



Universidade de Brasília

Instituto de Ciências Exatas
Departamento de Ciência da Computação

**Método para reconstrução in silico de redes
metabólicas de fungos: um estudo de caso para o
Paracoccidioides lutzii**

Waldeyr Mendes Cordeiro da Silva

Dissertação apresentada como requisito parcial
para conclusão do Mestrado em Informática

Orientadora

Prof.^a Dr.^a Maria Emilia Machado Telles Walter

Brasília
2014

Universidade de Brasília — UnB
Instituto de Ciências Exatas
Departamento de Ciência da Computação
Mestrado em Informática

Coordenadora: Prof.^a Dr.^a Alba Cristina Magalhães Alves de Melo

Banca examinadora composta por:

Prof.^a Dr.^a Maria Emilia Machado Telles Walter (Orientadora) — CIC/UnB
Prof. Dr.^a Célia Ghedini Ralha — CIC/UnB
Prof. Dr.^a Maria Sueli Soares Felipe — IB/UnB

CIP — Catalogação Internacional na Publicação

Silva, Waldeyr Mendes Cordeiro da.

Método para reconstrução in silico de redes metabólicas de fungos:
um estudo de caso para o *Paracoccidioides lutzii* / Waldeyr Mendes
Cordeiro da Silva. Brasília : UnB, 2014.

104 p. : il. ; 29,5 cm.

Dissertação (Mestrado) — Universidade de Brasília, Brasília, 2014.

1. Bioinformática, 2. rede metabólica, 3. fungos, 4. *Paracoccidioides lutzii*

CDU 004

Endereço: Universidade de Brasília
Campus Universitário Darcy Ribeiro — Asa Norte
CEP 70910-900
Brasília-DF — Brasil



Universidade de Brasília

Instituto de Ciências Exatas
Departamento de Ciência da Computação

Método para reconstrução *in silico* de redes metabólicas de fungos: um estudo de caso para o *Paracoccidioides lutzii*

Waldeyr Mendes Cordeiro da Silva

Dissertação apresentada como requisito parcial
para conclusão do Mestrado em Informática

Prof.^a Dr.^a Maria Emilia Machado Telles Walter (Orientadora)
CIC/UnB

Prof. Dr.^a Célia Ghedini Ralha Prof. Dr.^a Maria Sueli Soares Felipe
CIC/UnB IB/UnB

Prof.^a Dr.^a Alba Cristina Magalhães Alves de Melo
Coordenadora do Mestrado em Informática

Brasília, 25 de julho de 2014

Dedicatória

Dedico este trabalho à minha família, especialmente minha esposa e filha.

Agradecimentos

Agradeço à Professora Dra. Maria Emília Machado Telles Walter. Sua competência inspira.

Agradeço aos meus amigos, colegas de trabalho e de UnB.

Resumo

Os fungos produzem uma variedade de metabólitos com aplicações na indústria e na medicina, ao mesmo tempo que podem ser patógenos humanos. O *Paracoccidioides lutzii* é um fungo dimórfico causador da paracoccidioidomicose (PCM), uma micose sistêmica autóctone da América Latina, onde cerca de 10 milhões de pessoas estão infectadas, inclusive no Brasil, o qual contribui com cerca de 80% dos casos, afetando principalmente indivíduos em áreas rurais. Neste trabalho, foi criado um método para identificar e organizar vias de metabolismo em fungos, com ênfase no metabolismo secundário. Para validar o método, foi realizado um estudo de caso para o fungo *P. lutzii*, cujos resultados obtidos identificaram 2.087 reações enzimáticas, 1.437 enzimas, 1.464 compostos e 335 vias metabólicas. As vias de metabolismo secundário representaram cerca de 4,5% do total de vias identificadas no *P. lutzii*. O método criado pode ser aplicado à reconstrução *in silico* de redes metabólicas de outros fungos.

Palavras-chave: Bioinformática, rede metabólica, fungos, *Paracoccidioides lutzii*

Abstract

Fungi produce various metabolites with applications in industry and medicine, furthermore, can be human pathogens. *Paracoccidioides lutzii* is a dimorphic fungus that causes paracoccidioidomycosis (PCM), a systemic mycosis in Latin America affecting approximately 10 million people, including Brazil, which contributes about 80% of cases, usually affecting individuals in rural areas. In this work, we create a method to identify and organize metabolic pathways in fungi, with emphasis on secondary metabolism. To validate this method, a case study was performed for the fungus *P. lutzii*. The results identified 2,087 enzymatic reactions, 1,437 enzymes, 1,464 compounds and 335 metabolic pathways. The secondary metabolism pathways represented around 4.5% of pathways identified in the *P. lutzii*. The method created can be applied to the *in silico* reconstruction of metabolic networks of other fungi.

Keywords: Bioinformatics, metabolic network, fungi, *Paracoccidioides lutzii*

Sumário

1	Introdução	1
1.1	Motivação	2
1.2	Problema	2
1.3	Objetivos	3
1.4	Descrição dos Capítulos	3
2	Metabolismo	4
2.1	Conceitos básicos de Bioquímica	4
2.2	Conceitos básicos de metabolismo	7
2.2.1	Vias metabólicas	9
2.2.2	Metabolismo primário	11
2.2.3	Metabolismo secundário	11
3	O fungo <i>Paracoccidioides lutzii</i>	15
3.1	<i>Paracoccidioides brasiliensis</i> e <i>Paracoccidioides lutzii</i>	15
3.2	A paracoccidioidomicose (PCM)	18
3.3	Rede metabólica do <i>P. lutzii</i> : uma primeira versão	19
4	Ferramentas computacionais	22
4.1	Bancos de dados	22
4.1.1	Bancos de dados de relacionais	22
4.1.2	UniProt	23
4.1.3	BRENDA	23
4.1.4	BioCyc	23
4.1.5	KEGG	24
4.2	Anotação	24
4.2.1	KAAS	25
4.2.2	CDD	26
4.2.3	InterPro	26
4.3	Metabolismo secundário	26

4.3.1	antiSMASH	26
4.3.2	SMURF	27
4.4	Reconstrução de redes metabólicas	28
4.4.1	PathwayTools	29
5	Método para reconstrução <i>in silico</i> de redes metabólicas em fungos	32
5.1	Descrição geral do método	32
5.2	Detalhes das etapas	32
6	Estudo de caso: rede metabólica do <i>P. lutzii</i>	36
6.1	Método aplicado ao <i>P. lutzii</i>	36
6.2	Dados da rede metabólica <i>in silico</i> do <i>P. lutzii</i>	40
6.3	Discussão	47
7	Conclusão	51
7.1	Contribuições	52
7.2	Trabalhos futuros	52
Referências		53
A	Tabela com as vias de metabolismo secundário identificadas	59
B	Tabela de reações com gaps e as vias metabólicas onde ocorrem	61
C	Tabela de transcritos em suas respectivas formas de ocorrência	70

Listas de Figuras

2.1	Ciclo do ATP [8]	6
2.2	Estágios do Catabolismo [8]	8
2.3	Via da gluconeogênese encontrada no fungo <i>Candida albicans SC5314</i> [58].	10
2.4	Expressão de metabólitos secundários em ágar como forma de inibição de colônias adjacentes [66].	11
2.5	Os principais grupos de metabólitos secundários de fungos. [47].	14
3.1	Distribuição geográfica do gênero <i>Paracoccidioides</i> e <i>Lacazia loboi</i> [77]. . .	17
3.2	Árvore filogenética de fungos dimórficos, obtida usando o critério <i>maximum likelihood RaxML</i> [17].	17
3.3	As duas formas do <i>Paracoccidioides lutzii</i> : forma leveduriforme característica com brotamentos múltiplos em forma de roda de leme [25] à esquerda e forma de micélio à direita.	18
3.4	Página com os dados da rede metabólica reconstruída por Alvarez [3]. . . .	20
3.5	Página com <i>links</i> para as vias metabólicas reconstruída por Alvarez [3]. . .	21
4.1	Metabolismo da pirimidina no <i>P. lutzii</i> [46]	25
4.2	Exemplo de arquivo no formatoFASTA, representando aminoácidos.	25
4.3	Pipeline da análise genômica de metabólitos secundários realizada pelo antiSMASH [57].	27
4.4	Trechos dos arquivos de entrada gerados a partir do banco de dados Plutzii: (A) genetic-elements.dat, (B) PAAG_0001T0.fsa, (C) PAAG_0001T0.pf .	31
5.1	Método para reconstrução <i>in silico</i> de redes metabólicas de fungos.	33
5.2	Modelo de Entidade-Relacionamento do banco de dados	34
6.1	Método para reconstrução <i>in silico</i> de redes metabólicas em fungos aplicado ao <i>P. lutzii</i>	37

6.2	Modelo de Entidade-Relacionamento (MER) do banco de dados adaptado para o <i>P. lutzii</i> . Este MER foi chamado de Plutzii e contém as tabelas e relacionamentos com os dados do <i>P. lutzii</i> e os dados da rede previamente reconstruída. É utilizado como ponto de partida para o método e ao longo de sua execução, os dados são atualizados com os resultados intermediários.	38
6.3	Funcionalidade do PathoLogic [42] para reconstrução automática da rede metabólica a partir dos arquivos com elementos genéticos gerados com informações do banco de dados Plutzii.	40
6.4	Funcionalidade do PathoLogic [42] para refinar dados de vias metabólicas reconstruídas.	40
6.5	Rede metabólica reconstruída <i>in silico</i> para o fungo <i>P. lutzii</i> , incluindo vias metabólicas completas e reações não atreladas a vias.	42
6.6	Distribuição (%) das vias metabólicas identificadas, de acordo com a classificação do BioCyc.	43
6.7	Distribuição de transcritos oriundos dos <i>contigs</i> e <i>singlets</i> do <i>P. lutzii</i> , sob as formas de levedura, micélio e ambas as formas.	45
6.8	Exemplo da página com dados gerais da rede metabólica do <i>P. lutzii</i> na web.	46
6.9	Exemplo de navegação pela rede metabólica através das classificações das vias metabólicas.	46
6.10	Exemplo de exibição da via metabólica de biossíntese de betaxantin.	47
6.11	Via metabólica do ciclo do glioxilato	48
6.12	Via metabólica da degradação de mannitol I	49
6.13	Via metabólica da biossíntese de Geranilgeranil di-fosfato	50

Lista de Tabelas

2.1	Funções biológicas das proteínas	5
6.1	Tabela comparativa das informações das redes metabólicas reconstruídas <i>in silico</i> para o <i>P. lutzii</i>	41
6.2	Possíveis aplicações de vias metabólicas secundárias identificadas <i>in silico</i> para o <i>P. lutzii</i>	44
6.3	Anotações propostas para enzimas antes anotadas como preditas ou hipó- téticas	45

Capítulo 1

Introdução

O DNA pode ser transcrito sob a forma RNAs e então traduzido em proteínas encontradas nos organismos [71]. A estrutura em dupla hélice do DNA foi proposta por Watson e Crick em 1953 [83] e desde então, o DNA vem sendo estudado através de diversas técnicas, tanto de laboratório quanto em computadores.

