>Carbon fixation in photosynthetic organisms</i>	5	<i>dehydrogenase (decarboxylating)</i>	ec:1.1.1.39	2	Cgr26.017340, Cgr26.010210	map00710

Via metabolic	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Carbon fixation pathways in prokaryotes</i>	8	<i>reductase [NAD(P)H]</i>	ec:1.5.1.20	2	Cgr26.034920, Cgr26.058060	map00720
<i>Carbon fixation pathways in prokaryotes</i>	8	<i>Hydratase</i>	ec:4.2.1.3	2	Cgr26.001430, Cgr26.038430	map00720
<i>Carbon fixation pathways in prokaryotes</i>	8	<i>ligase (ADP-forming)</i>	ec:6.2.1.5	1	Cgr26.004090	map00720
<i>Carbon fixation pathways in prokaryotes</i>	8	<i>Ligase</i>	ec:6.2.1.1	1	Cgr26.057710	map00720
<i>Carbon fixation pathways in prokaryotes</i>	8	<i>Dehydrogenase</i>	ec:1.1.1.35	1	Cgr26.017520	map00720
<i>Carbon fixation pathways in prokaryotes</i>	8	<i>dehydrogenase (NADP+)</i>	ec:1.1.1.42	1	Cgr26.014330	map00720
<i>Carbon fixation pathways in prokaryotes</i>	8	<i>citrate synthase</i>	ec:2.3.3.8	1	Cgr26.004090	map00720
<i>Citrate cycle (TCA cycle)</i>	11	<i>Dehydrogenase</i>	ec:1.8.1.4	1	Cgr26.033850	map00020
<i>Citrate cycle (TCA cycle)</i>	11	<i>succinyltransferase</i>	ec:2.3.1.61	1	Cgr26.042200	map00020
<i>Citrate cycle (TCA cycle)</i>	11	<i>carboxykinase (ATP)</i>	ec:4.1.1.49	1	Cgr26.018450	map00020
<i>Citrate cycle (TCA cycle)</i>	11	<i>Hydratase</i>	ec:4.2.1.3	2	Cgr26.001430, Cgr26.038430	map00020
<i>Citrate cycle (TCA cycle)</i>	11	<i>ligase (ADP-forming)</i>	ec:6.2.1.5	1	Cgr26.004090	map00020
<i>Citrate cycle (TCA cycle)</i>	11	<i>ligase (GDP-forming)</i>	ec:6.2.1.4	1	Cgr26.004090	map00020
<i>Citrate cycle (TCA cycle)</i>	11	<i>acetyltransferase</i>	ec:2.3.1.12	1	Cgr26.038250	map00020
<i>Citrate cycle (TCA cycle)</i>	11	<i>dehydrogenase (NADP+)</i>	ec:1.1.1.42	1	Cgr26.014330	map00020
<i>Citrate cycle (TCA cycle)</i>	11	<i>dehydrogenase (NAD+)</i>	ec:1.1.1.41	2	Cgr26.011400, Cgr26.060660	map00020
<i>Citrate cycle (TCA cycle)</i>	11	<i>citrate synthase</i>	ec:2.3.3.8	1	Cgr26.004090	map00020
<i>Citrate cycle (TCA cycle)</i>	11	<i>dehydrogenase (succinyl-transferring)</i>	ec:1.2.4.2	1	Cgr26.042940	map00020
<i>Cyanoamino acid metabolism</i>	4	<i>glutamyl transpeptidase</i>	ec:2.3.2.2	3	Cgr26.055200, Cgr26.037410, Cgr26.006340	map00460
<i>Cyanoamino acid metabolism</i>	4	<i>Gentiobiase</i>	ec:3.2.1.21	1	Cgr26.026880	map00460
<i>Cysteine and methionine metabolism</i>	6	<i>S-methyltransferase</i>	ec:2.1.1.10	1	Cgr26.011830	map00270
<i>Cysteine and methionine metabolism</i>	6	<i>Kinase</i>	ec:2.7.2.4	1	Cgr26.017030	map00270
<i>Cysteine and methionine metabolism</i>	6	<i>Ligase</i>	ec:6.3.2.2	1	Cgr26.051870	map00270
<i>Cysteine and methionine metabolism</i>	6	<i>dioxygenase [iron(II)-requiring]</i>	ec:1.13.11.54	1	Cgr26.012220	map00270
<i>Cysteine and methionine metabolism</i>	6	<i>O-acetyltransferase</i>	ec:2.3.1.31	2	Cgr26.031080, Cgr26.035110	map00270
<i>D-Arginine and D-ornithine metabolism</i>	2	<i>Oxidase</i>	ec:1.4.3.3	2	Cgr26.024470, Cgr26.020460	map00472

Via metabolic	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Drug metabolism - other enzymes</i>	1	<i>Dehydrogenase</i>	ec:1.1.1.205	1	Cgr26.061010	map00983
<i>Ether lipid metabolism</i>	3	<i>D</i>	ec:3.1.4.4	2	Cgr26.037240, Cgr26.054790	map00565
<i>Ether lipid metabolism</i>	3	<i>Esterase</i>	ec:3.1.1.47	1	Cgr26.003010	map00565
<i>Fatty acid biosynthesis</i>	1	<i>Synthase</i>	ec:2.3.1.85	1	Cgr26.048510	map00061
<i>Fatty acid degradation</i>	1	<i>Dehydrogenase</i>	ec:1.1.1.35	1	Cgr26.017520	map00071
<i>Fatty acid elongation</i>	1	<i>Dehydrogenase</i>	ec:1.1.1.35	1	Cgr26.017520	map00062
<i>Folate biosynthesis</i>	2	<i>Synthase</i>	ec:6.3.2.17	1	Cgr26.019520	map00790
<i>Folate biosynthesis</i>	2	<i>cyclohydrolase I</i>	ec:3.5.4.16	1	Cgr26.013500	map00790
<i>Fructose and mannose metabolism</i>	5	<i>phosphofructokinase 2</i>	ec:2.7.1.105	2	Cgr26.059440, Cgr26.018410	map00051
<i>Fructose and mannose metabolism</i>	5	<i>2-phosphatase</i>	ec:3.1.3.46	2	Cgr26.059440, Cgr26.018410	map00051
<i>Fructose and mannose metabolism</i>	5	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	1	Cgr26.020830	map00051
<i>Fructose and mannose metabolism</i>	5	<i>phosphohexokinase</i>	ec:2.7.1.11	1	Cgr26.041180	map00051
<i>Fructose and mannose metabolism</i>	5	<i>Synthase</i>	ec:1.1.1.271	1	Cgr26.058270	map00051
<i>Galactose metabolism</i>	3	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	1	Cgr26.020830	map00052
<i>Galactose metabolism</i>	3	<i>phosphohexokinase</i>	ec:2.7.1.11	1	Cgr26.041180	map00052
<i>Galactose metabolism</i>	3	<i>Maltase</i>	ec:3.2.1.20	1	Cgr26.020940	map00052
<i>Geraniol degradation</i>	1	<i>Dehydrogenase</i>	ec:1.1.1.35	1	Cgr26.017520	map00281
<i>Glutathione metabolism</i>	8	<i>Reductase</i>	ec:1.17.4.1	1	Cgr26.049540	map00480
<i>Glutathione metabolism</i>	8	<i>Ligase</i>	ec:6.3.2.2	1	Cgr26.051870	map00480
<i>Glutathione metabolism</i>	8	<i>dehydrogenase (NADP+-dependent, decarboxylating)</i>	ec:1.1.1.44	1	Cgr26.028980	map00480
<i>Glutathione metabolism</i>	8	<i>dehydrogenase (NADP+)</i>	ec:1.1.1.42	1	Cgr26.014330	map00480
<i>Glutathione metabolism</i>	8	<i>glutamyl transpeptidase</i>	ec:2.3.2.2	3	Cgr26.055200, Cgr26.037410, Cgr26.006340	map00480
<i>Glutathione metabolism</i>	8	<i>thioredoxin peroxidase</i>	ec:1.11.1.15	1	Cgr26.027010	map00480
<i>Glycerolipid metabolism</i>	4	<i>kinase (ATP)</i>	ec:2.7.1.107	2	Cgr26.006840, Cgr26.038710	map00561
<i>Glycerolipid metabolism</i>	4	<i>O-acyltransferase</i>	ec:2.3.1.51	2	Cgr26.038120, Cgr26.022290	map00561

Via metabolic	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Glycerophospholipid metabolism</i>	12	<i>kinase (ATP)</i>	ec:2.7.1.107	2	Cgr26.006840, Cgr26.038710	map00564
<i>Glycerophospholipid metabolism</i>	12	<i>cytidyltransferase</i>	ec:2.7.7.41	1	Cgr26.036880	map00564
<i>Glycerophospholipid metabolism</i>	12	<i>N-methyltransferase</i>	ec:2.1.1.17	1	Cgr26.006230	map00564
<i>Glycerophospholipid metabolism</i>	12	<i>Decarboxylase</i>	ec:4.1.1.65	2	Cgr26.004490, Cgr26.011510	map00564
<i>Glycerophospholipid metabolism</i>	12	<i>D</i>	ec:3.1.4.4	2	Cgr26.037240, Cgr26.054790	map00564
<i>Glycerophospholipid metabolism</i>	12	<i>Dehydrogenase</i>	ec:1.1.5.3	1	Cgr26.054500	map00564
<i>Glycerophospholipid metabolism</i>	12	<i>lecithinase B</i>	ec:3.1.1.5	1	Cgr26.040010	map00564
<i>Glycerophospholipid metabolism</i>	12	<i>O-acyltransferase</i>	ec:2.3.1.51	2	Cgr26.038120, Cgr26.022290	map00564
<i>Glycine, serine and threonine metabolism</i>	10	<i>Dehydrogenase</i>	ec:1.8.1.4	1	Cgr26.033850	map00260
<i>Glycine, serine and threonine metabolism</i>	10	<i>S-aminomethyldihydrolipoylprotein: (6S)-tetrahydrofolate aminomethyltransferase (ammonia-forming)</i>	ec:2.1.2.10	1	Cgr26.054540	map00260
<i>Glycine, serine and threonine metabolism</i>	10	<i>Oxidase</i>	ec:1.4.3.3	2	Cgr26.024470, Cgr26.020460	map00260
<i>Glycine, serine and threonine metabolism</i>	10	<i>dehydrogenase (aminomethyl-transferring)</i>	ec:1.4.4.2	1	Cgr26.028050	map00260
<i>Glycine, serine and threonine metabolism</i>	10	<i>Dehydrogenase</i>	ec:1.1.1.95	1	Cgr26.057930	map00260
<i>Glycine, serine and threonine metabolism</i>	10	<i>Kinase</i>	ec:2.7.2.4	1	Cgr26.017030	map00260
<i>Glycine, serine and threonine metabolism</i>	10	<i>Oxidase</i>	ec:1.5.3.1	1	Cgr26.056140	map00260
<i>Glycine, serine and threonine metabolism</i>	10	<i>Dehydrogenase</i>	ec:1.1.99.1	1	Cgr26.022400	map00260
<i>Glycine, serine and threonine metabolism</i>	10	<i>Oxidase</i>	ec:1.4.3.21	1	Cgr26.035190	map00260
<i>Glycolysis / Gluconeogenesis</i>	8	<i>Dehydrogenase</i>	ec:1.8.1.4	1	Cgr26.033850	map00010
<i>Glycolysis / Gluconeogenesis</i>	8	<i>Isomerase</i>	ec:5.3.1.9	1	Cgr26.014370	map00010
<i>Glycolysis / Gluconeogenesis</i>	8	<i>carboxykinase (ATP)</i>	ec:4.1.1.49	1	Cgr26.018450	map00010
<i>Glycolysis / Gluconeogenesis</i>	8	<i>Kinase</i>	ec:2.7.2.3	1	Cgr26.004250	map00010

Via metabolic	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Glycolysis / Gluconeogenesis</i>	8	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	1	Cgr26.020830	map00010
<i>Glycolysis / Gluconeogenesis</i>	8	<i>phosphohexokinase</i>	ec:2.7.1.11	1	Cgr26.041180	map00010
<i>Glycolysis / Gluconeogenesis</i>	8	<i>Ligase</i>	ec:6.2.1.1	1	Cgr26.057710	map00010
<i>Glycolysis / Gluconeogenesis</i>	8	<i>acetyltransferase</i>	ec:2.3.1.12	1	Cgr26.038250	map00010
<i>Histidine metabolism</i>	2	<i>histidinol phosphate phosphatase</i>	ec:3.1.3.15	1	Cgr26.054950	map00340
<i>Histidine metabolism</i>	2	<i>Diphosphatase</i>	ec:3.6.1.31	1	Cgr26.034090	map00340
<i>Histidine metabolism</i>	2	<i>Dehydrogenase</i>	ec:1.1.1.23	1	Cgr26.034090	map00340
<i>Histidine metabolism</i>	2	<i>Cyclohydrolase</i>	ec:3.5.4.19	1	Cgr26.034090	map00340
<i>Inositol phosphate metabolism</i>	4	<i>3-kinase</i>	ec:2.7.1.127	1	Cgr26.055300	map00562
<i>Inositol phosphate metabolism</i>	4	<i>4-kinase</i>	ec:2.7.1.67	1	Cgr26.002610	map00562
<i>Inositol phosphate metabolism</i>	4	<i>Synthase</i>	ec:5.5.1.4	1	Cgr26.028540	map00562
<i>Inositol phosphate metabolism</i>	4	<i>3-kinase</i>	ec:2.7.1.137	1	Cgr26.061700	map00562
<i>Isoquinoline alkaloid biosynthesis</i>	1	<i>Oxidase</i>	ec:1.4.3.21	1	Cgr26.035190	map00950
<i>Lipoic acid metabolism</i>	1	<i>Synthase</i>	ec:2.8.1.8	1	Cgr26.034490	map00785
<i>Lysine biosynthesis</i>	2	<i>Synthase</i>	ec:2.3.3.14	1	Cgr26.037080	map00300
<i>Lysine biosynthesis</i>	2	<i>Kinase</i>	ec:2.7.2.4	1	Cgr26.017030	map00300
<i>Lysine degradation</i>	5	<i>N-methyltransferase</i>	ec:2.1.1.43	2	Cgr26.039440, Cgr26.024240	map00310
<i>Lysine degradation</i>	5	<i>succinyltransferase</i>	ec:2.3.1.61	1	Cgr26.042200	map00310
<i>Lysine degradation</i>	5	<i>Dehydrogenase</i>	ec:1.1.1.35	1	Cgr26.017520	map00310
<i>Lysine degradation</i>	5	<i>dehydrogenase (succinyl-transferring)</i>	ec:1.2.4.2	1	Cgr26.042940	map00310
<i>Methane metabolism</i>	3	<i>Dehydrogenase</i>	ec:1.1.1.95	1	Cgr26.057930	map00680
<i>Methane metabolism</i>	3	<i>phosphohexokinase</i>	ec:2.7.1.11	1	Cgr26.041180	map00680
<i>Methane metabolism</i>	3	<i>Ligase</i>	ec:6.2.1.1	1	Cgr26.057710	map00680
<i>Monobactam biosynthesis</i>	1	<i>Kinase</i>	ec:2.7.2.4	1	Cgr26.017030	map00261
<i>mTOR signaling pathway</i>	4	<i>protein kinase</i>	ec:2.7.11.24	4	Cgr26.017780, Cgr26.035650, Cgr26.015610, Cgr26.033440	map04150