O Projeto Genoma Humano (PGH) [37], iniciado em 1990, teve o objetivo de determinar a sequência de DNA do genoma humano, tendo sido concluído com sucesso em 2001 [37, 81]. Além disso, levou a significativos e rápidos avanços nas tecnologias utilizadas para sequenciar DNA [14]. O PGH identificou cerca de 20.000 a 25.000 genes do DNA humano e determinou a sequência dos 3 bilhões de pares de nucleotídeos que compõem os cromossomos humanos. O PGH inaugurou uma nova era nas áreas de saúde, agricultura, pecuária e indústria, entre outras. Além disso, houve transferência de tecnologia entre o setor público e o setor privado e a discussão de questões éticas e sociais [33].

Milhares de projetos de sequenciamento de genoma foram concluídos ou estão em andamento. Segundo o GOLD [29], são conhecidos 3.450 projetos genoma, 2.173 deles terminados, sendo 157 de archeas, 3.118 de bactérias e 175 de eucariotos. Neste contexto, surgiu uma nova área de pesquisa e desenvolvimento, chamada de Bioinformática, que se propõe a resolver problemas de Biologia Molecular usando ferramentas e métodos de Computação, Matemática e Estatística.

No Brasil, o primeiro projeto genoma concluído com sucesso foi o da bactéria *Xylella fastidiosa* [73], responsável por infectar cerca de 30% dos laranjais do estado de São Paulo com uma doença conhecida por *amarelinho*, que provoca imensos prejuízos financeiros aos produtores de laranja.

Em seguida, no âmbito dos projetos genoma induzidos pelo CNPq e MCT, foi criada a Rede Genoma Centro-Oeste como uma das oito unidades regionais brasileiras para o sequenciamento de genomas. O primeiro transcriptoma sequenciado pela Rede Genoma Centro-Oeste foi o do fungo *Paracoccidioides brasiliensis* (*Pb*), entre 2001 e 2003 [21, 22],

que é o patógeno humano causador da paracoccidioidomicose (PCM). Anos mais tarde, o Instituto Broad dos Estados Unidos, sequenciou os genomas dos isolados *Pb03* e *Pb18* e *Pb01*. Finalmente, Teixeira *et al* [76] descobriram que o isolado *Pb01*, é na verdade, uma nova espécie pertencente ao gênero *Paracoccidioides*, denominada *Paracoccidioides lutzii* (*P. lutzii*), cujo nome é uma homenagem ao descobridor do *P. brasiliensis* em 1908, Adolf Lutz. Os dois outros isolados, *Pb03* and *Pb18*, pertencem à espécie *Paracoccidioides brasiliensis*.

A Rede Genoma Centro-Oeste, participou apoiando a bioinformática de outros projetos como o Genoma Guaraná [27, 86] e o Genoma Anaplasma [4, 39]. Houve ainda outro projeto que integra algumas instituições de pesquisa dessa região chamado BIOFOCO, que teve como principal objetivo desenvolver ferramentas de Bioinformática em ambiente computacional distribuído para uso público [11].

1.1 Motivação

O Projeto Genoma *Pb* trouxe significativos avanços ao conhecimento da biologia molecular do *Paracoccidioides*. Os dados gerados neste projeto serviram de base para várias outros, sobretudo na UnB e UFG. Um desses projetos, foi a definição de um método para reconstrução *in silico* de redes metabólicas de fungos contemplando um *pipeline* para metabolismo primário e secundário, tendo como estudo de caso o *P. lutzii*. A reconstrução da rede metabólica *in silico* permite identificar importantes interações, bem como proporciona uma visão sistêmica dos processos bioquímicos do *P. lutzii*, especialmente para aqueles genes relacionados à interação patógeno-hospedeiro e à doença PCM, possibilitando explorá-los em busca de novos alvos moleculares visando o desenvolvimento de novas drogas anti-fúngicas.

1.2 Problema

Uma rede de metabolismo do *P. lutzii* foi reconstruída *in silico* por Alvarez [3], mas não foi feito um estudo aprofundado das vias do metabolismo secundário, que ainda não foram completamente reportadas.

1.3 Objetivos

Objetivo principal

Desenvolver um método *in silico* para identificar e organizar vias de metabolismo, com ênfase no metabolismo secundário, e reconstruir redes metabólicas em fungos.

Objetivos específicos

- Organizar os dados do fungo em um banco de dados centralizador das entradas e saídas das pipelines que compõem o método.
- Utilizar ferramentas de bioinformática para predição de metabolismo secundário;
- Realizar um estudo de caso, reconstruindo *in silico* a rede metabólica do fungo *P. lutzii*;
- Gerar uma visualização da rede metabólica reconstruída;
- Disponibilizar os resultados na web.

1.4 Descrição dos Capítulos

Este trabalho está organizado sob a forma de capítulos. O Capítulo 2 apresenta conceitos básicos de bioquímica, metabolismo, vias e redes metabólicas, necessários ao entendimento deste trabalho. O Capítulo 3 apresenta uma visão geral do organismo alvo do estudo de caso, o *Paracoccidioides lutzii*. O Capítulo 4 apresenta as ferramentas e métodos utilizados neste trabalho, a saber, ferramentas de predição de *clusters* de genes, bancos de dados relacionais, bancos de dados de proteínas, PGDBs (*Pathway/Genome Databases*), o software PathwayTools e os métodos de reconstrução de redes metabólicas empregados. O Capítulo 5 apresenta a proposta metodológica de pesquisa organizando os *pipelines* em um método definido. O Capítulo 6 apresenta o estudo de caso para o fungo *P. lutzii* e os resultados obtidos, além de uma breve discussão sobre três vias metabólicas identificadas no estudo de caso do *P. lutzii*. O Capítulo 7 traz as conclusões e perspectivas.

Capítulo 2

Metabolismo

Este Capítulo apresenta conceitos básicos de Bioquímica e de metabolismo, necessários para o entendimento deste trabalho. Na Seção 2.1 são descritos conceitos de Bioquímica relacionados ao metabolismo. Na Seção 2.2 são definidos metabolismo, vias metabólicas, metabolismo primário e secundário, este último, com ênfase na sua especial importância em fungos.

2.1 Conceitos básicos de Bioquímica

Uma reação química é a transformação de um conjunto de substâncias químicas, chamadas de reagentes ou substratos, gerando um conjunto de outras substâncias químicas, chamadas de produtos da reação [8].

A Bioquímica é o estudo da química dos processos vitais [8]. Estes processos envolvem um intercâmbio de duas classes diferentes de moléculas: moléculas grandes, tais como proteínas e ácidos nucléicos, chamadas de *macromoléculas biológicas*, e moléculas de baixa massa molecular, tais como glicose e glicerol, chamadas metabólitos [8]. Os metabólitos participam das reações químicas que compõem o metabolismo de um organismo, sendo quimicamente transformados em processos biológicos, podendo ser tanto substratos quanto produtos de reações [60].

As proteínas são polímeros de resíduos de aminoácidos. Os aminoácidos são também encontrados nos organismos, em sua forma monomérica, como transportadores de nitrogênio, neurotransmissores ou precursores de outras biomoléculas, como nucleotídeos, glicose, heme e creatina. O conjunto de aminoácidos unidos por ligações peptídicas forma a cadeia polipeptídica de uma proteína, também chamada de estrutura primária. As estruturas espaciais das proteínas costumam ser bastante complexas, tendo sido divididas em níveis organizacionais, de primária até quaternária [23]. Essa estrutura espacial é essencial para a atividade catalítica [63].

Proteínas e peptídeos participam de praticamente todos os processos que ocorrem nos seres vivos. A diversidade de suas funções é apresentada na Tabela 2.1.

Tabela 2.1: Funções biológicas das proteínas

Função	Exemplos
Catalise	DNA polimerases, lactato dehydrogenase, tripsina
Transporte	Hemoglobina
Armazenamento	Ovalbumina
Contração	Miosina, actina
Suporte estrutural e mecânico	Colágeno, elastina
Defesa	Anticorpos, fatores anticoagulantes do sangue
Transdução de sinal	Receptores, canais de íons, proteínas de vias de tradução de sinais
Controle de crescimento, diferenciação e metabolismo	Fatores de crescimento, peptídeos de hormônios
Toxinas	Venenos de cobra, toxinas de cólera

Exceto por um pequeno grupo de moléculas de RNA catalíticas, todas as enzimas são proteínas [63]. As enzimas são classificadas e catalogadas pelo IUBMB (*International Union of Biochemistry and Molecular Biology*) [38]. Para cada enzima, é atribuído um EC (*Enzyme Code Number*), um código hierárquico composto de quatro números separados por pontos, na forma $X.Y.Z.W$, onde X define o tipo de reação catalisada, Y e Z definem, respectivamente, a classe e a sub-classe da enzima e Z é um identificador sequencial da enzima na categoria indicada pelos outros três números. A enzima 1.6.99.3, por exemplo, de acordo com o IUBMB, tem o seguinte significado:

- Quanto ao tipo de reação: 1 = Oxiredutases;
- Quanto à classe da enzima: 6 = Atua sobre NADH ou NADPH;
- Quanto à subclasse da enzima: 99 = Outros receptores;
- O identificador sequencial da enzima é 3 = NADH dehydrogenase.

Algumas pequenas moléculas especiais, chamadas co-fatores, são essenciais para a ação das enzimas. Elas se ligam às enzimas de maneira não-covalente, aumentando sua atividade (ativadores) ou diminuindo sua atividade (inibidores). Para que uma reação química ocorra, algumas condições devem ser satisfeitas [60]:

- Todos os reagentes devem estar presentes na célula, em quantidades necessárias para que a reação ocorra;
- A enzima que catalisa a reação deve estar presente;
- Co-fatores que interferem na atividade enzimática, em alguns casos devem estar presentes enquanto em outros não podem estar presentes.

Sobre o comportamento dos sistemas bioquímicos, é importante saber que obedecem às leis da termodinâmica. A primeira lei da termodinâmica diz que a energia total de um sistema e de sua vizinhança é constante, ou seja, a energia não pode ser criada nem destruída, apenas assume diferentes formas. Uma forma alternativa é a energia potencial, aquela armazenada em moléculas, que pode ser liberada na ocorrência de algum processo. Todos os seres vivos, sejam unicelulares ou multicelulares, teciduais ou não, precisam de energia [8].

O processo de transformação de energia ocorre através do metabolismo, uma rede integrada de reações químicas. Basicamente há dois tipos de reações químicas que compõem o metabolismo:

- Reações catabólicas: reações que extraem energia de alguma fonte, como alimentos;
- Reações anabólicas: reações que utilizam energia para biossíntese.

Processos metabólicos são comuns a muitos organismos. Por exemplo, o conjunto de transformações químicas que converte glicose e oxigênio em dióxido de carbono e água é essencialmente idêntico em bactérias como a *Escherichia coli* (*E. coli*) e em organismos complexos como o humano. Mesmo processos que parecem muitos distintos, têm características comuns em nível bioquímico [8, 63].

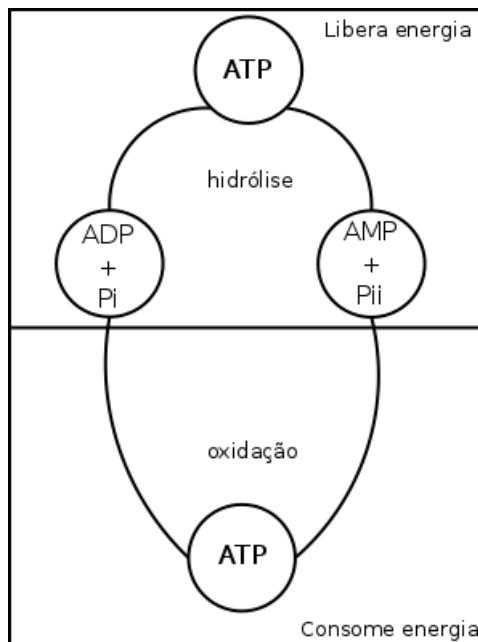


Figura 2.1: Ciclo do ATP [8]

A maior parte do catabolismo é constituída de reações que extraem energia das moléculas energéticas - como glicídeos e lipídeos - e as converte em *ATP* (*Adenosina trifosfato*).

O ATP é a moeda corrente universal de energia livre em sistemas biológicos, e é uma molécula rica em energia porque seu trifosfato contém duas ligações de anidrido fosfórico.

A energia é liberada - reação exergônica - quando o ATP sofre hidrólise para adenosina difosfato (ADP) e ortofosfato (P_i) ou para adenosina monofosfato (AMP) e pirofosfato (P_{ii}). A variação de energia livre necessária para esta reação depende da força iônica do meio e das concentrações de Mg^{2+} ou de outros íons¹ metálicos. A forma ativa do ATP é geralmente um complexo de ATP com Mg^{2+} ou Mn^{2+} [8, 63].

O ATP forma-se a partir do ADP e P_i , quando as moléculas energéticas são oxidadas nos seres quimiotróficos ou quando a luz é capturada pelos fototróficos. Este ciclo $ATP - ADP$, cujo esquema é mostrado na Figura 2.1, é fundamental na troca de energia em sistemas biológicos.

Há três estágios na extração de energia dos alimentos pelos organismos aeróbicos. No primeiro estágio, grandes moléculas são quebradas em moléculas menores. As proteínas são hidrolisadas para os seus 20 tipos de aminoácidos constituintes, os polissídeos hidrolisados a *oses*, como a glicose, os lipídios a glicerol e ácidos graxos. Neste primeiro estágio, nenhuma energia utilizável é capturada. No segundo estágio, essas moléculas pequenas são degradadas a poucas unidades simples que desempenham um papel fundamental no metabolismo. A maioria delas é transformada na unidade acetila da *acetilCoA*. Algum ATP é gerado neste estágio, mas a quantidade é inferior ao próximo estágio. No terceiro estágio, é produzido ATP pela oxidação completa da unidade acetil da *acetilCoA*. Este terceiro estágio é conhecido como ciclo do ácido cítrico, ou ciclo de Krebs e da fosforilação oxidativa, que são as vias finais comuns das moléculas alimentares [8, 63]. Os três estágios descritos são mostrados na Figura 2.2.

2.2 Conceitos básicos de metabolismo

O metabolismo é a soma de todas as transformações químicas que ocorrem em uma célula ou em um organismo através de uma série de reações catalisadas por enzimas que constituem as vias metabólicas [59, 63].

As enzimas catalisam a conversão dos substratos em produtos. A ocorrência deste processo se dá quando, inicialmente, ocorre a associação da enzima E com os substratos S , produzindo um complexo intermediário enzima-substrato: $ES : E + S \rightleftharpoons ES$. A seguir, esse complexo ES passa por um estágio de transição, que é denotado por EA , transformando-se em um complexo enzima-produto: $EP : ES \rightleftharpoons EA \rightleftharpoons EP$. Finalmente, ocorre a quebra do complexo enzima-produto EP , liberando as enzimas e os produtos resultantes da reação: $EP \rightleftharpoons E + P$, conforme exemplificado a seguir [63]:

¹íons eletricamente carregados, positiva ou negativamente.

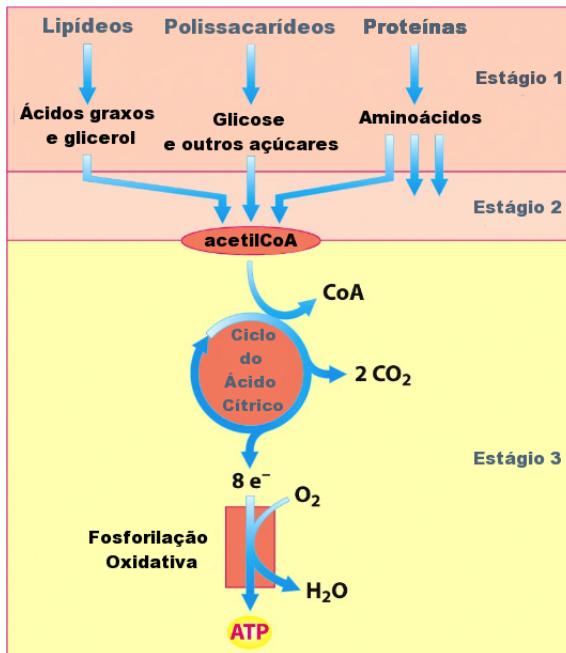


Figura 2.2: Estágios do Catabolismo [8]



Os processos metabólicos são regulados por três métodos principais: controle das quantidades de enzimas, controle da atividade catalítica e controle de acesso aos substratos [8, 59].

O controle das quantidades de enzimas ocorre através do ajuste do nível de transcrição dos genes que as codificam².

Quanto à atividade catalítica das enzimas, pode haver um controle alostérico, que consiste numa alteração na estrutura terciária ou quaternária de uma enzima protéica induzida pela ação de uma molécula ligante, que pode ser um ativador, um inibidor, um substrato, ou os três. Pode ocorrer também uma inibição por *feedback* (ou inibição retroativa), ocorre quando o produto final de uma reação de biossíntese influencia o início de outra [59, 63].

O controle de acessibilidade aos substratos pode ocorrer de duas formas: controle do fluxo dos substratos ou compartmentalização. Nos seres eucariotos, a regulação e a flexibilidade metabólicas são potencializadas pela compartmentalização. Por exemplo, a degradação por oxidação de ácidos graxos ocorre nas mitocôndrias, enquanto a síntese desses mesmos ácidos ocorre no citoplasma. Essa compartmentalização segregá reações opostas [8, 59, 63].

²Por exemplo, na *E. coli*, a presença de lactose induz dentro de minutos um aumento de cerca de 50 vezes a velocidade de síntese da galactosidase β , uma enzima necessária à degradação deste dióisídeo [8, 59]

2.2.1 Vias metabólicas

O metabolismo de um organismo é formado por vias metabólicas, que compõem a sua rede metabólica. Vias metabólicas são séries de reações químicas sobre compostos que ocorrem nas células dos organismos vivos. Essas reações podem ser catalisadas por enzimas que muitas vezes precisam ainda de minerais, vitaminas e outros co-fatores [59]. Os produtos de algumas reações são utilizados como reagentes de outras reações, esse encadeamento forma vias, que por sua vez, formam redes de processamento de compostos químicos em uma célula [3]. As vias metabólicas são geralmente mostradas graficamente. A figura 2.3, é uma representação gráfica da via da gluconeogênese encontrada no fungo *Candida albicans SC5314*.