Via metabolic	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>N-Glycan biosynthesis</i>	6	<i>Kinase</i>	ec:2.7.1.108	2	Cgr26.012830, Cgr26.012840	map00510
<i>N-Glycan biosynthesis</i>	6	<i>1,2-alpha-mannosidase</i>	ec:3.2.1.113	2	Cgr26.005290, Cgr26.023300	map00510
<i>N-Glycan biosynthesis</i>	6	<i>N-acetylglucosaminyltransferase</i>	ec:2.4.1.141	1	Cgr26.013050	map00510
<i>N-Glycan biosynthesis</i>	6	<i>Glucosidase</i>	ec:3.2.1.106	1	Cgr26.048320	map00510
<i>Nicotinate and nicotinamide metabolism</i>	5	<i>Kinase</i>	ec:2.7.1.23	4	Cgr26.017440, Cgr26.006840, Cgr26.017460, Cgr26.038710	map00760
<i>Nicotinate and nicotinamide metabolism</i>	5	<i>diphosphorylase (carboxylating)</i>	ec:2.4.2.19	1	Cgr26.054170	map00760
<i>Nitrogen metabolism</i>	1	<i>synthase (NADH)</i>	ec:1.4.1.14	1	Cgr26.012120	map00910
<i>Nitrogen metabolism</i>	1	<i>synthase (NADPH)</i>	ec:1.4.1.13	1	Cgr26.012120	map00910
<i>One carbon pool by folate</i>	3	<i>S-aminomethyldihydropolyprotein: (6S)-tetrahydrofolate aminomethyltransferase (ammonia-forming)</i>	ec:2.1.2.10	1	Cgr26.054540	map00670
<i>One carbon pool by folate</i>	3	<i>reductase [NAD(P)H]</i>	ec:1.5.1.20	2	Cgr26.034920, Cgr26.058060	map00670
<i>Other glycan degradation</i>	1	<i>psychosine hydrolase</i>	ec:3.2.1.45	1	Cgr26.059390	map00511
<i>Other types of O-glycan biosynthesis</i>	1	<i>mannosyltransferase</i>	ec:2.4.1.109	1	Cgr26.034110	map00514
<i>Pantothenate and CoA biosynthesis</i>	7	<i>hydroxymethyltransferase</i>	ec:2.1.2.11	1	Cgr26.041850	map00770
<i>Pantothenate and CoA biosynthesis</i>	7	<i>Synthase</i>	ec:2.2.1.6	2	Cgr26.045080, Cgr26.050780	map00770
<i>Pantothenate and CoA biosynthesis</i>	7	<i>Kinase</i>	ec:2.7.1.33	1	Cgr26.054980	map00770
<i>Pantothenate and CoA biosynthesis</i>	7	<i>reductoisomerase (NADP+)</i>	ec:1.1.1.86	1	Cgr26.009090	map00770
<i>Pantothenate and CoA biosynthesis</i>	7	<i>Synthase</i>	ec:2.7.8.7	2	Cgr26.018890, Cgr26.048510	map00770
<i>Penicillin and cephalosporin biosynthesis</i>	2	<i>Oxidase</i>	ec:1.4.3.3	2	Cgr26.024470, Cgr26.020460	map00311
<i>Pentose phosphate pathway</i>	5	<i>Isomerase</i>	ec:5.3.1.9	1	Cgr26.014370	map00030
<i>Pentose phosphate pathway</i>	5	<i>dihydroxyacetone transferase</i>	ec:2.2.1.2	1	Cgr26.013380	map00030
<i>Pentose phosphate pathway</i>	5	<i>glycolaldehydetransferase</i>	ec:2.2.1.1	1	Cgr26.001390	map00030
<i>Pentose phosphate pathway</i>	5	<i>Phosphohexokinase</i>	ec:2.7.1.11	1	Cgr26.041180	map00030

Via metabolic	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Pentose phosphate pathway</i>	5	<i>dehydrogenase (NADP+-dependent, decarboxylating)</i>	ec:1.1.1.44	1	Cgr26.028980	map00030
<i>Phenylalanine metabolism</i>	1	<i>Oxidase</i>	ec:1.4.3.21	1	Cgr26.035190	map00360
<i>Phenylalanine, tyrosine and tryptophan biosynthesis</i>	2	<i>Kinase</i>	ec:2.7.1.71	1	Cgr26.043160	map00400
<i>Phenylalanine, tyrosine and tryptophan biosynthesis</i>	2	<i>Dehydratase</i>	ec:4.2.1.10	1	Cgr26.043160	map00400
<i>Phenylalanine, tyrosine and tryptophan biosynthesis</i>	2	<i>Synthase</i>	ec:2.5.1.54	1	Cgr26.054350	map00400
<i>Phenylalanine, tyrosine and tryptophan biosynthesis</i>	2	<i>Synthase</i>	ec:4.2.3.4	1	Cgr26.043160	map00400
<i>Phenylalanine, tyrosine and tryptophan biosynthesis</i>	2	<i>1-carboxyvinyltransferase</i>	ec:2.5.1.19	1	Cgr26.043160	map00400
<i>Phenylalanine, tyrosine and tryptophan biosynthesis</i>	2	<i>Dehydrogenase</i>	ec:1.1.1.25	1	Cgr26.043160	map00400
<i>Phenylpropanoid biosynthesis</i>	2	<i>Lactoperoxidase</i>	ec:1.11.1.7	1	Cgr26.027010	map00940
<i>Phenylpropanoid biosynthesis</i>	2	<i>Gentiobiase</i>	ec:3.2.1.21	1	Cgr26.026880	map00940
<i>Phosphatidylinositol signaling system</i>	6	<i>kinase (ATP)</i>	ec:2.7.1.107	2	Cgr26.006840, Cgr26.038710	map04070
<i>Phosphatidylinositol signaling system</i>	6	<i>3-kinase</i>	ec:2.7.1.127	1	Cgr26.055300	map04070
<i>Phosphatidylinositol signaling system</i>	6	<i>4-kinase</i>	ec:2.7.1.67	1	Cgr26.002610	map04070
<i>Phosphatidylinositol signaling system</i>	6	<i>Cytidylyltransferase</i>	ec:2.7.7.41	1	Cgr26.036880	map04070
<i>Phosphatidylinositol signaling system</i>	6	<i>3-kinase</i>	ec:2.7.1.137	1	Cgr26.061700	map04070
<i>Porphyrin and chlorophyll metabolism</i>	3	<i>Synthase</i>	ec:4.4.1.17	1	Cgr26.007610	map00860
<i>Porphyrin and chlorophyll metabolism</i>	3	<i>Dehydrogenase</i>	ec:1.3.1.76	1	Cgr26.036730	map00860
<i>Porphyrin and chlorophyll metabolism</i>	3	<i>Oxidase</i>	ec:1.3.3.3	1	Cgr26.032810	map00860
<i>Primary bile acid biosynthesis</i>	1	<i>Dehydrogenase</i>	ec:1.1.1.35	1	Cgr26.017520	map00120
<i>Propanoate metabolism</i>	4	<i>Dehydrogenase</i>	ec:1.8.1.4	1	Cgr26.033850	map00640

Via metabolic	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Propanoate metabolism</i>	4	<i>Transaminase</i>	ec:2.6.1.19	1	Cgr26.031970	map00640
<i>Propanoate metabolism</i>	4	<i>ligase (ADP-forming)</i>	ec:6.2.1.5	1	Cgr26.004090	map00640
<i>Propanoate metabolism</i>	4	<i>ligase (GDP-forming)</i>	ec:6.2.1.4	1	Cgr26.004090	map00640
<i>Propanoate metabolism</i>	4	<i>Ligase</i>	ec:6.2.1.1	1	Cgr26.057710	map00640
<i>Purine metabolism</i>	52	<i>Cyclase</i>	ec:4.6.1.1	1	Cgr26.005600	map00230
<i>Purine metabolism</i>	52	<i>Kinase</i>	ec:2.7.4.3	1	Cgr26.030800	map00230
<i>Purine metabolism</i>	52	<i>Adenylypyrophosphatase</i>	ec:3.6.1.3	13	Cgr26.048160, Cgr26.055120, Cgr26.030700, Cgr26.040850, Cgr26.046610, Cgr26.020950, Cgr26.038230, Cgr26.002920, Cgr26.039690, Cgr26.058120, Cgr26.059100, Cgr26.017500, Cgr26.011140	map00230
<i>Purine metabolism</i>	52	<i>Adenylyltransferase</i>	ec:2.7.7.53	1	Cgr26.022090	map00230
<i>Purine metabolism</i>	52	<i>ec:3.5.3.4 allantoicase</i>	ec:3.5.3.4	1	Cgr26.038130	map00230
<i>Purine metabolism</i>	52	<i>Phosphodiesterase</i>	ec:3.1.4.17	1	Cgr26.042380	map00230
<i>Purine metabolism</i>	52	<i>Reductase</i>	ec:1.17.4.1	1	Cgr26.049540	map00230
<i>Purine metabolism</i>	52	<i>Deaminase</i>	ec:3.5.4.6	2	Cgr26.053810, Cgr26.053800	map00230
<i>Purine metabolism</i>	52	<i>Synthase</i>	ec:6.3.2.6	1	Cgr26.046250	map00230
<i>Purine metabolism</i>	52	<i>Dehydrogenase</i>	ec:1.1.1.205	1	Cgr26.061010	map00230
<i>Purine metabolism</i>	52	<i>Kinase</i>	ec:2.7.1.25	1	Cgr26.049490	map00230
<i>Purine metabolism</i>	52	<i>RNA polymerase</i>	ec:2.7.7.6	15	Cgr26.033830, Cgr26.057750, Cgr26.052290, Cgr26.025070, Cgr26.049140, Cgr26.007040, Cgr26.047870, Cgr26.010180, Cgr26.061370, Cgr26.038960, Cgr26.015620, Cgr26.038860, Cgr26.010470, Cgr26.016870, Cgr26.017740	map00230

Via metabolic	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Purine metabolism</i>	52	<i>DNA polymerase</i>	ec:2.7.7.7	12	Cgr26.045900, Cgr26.014700, Cgr26.010500, Cgr26.035170, Cgr26.049200, Cgr26.001120, Cgr26.034630, Cgr26.025000, Cgr26.016650, Cgr26.027620, Cgr26.025750, Cgr26.052960	map00230
<i>Purine metabolism</i>	52	<i>phosphoribosylidiphosphate amidotransferase</i>	5- ec:2.4.2.14	1	Cgr26.054870	map00230
<i>Purine metabolism</i>	52	<i>Kinase</i>	ec:2.7.4.10	1	Cgr26.030800	map00230
<i>Pyrimidine metabolism</i>	34	<i>carbamoylaspartic dehydrase</i>	ec:3.5.2.3	1	Cgr26.058330	map00240
<i>Pyrimidine metabolism</i>	34	<i>synthase (glutamine hydrolysing)</i>	ec:6.3.4.2	1	Cgr26.056860	map00240
<i>Pyrimidine metabolism</i>	34	<i>Reductase</i>	ec:1.17.4.1	1	Cgr26.049540	map00240
<i>Pyrimidine metabolism</i>	34	<i>carbamoyltransferase</i>	ec:2.1.3.2	1	Cgr26.059420	map00240
<i>Pyrimidine metabolism</i>	34	<i>Synthase</i>	ec:4.2.1.70	1	Cgr26.027780	map00240
<i>Pyrimidine metabolism</i>	34	<i>Diphosphatase</i>	ec:3.6.1.23	1	Cgr26.012450	map00240
<i>Pyrimidine metabolism</i>	34	<i>RNA polymerase</i>	ec:2.7.7.6	15	Cgr26.033830, Cgr26.057750, Cgr26.052290, Cgr26.025070, Cgr26.049140, Cgr26.007040, Cgr26.047870, Cgr26.010180, Cgr26.061370, Cgr26.038960, Cgr26.015620, Cgr26.038860, Cgr26.010470, Cgr26.016870, Cgr26.017740	map00240
<i>Pyrimidine metabolism</i>	34	<i>DNA polymerase</i>	ec:2.7.7.7	12	Cgr26.045900, Cgr26.014700, Cgr26.010500, Cgr26.035170, Cgr26.049200, Cgr26.001120, Cgr26.034630, Cgr26.025000, Cgr26.016650, Cgr26.027620, Cgr26.025750, Cgr26.052960	map00240
<i>Pyrimidine metabolism</i>	34	<i>Kinase</i>	ec:2.7.4.14	1	Cgr26.007060	map00240
<i>Pyruvate metabolism</i>	9	<i>Dehydrogenase</i>	ec:1.8.1.4	1	Cgr26.033850	map00620
<i>Pyruvate metabolism</i>	9	<i>Synthase</i>	ec:2.3.3.14	1	Cgr26.037080	map00620
<i>Pyruvate metabolism</i>	9	<i>Synthase</i>	ec:2.3.3.13	1	Cgr26.020750	map00620
<i>Pyruvate metabolism</i>	9	<i>carboxykinase (ATP)</i>	ec:4.1.1.49	1	Cgr26.018450	map00620
<i>Pyruvate metabolism</i>	9	<i>Hydrolase</i>	ec:3.1.2.6	1	Cgr26.038340	map00620

Via metabolic	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Pyruvate metabolism</i>	9	<i>Ligase</i>	ec:6.2.1.1	1	Cgr26.057710	map00620
<i>Pyruvate metabolism</i>	9	<i>dehydrogenase (decarboxylating)</i>	ec:1.1.1.39	2	Cgr26.017340, Cgr26.010210	map00620
<i>Pyruvate metabolism</i>	9	<i>acetyltransferase</i>	ec:2.3.1.12	1	Cgr26.038250	map00620
<i>Riboflavin metabolism</i>	5	<i>Synthase</i>	ec:2.5.1.9	1	Cgr26.056330	map00740
<i>Riboflavin metabolism</i>	5	<i>Kinase</i>	ec:2.7.1.26	1	Cgr26.056060	map00740
<i>Riboflavin metabolism</i>	5	<i>cyclohydrolase II</i>	ec:3.5.4.25	2	Cgr26.059800, Cgr26.000130	map00740
<i>Riboflavin metabolism</i>	5	<i>reductase (NADPH)</i>	ec:1.5.1.30	1	Cgr26.009260	map00740
<i>Selenocompound metabolism</i>	1	<i>Ligase</i>	ec:6.1.1.10	1	Cgr26.040070	map00450
<i>Sesquiterpenoid and triterpenoid biosynthesis</i>	1	<i>Synthase</i>	ec:2.5.1.21	1	Cgr26.055880	map00909
<i>Sphingolipid metabolism</i>	3	<i>psychosine hydrolase</i>	ec:3.2.1.45	1	Cgr26.059390	map00600
<i>Sphingolipid metabolism</i>	3	<i>Reductase</i>	ec:1.1.1.102	1	Cgr26.002280	map00600
<i>Sphingolipid metabolism</i>	3	<i>Sulfatase</i>	ec:3.1.6.1	1	Cgr26.028930	map00600
<i>Starch and sucrose metabolism</i>	12	<i>endo-1,4-beta-D-glucanase</i>	ec:3.2.1.4	1	Cgr26.021440	map00500
<i>Starch and sucrose metabolism</i>	12	<i>Glycogenase</i>	ec:3.2.1.1	1	Cgr26.023740	map00500
<i>Starch and sucrose metabolism</i>	12	<i>Isomerase</i>	ec:5.3.1.9	1	Cgr26.014370	map00500
<i>Starch and sucrose metabolism</i>	12	<i>Decarboxylase</i>	ec:4.1.1.35	1	Cgr26.004580	map00500
<i>Starch and sucrose metabolism</i>	12	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	1	Cgr26.020830	map00500
<i>Starch and sucrose metabolism</i>	12	<i>Maltase</i>	ec:3.2.1.20	1	Cgr26.020940	map00500
<i>Starch and sucrose metabolism</i>	12	<i>Gentiobiase</i>	ec:3.2.1.21	1	Cgr26.026880	map00500
<i>Starch and sucrose metabolism</i>	12	<i>Synthase</i>	ec:2.4.1.34	1	Cgr26.050010	map00500
<i>Starch and sucrose metabolism</i>	12	<i>Trehalase</i>	ec:3.2.1.28	1	Cgr26.043980	map00500
<i>Starch and sucrose metabolism</i>	12	<i>amyl-1,6-glycosidase</i>	ec:3.2.1.33	1	Cgr26.022610	map00500
<i>Starch and sucrose metabolism</i>	12	<i>synthase (UDP-forming)</i>	ec:2.4.1.15	1	Cgr26.012360	map00500
<i>Starch and sucrose metabolism</i>	12	<i>Synthase</i>	ec:2.4.1.11	1	Cgr26.041670	map00500
<i>Starch and sucrose metabolism</i>	12	<i>disproportionating enzyme</i>	ec:2.4.1.25	1	Cgr26.022610	map00500

Via metabolic	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Steroid biosynthesis</i>	1	<i>Synthase</i>	ec:2.5.1.21	1	Cgr26.055880	map00100
<i>Steroid hormone biosynthesis</i>	1	<i>Sulfatase</i>	ec:3.1.6.1	1	Cgr26.028930	map00140
<i>Streptomycin biosynthesis</i>	2	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	1	Cgr26.020830	map00521
<i>Streptomycin biosynthesis</i>	2	<i>Synthase</i>	ec:5.5.1.4	1	Cgr26.028540	map00521
<i>Sulfur metabolism</i>	2	<i>Kinase</i>	ec:2.7.1.25	1	Cgr26.049490	map00920
<i>Sulfur metabolism</i>	2	<i>reductase (thioredoxin)</i>	ec:1.8.4.8	1	Cgr26.014540	map00920
<i>Synthesis and degradation of ketone bodies</i>	1	<i>CoA-transferase</i>	ec:2.8.3.5	1	Cgr26.035940	map00072
<i>T cell receptor signaling pathway</i>	5	<i>Phosphatase</i>	ec:3.1.3.16	5	Cgr26.039110, Cgr26.019290, Cgr26.011540, Cgr26.019360, Cgr26.037650	map04660
<i>Taurine and hypotaurine metabolism</i>	4	<i>glutamyl transpeptidase</i>	ec:2.3.2.2	3	Cgr26.055200, Cgr26.037410, Cgr26.006340	map00430
<i>Taurine and hypotaurine metabolism</i>	4	<i>Decarboxylase</i>	ec:4.1.1.15	1	Cgr26.022420	map00430
<i>Terpenoid backbone biosynthesis</i>	2	<i>Decarboxylase</i>	ec:4.1.1.33	1	Cgr26.035080	map00900
<i>Terpenoid backbone biosynthesis</i>	2	<i>reductase (NADPH)</i>	ec:1.1.1.34	1	Cgr26.011130	map00900
<i>Thiamine metabolism</i>	1	<i>Desulfurase</i>	ec:2.8.1.7	1	Cgr26.033640	map00730
<i>Toluene degradation</i>	1	<i>Dehydrogenase</i>	ec:1.1.1.35	1	Cgr26.017520	map00623
<i>Tropane, piperidine and pyridine alkaloid biosynthesis</i>	1	<i>Oxidase</i>	ec:1.4.3.21	1	Cgr26.035190	map00960
<i>Tryptophan metabolism</i>	5	<i>3-monooxygenase</i>	ec:1.14.13.9	1	Cgr26.054320	map00380
<i>Tryptophan metabolism</i>	5	<i>Equilase</i>	ec:1.11.1.6	2	Cgr26.036390, Cgr26.036100	map00380
<i>Tryptophan metabolism</i>	5	<i>Dehydrogenase</i>	ec:1.1.1.35	1	Cgr26.017520	map00380
<i>Tryptophan metabolism</i>	5	<i>dehydrogenase (succinyl-transferring)</i>	ec:1.2.4.2	1	Cgr26.042940	map00380
<i>Tyrosine metabolism</i>	1	<i>Oxidase</i>	ec:1.4.3.21	1	Cgr26.035190	map00350
<i>Valine, leucine and isoleucine biosynthesis</i>	5	<i>Synthase</i>	ec:2.3.3.13	1	Cgr26.020750	map00290
<i>Valine, leucine and isoleucine biosynthesis</i>	5	<i>Synthase</i>	ec:2.2.1.6	2	Cgr26.045080, Cgr26.050780	map00290

Via metabolic	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Valine, leucine and isoleucine biosynthesis</i>	5	<i>reductoisomerase (NADP+)</i>	ec:1.1.1.86	1	Cgr26.009090	map00290
<i>Valine, leucine and isoleucine biosynthesis</i>	5	<i>Dehydrogenase</i>	ec:1.1.1.85	1	Cgr26.060940	map00290
<i>Valine, leucine and isoleucine degradation</i>	5	<i>Dehydrogenase</i>	ec:1.8.1.4	1	Cgr26.033850	map00280
<i>Valine, leucine and isoleucine degradation</i>	5	<i>Dehydrogenase</i>	ec:1.1.1.35	1	Cgr26.017520	map00280
<i>Valine, leucine and isoleucine degradation</i>	5	<i>Dehydrogenase</i>	ec:1.1.1.31	1	Cgr26.028980	map00280
<i>Valine, leucine and isoleucine degradation</i>	5	<i>Ligase</i>	ec:6.2.1.16	1	Cgr26.048040	map00280
<i>Valine, leucine and isoleucine degradation</i>	5	<i>CoA-transferase</i>	ec:2.8.3.5	1	Cgr26.035940	map00280
<i>Various types of N-glycan biosynthesis</i>	3	<i>1,2-alpha-mannosidase</i>	ec:3.2.1.113	2	Cgr26.005290, Cgr26.023300	map00513
<i>Various types of N-glycan biosynthesis</i>	3	<i>N-acetylglucosaminyltransferase</i>	ec:2.4.1.141	1	Cgr26.013050	map00513
<i>Vitamin B6 metabolism</i>	1	<i>Kinase</i>	ec:2.7.1.35	1	Cgr26.034430	map00750

Tabela C. 3: Vias metabólicas mapeadas em *C. gattii*, linhagem WM276. Todas as proteínas mapeadas referem-se às IDPs e IDRs identificadas neste estudo.

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Alanine, aspartate and glutamate metabolism</i>	5	<i>Transaminase</i>	ec:2.6.1.19	1	Cgwm2.031420	map00250
<i>Alanine, aspartate and glutamate metabolism</i>	5	<i>synthase (NADH)</i>	ec:1.4.1.14	1	Cgwm2.051890	map00250
<i>Alanine, aspartate and glutamate metabolism</i>	5	<i>carbamoyltransferase</i>	ec:2.1.3.2	1	Cgwm2.005210	map00250
<i>Alanine, aspartate and glutamate metabolism</i>	5	<i>synthase (glutamine-hydrolysing)</i>	ec:6.3.5.4	1	Cgwm2.057550	map00250
<i>Alanine, aspartate and glutamate metabolism</i>	5	<i>phosphoribosylidiphosphate amidotransferase</i>	5- ec:2.4.2.14	1	Cgwm2.019260	map00250
<i>Amino sugar and nucleotide sugar metabolism</i>	7	<i>Decarboxylase</i>	ec:4.1.1.35	1	Cgwm2.041700	map00520
<i>Amino sugar and nucleotide sugar metabolism</i>	7	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	1	Cgwm2.059230	map00520
<i>Amino sugar and nucleotide sugar metabolism</i>	7	<i>Synthase</i>	ec:2.4.1.16	3	Cgwm2.041740, Cgwm2.021440, Cgwm2.065400	map00520
<i>Amino sugar and nucleotide sugar metabolism</i>	7	<i>Chitodextrinase</i>	ec:3.2.1.14	2	Cgwm2.047620, Cgwm2.057120	map00520
<i>Aminoacyl-tRNA biosynthesis</i>	18	<i>synthase (glutamine-hydrolysing)</i>	ec:6.3.5.7	1	Cgwm2.050510	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	18	<i>Ligase</i>	ec:6.1.1.18	1	Cgwm2.008330	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	18	<i>Ligase</i>	ec:6.1.1.19	1	Cgwm2.043000	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	18	<i>Ligase</i>	ec:6.1.1.14	1	Cgwm2.038620	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	18	<i>Ligase</i>	ec:6.1.1.16	1	Cgwm2.013280	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	18	<i>Ligase</i>	ec:6.1.1.10	1	Cgwm2.027370	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	18	<i>Ligase</i>	ec:6.1.1.12	2	Cgwm2.043960, Cgwm2.049180	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	18	<i>Ligase</i>	ec:6.1.1.6	1	Cgwm2.048290	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	18	<i>Ligase</i>	ec:6.1.1.7	1	Cgwm2.036860	map00970

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Aminoacyl-tRNA biosynthesis</i>	18	<i>Ligase</i>	ec:6.1.1.9	1	Cgwm2.012270	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	18	<i>Ligase</i>	ec:6.1.1.2	1	Cgwm2.029150	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	18	<i>Ligase</i>	ec:6.1.1.3	1	Cgwm2.013700	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	18	<i>Ligase</i>	ec:6.1.1.5	1	Cgwm2.009790	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	18	<i>Ligase</i>	ec:6.1.1.1	1	Cgwm2.008900	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	18	<i>Ligase</i>	ec:6.1.1.20	1	Cgwm2.051320	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	18	<i>Ligase</i>	ec:6.1.1.22	2	Cgwm2.030920, Cgwm2.027020	map00970
<i>Aminobenzoate degradation</i>	2	<i>Phosphatase</i>	ec:3.1.3.2	2	Cgwm2.044270, Cgwm2.033300	map00627
<i>Arginine and proline metabolism</i>	1	<i>5-kinase</i>	ec:2.7.2.11	1	Cgwm2.058350	map00330
<i>Arginine biosynthesis</i>	1	<i>Kinase</i>	ec:2.7.2.8	1	Cgwm2.037160	map00220
<i>Arginine biosynthesis</i>	1	<i>Reductase</i>	ec:1.2.1.38	1	Cgwm2.037160	map00220
<i>beta-Alanine metabolism</i>	3	<i>Transaminase</i>	ec:2.6.1.19	1	Cgwm2.031420	map00410
<i>beta-Alanine metabolism</i>	3	<i>Oxidase</i>	ec:1.4.3.21	2	Cgwm2.023190, Cgwm2.043180	map00410
<i>Biosynthesis of antibiotics</i>	34	<i>Dehydrogenase</i>	ec:1.8.1.4	1	Cgwm2.049660	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>Synthase</i>	ec:2.3.3.14	1	Cgwm2.024330	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>succinyltransferase</i>	ec:2.3.1.61	1	Cgwm2.013190	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>carboxykinase (ATP)</i>	ec:4.1.1.49	1	Cgwm2.047890	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>Decarboxylase</i>	ec:4.1.1.33	1	Cgwm2.023310	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>Hydratase</i>	ec:4.2.1.3	2	Cgwm2.062620, Cgwm2.025710	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>Dehydrogenase</i>	ec:1.1.1.95	1	Cgwm2.006750	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>Synthase</i>	ec:2.2.1.6	1	Cgwm2.010000	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>Equilase</i>	ec:1.11.1.6	1	Cgwm2.022230	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>Synthase</i>	ec:2.5.1.54	1	Cgwm2.018730	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>Kinase</i>	ec:2.7.2.8	1	Cgwm2.037160	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>Kinase</i>	ec:2.7.2.4	1	Cgwm2.046660	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>Kinase</i>	ec:2.7.2.3	1	Cgwm2.042010	map01130

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Biosynthesis of antibiotics</i>	34	<i>Reductase</i>	ec:1.2.1.38	1	Cgwm2.037160	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>5-kinase</i>	ec:2.7.2.11	1	Cgwm2.058350	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>Deaminase</i>	ec:3.5.4.6	1	Cgwm2.018130	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>Synthase</i>	ec:6.3.2.6	1	Cgwm2.008900	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>synthase (NADH)</i>	ec:1.4.1.14	1	Cgwm2.051890	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>ammonia-lyase</i>	ec:4.3.1.19	1	Cgwm2.056570	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	1	Cgwm2.059230	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>reductoisomerase (NADP+)</i>	ec:1.1.1.86	1	Cgwm2.036890	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>Hydratase</i>	ec:4.2.1.36	1	Cgwm2.057410	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>phosphohexokinase</i>	ec:2.7.1.11	1	Cgwm2.028480	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>Synthase</i>	ec:5.5.1.4	1	Cgwm2.014980	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>Ligase</i>	ec:6.2.1.1	1	Cgwm2.006980	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>reductase (NADPH)</i>	ec:1.1.1.34	1	Cgwm2.034750	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>acetyltransferase</i>	ec:2.3.1.12	1	Cgwm2.025530	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>dehydrogenase (NADP+-dependent, decarboxylating)</i>	ec:1.1.1.44	1	Cgwm2.014520	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>dehydrogenase (NADP+)</i>	ec:1.1.1.42	1	Cgwm2.001680	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>dehydrogenase (NAD+)</i>	ec:1.1.1.41	1	Cgwm2.003790	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>Oxidase</i>	ec:1.8.3.5	1	Cgwm2.056040	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>O-acetyltransferase</i>	ec:2.3.1.31	2	Cgwm2.032350, Cgwm2.023280	map01130
<i>Biosynthesis of antibiotics</i>	34	<i>phosphoribosylidiphosphate 5-amidotransferase</i>	ec:2.4.2.14	1	Cgwm2.019260	map01130
<i>Biotin metabolism</i>	2	<i>Synthase</i>	ec:2.8.1.6	1	Cgwm2.035550	map00780
<i>Biotin metabolism</i>	2	<i>Ligase</i>	ec:6.3.4.15	1	Cgwm2.015240	map00780
<i>Butanoate metabolism</i>	4	<i>Transaminase</i>	ec:2.6.1.19	1	Cgwm2.031420	map00650
<i>Butanoate metabolism</i>	4	<i>Synthase</i>	ec:2.2.1.6	1	Cgwm2.010000	map00650
<i>Butanoate metabolism</i>	4	<i>Ligase</i>	ec:6.2.1.16	1	Cgwm2.034400	map00650

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Butanoate metabolism</i>	4	<i>CoA-transferase</i>	ec:2.8.3.5	1	Cgwm2.022390	map00650
<i>Butirosin and neomycin biosynthesis</i>	1	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	1	Cgwm2.059230	map00524
<i>C5-Branched dibasic acid metabolism</i>	2	<i>Synthase</i>	ec:2.2.1.6	1	Cgwm2.010000	map00660
<i>C5-Branched dibasic acid metabolism</i>	2	<i>Dehydrogenase</i>	ec:1.1.1.85	1	Cgwm2.003510	map00660
<i>Carbapenem biosynthesis</i>	1	<i>5-kinase</i>	ec:2.7.2.11	1	Cgwm2.058350	map00332
<i>Carbon fixation in photosynthetic organisms</i>	4	<i>carboxykinase (ATP)</i>	ec:4.1.1.49	1	Cgwm2.047890	map00710
<i>Carbon fixation in photosynthetic organisms</i>	4	<i>Kinase</i>	ec:2.7.2.3	1	Cgwm2.042010	map00710
<i>Carbon fixation in photosynthetic organisms</i>	4	<i>dehydrogenase (decarboxylating)</i>	ec:1.1.1.39	2	Cgwm2.046860, Cgwm2.035730	map00710
<i>Carbon fixation pathways in prokaryotes</i>	7	<i>Ligase</i>	ec:6.3.4.3	1	Cgwm2.043760	map00720
<i>Carbon fixation pathways in prokaryotes</i>	7	<i>reductase [NAD(P)H]</i>	ec:1.5.1.20	2	Cgwm2.023450, Cgwm2.006620	map00720
<i>Carbon fixation pathways in prokaryotes</i>	7	<i>Hydratase</i>	ec:4.2.1.3	2	Cgwm2.062620, Cgwm2.025710	map00720
<i>Carbon fixation pathways in prokaryotes</i>	7	<i>Ligase</i>	ec:6.2.1.1	1	Cgwm2.006980	map00720
<i>Carbon fixation pathways in prokaryotes</i>	7	<i>dehydrogenase (NADP+)</i>	ec:1.1.1.42	1	Cgwm2.001680	map00720
<i>Carbon fixation pathways in prokaryotes</i>	7	<i>dehydrogenase (NADP+)</i>	ec:1.5.1.5	1	Cgwm2.043760	map00720
<i>Citrate cycle (TCA cycle)</i>	8	<i>Dehydrogenase</i>	ec:1.8.1.4	1	Cgwm2.049660	map00020
<i>Citrate cycle (TCA cycle)</i>	8	<i>succinyltransferase</i>	ec:2.3.1.61	1	Cgwm2.013190	map00020
<i>Citrate cycle (TCA cycle)</i>	8	<i>carboxykinase (ATP)</i>	ec:4.1.1.49	1	Cgwm2.047890	map00020
<i>Citrate cycle (TCA cycle)</i>	8	<i>Hydratase</i>	ec:4.2.1.3	2	Cgwm2.062620, Cgwm2.025710	map00020
<i>Citrate cycle (TCA cycle)</i>	8	<i>acetyltransferase</i>	ec:2.3.1.12	1	Cgwm2.025530	map00020
<i>Citrate cycle (TCA cycle)</i>	8	<i>dehydrogenase (NADP+)</i>	ec:1.1.1.42	1	Cgwm2.001680	map00020
<i>Citrate cycle (TCA cycle)</i>	8	<i>dehydrogenase (NAD+)</i>	ec:1.1.1.41	1	Cgwm2.003790	map00020
<i>Cyanoamino acid metabolism</i>	3	<i>glutamyl transpeptidase</i>	ec:2.3.2.2	3	Cgwm2.019590, Cgwm2.039820, Cgwm2.024670	map00460
<i>Cysteine and methionine metabolism</i>	5	<i>S-methyltransferase</i>	ec:2.1.1.10	1	Cgwm2.051560	map00270

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Cysteine and methionine metabolism</i>	5	<i>Kinase</i>	ec:2.7.2.4	1	Cgwm2.046660	map00270
<i>Cysteine and methionine metabolism</i>	5	<i>Ligase</i>	ec:6.3.2.2	1	Cgwm2.063660	map00270
<i>Cysteine and methionine metabolism</i>	5	<i>O-acetyltransferase</i>	ec:2.3.1.31	2	Cgwm2.032350, Cgwm2.023280	map00270
<i>Drug metabolism - cytochrome P450</i>	1	<i>monooxygenase</i>	ec:1.14.13.8	1	Cgwm2.014410	map00982
<i>Drug metabolism - other enzymes</i>	2	<i>Kinase</i>	ec:2.7.1.48	1	Cgwm2.042090	map00983
<i>Drug metabolism - other enzymes</i>	2	<i>Dehydrogenase</i>	ec:1.1.1.205	1	Cgwm2.003410	map00983
<i>Ether lipid metabolism</i>	2	<i>D</i>	ec:3.1.4.4	1	Cgwm2.019180	map00565
<i>Ether lipid metabolism</i>	2	<i>Esterase</i>	ec:3.1.1.47	1	Cgwm2.043300	map00565
<i>Folate biosynthesis</i>	2	<i>Synthase</i>	ec:6.3.2.17	1	Cgwm2.060530	map00790
<i>Folate biosynthesis</i>	2	<i>cyclohydrolase I</i>	ec:3.5.4.16	1	Cgwm2.000830	map00790
<i>Fructose and mannose metabolism</i>	4	<i>phosphofructokinase 2</i>	ec:2.7.1.105	2	Cgwm2.005190, Cgwm2.047850	map00051
<i>Fructose and mannose metabolism</i>	4	<i>2-phosphatase</i>	ec:3.1.3.46	2	Cgwm2.005190, Cgwm2.047850	map00051
<i>Fructose and mannose metabolism</i>	4	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	1	Cgwm2.059230	map00051
<i>Fructose and mannose metabolism</i>	4	<i>phosphohexokinase</i>	ec:2.7.1.11	1	Cgwm2.028480	map00051
<i>Galactose metabolism</i>	2	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	1	Cgwm2.059230	map00052
<i>Galactose metabolism</i>	2	<i>phosphohexokinase</i>	ec:2.7.1.11	1	Cgwm2.028480	map00052
<i>Glutathione metabolism</i>	8	<i>Reductase</i>	ec:1.17.4.1	1	Cgwm2.032850	map00480
<i>Glutathione metabolism</i>	8	<i>Ligase</i>	ec:6.3.2.2	1	Cgwm2.063660	map00480
<i>Glutathione metabolism</i>	8	<i>dehydrogenase (NADP+-dependent, decarboxylating)</i>	ec:1.1.1.44	1	Cgwm2.014520	map00480
<i>Glutathione metabolism</i>	8	<i>dehydrogenase (NADP+)</i>	ec:1.1.1.42	1	Cgwm2.001680	map00480
<i>Glutathione metabolism</i>	8	<i>glutamyl transpeptidase</i>	ec:2.3.2.2	3	Cgwm2.019590, Cgwm2.039820, Cgwm2.024670	map00480
<i>Glutathione metabolism</i>	8	<i>thioredoxin peroxidase</i>	ec:1.11.1.15	1	Cgwm2.057820	map00480
<i>Glycerolipid metabolism</i>	3	<i>kinase (ATP)</i>	ec:2.7.1.107	1	Cgwm2.025990	map00561
<i>Glycerolipid metabolism</i>	3	<i>O-acyltransferase</i>	ec:2.3.1.51	2	Cgwm2.025420, Cgwm2.053200	map00561
<i>Glycerophospholipid metabolism</i>	10	<i>kinase (ATP)</i>	ec:2.7.1.107	1	Cgwm2.025990	map00564