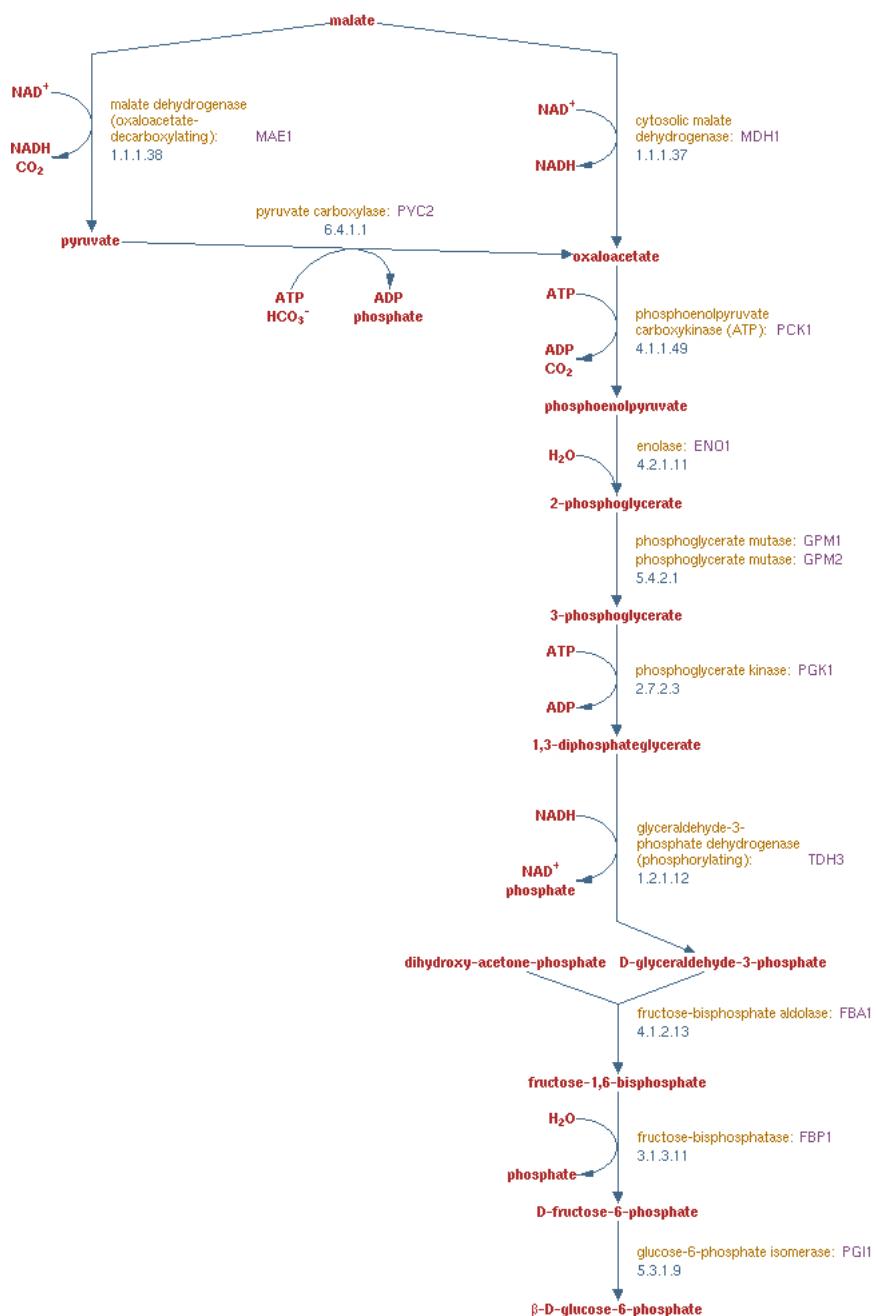


Figura 2.3: Via da gluconeogênese encontrada no fungo *Candida albicans* SC5314 [58].

2.2.2 Metabolismo primário

O metabolismo pode classificado em primário e secundário. O metabolismo primário é o conjunto de processos metabólicos que desempenham funções essenciais ao organismo, como a divisão e crescimento celular, respiração, estocagem de energia e reprodução [59, 63]. Muitos sistemas enzimáticos cooperam para obter energia química capturando energia solar ou degradando nutrientes ricos em energia obtidos do ambiente, ou ainda convertendo moléculas dos nutrientes obtidos em outras moléculas, próprias de cada célula [63].

2.2.3 Metabolismo secundário

Metabolismo secundário é aquele relativo aos compostos secundários, que diferentemente dos primários, não são essenciais para o crescimento, desenvolvimento ou reprodução. Apesar dos compostos secundários fazerem parte do metabolismo normal de um organismo, são frequentemente produzidos em células especializadas, e tendem a ser mais complexos do que os compostos primários [49, 63].

Os metabólitos secundários são particularmente abundantes em fungos filamentosos que vivem no solo [49]. Muitos fungos produzem metabólitos secundários por influência de competição ambiental e nestes casos os compostos são produzidos juntos com enzimas para digestão extracelular. Eles podem oferecer proteção contra estresses ambientais ou interações com outros organismos [49], como pode ser visto na Figura 2.4.

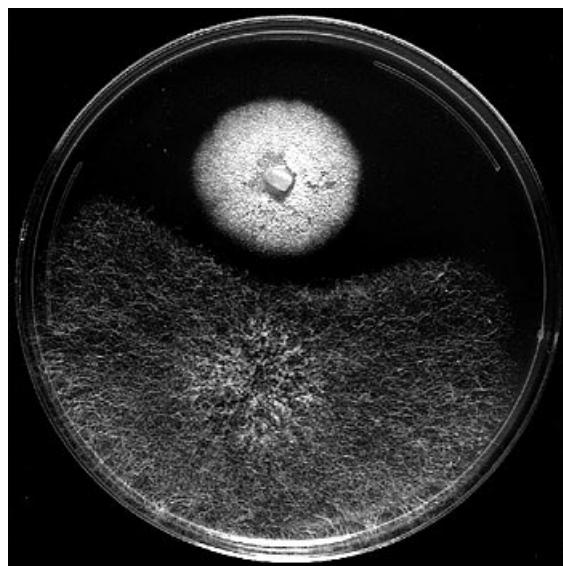


Figura 2.4: Expressão de metabólitos secundários em ágar como forma de inibição de colônias adjacentes [66].

Metabólitos secundários são geralmente produzidos após um crescimento ativo e alguns metabólitos, frequentemente são encontrados em fungos relacionados, enquanto outros são encontrados apenas em uma ou algumas espécies [66].

Existem razões claras para estudar metabólitos secundários, pois há aplicações na indústria e na medicina. Seis dos vinte medicamentos mais comumente prescritos para os seres humanos são de origem fúngica [66]. A importância dos metabólitos secundários de fungos cresceu a partir da década de 1940 com o enorme impacto da penicilina sobre a saúde humana. Desde então, diversos outros compostos foram descobertos, como agentes imunossupressores, drogas para baixar o colesterol, drogas antivirais e anti-tumorais [49, 57]. A disponibilidade de dados de genomas de fungos têm facilitado a descoberta e caracterização de novos componentes e suas vias metabólicas [49].

Metabólitos secundários importantes são sintetizados a partir de uma ou da combinação de três das vias biossintéticas [34] de:

- Policitídeos decorrentes de Acetyl CoA;
- Via do mevalonato, também decorrente de Acetyl CoA;
- Aminoácidos.

Além disso, os genes para a síntese de alguns metabólitos secundários encontram-se agrupados, e a expressão do *cluster* parece ser induzida por um ou alguns reguladores globais [34, 66].

Classes de metabólitos secundários

O metabolismo secundário está relacionado, em grande parte, ao domínio da química orgânica. Metabólitos secundários são muitas vezes bioativos, geralmente de baixa massa molecular, e são produzidos como famílias de compostos relacionados a áreas restritas do ciclo de vida, com a produção, muitas vezes correlacionada com um estágio específico de diferenciação morfológica [47]. Existem várias classes de metabólitos secundários, conforme mostrados na Figura 2.5.

- **Policitídeos (PK):** são os mais abundantes metabólitos secundários de fungos. Eles são sintetizados por policitídeo-sintases (PKS), que são proteínas de vários domínios que estão relacionadas com as sintases de ácidos graxos em eucariotos e contêm estruturas de domínio semelhante [47].;
- **Peptídeos não-ribossomais (NRP):** são derivados, ao mesmo tempo de aminoácidos proteinogênicos (aqueles que constituem as proteínas) ou aminoácidos não-

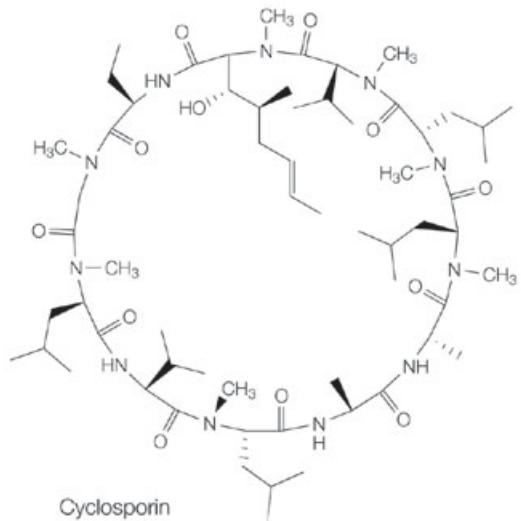
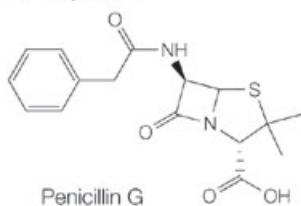
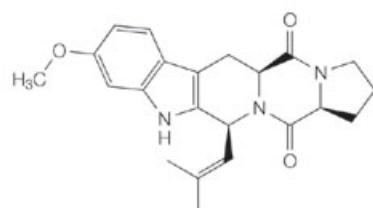
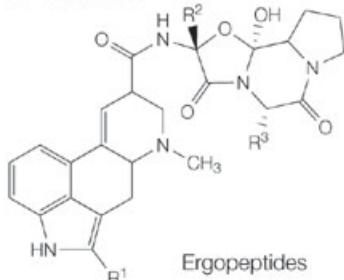
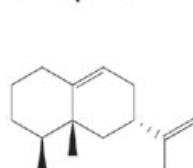
proteinogênicos multidomínio³, e enzimas multimodulares chamadas não-ribossomais peptídeo sintases(NRPs). Cada módulo em um NRPS contém vários domínios que permitem o reconhecimento, a ativação e a ligação covalente de um aminoácido específico do módulo. Subsequentemente, as ligações peptídicas são formados entre os aminoácidos ligados. O peptídeo resultante é então liberado por um domínio do tipo tioesterase⁴ geralmente localizada na extremidade *C* do módulo final. A diversidade entre os peptídeos não ribossomais surge do comprimento de peptídeos produzidos [47];

- **Alcalóides:** Constituem um grupo heterogêneo de substâncias nitrogenadas, geralmente de origem vegetal, de caráter básico e que apresentam acentuada ação farmacológica em animais. São encontrados predominantemente na forma combinada, com ácidos orgânicos, e em concentração menor, na forma livre [63];
- **Terpenos:** Podem ser definidos como um grupo de moléculas cuja estrutura está baseada em um número definido de unidades de isoprénicas⁵. Os terpenos podem ser usados como substratos para obtenção de energia por respiração celular, ou para criar outras substâncias como aminoácidos e novos tecidos [9].