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Glycerophospholipid metabolism</i>	10	<i>cytidyltransferase</i>	ec:2.7.7.41	1	Cgwm2.024130	map00564
<i>Glycerophospholipid metabolism</i>	10	<i>N-methyltransferase</i>	ec:2.1.1.17	1	Cgwm2.039920	map00564
<i>Glycerophospholipid metabolism</i>	10	<i>Decarboxylase</i>	ec:4.1.1.65	2	Cgwm2.051240, Cgwm2.041790	map00564
<i>Glycerophospholipid metabolism</i>	10	<i>D</i>	ec:3.1.4.4	1	Cgwm2.019180	map00564
<i>Glycerophospholipid metabolism</i>	10	<i>Dehydrogenase</i>	ec:1.1.5.3	1	Cgwm2.018890	map00564
<i>Glycerophospholipid metabolism</i>	10	<i>lecithinase B</i>	ec:3.1.1.5	1	Cgwm2.027320	map00564
<i>Glycerophospholipid metabolism</i>	10	<i>O-acyltransferase</i>	ec:2.3.1.51	2	Cgwm2.025420, Cgwm2.053200	map00564
<i>Glycine, serine and threonine metabolism</i>	8	<i>Dehydrogenase</i>	ec:1.8.1.4	1	Cgwm2.049660	map00260
<i>Glycine, serine and threonine metabolism</i>	8	<i>dehydrogenase (aminomethyl-transferring)</i>	ec:1.4.4.2	1	Cgwm2.015490	map00260
<i>Glycine, serine and threonine metabolism</i>	8	<i>Dehydrogenase</i>	ec:1.1.1.95	1	Cgwm2.006750	map00260
<i>Glycine, serine and threonine metabolism</i>	8	<i>Kinase</i>	ec:2.7.2.4	1	Cgwm2.046660	map00260
<i>Glycine, serine and threonine metabolism</i>	8	<i>ammonia-lyase</i>	ec:4.3.1.19	1	Cgwm2.056570	map00260
<i>Glycine, serine and threonine metabolism</i>	8	<i>Dehydrogenase</i>	ec:1.1.99.1	1	Cgwm2.053090	map00260
<i>Glycine, serine and threonine metabolism</i>	8	<i>Oxidase</i>	ec:1.4.3.21	2	Cgwm2.023190, Cgwm2.043180	map00260
<i>Glycolysis / Gluconeogenesis</i>	7	<i>Dehydrogenase</i>	ec:1.8.1.4	1	Cgwm2.049660	map00010
<i>Glycolysis / Gluconeogenesis</i>	7	<i>carboxykinase (ATP)</i>	ec:4.1.1.49	1	Cgwm2.047890	map00010
<i>Glycolysis / Gluconeogenesis</i>	7	<i>Kinase</i>	ec:2.7.2.3	1	Cgwm2.042010	map00010
<i>Glycolysis / Gluconeogenesis</i>	7	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	1	Cgwm2.059230	map00010
<i>Glycolysis / Gluconeogenesis</i>	7	<i>phosphohexokinase</i>	ec:2.7.1.11	1	Cgwm2.028480	map00010
<i>Glycolysis / Gluconeogenesis</i>	7	<i>Ligase</i>	ec:6.2.1.1	1	Cgwm2.006980	map00010
<i>Glycolysis / Gluconeogenesis</i>	7	<i>acetyltransferase</i>	ec:2.3.1.12	1	Cgwm2.025530	map00010
<i>Glyoxylate and dicarboxylate metabolism</i>	3	<i>Hydratase</i>	ec:4.2.1.3	2	Cgwm2.062620, Cgwm2.025710	map00630
<i>Glyoxylate and dicarboxylate metabolism</i>	3	<i>Equilase</i>	ec:1.11.1.6	1	Cgwm2.022230	map00630
<i>Histidine metabolism</i>	2	<i>histidinol phosphate phosphatase</i>	ec:3.1.3.15	1	Cgwm2.019340	map00340
<i>Histidine metabolism</i>	2	<i>Diphosphatase</i>	ec:3.6.1.31	1	Cgwm2.049350	map00340

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Histidine metabolism</i>	2	<i>Dehydrogenase</i>	ec:1.1.1.23	1	Cgwm2.049350	map00340
<i>Histidine metabolism</i>	2	<i>Cyclohydrolase</i>	ec:3.5.4.19	1	Cgwm2.049350	map00340
<i>Inositol phosphate metabolism</i>	4	<i>3-kinase</i>	ec:2.7.1.127	1	Cgwm2.019660	map00562
<i>Inositol phosphate metabolism</i>	4	<i>4-kinase</i>	ec:2.7.1.67	1	Cgwm2.043740	map00562
<i>Inositol phosphate metabolism</i>	4	<i>Synthase</i>	ec:5.5.1.4	1	Cgwm2.014980	map00562
<i>Inositol phosphate metabolism</i>	4	<i>3-kinase</i>	ec:2.7.1.137	1	Cgwm2.002640	map00562
<i>Isoquinoline alkaloid biosynthesis</i>	2	<i>Oxidase</i>	ec:1.4.3.21	2	Cgwm2.023190, Cgwm2.043180	map00950
<i>Lipoic acid metabolism</i>	1	<i>Synthase</i>	ec:2.8.1.8	1	Cgwm2.023930	map00785
<i>Lysine biosynthesis</i>	3	<i>Synthase</i>	ec:2.3.3.14	1	Cgwm2.024330	map00300
<i>Lysine biosynthesis</i>	3	<i>Kinase</i>	ec:2.7.2.4	1	Cgwm2.046660	map00300
<i>Lysine biosynthesis</i>	3	<i>Hydratase</i>	ec:4.2.1.36	1	Cgwm2.057410	map00300
<i>Lysine degradation</i>	4	<i>succinyltransferase</i>	ec:2.3.1.61	1	Cgwm2.013190	map00310
<i>Lysine degradation</i>	4	<i>N-methyltransferase</i>	ec:2.1.1.43	3	Cgwm2.058520, Cgwm2.055760, Cgwm2.058710	map00310
<i>Methane metabolism</i>	3	<i>Dehydrogenase</i>	ec:1.1.1.95	1	Cgwm2.006750	map00680
<i>Methane metabolism</i>	3	<i>phosphohexokinase</i>	ec:2.7.1.11	1	Cgwm2.028480	map00680
<i>Methane metabolism</i>	3	<i>Ligase</i>	ec:6.2.1.1	1	Cgwm2.006980	map00680
<i>Monobactam biosynthesis</i>	1	<i>Kinase</i>	ec:2.7.2.4	1	Cgwm2.046660	map00261
<i>mTOR signaling pathway</i>	3	<i>protein kinase</i>	ec:2.7.11.24	3	Cgwm2.029870, Cgwm2.045100, Cgwm2.047290	map04150
<i>N-Glycan biosynthesis</i>	4	<i>Kinase</i>	ec:2.7.1.108	1	Cgwm2.000150	map00510
<i>N-Glycan biosynthesis</i>	4	<i>1,2-alpha-mannosidase</i>	ec:3.2.1.113	2	Cgwm2.040880, Cgwm2.054740	map00510
<i>N-Glycan biosynthesis</i>	4	<i>Glucosidase</i>	ec:3.2.1.106	1	Cgwm2.034080	map00510
<i>Nicotinate and nicotinamide metabolism</i>	4	<i>Kinase</i>	ec:2.7.1.23	3	Cgwm2.046990, Cgwm2.025990, Cgwm2.046970	map00760
<i>Nicotinate and nicotinamide metabolism</i>	4	<i>diphosphorylase (carboxylating)</i>	ec:2.4.2.19	1	Cgwm2.018560	map00760
<i>Nitrogen metabolism</i>	1	<i>synthase (NADH)</i>	ec:1.4.1.14	1	Cgwm2.051890	map00910

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>One carbon pool by folate</i>	4	<i>Synthase</i>	ec:2.1.1.45	1	Cgwm2.028320	map00670
<i>One carbon pool by folate</i>	4	<i>Ligase</i>	ec:6.3.4.3	1	Cgwm2.043760	map00670
<i>One carbon pool by folate</i>	4	<i>reductase [NAD(P)H]</i>	ec:1.5.1.20	2	Cgwm2.023450, Cgwm2.006620	map00670
<i>One carbon pool by folate</i>	4	<i>dehydrogenase (NADP+)</i>	ec:1.5.1.5	1	Cgwm2.043760	map00670
<i>Other types of O-glycan biosynthesis</i>	1	<i>mannosyltransferase</i>	ec:2.4.1.109	1	Cgwm2.049320	map00514
<i>Oxidative phosphorylation</i>	1	<i>reductase (H+-translocating)</i>	ec:1.6.5.3	1	Cgwm2.032410	map00190
<i>Pantothenate and CoA biosynthesis</i>	7	<i>hydroxymethyltransferase</i>	ec:2.1.2.11	1	Cgwm2.029200	map00770
<i>Pantothenate and CoA biosynthesis</i>	7	<i>Synthase</i>	ec:2.2.1.6	1	Cgwm2.010000	map00770
<i>Pantothenate and CoA biosynthesis</i>	7	<i>Kinase</i>	ec:2.7.1.33	1	Cgwm2.019370	map00770
<i>Pantothenate and CoA biosynthesis</i>	7	<i>reductoisomerase (NADP+)</i>	ec:1.1.1.86	1	Cgwm2.036890	map00770
<i>Pantothenate and CoA biosynthesis</i>	7	<i>Kinase</i>	ec:2.7.1.24	1	Cgwm2.018640	map00770
<i>Pantothenate and CoA biosynthesis</i>	7	<i>Synthase</i>	ec:2.7.8.7	2	Cgwm2.048400, Cgwm2.033880	map00770
<i>Pentose phosphate pathway</i>	2	<i>phosphohexokinase</i>	ec:2.7.1.11	1	Cgwm2.028480	map00030
<i>Pentose phosphate pathway</i>	2	<i>dehydrogenase (NADP+-dependent, decarboxylating)</i>	ec:1.1.1.44	1	Cgwm2.014520	map00030
<i>Phenylalanine metabolism</i>	2	<i>Oxidase</i>	ec:1.4.3.21	2	Cgwm2.023190, Cgwm2.043180	map00360
<i>Phenylalanine, tyrosine and tryptophan biosynthesis</i>	1	<i>Synthase</i>	ec:2.5.1.54	1	Cgwm2.018730	map00400
<i>Phenylpropanoid biosynthesis</i>	1	<i>Lactoperoxidase</i>	ec:1.11.1.7	1	Cgwm2.057820	map00940
<i>Phosphatidylinositol signaling system</i>	5	<i>kinase (ATP)</i>	ec:2.7.1.107	1	Cgwm2.025990	map04070
<i>Phosphatidylinositol signaling system</i>	5	<i>3-kinase</i>	ec:2.7.1.127	1	Cgwm2.019660	map04070
<i>Phosphatidylinositol signaling system</i>	5	<i>4-kinase</i>	ec:2.7.1.67	1	Cgwm2.043740	map04070
<i>Phosphatidylinositol signaling system</i>	5	<i>cytidyltransferase</i>	ec:2.7.7.41	1	Cgwm2.024130	map04070
<i>Phosphatidylinositol signaling system</i>	5	<i>3-kinase</i>	ec:2.7.1.137	1	Cgwm2.002640	map04070
<i>Porphyryn and chlorophyll metabolism</i>	4	<i>Synthase</i>	ec:4.4.1.17	1	Cgwm2.038380	map00860
<i>Porphyryn and chlorophyll metabolism</i>	4	<i>Dehydrogenase</i>	ec:1.3.1.76	1	Cgwm2.021590	map00860
<i>Porphyryn and chlorophyll metabolism</i>	4	<i>ferro-protoporphyrin chelatase</i>	ec:4.99.1.1	1	Cgwm2.037700	map00860

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Porphyrin and chlorophyll metabolism</i>	4	<i>Oxidase</i>	ec:1.3.3.4	1	Cgwm2.040400	map00860
<i>Propanoate metabolism</i>	3	<i>Dehydrogenase</i>	ec:1.8.1.4	1	Cgwm2.049660	map00640
<i>Propanoate metabolism</i>	3	<i>Transaminase</i>	ec:2.6.1.19	1	Cgwm2.031420	map00640
<i>Propanoate metabolism</i>	3	<i>Ligase</i>	ec:6.2.1.1	1	Cgwm2.006980	map00640
<i>Purine metabolism</i>	44	<i>Cyclase</i>	ec:4.6.1.1	1	Cgwm2.040540	map00230
<i>Purine metabolism</i>	44	<i>adenylpyrophosphatase</i>	ec:3.6.1.3	9	Cgwm2.059120, Cgwm2.025510, Cgwm2.008540, Cgwm2.006560, Cgwm2.019510, Cgwm2.034740, Cgwm2.054050, Cgwm2.028190, Cgwm2.034280	map00230
<i>Purine metabolism</i>	44	<i>adenyltransferase</i>	ec:2.7.7.53	1	Cgwm2.053410	map00230
<i>Purine metabolism</i>	44	<i>ec:3.5.3.4 allantoicase</i>	ec:3.5.3.4	1	Cgwm2.025430	map00230
<i>Purine metabolism</i>	44	<i>phosphodiesterase</i>	ec:3.1.4.17	1	Cgwm2.013000	map00230
<i>Purine metabolism</i>	44	<i>Reductase</i>	ec:1.17.4.1	1	Cgwm2.032850	map00230
<i>Purine metabolism</i>	44	<i>Deaminase</i>	ec:3.5.4.6	1	Cgwm2.018130	map00230
<i>Purine metabolism</i>	44	<i>Synthase</i>	ec:6.3.2.6	1	Cgwm2.008900	map00230
<i>Purine metabolism</i>	44	<i>Dehydrogenase</i>	ec:1.1.1.205	1	Cgwm2.003410	map00230
<i>Purine metabolism</i>	44	<i>RNA polymerase</i>	ec:2.7.7.6	14	Cgwm2.035450, Cgwm2.033250, Cgwm2.013360, Cgwm2.034650, Cgwm2.035760, Cgwm2.038970, Cgwm2.034620, Cgwm2.012820, Cgwm2.047250, Cgwm2.026130, Cgwm2.026240, Cgwm2.026650, Cgwm2.046420, Cgwm2.005220	map00230

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Purine metabolism</i>	44	<i>DNA polymerase</i>	ec:2.7.7.7	12	Cgwm2.046150, Cgwm2.015970, Cgwm2.062240, Cgwm2.023210, Cgwm2.035410, Cgwm2.002030, Cgwm2.009220, Cgwm2.023780, Cgwm2.033200, Cgwm2.057430, Cgwm2.056560, Cgwm2.013830	map00230
<i>Purine metabolism</i>	44	<i>phosphoribosylidiphosphate 5-amidotransferase</i>	ec:2.4.2.14	1	Cgwm2.019260	map00230
<i>Pyrimidine metabolism</i>	34	<i>Synthase</i>	ec:2.1.1.45	1	Cgwm2.028320	map00240
<i>Pyrimidine metabolism</i>	34	<i>carbamoylaspartic dehydrase</i>	ec:3.5.2.3	1	Cgwm2.006360	map00240
<i>Pyrimidine metabolism</i>	34	<i>synthase (glutamine hydrolysing)</i>	ec:6.3.4.2	1	Cgwm2.021280	map00240
<i>Pyrimidine metabolism</i>	34	<i>Decarboxylase</i>	ec:4.1.1.23	1	Cgwm2.021760	map00240
<i>Pyrimidine metabolism</i>	34	<i>Kinase</i>	ec:2.7.1.48	1	Cgwm2.042090	map00240
<i>Pyrimidine metabolism</i>	34	<i>Reductase</i>	ec:1.17.4.1	1	Cgwm2.032850	map00240
<i>Pyrimidine metabolism</i>	34	<i>carbamoyltransferase</i>	ec:2.1.3.2	1	Cgwm2.005210	map00240
<i>Pyrimidine metabolism</i>	34	<i>RNA polymerase</i>	ec:2.7.7.6	14	Cgwm2.035450, Cgwm2.033250, Cgwm2.013360, Cgwm2.034650, Cgwm2.035760, Cgwm2.038970, Cgwm2.034620, Cgwm2.012820, Cgwm2.047250, Cgwm2.026130, Cgwm2.026240, Cgwm2.026650, Cgwm2.046420, Cgwm2.005220, Cgwm2.046150, Cgwm2.015970, Cgwm2.062240, Cgwm2.023210, Cgwm2.035410, Cgwm2.002030, Cgwm2.009220, Cgwm2.023780, Cgwm2.033200, Cgwm2.057430, Cgwm2.056560, Cgwm2.013830	map00240
<i>Pyrimidine metabolism</i>	34	<i>DNA polymerase</i>	ec:2.7.7.7	12	Cgwm2.046150, Cgwm2.015970, Cgwm2.062240, Cgwm2.023210, Cgwm2.035410, Cgwm2.002030, Cgwm2.009220, Cgwm2.023780, Cgwm2.033200, Cgwm2.057430, Cgwm2.056560, Cgwm2.013830	map00240

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Pyrimidine metabolism</i>	34	<i>Kinase</i>	ec:2.7.4.14	1	Cgwm2.038960	map00240
<i>Pyruvate metabolism</i>	8	<i>Dehydrogenase</i>	ec:1.8.1.4	1	Cgwm2.049660	map00620
<i>Pyruvate metabolism</i>	8	<i>Synthase</i>	ec:2.3.3.14	1	Cgwm2.024330	map00620
<i>Pyruvate metabolism</i>	8	<i>Synthase</i>	ec:2.3.3.13	1	Cgwm2.059300	map00620
<i>Pyruvate metabolism</i>	8	<i>carboxykinase (ATP)</i>	ec:4.1.1.49	1	Cgwm2.047890	map00620
<i>Pyruvate metabolism</i>	8	<i>Ligase</i>	ec:6.2.1.1	1	Cgwm2.006980	map00620
<i>Pyruvate metabolism</i>	8	<i>dehydrogenase (decarboxylating)</i>	ec:1.1.1.39	2	Cgwm2.046860, Cgwm2.035730	map00620
<i>Pyruvate metabolism</i>	8	<i>acetyltransferase</i>	ec:2.3.1.12	1	Cgwm2.025530	map00620
<i>Riboflavin metabolism</i>	4	<i>Synthase</i>	ec:2.5.1.9	1	Cgwm2.020700	map00740
<i>Riboflavin metabolism</i>	4	<i>cyclohydrolase II</i>	ec:3.5.4.25	2	Cgwm2.004830, Cgwm2.061170	map00740
<i>Riboflavin metabolism</i>	4	<i>reductase (NADPH)</i>	ec:1.5.1.30	1	Cgwm2.036690	map00740
<i>Selenocompound metabolism</i>	1	<i>Ligase</i>	ec:6.1.1.10	1	Cgwm2.027370	map00450
<i>Sphingolipid metabolism</i>	1	<i>Sulfatase</i>	ec:3.1.6.1	1	Cgwm2.014600	map00600
<i>Starch and sucrose metabolism</i>	8	<i>endo-1,4-beta-D-glucanase</i>	ec:3.2.1.4	1	Cgwm2.058580	map00500
<i>Starch and sucrose metabolism</i>	8	<i>Glycogenase</i>	ec:3.2.1.1	1	Cgwm2.055190	map00500
<i>Starch and sucrose metabolism</i>	8	<i>Decarboxylase</i>	ec:4.1.1.35	1	Cgwm2.041700	map00500
<i>Starch and sucrose metabolism</i>	8	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	1	Cgwm2.059230	map00500
<i>Starch and sucrose metabolism</i>	8	<i>Synthase</i>	ec:2.4.1.34	1	Cgwm2.065610	map00500
<i>Starch and sucrose metabolism</i>	8	<i>Trehalase</i>	ec:3.2.1.28	2	Cgwm2.043650, Cgwm2.011310	map00500
<i>Starch and sucrose metabolism</i>	8	<i>amylase-1,6-glucosidase</i>	ec:3.2.1.33	1	Cgwm2.052870	map00500
<i>Starch and sucrose metabolism</i>	8	<i>disproportionating enzyme</i>	ec:2.4.1.25	1	Cgwm2.052870	map00500
<i>Steroid hormone biosynthesis</i>	1	<i>Sulfatase</i>	ec:3.1.6.1	1	Cgwm2.014600	map00140
<i>Streptomycin biosynthesis</i>	2	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	1	Cgwm2.059230	map00521
<i>Streptomycin biosynthesis</i>	2	<i>Synthase</i>	ec:5.5.1.4	1	Cgwm2.014980	map00521
<i>Sulfur metabolism</i>	1	<i>reductase (thioredoxin)</i>	ec:1.8.4.8	1	Cgwm2.001890	map00920

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Synthesis and degradation of ketone bodies</i>	1	<i>CoA-transferase</i>	ec:2.8.3.5	1	Cgwm2.022390	map00072
<i>T cell receptor signaling pathway</i>	3	<i>Phosphatase</i>	ec:3.1.3.16	3	Cgwm2.026400, Cgwm2.048850, Cgwm2.051280	map04660
<i>Taurine and hypotaurine metabolism</i>	3	<i>glutamyl transpeptidase</i>	ec:2.3.2.2	3	Cgwm2.019590, Cgwm2.039820, Cgwm2.024670	map00430
<i>Terpenoid backbone biosynthesis</i>	3	<i>Decarboxylase</i>	ec:4.1.1.33	1	Cgwm2.023310	map00900
<i>Terpenoid backbone biosynthesis</i>	3	<i>reductase (NADPH)</i>	ec:1.1.1.34	1	Cgwm2.034750	map00900
<i>Terpenoid backbone biosynthesis</i>	3	<i>Oxidase</i>	ec:1.8.3.5	1	Cgwm2.056040	map00900
<i>Tropane, piperidine and pyridine alkaloid biosynthesis</i>	2	<i>Oxidase</i>	ec:1.4.3.21	2	Cgwm2.023190, Cgwm2.043180	map00960
<i>Tryptophan metabolism</i>	1	<i>Equilase</i>	ec:1.11.1.6	1	Cgwm2.022230	map00380
<i>Tyrosine metabolism</i>	2	<i>Oxidase</i>	ec:1.4.3.21	2	Cgwm2.023190, Cgwm2.043180	map00350
<i>Valine, leucine and isoleucine biosynthesis</i>	5	<i>Synthase</i>	ec:2.3.3.13	1	Cgwm2.059300	map00290
<i>Valine, leucine and isoleucine biosynthesis</i>	5	<i>Synthase</i>	ec:2.2.1.6	1	Cgwm2.010000	map00290
<i>Valine, leucine and isoleucine biosynthesis</i>	5	<i>ammonia-lyase</i>	ec:4.3.1.19	1	Cgwm2.056570	map00290
<i>Valine, leucine and isoleucine biosynthesis</i>	5	<i>reductoisomerase (NADP+)</i>	ec:1.1.1.86	1	Cgwm2.036890	map00290
<i>Valine, leucine and isoleucine biosynthesis</i>	5	<i>Dehydrogenase</i>	ec:1.1.1.85	1	Cgwm2.003510	map00290
<i>Valine, leucine and isoleucine degradation</i>	4	<i>Dehydrogenase</i>	ec:1.8.1.4	1	Cgwm2.049660	map00280
<i>Valine, leucine and isoleucine degradation</i>	4	<i>Dehydrogenase</i>	ec:1.1.1.31	1	Cgwm2.014520	map00280

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Valine, leucine and isoleucine degradation</i>	4	<i>Ligase</i>	ec:6.2.1.16	1	Cgwm2.034400	map00280
<i>Valine, leucine and isoleucine degradation</i>	4	<i>CoA-transferase</i>	ec:2.8.3.5	1	Cgwm2.022390	map00280
<i>Various types of N-glycan biosynthesis</i>	2	<i>1,2-alpha-mannosidase</i>	ec:3.2.1.113	2	Cgwm2.040880, Cgwm2.054740	map00513
<i>Vitamin B6 metabolism</i>	1	<i>5'-phosphate synthase</i>	ec:1.4.3.5	1	Cgwm2.058860	map00750

Tabela C. 4: Vias metabólicas mapeadas em *C. neoformans* var. *grubii*, linhagem CBS 7750. Todas as proteínas mapeadas referem-se às IDPs e IDRs identificadas neste estudo.

<i>Via metabólica</i>	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Amino sugar and nucleotide sugar metabolism</i>	14	<i>end alpha-L-arabinofuranosidase</i>	ec:3.2.1.55	2	cngru.020450, cngru.020460	map00520
<i>Amino sugar and nucleotide sugar metabolism</i>	14	<i>uridylyltransferase</i>	ec:2.7.7.12	1	cngru.012950	map00520
<i>Amino sugar and nucleotide sugar metabolism</i>	14	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	2	cngru.018670, cngru.018680	map00520
<i>Amino sugar and nucleotide sugar metabolism</i>	14	<i>Synthase</i>	ec:2.4.1.16	6	cngru.014040, cngru.058070, cngru.026620, cngru.072980, cngru.043370, cngru.070470	map00520
<i>Amino sugar and nucleotide sugar metabolism</i>	14	<i>Chitodextrinase</i>	ec:3.2.1.14	3	cngru.076510, cngru.076520, cngru.038020	map00520
<i>Aminoacyl-tRNA biosynthesis</i>	15	<i>synthase (glutamine-hydrolysing)</i>	ec:6.3.5.7	1	cngru.053960	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	15	<i>Ligase</i>	ec:6.1.1.18	1	cngru.021550	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	15	<i>Ligase</i>	ec:6.1.1.17	1	cngru.038220	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	15	<i>Ligase</i>	ec:6.1.1.14	1	cngru.062730	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	15	<i>Ligase</i>	ec:6.1.1.16	1	cngru.035850	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	15	<i>Ligase</i>	ec:6.1.1.10	1	cngru.053100	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	15	<i>Ligase</i>	ec:6.1.1.12	2	cngru.055520, cngru.067800	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	15	<i>Ligase</i>	ec:6.1.1.11	1	cngru.036430	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	15	<i>Ligase</i>	ec:6.1.1.6	1	cngru.077240	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	15	<i>Ligase</i>	ec:6.1.1.9	1	cngru.034660	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	15	<i>Ligase</i>	ec:6.1.1.2	1	cngru.000650	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	15	<i>Ligase</i>	ec:6.1.1.3	1	cngru.036330	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	15	<i>Ligase</i>	ec:6.1.1.1	1	cngru.022170	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	15	<i>Ligase</i>	ec:6.1.1.22	1	cngru.060180	map00970
<i>Aminobenzoate degradation</i>	2	<i>Phosphatase</i>	ec:3.1.3.2	2	cngru.057450, cngru.067440	map00627
<i>Arginine and proline metabolism</i>	4	<i>Oxidase</i>	ec:1.4.3.3	2	cngru.018210, cngru.039280	map00330

<i>Via metabólica</i>	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Arginine and proline metabolism</i>	4	<i>5-kinase</i>	ec:2.7.2.11	2	cngru.019810, cngru.019820	map00330
<i>Arginine biosynthesis</i>	4	<i>Synthase</i>	ec:6.3.4.5	1	cngru.070130	map00220
<i>Arginine biosynthesis</i>	4	<i>Kinase</i>	ec:2.7.2.8	2	cngru.064420, cngru.064430	map00220
<i>Arginine biosynthesis</i>	4	<i>Reductase</i>	ec:1.2.1.38	2	cngru.064420, cngru.064430	map00220
<i>Arginine biosynthesis</i>	4	<i>N-acetyltransferase</i>	ec:2.3.1.1	1	cngru.040520	map00220
<i>Benzoate degradation</i>	1	<i>Dehydrogenase</i>	ec:1.1.1.35	1	cngru.075850	map00362
<i>beta-Alanine metabolism</i>	2	<i>Transaminase</i>	ec:2.6.1.19	1	cngru.059650	map00410
<i>beta-Alanine metabolism</i>	2	<i>Oxidase</i>	ec:1.4.3.21	1	cngru.045940	map00410
<i>Biosynthesis of antibiotics</i>	45	<i>Dehydrogenase</i>	ec:1.8.1.4	1	cngru.054820	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>Synthase</i>	ec:2.3.3.14	1	cngru.049540	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>succinyltransferase</i>	ec:2.3.1.61	1	cngru.035700	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>Kinase</i>	ec:2.7.4.3	2	cngru.048390, cngru.048380	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>carboxykinase (ATP)</i>	ec:4.1.1.49	1	cngru.076840	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>Synthase</i>	ec:6.3.4.5	1	cngru.070130	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>Decarboxylase</i>	ec:4.1.1.33	1	cngru.045830	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>Oxidase</i>	ec:1.4.3.3	2	cngru.018210, cngru.039280	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>Hydratase</i>	ec:4.2.1.3	2	cngru.011530, cngru.051190	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>Hydratase</i>	ec:4.2.1.2	1	cngru.008320	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>Dehydrogenase</i>	ec:1.1.1.95	1	cngru.029130	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>glycolaldehydetransferase</i>	ec:2.2.1.1	1	cngru.011570	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>Synthase</i>	ec:2.2.1.6	2	cngru.014720, cngru.023570	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>Equilase</i>	ec:1.11.1.6	1	cngru.047130	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>Synthase</i>	ec:2.5.1.54	1	cngru.040170	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>Kinase</i>	ec:2.7.2.8	2	cngru.064420, cngru.064430	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>Kinase</i>	ec:2.7.2.4	1	cngru.075440	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>Kinase</i>	ec:2.7.2.3	1	cngru.070130	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>Reductase</i>	ec:1.2.1.38	2	cngru.064420, cngru.064430	map01130

<i>Via metabólica</i>	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Biosynthesis of antibiotics</i>	45	<i>5-kinase</i>	ec:2.7.2.11	2	cngru.019810, cngru.019820	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>Deaminase</i>	ec:3.5.4.6	1	cngru.039500	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>Synthase</i>	ec:6.3.2.6	1	cngru.022170	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>synthase (NADH)</i>	ec:1.4.1.14	2	cngru.003460, cngru.003450	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>synthase (NADPH)</i>	ec:1.4.1.13	2	cngru.003460, cngru.003450	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	2	cngru.018670, cngru.018680	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>reductoisomerase (NADP+)</i>	ec:1.1.1.86	1	cngru.064790	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>phosphohexokinase</i>	ec:2.7.1.11	1	cngru.001390	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>N-acetyltransferase</i>	ec:2.3.1.1	1	cngru.040520	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>ligase (ADP-forming)</i>	ec:6.2.1.5	1	cngru.069920	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>Synthase</i>	ec:5.5.1.4	1	cngru.009630	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>ligase (GDP-forming)</i>	ec:6.2.1.4	1	cngru.069920	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>Ligase</i>	ec:6.2.1.1	1	cngru.029370	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>Dehydrogenase</i>	ec:1.1.1.35	1	cngru.075850	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>reductase (NADPH)</i>	ec:1.1.1.34	1	cngru.067190	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>acetyltransferase</i>	ec:2.3.1.12	1	cngru.050980	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>dehydrogenase (NADP+-dependent, decarboxylating)</i>	ec:1.1.1.44	1	cngru.010150	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>dehydrogenase (NAD+)</i>	ec:1.1.1.41	2	cngru.025870, cngru.002500	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>Oxidase</i>	ec:1.8.3.5	1	cngru.039260	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>O-acetyltransferase</i>	ec:2.3.1.31	1	cngru.045860	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>citrate synthase</i>	ec:2.3.3.8	1	cngru.069920	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>phosphoribosyl(diphosphate amidotransferase</i>	5- ec:2.4.2.14	1	cngru.040830	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>Kinase</i>	ec:2.7.4.10	2	cngru.048390, cngru.048380	map01130
<i>Biosynthesis of unsaturated fatty acids</i>	1	<i>9-desaturase</i>	ec:1.14.19.1	1	cngru.001510	map01040
<i>Biotin metabolism</i>	2	<i>Synthase</i>	ec:2.8.1.6	1	cngru.066230	map00780