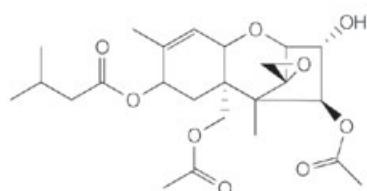
³Regiões com estruturas terciárias semi-independentes ligadas entre si por segmentos lineares da cadeia polipeptídica, ou seja, unidades funcionais e de estrutura tridimensional de uma proteína.

⁴As hidrolases incluem enzimas de baixa especificidade, como esterases e tioesterases, que hidrolisam um número muito grande de ésteres e tioésteres, embora com velocidades diferentes.

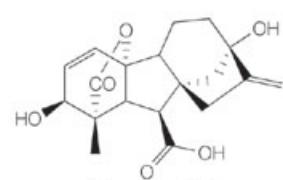
⁵metil-but-1,3-dieno, com 5 átomos de carbono

a Peptides**b Alkaloids****c Terpenes**

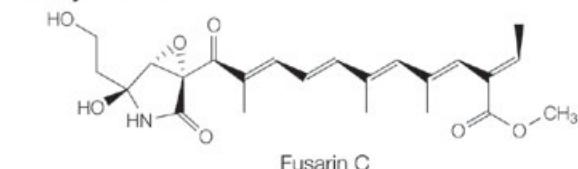
Aristolochene



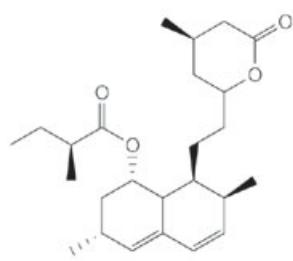
Trichothecene T2 toxin



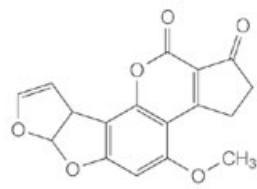
Gibberellin GA3

d Polyketides

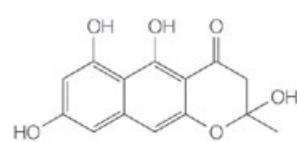
Fusarin C



Lovastatin



Aflatoxin B1



WA

Copyright © 2005 Nature Publishing Group
Nature Reviews | Microbiology

Figura 2.5: Os principais grupos de metabólitos secundários de fungos. [47].

Capítulo 3

O fungo *Paracoccidioides lutzii*

Este capítulo apresenta uma visão geral sobre o *Paracoccidioides lutzii*. Na Seção 3.1 é descrito o *P. lutzii*, o sequenciamento de seu genoma funcional e diferencial pela Rede Genoma Centro-Oeste e o sequenciamento do genoma dos isolados *Pb03*, *Pb18*, *Pb01* pelo BROAD Institute. A Seção 3.2 Apresenta uma visão geral sobre a paracoccidioidomíose. Finalmente, a Seção 3.3 apresenta a rede metabólica previamente reconstruída por Alvarez [3].

3.1 *Paracoccidioides brasiliensis* e *Paracoccidioides lutzii*

O *P. brasiliensis* é um fungo que vem sendo estudado desde 1908, quando foi identificado por Adolfo Lutz, a partir de amostras colhidas da mucosa bucal em pacientes da cidade de São Paulo [52]. Mais tarde, em 1912, Splendore estabeleceu as condições para seu cultivo *in vitro*, tendo classificado-o como *Zymonema brasiliensis* [75]. Alguns anos depois, em 1930, Almeida propôs o gênero *Paracoccidioides* em substituição à classificação anterior de Splendore, como citado em [5]. Em meados de 2000, Bailek [7] demonstrou a proximidade filogenética entre o *P. brasiliensis*, com os fungos dos gêneros *Histoplasma*, *Emmonsia* e *Blastomyces*, utilizando análise ribossomal parcial do RNA ribossomal 26S [7].

O Projeto Genoma Funcional e Diferencial do *Paracoccidioides*, iniciado em 2001 pela Rede Genoma Centro-Oeste, propunha-se a identificar os genes de expressão de micélio e levedura e os genes de expressão diferencial, que potencialmente exercem funções relacionadas à adaptação do fungo ao hospedeiro, à manutenção do estado diferenciado, bem como com à patogenicidade deste fungo [28]. A metodologia proposta utilizou técnicas que possibilitaram o sequenciamento de ESTs das formas de micélio e de levedura, e a

identificação de cDNAs (ESTs) diferenciais estágio-específicas, de uma das duas formas (micélio ou levedura) do *P. brasiliensis*; identificação de ESTs diferenciais por DDRT-PCR; e identificação de ESTs diferenciais por microarranjos [22, 28, 65].

O sequenciamento do transcriptoma do isolado *Pb01*, até então conhecido como *P. brasiliensis*, foi realizado com a técnica *Sanger* e foram gerados 19.718 ESTs (*Expressed Sequence Tags*) de alta qualidade. Após a montagem, foram obtidos 2.655 *contigs* e 3.367 *singlets*, constituindo um total 6.022 ESTs montadas. Esses dados sugeriram um tamanho de genoma estimado em aproximadamente 30 Mb, contendo entre 7.500 até 9.000 genes [21, 22]. O projeto foi finalizado em 2003 [21].

A identificação de genes de expressão diferencial, requeridos para a transição dimórfica, permitiu aprofundar o conhecimento de possíveis causas de estabelecimento de infecção e da doença PCM. Vários genes que participam do processo de dimorfismo e de virulência já foram descritos também para outros fungos patogênicos como o *Histoplasma capsulatum* e o *Candida albicans*.

Vários genes potencialmente envolvidos no dimorfismo ou na patogenicidade foram descritos para o *P. lutzii* [22, 28]. Em 2003, durante um período entre 5 e 120 horas, a expressão gênica do *P. brasiliensis* foi monitorada durante a transição micélio-levedura. Esse experimento revelou mais de 2 mil genes que apresentaram modulações estatisticamente significativas [65].

O Instituto Broad nos Estados Unidos sequenciou os genomas dos isolados *Pb03* e *Pb18* e do isolado *Pb01*. O tamanho obtido para estes genomas foi de, respectivamente 29,1 Mb em 7.610 genes, 30 Mb em 8.042 genes e 32,9 Mb em 8.130 genes [17, 36].

Teixeira *et al* [76] revelaram que o isolado *Pb01*, é de fato, uma nova espécie pertencente ao gênero *Paracoccidioides*, denominada *Paracoccidioides lutzii* em homenagem ao seu descobridor. Os dois outros isolados, *Pb03* and *Pb18*, permaneceram pertencentes à espécie *Paracoccidioides brasiliensis*. O *P. lutzii* ocorre predominantemente, mas não exclusivamente, na Região Centro-Oeste do Brasil, conforme pode ser observado na Figura 3.1. A Figura 3.2 mostra a filogenia e ortologia entre alguns fungos dimórficos.

O *P. lutzii*, (Figura 3.3) pertence ao Reino Fungi; Divisão *Ascomycota*; Subdivisão *Euscomycetes*; Classe *Plectomyceto*; Ordem *Onygenales*; Família *Onygenacea*; Subfamília *Onygenaceae anamorficos*; Gênero *Paracoccidioides*; Espécie *Paracoccidioides lutzii*. É um fungo dimórfico, ou seja, pode ocorrer sob duas formas. É encontrado na natureza sob a forma miceliana à temperatura de aproximadamente 25°C, que sofre uma transição da forma miceliana (24-26°C) para a forma levedura (37°C). Esse processo pode ser controlado *in vitro* por alteração da temperatura de incubação [5].

O *P. lutzii* vive saprofiticamente no solo sob a forma miceliana, que apresenta uma organização pluricelular formada por filamentos de células cilíndricas, com ramificações

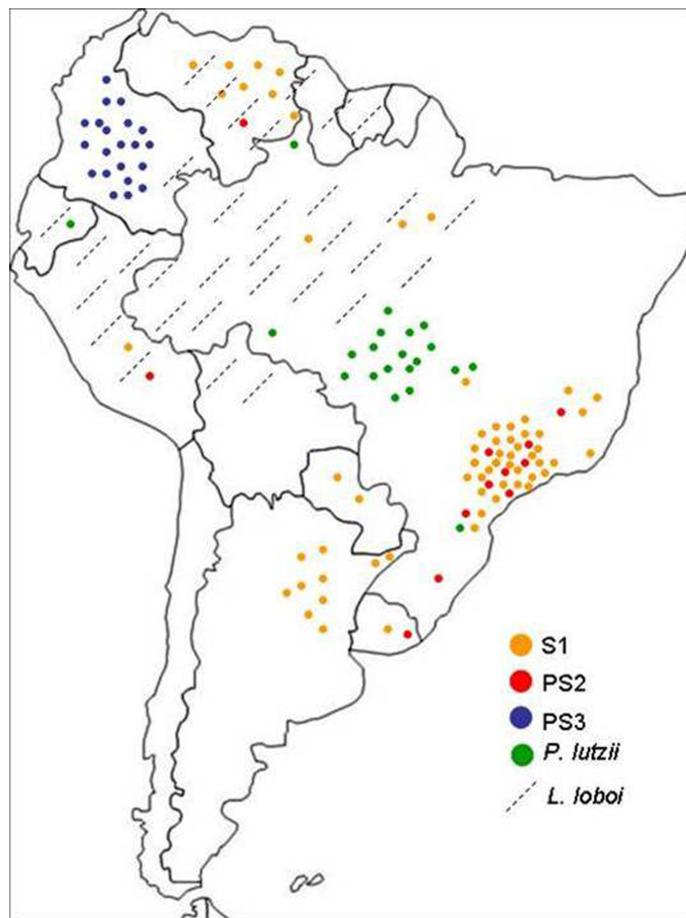


Figura 3.1: Distribuição geográfica do gênero *Paracoccidioides* e *Lacazia loboi* [77].

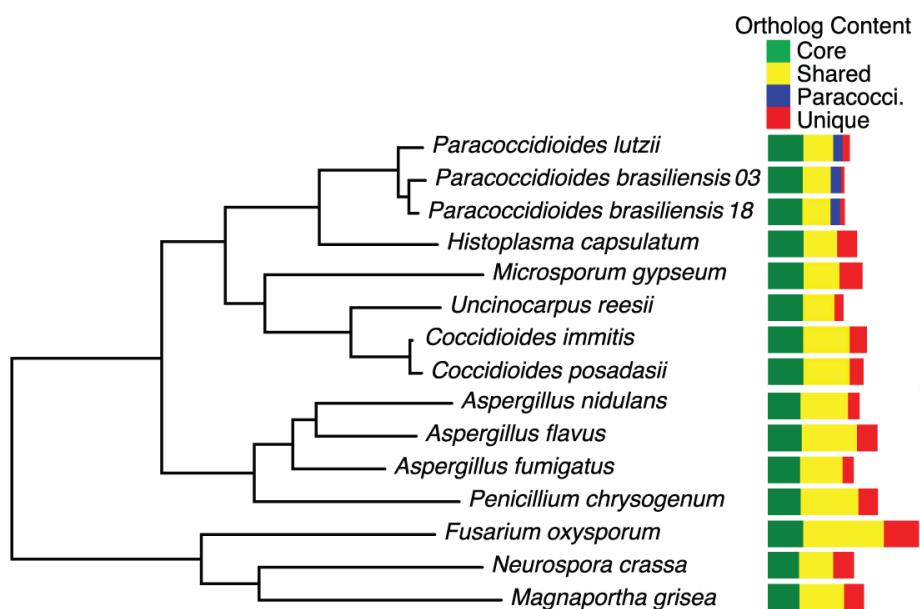


Figura 3.2: Árvore filogenética de fungos dimórficos, obtida usando o critério *maximum likelihood* RaxML [17].

denominadas hifas. As hifas são multinucleadas, finas e septadas, podendo ter esporos terminais ou intercalares. As células de levedura são multinucleadas, arredondadas e algumas exibem múltiplos brotamentos, que conferem ao fungo uma estrutura típica, usada inclusive para diagnóstico, denominada roda de leme [5].

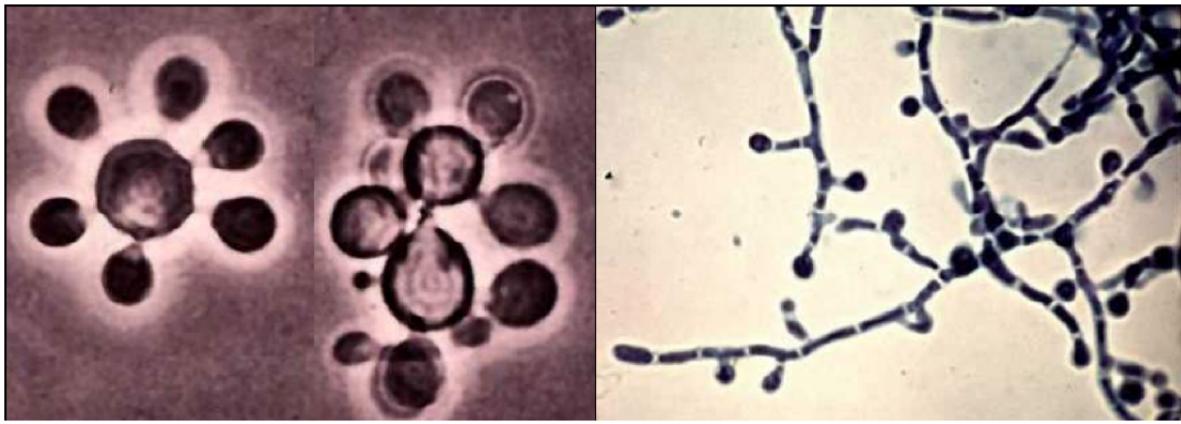


Figura 3.3: As duas formas do *Paracoccidioides lutzii*: forma leveduriforme característica com brotamentos múltiplos em forma de roda de leme [25] à esquerda e forma de micélio à direita.

3.2 A paracoccidioidomicose (PCM)

A paracoccidioidomicose (PCM), doença causada pelo *P. lutzii*, é uma micose sistêmica, autóctone da América Latina, onde cerca de 10 milhões de pessoas estão infectadas, com maior ocorrência na Argentina, Colômbia, Venezuela e no Brasil, o qual contribui com cerca de 80% dos casos, afetando principalmente indivíduos em áreas rurais [62].

Uma vez que a PCM não é doença de notificação compulsória, não há dados precisos sobre sua incidência no país [62]. O conhecimento de áreas endêmicas e dos dados de prevalência, incidência e morbidade baseia-se em relatos de casos clínicos e de inquéritos intradérmicos. A maioria dos casos tem sido reportada nas regiões Sul, Sudeste e Centro-Oeste. A PCM apresenta caráter endêmico [5] entre as populações de zona rural. Sua ocorrência também tem sido relatada em áreas de colonização mais recente, submetidas a desmatamento, como em partes da Amazônia, atingindo áreas dos estados do Maranhão, Tocantins, Pará, Mato Grosso, Rondônia, Acre e Amazonas, onde a PCM, pela falta de informações de ocorrência prévia, tem sido considerada uma micose sistêmica emergente [62]. A infecção da PCM não privilegia etnia, já a PCM doença, afeta principalmente trabalhadores rurais adultos, do sexo masculino, em contato constante com a vegetação e solo, no período mais produtivo de suas vidas, o que implica em importante repercussão econômica para pacientes e dependentes [25].

A PCM representa um problema de Saúde Pública, pelo seu alto potencial incapacitante e pela quantidade de mortes prematuras que provoca, quando não diagnosticada e tratada oportunamente. A doença também é conhecida como *Blastomicose sul-americana* ou *moléstia de Lutz-Splendore e Almeida* [62].

O período de incubação é desconhecido, podendo levar anos até que se manifeste a doença. Não há transmissão homem a homem e não há relatos de contágio de animais para o homem. O controle da infecção depende da resposta imune celular associada ao padrão *tipo 1* da resposta imunológica, resultando na formação de granulomas compactos [62].

A infecção humana ocorre por inalação de conídios e fragmentos micelianos do *P. brasiliensis*, que alcançam os bronquíolos terminais e alvéolos pulmonares, onde se transformam em células leveduriformes, produzindo a infecção que pode disseminar-se para os demais tecidos por via linfática e hematogênica [25]. As manifestações clínicas variam em diferentes indivíduos, dependendo tanto de aspectos específicos do parasita como do hospedeiro [5]. A doença pode se apresentar sob duas formas básicas: aguda e crônica. A primeira acomete ambos os sexos, geralmente pacientes jovens, e origina-se a partir de uma lesão primária não detectada que progride rapidamente com disseminação para o baço, fígado, linfonodos e medula óssea. A forma crônica tem sua origem em um foco primário ou em uma lesão antiga, contendo formas viáveis de *P. brasiliensis*, podendo estar restrita a um órgão, sistema ou podendo ainda atingir múltiplos órgãos, como intestinos, ossos e o sistema nervoso central.

3.3 Rede metabólica do *P. lutzii*: uma primeira versão

Alvarez [3] reconstruiu *in silico* uma primeira versão em larga escala da rede metabólica do fungo *Paracoccidioides lutzii* e disponibilizou uma página (Figuras 3.4 e 3.5) com diversas informações sobre o fungo.

Foram identificadas vias metabólicas do *P. lutzii* usando os bancos de dados do KEGG [46] dos organismos *Saccharomyces cerevisiae*, *Coccidioides immitis*, *Aspergillus fumigatus*, *Aspergillus nidulans* e *Escherichia coli*, e o banco de dados Biocyc [10] com os organismos *S. cerevisiae* e *E. coli*. A rede reconstruída identificou 295 vias metabólicas, 1.743 reações catalisadas por 1.413 enzimas, além de 9 reações de transporte, as reações atuam sobre um total de 1.202 compostos. Foi construída uma visualização da rede com o PathwayTools [43]. Entre as perspectivas do trabalho de Alvarez [3], estava o estudo mais aprofundado do metabolismo secundário.



The Pb01 Metabolic Network Project



About *Paracoccidioides brasiliensis*

Data sets

Metabolic pathways

Methodology

Team

Metabolic network picture

Paracoccidioides brasiliensis is the etiological agent of paracoccidioidomycosis (PCM) a human systemic mycosis that mainly affects rural communities in South and Central America. PCM is manifested in different clinical forms, which are tightly dependent on the immune competence of the host. In individuals displaying the localized form of the disease it is observed the formation of granulomas and the presence of competent T lymphocytes; on the other hand, in the dispersed form of the disease, the granulomas are not formed and the T lymphocytes are hyporeactive. It is believed that the cellular immune response is important in the control of the disease and it has recently been demonstrated that in mice a peptide (P10) from the gp43 protein (the main antigen used for immunodiagnose) is capable of triggering a protective cellular immune response.

The dimorphic fungus *P. brasiliensis* is found as a mycelium (M) or as a spore when grown *in vitro* at room temperature or in its natural habitat. When it is shifted to temperatures close to 37°C it is found as a yeast, the cell form predominantly found in tissues of patients with PCM. Infection occurs mainly when spores or even fragments of mycelium are inhaled. Upon host infection, *P. brasiliensis* undergoes a dimorphic transition from the mycelium form (or spore) to the yeast form probably triggered by the temperature shift. These observations strongly suggest that the dimorphic transition is an important event in the establishment of the infection as it has also been observed in other pathogenic fungi such as *Candida albicans* and *Histoplasma capsulatum*.

Figura 3.4: Página com os dados da rede metabólica reconstruída por Alvarez [3].



The Pb01 Metabolic Network Project



- ▶ About *Paracoccidioides brasiliensis*
- ▶ Data sets
- ▶ Metabolic pathways
- ▶ Methodology
- ▶ Team
- ▶ Metabolic network picture

Metabolic pathways with enzyme codes of Pb01

Here are links for the various metabolic pathways of *Saccharomyces cerevisiae* in which enzyme codes of Pb01 were located. Enzyme codes of *S. cerevisiae* are represented as boxes with a green background, whilst enzyme codes common to both *S. cerevisiae* and Pb01 are represented as green boxes with a red border.