<i>Via metabólica</i>	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Biotin metabolism</i>	2	<i>Ligase</i>	ec:6.3.4.15	1	cngru.009350	map00780
<i>Butanoate metabolism</i>	6	<i>Transaminase</i>	ec:2.6.1.19	1	cngru.059650	map00650
<i>Butanoate metabolism</i>	6	<i>Synthase</i>	ec:2.2.1.6	2	cngru.014720, cngru.023570	map00650
<i>Butanoate metabolism</i>	6	<i>Dehydrogenase</i>	ec:1.1.1.35	1	cngru.075850	map00650
<i>Butanoate metabolism</i>	6	<i>Ligase</i>	ec:6.2.1.16	1	cngru.056190	map00650
<i>Butanoate metabolism</i>	6	<i>CoA-transferase</i>	ec:2.8.3.5	1	cngru.046890	map00650
<i>Butirosin and neomycin biosynthesis</i>	2	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	2	cngru.018670, cngru.018680	map00524
<i>C5-Branched dibasic acid metabolism</i>	4	<i>Synthase</i>	ec:2.2.1.6	2	cngru.014720, cngru.023570	map00660
<i>C5-Branched dibasic acid metabolism</i>	4	<i>Dehydrogenase</i>	ec:1.1.1.85	1	cngru.025550	map00660
<i>C5-Branched dibasic acid metabolism</i>	4	<i>ligase (ADP-forming)</i>	ec:6.2.1.5	1	cngru.069920	map00660
<i>Carbapenem biosynthesis</i>	2	<i>5-kinase</i>	ec:2.7.2.11	2	cngru.019810, cngru.019820	map00332
<i>Carbon fixation in photosynthetic organisms</i>	5	<i>carboxykinase (ATP)</i>	ec:4.1.1.49	1	cngru.076840	map00710
<i>Carbon fixation in photosynthetic organisms</i>	5	<i>glycolaldehydetransferase</i>	ec:2.2.1.1	1	cngru.011570	map00710
<i>Carbon fixation in photosynthetic organisms</i>	5	<i>Kinase</i>	ec:2.7.2.3	1	cngru.070130	map00710
<i>Carbon fixation in photosynthetic organisms</i>	5	<i>dehydrogenase (decarboxylating)</i>	ec:1.1.1.39	2	cngru.066050, cngru.075660	map00710
<i>Carbon fixation pathways in prokaryotes</i>	8	<i>reductase [NAD(P)H]</i>	ec:1.5.1.20	2	cngru.029000, cngru.045670	map00720
<i>Carbon fixation pathways in prokaryotes</i>	8	<i>Hydratase</i>	ec:4.2.1.3	2	cngru.011530, cngru.051190	map00720
<i>Carbon fixation pathways in prokaryotes</i>	8	<i>Hydratase</i>	ec:4.2.1.2	1	cngru.008320	map00720
<i>Carbon fixation pathways in prokaryotes</i>	8	<i>ligase (ADP-forming)</i>	ec:6.2.1.5	1	cngru.069920	map00720
<i>Carbon fixation pathways in prokaryotes</i>	8	<i>Ligase</i>	ec:6.2.1.1	1	cngru.029370	map00720
<i>Carbon fixation pathways in prokaryotes</i>	8	<i>Dehydrogenase</i>	ec:1.1.1.35	1	cngru.075850	map00720
<i>Carbon fixation pathways in prokaryotes</i>	8	<i>citrate synthase</i>	ec:2.3.3.8	1	cngru.069920	map00720
<i>Citrate cycle (TCA cycle)</i>	10	<i>Dehydrogenase</i>	ec:1.8.1.4	1	cngru.054820	map00020

<i>Via metabólica</i>	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Citrate cycle (TCA cycle)</i>	10	<i>succinyltransferase</i>	ec:2.3.1.61	1	cngru.035700	map00020
<i>Citrate cycle (TCA cycle)</i>	10	<i>carboxykinase (ATP)</i>	ec:4.1.1.49	1	cngru.076840	map00020
<i>Citrate cycle (TCA cycle)</i>	10	<i>Hydratase</i>	ec:4.2.1.3	2	cngru.011530, cngru.051190	map00020
<i>Citrate cycle (TCA cycle)</i>	10	<i>Hydratase</i>	ec:4.2.1.2	1	cngru.008320	map00020
<i>Citrate cycle (TCA cycle)</i>	10	<i>ligase (ADP-forming)</i>	ec:6.2.1.5	1	cngru.069920	map00020
<i>Citrate cycle (TCA cycle)</i>	10	<i>ligase (GDP-forming)</i>	ec:6.2.1.4	1	cngru.069920	map00020
<i>Citrate cycle (TCA cycle)</i>	10	<i>acetyltransferase</i>	ec:2.3.1.12	1	cngru.050980	map00020
<i>Citrate cycle (TCA cycle)</i>	10	<i>dehydrogenase (NAD+)</i>	ec:1.1.1.41	2	cngru.025870, cngru.002500	map00020
<i>Citrate cycle (TCA cycle)</i>	10	<i>citrate synthase</i>	ec:2.3.3.8	1	cngru.069920	map00020
<i>Cyanoamino acid metabolism</i>	5	<i>glutamyl transpeptidase</i>	ec:2.3.2.2	4	cngru.041230, cngru.049940, cngru.072660, cngru.072670	map00460
<i>Cyanoamino acid metabolism</i>	5	<i>Gentiobiase</i>	ec:3.2.1.21	1	cngru.036970	map00460
<i>Cysteine and methionine metabolism</i>	5	<i>S-methyltransferase</i>	ec:2.1.1.10	1	cngru.003070	map00270
<i>Cysteine and methionine metabolism</i>	5	<i>Kinase</i>	ec:2.7.2.4	1	cngru.075440	map00270
<i>Cysteine and methionine metabolism</i>	5	<i>Ligase</i>	ec:6.3.2.2	2	cngru.016110, cngru.016120	map00270
<i>Cysteine and methionine metabolism</i>	5	<i>O-acetyltransferase</i>	ec:2.3.1.31	1	cngru.045860	map00270
<i>D-Arginine and D-ornithine metabolism</i>	2	<i>Oxidase</i>	ec:1.4.3.3	2	cngru.018210, cngru.039280	map00472
<i>Drug metabolism - cytochrome P450</i>	1	<i>monooxygenase</i>	ec:1.14.13.8	1	cngru.010290	map00982
<i>Drug metabolism - other enzymes</i>	1	<i>Kinase</i>	ec:2.7.1.48	1	cngru.070020	map00983
<i>Ether lipid metabolism</i>	1	<i>D</i>	ec:3.1.4.4	1	cngru.040740	map00565
<i>Fatty acid degradation</i>	1	<i>Dehydrogenase</i>	ec:1.1.1.35	1	cngru.075850	map00071
<i>Fatty acid elongation</i>	1	<i>Dehydrogenase</i>	ec:1.1.1.35	1	cngru.075850	map00062
<i>Geraniol degradation</i>	1	<i>Dehydrogenase</i>	ec:1.1.1.35	1	cngru.075850	map00281
<i>Glutathione metabolism</i>	8	<i>Reductase</i>	ec:1.17.4.1	1	cngru.057970	map00480
<i>Glutathione metabolism</i>	8	<i>Ligase</i>	ec:6.3.2.2	2	cngru.016110, cngru.016120	map00480
<i>Glutathione metabolism</i>	8	<i>dehydrogenase (NADP+-dependent, decarboxylating)</i>	ec:1.1.1.44	1	cngru.010150	map00480

<i>Via metabólica</i>	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Glutathione metabolism</i>	8	<i>glutamyl transpeptidase</i>	ec:2.3.2.2	4	cngru.041230, cngru.049940, cngru.072660, cngru.072670	map00480
<i>Glycerolipid metabolism</i>	6	<i>kinase (ATP)</i>	ec:2.7.1.107	2	cngru.061990, cngru.051480	map00561
<i>Glycerolipid metabolism</i>	6	<i>Lípase</i>	ec:3.1.1.3	2	cngru.008970, cngru.058660	map00561
<i>Glycerolipid metabolism</i>	6	<i>O-acyltransferase</i>	ec:2.3.1.51	2	cngru.044600, cngru.050850	map00561
<i>Glycerophospholipid metabolism</i>	12	<i>phosphodiesterase</i>	ec:3.1.4.46	1	cngru.066290	map00564
<i>Glycerophospholipid metabolism</i>	12	<i>kinase (ATP)</i>	ec:2.7.1.107	2	cngru.061990, cngru.051480	map00564
<i>Glycerophospholipid metabolism</i>	12	<i>cytidyltransferase</i>	ec:2.7.7.41	1	cngru.049270	map00564
<i>Glycerophospholipid metabolism</i>	12	<i>N-methyltransferase</i>	ec:2.1.1.17	1	cngru.072540	map00564
<i>Glycerophospholipid metabolism</i>	12	<i>Decarboxylase</i>	ec:4.1.1.65	2	cngru.002690, cngru.070420	map00564
<i>Glycerophospholipid metabolism</i>	12	<i>D</i>	ec:3.1.4.4	1	cngru.040740	map00564
<i>Glycerophospholipid metabolism</i>	12	<i>Dehydrogenase</i>	ec:1.1.5.3	1	cngru.040380	map00564
<i>Glycerophospholipid metabolism</i>	12	<i>lecithinase B</i>	ec:3.1.1.5	1	cngru.053010	map00564
<i>Glycerophospholipid metabolism</i>	12	<i>O-acyltransferase</i>	ec:2.3.1.51	2	cngru.044600, cngru.050850	map00564
<i>Glycine, serine and threonine metabolism</i>	9	<i>Dehydrogenase</i>	ec:1.8.1.4	1	cngru.054820	map00260
<i>Glycine, serine and threonine metabolism</i>	9	<i>Oxidase</i>	ec:1.4.3.3	2	cngru.018210, cngru.039280	map00260
<i>Glycine, serine and threonine metabolism</i>	9	<i>dehydrogenase (aminomethyl-transferring)</i>	ec:1.4.4.2	1	cngru.009040	map00260
<i>Glycine, serine and threonine metabolism</i>	9	<i>Dehydrogenase</i>	ec:1.1.1.95	1	cngru.029130	map00260
<i>Glycine, serine and threonine metabolism</i>	9	<i>Kinase</i>	ec:2.7.2.4	1	cngru.075440	map00260
<i>Glycine, serine and threonine metabolism</i>	9	<i>Oxidase</i>	ec:1.5.3.1	1	cngru.042240	map00260
<i>Glycine, serine and threonine metabolism</i>	9	<i>Dehydrogenase</i>	ec:1.1.99.1	1	cngru.044490	map00260
<i>Glycine, serine and threonine metabolism</i>	9	<i>Oxidase</i>	ec:1.4.3.21	1	cngru.045940	map00260
<i>Glycolysis / Gluconeogenesis</i>	8	<i>Dehydrogenase</i>	ec:1.8.1.4	1	cngru.054820	map00010
<i>Glycolysis / Gluconeogenesis</i>	8	<i>carboxykinase (ATP)</i>	ec:4.1.1.49	1	cngru.076840	map00010
<i>Glycolysis / Gluconeogenesis</i>	8	<i>Kinase</i>	ec:2.7.2.3	1	cngru.070130	map00010
<i>Glycolysis / Gluconeogenesis</i>	8	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	2	cngru.018670, cngru.018680	map00010

<i>Via metabólica</i>	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Glycolysis / Gluconeogenesis</i>	8	<i>phosphohexokinase</i>	ec:2.7.1.11	1	cngru.001390	map00010
<i>Glycolysis / Gluconeogenesis</i>	8	<i>Ligase</i>	ec:6.2.1.1	1	cngru.029370	map00010
<i>Glycolysis / Gluconeogenesis</i>	8	<i>acetyltransferase</i>	ec:2.3.1.12	1	cngru.050980	map00010
<i>Inositol phosphate metabolism</i>	6	<i>3-kinase</i>	ec:2.7.1.127	1	cngru.041330	map00562
<i>Inositol phosphate metabolism</i>	6	<i>4-kinase</i>	ec:2.7.1.67	1	cngru.068110	map00562
<i>Inositol phosphate metabolism</i>	6	<i>Synthase</i>	ec:5.5.1.4	1	cngru.009630	map00562
<i>Inositol phosphate metabolism</i>	6	<i>3-kinase</i>	ec:2.7.1.137	1	cngru.033200	map00562
<i>Inositol phosphate metabolism</i>	6	<i>2-kinase</i>	ec:2.7.1.158	2	cngru.052940, cngru.052930	map00562
<i>Isoquinoline alkaloid biosynthesis</i>	1	<i>Oxidase</i>	ec:1.4.3.21	1	cngru.045940	map00950
<i>Lipoic acid metabolism</i>	1	<i>Synthase</i>	ec:2.8.1.8	1	cngru.045050	map00785
<i>Lysine biosynthesis</i>	2	<i>Synthase</i>	ec:2.3.3.14	1	cngru.049540	map00300
<i>Lysine biosynthesis</i>	2	<i>Kinase</i>	ec:2.7.2.4	1	cngru.075440	map00300
<i>Lysine degradation</i>	7	<i>N-methyltransferase</i>	ec:2.1.1.43	5	cngru.005830, cngru.052310, cngru.019590, cngru.019570, cngru.019580	map00310
<i>Lysine degradation</i>	7	<i>succinyltransferase</i>	ec:2.3.1.61	1	cngru.035700	map00310
<i>Lysine degradation</i>	7	<i>Dehydrogenase</i>	ec:1.1.1.35	1	cngru.075850	map00310
<i>Methane metabolism</i>	3	<i>Dehydrogenase</i>	ec:1.1.1.95	1	cngru.029130	map00680
<i>Methane metabolism</i>	3	<i>phosphohexokinase</i>	ec:2.7.1.11	1	cngru.001390	map00680
<i>Methane metabolism</i>	3	<i>Ligase</i>	ec:6.2.1.1	1	cngru.029370	map00680
<i>Monobactam biosynthesis</i>	1	<i>Kinase</i>	ec:2.7.2.4	1	cngru.075440	map00261
<i>N-Glycan biosynthesis</i>	7	<i>Kinase</i>	ec:2.7.1.108	2	cngru.030270, cngru.030280	map00510
<i>N-Glycan biosynthesis</i>	7	<i>1,2-alpha-mannosidase</i>	ec:3.2.1.113	3	cngru.004730, cngru.056600, cngru.071380	map00510
<i>N-Glycan biosynthesis</i>	7	<i>alpha-1,3-mannosyltransferase</i>	ec:2.4.1.132	1	cngru.048970	map00510
<i>N-Glycan biosynthesis</i>	7	<i>Glucosidase</i>	ec:3.2.1.106	1	cngru.056570	map00510
<i>Nicotinate and nicotinamide metabolism</i>	10	<i>uridine 5'-nucleotidase</i>	ec:3.1.3.5	1	cngru.022370	map00760

<i>Via metabólica</i>	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Nicotinate and nicotinamide metabolism</i>	10	<i>Kinase</i>	ec:2.7.1.23	9	cngru.075790, cngru.040930, cngru.040910, cngru.061990, cngru.040920, cngru.075760, cngru.040900, cngru.075780, cngru.051480	map00760
<i>Nitrogen metabolism</i>	2	<i>synthase (NADH)</i>	ec:1.4.1.14	2	cngru.003460, cngru.003450	map00910
<i>Nitrogen metabolism</i>	2	<i>synthase (NADPH)</i>	ec:1.4.1.13	2	cngru.003460, cngru.003450	map00910
<i>One carbon pool by folate</i>	3	<i>reductase [NAD(P)H]</i>	ec:1.5.1.20	2	cngru.029000, cngru.045670	map00670
<i>One carbon pool by folate</i>	3	<i>Reductase</i>	ec:1.5.1.3	1	cngru.002340	map00670
<i>Other glycan degradation</i>	1	<i>psychosine hydrolase</i>	ec:3.2.1.45	1	cngru.027440	map00511
<i>Other types of O-glycan biosynthesis</i>	1	<i>mannosyltransferase</i>	ec:2.4.1.109	1	cngru.055340	map00514
<i>Oxidative phosphorylation</i>	1	<i>reductase (H+-translocating)</i>	ec:1.6.5.3	1	cngru.035340	map00190
<i>Pantothenate and CoA biosynthesis</i>	9	<i>hydroxymethyltransferase</i>	ec:2.1.2.11	1	cngru.000600	map00770
<i>Pantothenate and CoA biosynthesis</i>	9	<i>Synthase</i>	ec:2.2.1.6	2	cngru.014720, cngru.023570	map00770
<i>Pantothenate and CoA biosynthesis</i>	9	<i>Kinase</i>	ec:2.7.1.33	1	cngru.041010	map00770
<i>Pantothenate and CoA biosynthesis</i>	9	<i>reductoisomerase (NADP+)</i>	ec:1.1.1.86	1	cngru.064790	map00770
<i>Pantothenate and CoA biosynthesis</i>	9	<i>Kinase</i>	ec:2.7.1.24	1	cngru.040080	map00770
<i>Pantothenate and CoA biosynthesis</i>	9	<i>Synthase</i>	ec:2.7.8.7	2	cngru.077370, cngru.056790	map00770
<i>Pantothenate and CoA biosynthesis</i>	9	<i>2-reductase</i>	ec:1.1.1.169	1	cngru.064560	map00770
<i>Penicillin and cephalosporin biosynthesis</i>	2	<i>Oxidase</i>	ec:1.4.3.3	2	cngru.018210, cngru.039280	map00311
<i>Pentose phosphate pathway</i>	3	<i>glycolaldehydetransferase</i>	ec:2.2.1.1	1	cngru.011570	map00030
<i>Pentose phosphate pathway</i>	3	<i>phosphohexokinase</i>	ec:2.7.1.11	1	cngru.001390	map00030
<i>Pentose phosphate pathway</i>	3	<i>dehydrogenase (NADP+-dependent, decarboxylating)</i>	ec:1.1.1.44	1	cngru.010150	map00030
<i>Phenylalanine metabolism</i>	1	<i>Oxidase</i>	ec:1.4.3.21	1	cngru.045940	map00360
<i>Phenylalanine, tyrosine and tryptophan biosynthesis</i>	1	<i>Synthase</i>	ec:2.5.1.54	1	cngru.040170	map00400
<i>Phosphatidylinositol signaling system</i>	8	<i>kinase (ATP)</i>	ec:2.7.1.107	2	cngru.061990, cngru.051480	map04070
<i>Phosphatidylinositol signaling system</i>	8	<i>3-kinase</i>	ec:2.7.1.127	1	cngru.041330	map04070