- * [sce01100 Metabolic pathways - *Saccharomyces cerevisiae* \(budding yeast\) \(570\)](#)
- * [sce01110 Biosynthesis of secondary metabolites - *Saccharomyces cerevisiae* \(budding yeast\) \(237\)](#)
- * [sce00190 Oxidative phosphorylation - *Saccharomyces cerevisiae* \(budding yeast\) \(73\)](#)
- * [sce00230 Purine metabolism - *Saccharomyces cerevisiae* \(budding yeast\) \(66\)](#)
- * [sce00010 Glycolysis / Gluconeogenesis - *Saccharomyces cerevisiae* \(budding yeast\) \(52\)](#)
- * [sce00240 Pyrimidine metabolism - *Saccharomyces cerevisiae* \(budding yeast\) \(41\)](#)
- * [sce00500 Starch and sucrose metabolism - *Saccharomyces cerevisiae* \(budding yeast\) \(40\)](#)
- * [sce00970 Aminoacyl-tRNA biosynthesis - *Saccharomyces cerevisiae* \(budding yeast\) \(37\)](#)
- * [sce00020 Citrate cycle \(TCA cycle\) - *Saccharomyces cerevisiae* \(budding yeast\) \(33\)](#)
- * [sce00620 Pyruvate metabolism - *Saccharomyces cerevisiae* \(budding yeast\) \(32\)](#)
- * [sce00250 Alanine, aspartate and glutamate metabolism - *Saccharomyces cerevisiae* \(budding yeast\) \(31\)](#)
- * [sce00330 Arginine and proline metabolism - *Saccharomyces cerevisiae* \(budding yeast\) \(29\)](#)

Figura 3.5: Página com *links* para as vias metabólicas reconstruída por Alvarez [3].

Capítulo 4

Ferramentas computacionais

Este capítulo apresenta o conceito de banco de dados relacional, as ferramentas para anotação de proteínas, predição de genes, bancos de dados biológicos e métodos de reconstrução de redes metabólicas. Na Seção 4.1 são apresentados conceitos de bancos de dados e também bancos de dados de sequências biológicas e de vias metabólicas e o modelo de banco de dados conhecido como *Pathway/Genome Database* (PGDB). Na Seção 4.2, são apresentadas as ferramentas utilizadas para anotação de ECs de enzimas e anotação de genes. Na Seção 4.3, são apresentadas as ferramentas utilizadas para predição de metabolismo secundário. Finalmente, na Seção 4.4 são apresentados os métodos e ferramentas utilizados neste trabalho para reconstrução e geração da visualização da rede metabólica reconstruída.

4.1 Bancos de dados

Nesta seção, é descrito o conceito de banco de dados e banco de dados relacional. Também são descritos os bancos de dados de metabolismo utilizados neste trabalho. Um banco de dados, segundo Date [16], é uma coleção de dados persistentes usado pelos sistemas das organizações. Persistência é a capacidade do dado, uma vez inserido num Sistema Gerenciador de Bancos de Dados (SGBD), permanecer armazenado até que seja explicitamente removido, exceto por algum erro computacional [16].

4.1.1 Bancos de dados de relacionais

Um banco de dados relacional tem seu alicerce numa teoria chamada modelo relacional de dados. Três aspectos descrevem este modelo [16]:

1. Estrutural: os bancos de dados são percebidos pelo usuário como tabelas;
2. Integridade: essas tabelas satisfazem a restrições de integridade;

3. Manipulável: existem operadores disponíveis para manipulação das atebals por parte do usuário.

Uma entidade é um objeto distingível que possa ser representado no banco de dados. As ligações entre essas entidades são os relacionamentos. Um arepresentação gráfica dessas entidades e seus relacionamentos é chamada de Modelo de Entidade-Relacionamento (MER) [16].

Um sistema gerenciador de bancos de dados (SGBD) é um software que trata de todo o acesso ao banco de dados. Este software deve ser capaz de aceitar definições de dados (esquemas e seus mapeamentos associados), manipulação de dados (inserção, busca, atualização e exclusão), otimização e desempenho, segurança e integridade, concorrência sobre os dados através de transações, além de fornecer um dicionário de dados [16].

4.1.2 UniProt

O UniProtKB (*UniProt Knowledgebase*) é um repositório centralizador de informações sobre proteínas, que fornece uma visão unificada de sequências protéicas e informações funcionais. O UniProtKB é composto de duas seções: UniProtKB/Swiss-Prot e UniProtKB/-TrEMBL. UniProtKB/Swiss-Prot é curado manualmente, o que significa que as informações de cada registro são analisadas e mantidas por um curador, enquanto os registros do UniProtKB/TrEMBL são gerados automaticamente e são enriquecidos com anotação automática e classificação [55]. Está disponível na web no endereço <http://www.uniprot.org>.

4.1.3 BRENDA

BRENDA (*BRaunschweig ENzyme Database*) é um sistema de informações de enzimas. A maior parte dos dados são manualmente extraídos da literatura. A cobertura de dados abrange função, estrutura, ocorrência, preparo, aplicação de enzimas, bem como propriedades de mutações e variações [70]. Está disponível na web no endereço <http://www.brenda-enzymes.org>.

4.1.4 BioCyc

O BioCyc [13] é um agregador de PGDBs (*Pathway/Genome Databases*). Há versões de PGDBs agregados que variam quanto à quantidade. Cada versão destas é chamada de Tier. O Tier1, por exemplo agrupa os PGDBs EcoCyc, MetaCyc, HumanCyc, AraCyc, YeastCyc, LeishCyc e TrypanoCyc. Há ainda o Tier 2, com maior número de PGDBs,

mas menor nível de cura¹. O Tier 3 com uma grande coleção de PGDBs gerados computacionalmente, mas não curado.

As ferramentas do BioCyc permitem navegação, visualização e análise dos dados, que incluem:

- Browser genômico;
- Visualização individual de vias metabólicas, com nível ajustável de detalhes;
- Ferramenta para análise comparativa;
- Mapeamento com *links* entre cada componente das vias metabólicas.

O MetaCyc é um banco de dados altamente curado e não redundante de metabolismo. Por este motivo ele foi escolhido para este trabalho. O MetaCyc contém vias metabólicas com dados de enzimas experimentalmente demonstrados pela literatura científica. As vias metabólicas e demais dados do MetaCyc são derivados dos demais bancos de dados do BioCyc [13]. Como o MetaCyc contém apenas dados com evidências experimentais, determinadas vias metabólicas e enzimas podem não estar disponíveis. O MetaCyc está disponível na web no endereço <http://metacyc.org>.

4.1.5 KEGG

KEGG (*Kyoto Encyclopedia of Genes and Genomes*) é uma base de conhecimento para análise sistemática de função genômica. A informação genômica é armazenada num banco de dados de genes, que é uma coleção de genes para genomas completamente sequenciados e alguns genomas parciais, com constante atualização de anotação de funções de genes [40, 41]. O acesso ao KEGG e suas ferramentas é possível através da página <http://www.genome.jp/kegg>. A Figura 4.1 foi extraída do KEGG e mostra a via metabólica de metabolismo da pirimidina no *P. lutzii*.

4.2 Anotação

A anotação é o procedimento por meio do qual uma função biológica é atribuída à sequência de resíduos de aminoácidos, no caso das proteínas, ou à sequência de nucleotídeos, no caso dos genes. Quanto ao formato dos dados, um dos mais conhecidos é o FASTA, onde os dados são apresentados em formato textual para representar tanto sequências de nucleotídeos quanto sequências de aminoácidos. Num arquivo FASTA, os

¹O termo "cura" é largamente utilizado na literatura especializada para designar uma verificação, assim algo curado é algo que passou por uma checagem do especialista

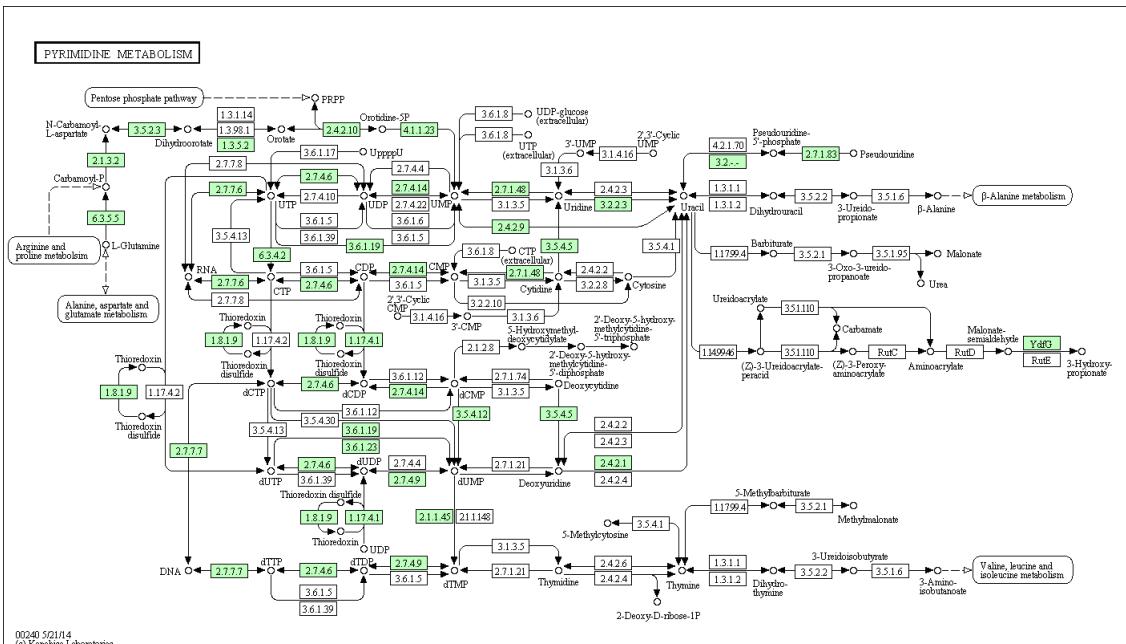


Figura 4.1: Metabolismo da pirimidina no *P. lutzii* [46].

nucleotídeos ou aminoácidos são representados usando códigos de uma única letra conforme mostrado na Figura 4.2.

Figura 4.2: Exemplo de arquivo no formatoFASTA, representando aminoácidos.