<i>Via metabólica</i>	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Phosphatidylinositol signaling system</i>	8	<i>4-kinase</i>	ec:2.7.1.67	1	cngnu.068110	map04070
<i>Phosphatidylinositol signaling system</i>	8	<i>cytidyltransferase</i>	ec:2.7.7.41	1	cngnu.049270	map04070
<i>Phosphatidylinositol signaling system</i>	8	<i>3-kinase</i>	ec:2.7.1.137	1	cngnu.033200	map04070
<i>Phosphatidylinositol signaling system</i>	8	<i>2-kinase</i>	ec:2.7.1.158	2	cngnu.052940, cngnu.052930	map04070
<i>Porphyrin and chlorophyll metabolism</i>	5	<i>Synthase</i>	ec:4.4.1.17	2	cngnu.020830, cngnu.063020	map00860
<i>Porphyrin and chlorophyll metabolism</i>	5	<i>Dehydrogenase</i>	ec:1.3.1.76	1	cngnu.047930	map00860
<i>Porphyrin and chlorophyll metabolism</i>	5	<i>Ligase</i>	ec:6.1.1.17	1	cngnu.038220	map00860
<i>Porphyrin and chlorophyll metabolism</i>	5	<i>Oxidase</i>	ec:1.3.3.4	1	cngnu.071980	map00860
<i>Primary bile acid biosynthesis</i>	1	<i>Dehydrogenase</i>	ec:1.1.1.35	1	cngnu.075850	map00120
<i>Purine metabolism</i>	55	<i>Cyclase</i>	ec:4.6.1.1	1	cngnu.071800	map00230
<i>Purine metabolism</i>	55	<i>Kinase</i>	ec:2.7.4.3	2	cngnu.048390, cngnu.048380	map00230
<i>Purine metabolism</i>	55	<i>uridine 5'-nucleotidase</i>	ec:3.1.3.5	1	cngnu.022370	map00230
<i>Purine metabolism</i>	55	<i>adenylpyrophosphatase</i>	ec:3.6.1.3	13	cngnu.027750, cngnu.028930, cngnu.050950, cngnu.052600, cngnu.041150, cngnu.048550, cngnu.062790, cngnu.067200, cngnu.018830, cngnu.021760, cngnu.075830, cngnu.056360, cngnu.001720	map00230
<i>Purine metabolism</i>	55	<i>ec:3.5.3.4 allantoicase</i>	ec:3.5.3.4	1	cngnu.050860	map00230
<i>Purine metabolism</i>	55	<i>phosphodiesterase</i>	ec:3.1.4.17	1	cngnu.035490	map00230
<i>Purine metabolism</i>	55	<i>Reductase</i>	ec:1.17.4.1	1	cngnu.057970	map00230
<i>Purine metabolism</i>	55	<i>Deaminase</i>	ec:3.5.4.6	1	cngnu.039500	map00230
<i>Purine metabolism</i>	55	<i>Synthase</i>	ec:6.3.2.6	1	cngnu.022170	map00230

<i>Via metabólica</i>	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Purine metabolism</i>	55	<i>RNA polymerase</i>	ec:2.7.7.6	20	cngru.055960, cngru.029330, cngru.057500, cngru.035920, cngru.052220, cngru.075160, cngru.052210, cngru.027420, cngru.055920, cngru.054840, cngru.051630, cngru.051740, cngru.076090, cngru.024990, cngru.066020, cngru.062360, cngru.035290, cngru.066320, cngru.073750, cngru.038540 cngru.036490, cngru.045920, cngru.008510, cngru.022560, cngru.037630, cngru.038620, cngru.032450, cngru.074960, cngru.066360, cngru.011920, cngru.057550, cngru.045250	map00230
<i>Purine metabolism</i>	55	<i>DNA polymerase</i>	ec:2.7.7.7	12		map00230
<i>Purine metabolism</i>	55	<i>phosphoribosylidiphosphate amidotransferase</i>	5- ec:2.4.2.14	1	cngru.040830	map00230
<i>Purine metabolism</i>	55	<i>Kinase</i>	ec:2.7.4.10	2	cngru.048390, cngru.048380	map00230
<i>Pyruvate metabolism</i>	9	<i>Dehydrogenase</i>	ec:1.8.1.4	1	cngru.054820	map00620
<i>Pyruvate metabolism</i>	9	<i>Synthase</i>	ec:2.3.3.14	1	cngru.049540	map00620
<i>Pyruvate metabolism</i>	9	<i>Synthase</i>	ec:2.3.3.13	1	cngru.018550	map00620
<i>Pyruvate metabolism</i>	9	<i>carboxykinase (ATP)</i>	ec:4.1.1.49	1	cngru.076840	map00620
<i>Pyruvate metabolism</i>	9	<i>Hydratase</i>	ec:4.2.1.2	1	cngru.008320	map00620
<i>Pyruvate metabolism</i>	9	<i>Ligase</i>	ec:6.2.1.1	1	cngru.029370	map00620
<i>Pyruvate metabolism</i>	9	<i>dehydrogenase (decarboxylating)</i>	ec:1.1.1.39	2	cngru.066050, cngru.075660	map00620
<i>Pyruvate metabolism</i>	9	<i>acetyltransferase</i>	ec:2.3.1.12	1	cngru.050980	map00620
<i>Sphingolipid metabolism</i>	1	<i>psychosine hydrolase</i>	ec:3.2.1.45	1	cngru.027440	map00600
<i>Starch and sucrose metabolism</i>	13	<i>Glycogenase</i>	ec:3.2.1.1	2	cngru.005230, cngru.005240	map00500
<i>Starch and sucrose metabolism</i>	13	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	2	cngru.018670, cngru.018680	map00500
<i>Starch and sucrose metabolism</i>	13	<i>Maltase</i>	ec:3.2.1.20	1	cngru.018820	map00500
<i>Starch and sucrose metabolism</i>	13	<i>Gentiobiase</i>	ec:3.2.1.21	1	cngru.036970	map00500
<i>Starch and sucrose metabolism</i>	13	<i>Synthase</i>	ec:2.4.1.34	1	cngru.013830	map00500

<i>Via metabólica</i>	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Starch and sucrose metabolism</i>	13	<i>Trehalase</i>	ec:3.2.1.28	2	cngru.033540, cngru.068210	map00500
<i>Starch and sucrose metabolism</i>	13	<i>amyl-1,6-glucosidase</i>	ec:3.2.1.33	1	cngru.003670	map00500
<i>Starch and sucrose metabolism</i>	13	<i>synthase (UDP-forming)</i>	ec:2.4.1.15	1	cngru.044010	map00500
<i>Starch and sucrose metabolism</i>	13	<i>Synthase</i>	ec:2.4.1.11	2	cngru.000860, cngru.000870	map00500
<i>Starch and sucrose metabolism</i>	13	<i>disproportionating enzyme</i>	ec:2.4.1.25	1	cngru.003670	map00500
<i>Streptomycin biosynthesis</i>	3	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	2	cngru.018670, cngru.018680	map00521
<i>Streptomycin biosynthesis</i>	3	<i>Synthase</i>	ec:5.5.1.4	1	cngru.009630	map00521
<i>Sulfur metabolism</i>	1	<i>reductase (thioredoxin)</i>	ec:1.8.4.8	1	cngru.032270	map00920
<i>Synthesis and degradation of ketone bodies</i>	1	<i>CoA-transferase</i>	ec:2.8.3.5	1	cngru.046890	map00072
<i>T cell receptor signaling pathway</i>	4	<i>Phosphatase</i>	ec:3.1.3.16	4	cngru.051950, cngru.051940, cngru.002730, cngru.050250	map04660
<i>Taurine and hypotaurine metabolism</i>	4	<i>glutamyl transpeptidase</i>	ec:2.3.2.2	4	cngru.041230, cngru.049940, cngru.072660, cngru.072670	map00430
<i>Terpenoid backbone biosynthesis</i>	3	<i>Decarboxylase</i>	ec:4.1.1.33	1	cngru.045830	map00900
<i>Terpenoid backbone biosynthesis</i>	3	<i>reductase (NADPH)</i>	ec:1.1.1.34	1	cngru.067190	map00900
<i>Terpenoid backbone biosynthesis</i>	3	<i>Oxidase</i>	ec:1.8.3.5	1	cngru.039260	map00900
<i>Toluene degradation</i>	1	<i>Dehydrogenase</i>	ec:1.1.1.35	1	cngru.075850	map00623
<i>Tropane, piperidine and pyridine alkaloid biosynthesis</i>	1	<i>Oxidase</i>	ec:1.4.3.21	1	cngru.045940	map00960
<i>Tryptophan metabolism</i>	2	<i>Equilase</i>	ec:1.11.1.6	1	cngru.047130	map00380
<i>Tryptophan metabolism</i>	2	<i>Dehydrogenase</i>	ec:1.1.1.35	1	cngru.075850	map00380
<i>Valine, leucine and isoleucine biosynthesis</i>	5	<i>Synthase</i>	ec:2.3.3.13	1	cngru.018550	map00290
<i>Valine, leucine and isoleucine biosynthesis</i>	5	<i>Synthase</i>	ec:2.2.1.6	2	cngru.014720, cngru.023570	map00290
<i>Valine, leucine and isoleucine biosynthesis</i>	5	<i>reductoisomerase (NADP+)</i>	ec:1.1.1.86	1	cngru.064790	map00290

<i>Via metabólica</i>	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Valine, leucine and isoleucine biosynthesis</i>	5	<i>Dehydrogenase</i>	ec:1.1.1.85	1	cngru.025550	map00290
<i>Valine, leucine and isoleucine degradation</i>	5	<i>Dehydrogenase</i>	ec:1.8.1.4	1	cngru.054820	map00280
<i>Valine, leucine and isoleucine degradation</i>	5	<i>Dehydrogenase</i>	ec:1.1.1.35	1	cngru.075850	map00280
<i>Valine, leucine and isoleucine degradation</i>	5	<i>Dehydrogenase</i>	ec:1.1.1.31	1	cngru.010150	map00280
<i>Valine, leucine and isoleucine degradation</i>	5	<i>Ligase</i>	ec:6.2.1.16	1	cngru.056190	map00280
<i>Valine, leucine and isoleucine degradation</i>	5	<i>CoA-transferase</i>	ec:2.8.3.5	1	cngru.046890	map00280
<i>Various types of N-glycan biosynthesis</i>	4	<i>1,2-alpha-mannosidase</i>	ec:3.2.1.113	3	cngru.004730, cngru.056600, cngru.071380	map00513
<i>Various types of N-glycan biosynthesis</i>	4	<i>alpha-1,6-mannosyltransferase</i>	ec:2.4.1.232	1	cngru.048970	map00513
<i>Various types of N-glycan biosynthesis</i>	4	<i>alpha-1,3-mannosyltransferase</i>	ec:2.4.1.132	1	cngru.048970	map00513
<i>Vitamin B6 metabolism</i>	1	<i>5'-phosphate synthase</i>	ec:1.4.3.5	1	cngru.019180	map00750

Tabela C. 5: Vias metabólicas mapeadas em *C. neoformans* var. *neoformans*, linhagem B3501A. Todas as proteínas mapeadas referem-se às IDPs e IDRs identificadas neste estudo.

<i>Via metabólica</i>	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Alanine, aspartate and glutamate metabolism</i>	1	<i>synthase (glutamine-hydrolysing)</i>	ec:6.3.5.5	1	Cnb35.005900	map00250
<i>Alanine, aspartate and glutamate metabolism</i>	1	<i>synthase (ammonia)</i>	ec:6.3.4.16	1	Cnb35.005900	map00250
<i>Alanine, aspartate and glutamate metabolism</i>	1	<i>carbamoyltransferase</i>	ec:2.1.3.2	1	Cnb35.005900	map00250
<i>Amino sugar and nucleotide sugar metabolism</i>	1	<i>Synthase</i>	ec:2.4.1.16	1	Cnb35.005180	map00520
<i>Aminoacyl-tRNA biosynthesis</i>	2	<i>Ligase</i>	ec:6.1.1.18	1	Cnb35.000760	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	2	<i>Ligase</i>	ec:6.1.1.5	1	Cnb35.001320	map00970
<i>Aminoacyl-tRNA biosynthesis</i>	2	<i>Ligase</i>	ec:6.1.1.1	1	Cnb35.001320	map00970
<i>Aminobenzoate degradation</i>	5	<i>nitrophenyl phosphatase</i>	ec:3.1.3.41	5	Cnb35.004710, Cnb35.004010, Cnb35.004670, Cnb35.010230, Cnb35.003660	map00627
<i>Arginine biosynthesis</i>	1	<i>synthase (ammonia)</i>	ec:6.3.4.16	1	Cnb35.005900	map00220
<i>Biosynthesis of antibiotics</i>	6	<i>Dehydrogenase</i>	ec:1.1.1.95	1	Cnb35.007450	map01130
<i>Biosynthesis of antibiotics</i>	6	<i>dihydroxyacetonetransferase</i>	ec:2.2.1.2	1	Cnb35.008970	map01130
<i>Biosynthesis of antibiotics</i>	6	<i>Synthase</i>	ec:2.2.1.6	1	Cnb35.002480	map01130
<i>Biosynthesis of antibiotics</i>	6	<i>Ligase</i>	ec:6.2.1.1	1	Cnb35.007690	map01130
<i>Biosynthesis of antibiotics</i>	6	<i>Synthase</i>	ec:6.3.2.6	1	Cnb35.001320	map01130
<i>Biosynthesis of antibiotics</i>	6	<i>dehydrogenase (NAD+)</i>	ec:1.1.1.41	1	Cnb35.004490	map01130
<i>Butanoate metabolism</i>	1	<i>Synthase</i>	ec:2.2.1.6	1	Cnb35.002480	map00650
<i>C5-Branched dibasic acid metabolism</i>	2	<i>Synthase</i>	ec:2.2.1.6	1	Cnb35.002480	map00660
<i>C5-Branched dibasic acid metabolism</i>	2	<i>dehydrogenase</i>	ec:1.1.1.85	1	Cnb35.004210	map00660
<i>Carbon fixation pathways in prokaryotes</i>	2	<i>Ligase</i>	ec:6.2.1.1	1	Cnb35.007690	map00720
<i>Carbon fixation pathways in prokaryotes</i>	2	<i>reductase [NAD(P)H]</i>	ec:1.5.1.20	1	Cnb35.007320	map00720
<i>Citrate cycle (TCA cycle)</i>	1	<i>dehydrogenase (NAD+)</i>	ec:1.1.1.41	1	Cnb35.004490	map00020
<i>Drug metabolism - cytochrome P450</i>	1	<i>monooxygenase</i>	ec:1.14.13.8	1	Cnb35.005130	map00982

<i>Via metabólica</i>	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Folate biosynthesis</i>	1	<i>cyclohydrolase I</i>	ec:3.5.4.16	1	Cnb35.009110	map00790
<i>Fructose and mannose metabolism</i>	2	<i>phosphofructokinase 2</i>	ec:2.7.1.105	2	Cnb35.005870, Cnb35.005880	map00051
<i>Glycine, serine and threonine metabolism</i>	1	<i>dehydrogenase</i>	ec:1.1.1.95	1	Cnb35.007450	map00260
<i>Glycolysis / Gluconeogenesis</i>	1	<i>Ligase</i>	ec:6.2.1.1	1	Cnb35.007690	map00010
<i>Inositol phosphate metabolism</i>	2	<i>3-kinase</i>	ec:2.7.1.137	1	Cnb35.010910	map00562
<i>Inositol phosphate metabolism</i>	2	<i>5-kinase</i>	ec:2.7.1.68	1	Cnb35.003630	map00562
<i>Lysine degradation</i>	1	<i>N-methyltransferase</i>	ec:2.1.1.43	1	Cnb35.001040	map00310
<i>Methane metabolism</i>	2	<i>dehydrogenase</i>	ec:1.1.1.95	1	Cnb35.007450	map00680
<i>Methane metabolism</i>	2	<i>Ligase</i>	ec:6.2.1.1	1	Cnb35.007690	map00680
<i>mTOR signaling pathway</i>	1	<i>protein kinase</i>	ec:2.7.11.24	1	Cnb35.010940	map04150
<i>N-Glycan biosynthesis</i>	1	<i>Kinase</i>	ec:2.7.1.108	1	Cnb35.008450	map00510
<i>Nitrogen metabolism</i>	1	<i>synthase (ammonia)</i>	ec:6.3.4.16	1	Cnb35.005900	map00910
<i>One carbon pool by folate</i>	1	<i>reductase [NAD(P)H]</i>	ec:1.5.1.20	1	Cnb35.007320	map00670
<i>Oxidative phosphorylation</i>	2	<i>dehydrogenase</i>	ec:1.6.99.3	2	Cnb35.004890, Cnb35.007590	map00190
<i>Oxidative phosphorylation</i>	2	<i>reductase (H⁺-translocating)</i>	ec:1.6.5.3	2	Cnb35.004890, Cnb35.007590	map00190
<i>Pantothenate and CoA biosynthesis</i>	1	<i>Synthase</i>	ec:2.2.1.6	1	Cnb35.002480	map00770
<i>Pentose phosphate pathway</i>	1	<i>dihydroxyacetone transferase</i>	ec:2.2.1.2	1	Cnb35.008970	map00030
<i>Phosphatidylinositol signaling system</i>	2	<i>3-kinase</i>	ec:2.7.1.137	1	Cnb35.010910	map04070
<i>Phosphatidylinositol signaling system</i>	2	<i>5-kinase</i>	ec:2.7.1.68	1	Cnb35.003630	map04070
<i>Propanoate metabolism</i>	1	<i>Ligase</i>	ec:6.2.1.1	1	Cnb35.007690	map00640
<i>Purine metabolism</i>	44	<i>phosphatase</i>	ec:3.6.1.15	39	Cnb35.009270, Cnb35.006080, Cnb35.006460, Cnb35.002380, Cnb35.004960, Cnb35.005830, Cnb35.000360, Cnb35.006940, Cnb35.007950, Cnb35.008720, Cnb35.006540, Cnb35.008300, Cnb35.008960, Cnb35.001610, Cnb35.000960, Cnb35.007000, Cnb35.007760, Cnb35.007680, Cnb35.009320, Cnb35.003580, Cnb35.006890,	map00230

<i>Via metabólica</i>	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
					Cnb35.008330, Cnb35.007260, Cnb35.007360, Cnb35.009580, Cnb35.001560, Cnb35.007820, Cnb35.003500, Cnb35.002790, Cnb35.006210, Cnb35.005520, Cnb35.006510, Cnb35.007640, Cnb35.001000, Cnb35.003540, Cnb35.005200, Cnb35.008730, Cnb35.000910, Cnb35.001900	
<i>Purine metabolism</i>	44	<i>Synthase</i>	ec:6.3.2.6	1	Cnb35.001320	map00230
					Cnb35.002380, Cnb35.005830, Cnb35.006940, Cnb35.007950, Cnb35.006540, Cnb35.008300, Cnb35.008960, Cnb35.001610, Cnb35.000960, Cnb35.007000, Cnb35.007680, Cnb35.003580,	
<i>Purine metabolism</i>	44	<i>adenylpyrophosphatase</i>	ec:3.6.1.3	27	Cnb35.008330, Cnb35.007260, Cnb35.007360, Cnb35.009580, Cnb35.001560, Cnb35.007820, Cnb35.003500, Cnb35.006210, Cnb35.006510, Cnb35.007640, Cnb35.001000, Cnb35.003540, Cnb35.005200, Cnb35.008730, Cnb35.001900	map00230
<i>Purine metabolism</i>	44	<i>RNA polymerase</i>	ec:2.7.7.6	2	Cnb35.007650, Cnb35.005910	map00230
<i>Purine metabolism</i>	44	<i>DNA polymerase</i>	ec:2.7.7.7	2	Cnb35.001670, Cnb35.010340	map00230
<i>Pyrimidine metabolism</i>	6	<i>synthase (glutamine-hydrolysing)</i>	ec:6.3.5.5	1	Cnb35.005900	map00240
<i>Pyrimidine metabolism</i>	6	<i>carbamoylaspartic dehydrase</i>	ec:3.5.2.3	2	Cnb35.005900, Cnb35.007030	map00240
<i>Pyrimidine metabolism</i>	6	<i>carbamoyltransferase</i>	ec:2.1.3.2	1	Cnb35.005900	map00240
<i>Pyrimidine metabolism</i>	6	<i>RNA polymerase</i>	ec:2.7.7.6	2	Cnb35.007650, Cnb35.005910	map00240
<i>Pyrimidine metabolism</i>	6	<i>DNA polymerase</i>	ec:2.7.7.7	2	Cnb35.001670, Cnb35.010340	map00240
<i>Pyruvate metabolism</i>	1	<i>Ligase</i>	ec:6.2.1.1	1	Cnb35.007690	map00620
<i>Riboflavin metabolism</i>	1	<i>cyclohydrolase II</i>	ec:3.5.4.25	1	Cnb35.005420	map00740
<i>Starch and sucrose metabolism</i>	1	<i>Trehalase</i>	ec:3.2.1.28	1	Cnb35.011220	map00500
<i>Sulfur metabolism</i>	1	<i>reductase (thioredoxin)</i>	ec:1.8.4.8	1	Cnb35.010200	map00920

<i>Via metabólica</i>	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>T cell receptor signaling pathway</i>	5	<i>Phosphatase</i>	ec:3.1.3.16	5	Cnb35.004710, Cnb35.004010, Cnb35.004670, Cnb35.010230, Cnb35.003660 Cnb35.009270, Cnb35.006080, Cnb35.006460, Cnb35.002380, Cnb35.004960, Cnb35.005830, Cnb35.000360, Cnb35.006940, Cnb35.007950, Cnb35.008720, Cnb35.006540, Cnb35.008300, Cnb35.008960, Cnb35.001610, Cnb35.000960, Cnb35.007000, Cnb35.007760, Cnb35.007680,	map04660
<i>Thiamine metabolism</i>	39	<i>Phosphatase</i>	ec:3.6.1.15	39	Cnb35.009320, Cnb35.003580, Cnb35.006890, Cnb35.008330, Cnb35.007260, Cnb35.007360, Cnb35.009580, Cnb35.001560, Cnb35.007820, Cnb35.003500, Cnb35.002790, Cnb35.006210, Cnb35.005520, Cnb35.006510, Cnb35.007640, Cnb35.001000, Cnb35.003540, Cnb35.005200, Cnb35.008730, Cnb35.000910, Cnb35.001900	map00730
<i>Valine, leucine and isoleucine biosynthesis</i>	2	<i>Synthase</i>	ec:2.2.1.6	1	Cnb35.002480	map00290
<i>Valine, leucine and isoleucine biosynthesis</i>	2	<i>dehydrogenase</i>	ec:1.1.1.85	1	Cnb35.004210	map00290
<i>Zeatin biosynthesis</i>	1	<i>dimethylallyltransferase</i>	ec:2.5.1.75	1	Cnb35.003390	map00908

Tabela C. 6: Vias metabólicas mapeadas em *C. neoformans* var. *neoformans*, linhagem JEC21. Todas as proteínas mapeadas referem-se às IDPs e IDRs identificadas neste estudo

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de seqüências enzimáticas	Proteínas	Mapa KEGG
<i>Aminobenzoate degradation</i>	3	<i>phosphatase</i>	ec:3.1.3.2	3	Cnj21.045540, Cnj21.051910, Cnj21.051900	map00627
<i>Arginine and proline metabolism</i>	3	<i>Oxidase</i>	ec:1.4.3.3	2	Cnj21.058690, Cnj21.013270	map00330
<i>Arginine and proline metabolism</i>	3	<i>5-kinase</i>	ec:2.7.2.11	1	Cnj21.059890	map00330
<i>Biosynthesis of antibiotics</i>	45	<i>dehydrogenase</i>	ec:1.8.1.4	1	Cnj21.041250	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>Synthase</i>	ec:2.3.3.14	1	Cnj21.036670	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>succinyltransferase</i>	ec:2.3.1.61	1	Cnj21.023690	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>Kinase</i>	ec:2.7.4.3	1	Cnj21.035680	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>carboxykinase (ATP)</i>	ec:4.1.1.49	1	Cnj21.063930	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>Synthase</i>	ec:6.3.4.5	2	Cnj21.054110, Cnj21.054120	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>decarboxylase</i>	ec:4.1.1.33	1	Cnj21.016030	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>Oxidase</i>	ec:1.4.3.3	2	Cnj21.058690, Cnj21.013270	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>Hydratase</i>	ec:4.2.1.3	2	Cnj21.019580, Cnj21.038060	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>dehydrogenase</i>	ec:1.1.1.3	1	Cnj21.010090	map01130
<i>Biosynthesis of antibiotics</i>	45	<i>dehydrogenase</i>	ec:1.1.1.95	1	Cnj21.007310	map01130
<i>Biosynthesis of unsaturated fatty acids</i>	1	<i>9-desaturase</i>	ec:1.14.19.1	1	Cnj21.009240	map01040
<i>Biotin metabolism</i>	3	<i>Synthase</i>	ec:2.8.1.6	1	Cnj21.050960	map00780
<i>Biotin metabolism</i>	3	<i>Ligase</i>	ec:6.3.4.15	2	Cnj21.029160, Cnj21.029170	map00780
<i>Butirosin and neomycin biosynthesis</i>	1	<i>hexokinase type IV glucokinase</i>	ec:2.7.1.1	1	Cnj21.059010	map00524
<i>Carbapenem biosynthesis</i>	1	<i>5-kinase</i>	ec:2.7.2.11	1	Cnj21.059890	map00332
<i>Carbon fixation pathways in prokaryotes</i>	7	<i>reductase [NAD(P)H]</i>	ec:1.5.1.20	2	Cnj21.007190, Cnj21.015890	map00720
<i>Carbon fixation pathways in prokaryotes</i>	7	<i>Hydratase</i>	ec:4.2.1.3	2	Cnj21.019580, Cnj21.038060	map00720
<i>Carbon fixation pathways in prokaryotes</i>	7	<i>ligase (ADP-forming)</i>	ec:6.2.1.5	1	Cnj21.053970	map00720
<i>Carbon fixation pathways in prokaryotes</i>	7	<i>Ligase</i>	ec:6.2.1.1	1	Cnj21.007530	map00720
<i>Carbon fixation pathways in prokaryotes</i>	7	<i>dehydrogenase</i>	ec:1.1.1.35	1	Cnj21.062770	map00720
<i>Carbon fixation pathways in prokaryotes</i>	7	<i>citrate synthase</i>	ec:2.3.3.8	1	Cnj21.053970	map00720
<i>Drug metabolism - cytochrome P450</i>	2	<i>monooxygenase</i>	ec:1.14.13.8	2	Cnj21.005020, Cnj21.035350	map00982

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Drug metabolism - other enzymes</i>	1	<i>Kinase</i>	ec:2.7.1.48	1	Cnj21.054040	map00983
<i>Lipoic acid metabolism</i>	1	<i>Synthase</i>	ec:2.8.1.8	1	Cnj21.015410	map00785
<i>Purine metabolism</i>	46	<i>Cyclase</i>	ec:4.6.1.1	1	Cnj21.055610	map00230
<i>Purine metabolism</i>	46	<i>Kinase</i>	ec:2.7.4.3	1	Cnj21.035680	map00230
<i>Purine metabolism</i>	46	<i>uridine 5'-nucleotidase</i>	ec:3.1.3.5	1	Cnj21.001540	map00230
<i>Purine metabolism</i>	46	<i>adenylpyrophosphatase</i>	ec:3.6.1.3	9	Cnj21.009430, Cnj21.006100, Cnj21.059110, Cnj21.035800, Cnj21.000990, Cnj21.033280, Cnj21.037860, Cnj21.051790, Cnj21.007130	map00230
<i>Purine metabolism</i>	46	<i>ec:3.5.3.4 allantoinase</i>	ec:3.5.3.4	1	Cnj21.037780	map00230
<i>Purine metabolism</i>	46	<i>phosphodiesterase</i>	ec:3.1.4.17	1	Cnj21.023870	map00230
<i>Purine metabolism</i>	46	<i>deaminase</i>	ec:3.5.4.3	1	Cnj21.051320	map00230
<i>Purine metabolism</i>	46	<i>Reductase</i>	ec:1.17.4.1	1	Cnj21.045110	map00230
<i>Purine metabolism</i>	46	<i>deaminase</i>	ec:3.5.4.6	1	Cnj21.031940	map00230
<i>Purine metabolism</i>	46	<i>Synthase</i>	ec:6.3.2.6	1	Cnj21.001350	map00230
<i>Purine metabolism</i>	46	<i>RNA polymerase</i>	ec:2.7.7.6	15	Cnj21.007500, Cnj21.003670, Cnj21.024050, Cnj21.045490, Cnj21.046880, Cnj21.047530, Cnj21.062980, Cnj21.023540, Cnj21.005800, Cnj21.062080, Cnj21.038560, Cnj21.038460, Cnj21.038950, Cnj21.051040, Cnj21.050760, Cnj21.001720, Cnj21.012710, Cnj21.019240, Cnj21.011870, Cnj21.029900, Cnj21.051090, Cnj21.016120, Cnj21.023010, Cnj21.061870, Cnj21.026360, Cnj21.015560, Cnj21.045440	map00230
<i>Purine metabolism</i>	46	<i>DNA polymerase</i>	ec:2.7.7.7	12	Cnj21.033050	map00230
<i>Purine metabolism</i>	46	<i>phosphoribosyldiphosphate amidotransferase</i>	5- ec:2.4.2.14	1	Cnj21.033050	map00230
<i>Purine metabolism</i>	46	<i>Kinase</i>	ec:2.7.4.10	1	Cnj21.035680	map00230
<i>Pyruvate metabolism</i>	9	<i>dehydrogenase</i>	ec:1.8.1.4	1	Cnj21.041250	map00620
<i>Pyruvate metabolism</i>	9	<i>Synthase</i>	ec:2.3.3.14	1	Cnj21.036670	map00620

Via metabólica	Número de Proteínas na Via	Enzima	Número da Enzima	Número de sequências enzimáticas	Proteínas	Mapa KEGG
<i>Pyruvate metabolism</i>	9	<i>Synthase</i>	ec:2.3.3.13	1	Cnj21.058950	map00620
<i>Pyruvate metabolism</i>	9	<i>carboxykinase (ATP)</i>	ec:4.1.1.49	1	Cnj21.063930	map00620
<i>Pyruvate metabolism</i>	9	<i>dehydrogenase (cytochrome)</i>	ec:1.1.2.3	1	Cnj21.059180	map00620
<i>Pyruvate metabolism</i>	9	<i>Ligase</i>	ec:6.2.1.1	1	Cnj21.007530	map00620
<i>Pyruvate metabolism</i>	9	<i>dehydrogenase (decarboxylating)</i>	ec:1.1.1.39	2	Cnj21.050790, Cnj21.062570	map00620
<i>Pyruvate metabolism</i>	9	<i>acetyltransferase</i>	ec:2.3.1.12	1	Cnj21.037890	map00620
<i>Steroid degradation</i>	1	<i>dehydrogenase</i>	ec:1.1.1.145	1	Cnj21.053730	map00984
<i>Sulfur metabolism</i>	2	<i>sulfite reductase (NADPH)</i>	ec:1.8.1.2	1	Cnj21.016570	map00920
<i>Sulfur metabolism</i>	2	<i>reductase (thioredoxin)</i>	ec:1.8.4.8	1	Cnj21.026510	map00920
<i>Taurine and hypotaurine metabolism</i>	5	<i>glutamyl transpeptidase</i>	ec:2.3.2.2	3	Cnj21.056250, Cnj21.037020, Cnj21.033350	map00430
<i>Taurine and hypotaurine metabolism</i>	5	<i>Decarboxylase</i>	ec:4.1.1.15	2	Cnj21.014780, Cnj21.056730	map00430
<i>Toluene degradation</i>	1	<i>dehydrogenase</i>	ec:1.1.1.35	1	Cnj21.062770	map00623
<i>Valine, leucine and isoleucine biosynthesis</i>	6	<i>Synthase</i>	ec:2.3.3.13	1	Cnj21.058950	map00290
<i>Valine, leucine and isoleucine biosynthesis</i>	6	<i>Synthase</i>	ec:2.2.1.6	2	Cnj21.021820, Cnj21.002560	map00290
<i>Valine, leucine and isoleucine biosynthesis</i>	6	<i>ammonia-lyase</i>	ec:4.3.1.19	1	Cnj21.012700	map00290
<i>Valine, leucine and isoleucine biosynthesis</i>	6	<i>reductoisomerase (NADP+)</i>	ec:1.1.1.86	1	Cnj21.049620	map00290
<i>Valine, leucine and isoleucine biosynthesis</i>	6	<i>dehydrogenase</i>	ec:1.1.1.85	1	Cnj21.004140	map00290