4.2.1 KAAS

KAAS (*KEGG Automatic Annotation Server*) é uma implementação web para anotação automática de genes. O KAAS é implementado sobre dois métodos: baseado em similaridade de sequências com *best hit* bidirecional e *best hit* unidirecional. Para um genoma completo, o método de *best hit* bidirecional obtém mais acurácia. Por outro lado, para um número limitado de ORFs ou ESTs, é recomendado o uso de *best hit* unidirecional. O método alcançou um elevado grau de precisão quando comparado com o

banco de dados KEGG curado manualmente [61]. Está disponível na web no endereço <http://www.genome.jp/kegg/kaas>.

4.2.2 CDD

O CDD (*Conserved Domain Database*) é utilizado para anotação de sequências de proteínas através da identificação de domínios conservados e sítios funcionais. É curado manualmente e utiliza modelos de estrutura tridimensional de proteínas para refinar os modelos de domínio e prover informação sobre o relacionamento sequência/estrutura/função [56]. Está disponível na web sob o endereço <http://www.ncbi.nlm.nih.gov/cdd>.

4.2.3 InterPro

O InterPro é um banco de dados que integra diversas informações sobre famílias de proteínas, domínios e sítios funcionais. Tem uma interface web amigável e *web service*. O Interpro foi construído a partir de uma análise em larga escala de genomas e metagenomas, além de caracterização individual de proteínas. O InterPro cobre cerca de 95% do UniProtKB/Swiss-Prot e cerca de 79% do UniProtKB/TrEMBL [35]. Está disponível na web no endereço <http://www.ebi.ac.uk/interpro>.

4.3 Metabolismo secundário

Nesta seção, são apresentadas as ferramentas para predição de metabolismo secundário em fungos. Os genes para a síntese de alguns metabólitos secundários encontram-se agrupados, e a expressão do *cluster* parece ser induzida por um ou alguns reguladores globais [34, 66]. Estas ferramentas fazem as suas predições de acordo com este critério.

4.3.1 antiSMASH

antiSMASH (*antibiotic and Secondary Metabolite Analysis Shell*) é um *pipeline* capaz de indentificar *locus* de biosíntese, cobrindo todo o genoma em busca de *clusters* de genes de metabolismo secundário das classes conhecidas (Policetídeos, Peptídeos não-ribossomais, Alcalóides e Terpenos). O antiSMASH alinha e identifica diferentes regiões do genoma em busca de *clusters* de genes. Como resultado, obtém genes de metabolismo secundário com *links* entre os componentes das vias metabólicas, numa vizualização interativa [57]. O esquema do *pipeline* pode ser observado na Figura 4.3. O antiSMASH está disponível na web no endereço <http://antismash.secondarymetabolites.org>.

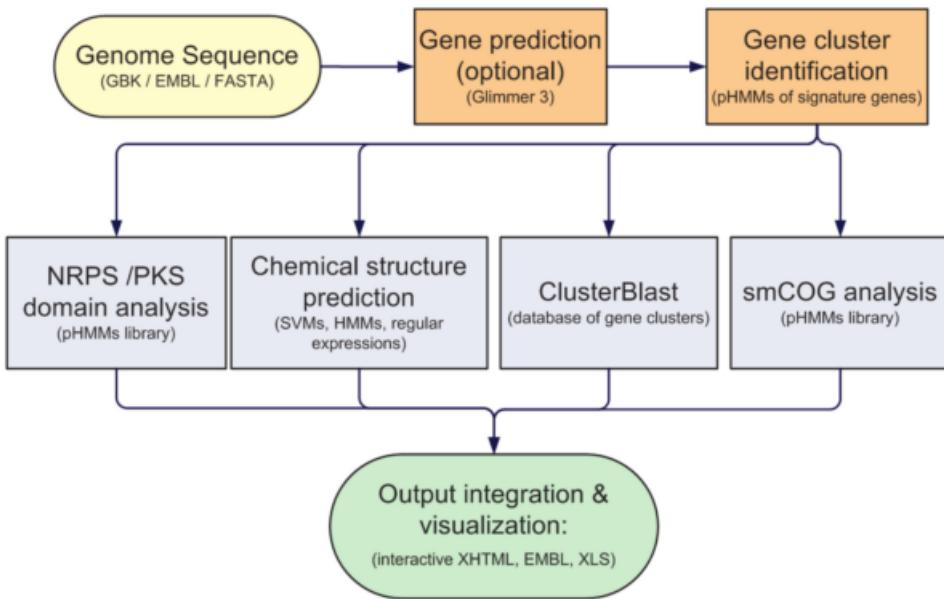


Figura 4.3: Pipeline da análise genômica de metabólitos secundários realizada pelo antiSMASH [57].

4.3.2 SMURF

SMURF (*Secondary Metabolite Unknown Regions Finder*) é um software disponível via web para o mapeamento sistemáticos de *clusters* de genes de metabolismo secundário. É baseado em três características de vias metabólicas secundárias de biossíntese: *backbone genes*, *clusters* e domínios conservados de proteínas. O SMURF baseia-se no Modelo Oculto de Markov (Hidden Markov Model - HMM) para detectar genes de *backbone* em sequências de genomas fúngicos [49].

O SMURF utiliza o mesmo algoritmo que o antiSMASH para identificar *backbone genes*, mas utiliza uma estratégia diferente para predizer os limites do *cluster*. Ambos, antiSMASH e SMURF, predizem genes que codificam *backbone* de enzimas com precisão próxima de 100%. O SMURF utiliza um conjunto de 27 domínios específicos de metabolismo secundário para tentar realizar uma previsão exata dos limites do *cluster* de genes. antiSMASH, por outro lado, utiliza uma margem de segurança entre 5 e 20kb ao redor do *backbone* de genes [19]. Genes além do *cluster* são difíceis de serem preditos, pois alguns deles contêm domínios desconhecidos, o que implica em previsões ruins [19]. O SMURF está disponível na web no endereço <http://www.jcvi.org/smurf>. É necessário um registro de usuário com endereço de e-mail.

4.4 Reconstrução de redes metabólicas

Existem diversos métodos para o processo de reconstrução de redes metabólicas *in silico*. A diversidade de métodos para modelar as vias e redes tem se expandido significativamente ao longo das últimas duas décadas. As abordagens modernas e tradicionais são igualmente importantes e recentes pesquisas têm possibilitado integrar as vantagens de ambas [68].

Uma destas abordagens que integra técnicas modernas e tradicionais é a Reconstrução de Redes Bioquímicas em Microorganismos, de Feist *et al* [20] que está organizado em etapas. A primeira etapa é a obtenção do genoma e a identificação de enzimas associadas às funções celulares através de informações de bancos de dados e literatura. Com essas informações é construída uma relação gene-proteína-reação (GPR). Ferramentas automatizadas de reconstrução, como o PathwayTools, estão disponíveis para ajudar neste processo e vários bancos de dados possuem as informações necessárias, mas as reconstruções resultantes exigem curadoria manual [20].

A segunda etapa é a cura do *draft* da rede reconstruída. A rede metabólica reconstruída automaticamente é incompleta, e pode ter lacunas ou reações equivocadamente incluídas que não deveriam ocorrer no organismo alvo. A cura manual acrescenta, retira e corrige informações que os procedimentos automáticos geram [20]. Enquanto a etapa de reconstrução automática é rápida, o processo de curadoria manual é trabalhoso e às vezes demorado.

A terceira etapa é a criação de um modelo computacional para a rede gerada. O quarto passo consiste em análises com ferramentas de alto rendimento, realizadas sobre o conjunto de dados em escala genômica para avaliar as interações genéticas e podem ser utilizadas para refinar e expandir o conteúdo de uma rede metabólica. Um exemplo particular desse tipo de análise é o GEM [45] que foi usado para anotar funções de genes desconhecidos em *E. coli*, por exemplo [20].

Outra abordagem é a análise biológica dos dados genômicos, bioquímicos e fisiológicos disponíveis. Para essa abordagem, existem várias *pipelines* que normalmente são específicas para o tipo de organismo alvo. Exemplos desta abordagem, são os trabalhos de Förster *et al.* [24] e Heavner [32], que respectivamente reconstruíram e expandiram a rede metabólica do fungo *Saccharomyces cerevisiae*.

A abordagem de Förster *et al.* [24] e Heavner [32], assim como a abordagem de Feist *et al* [20], utilizam numa primeira etapa o genoma anotado e informações de bancos de dados de vias metabólicas e revisão de literatura. Através dessas informações, é criado um modelo metabólico do organismo que permite predições das vias metabólicas que foram sua rede metabólica [24], mas a rede gerada também depende de cura manual.

4.4.1 PathwayTools

O PathwayTools [43] é um ambiente de software para a criação de um PGDB (*Pathway/genome database*), tal como o MetaCyc[58]. O PathwayTools é também um ambiente de gestão, análise e visualização de coleções integradas de genoma, vias metabólicas e dados de regulação. Ele permite a edição interativa de curadores e tem suporte web para publicação de PGDBs, e fornece um grande número de ferramentas de consulta e visualização. O PathwayTools também suporta análises comparativas de PGDBs, e fornece análises de vários sistemas biológicos dos PGDBs incluindo a análise de acessibilidade de redes metabólicas e traçado interativo de metabólitos através de uma rede metabólica [43]. Atualmente está em sua versão 18.0.

Além das vias metabólicas, o PathwayTools lida com muitos tipos de informação gênica e seus recursos são muito extensos. O software está em desenvolvimento contínuo pelo *SRI International* desde o início da década de 1990 [42]. Entre as tarefas que o PathwayTools provê, citamos:

- Apoio ao desenvolvimento de bancos de dados específicas de organismo (PGDBs);
- Suporte à visualização científica, publicação na web e disseminação desses bancos de dados específicos do organismo;
- Realização de inferências computacionais, incluindo a previsão de vias metabólicas, a previsão de preenchimento de *gaps* em vias metabólicas;
- Fornecimento de ferramentas visuais para análise de conjuntos de dados genômicos;
- Fornecimento de ferramentas para análise de redes biológicas;
- Fornecimento de análises comparativas de PGDBs.

O algoritmo de predição do PahwayTools consiste de duas fases. A primeira fase é a inferência do reactoma, ou seja, o conjunto de reações catalisadas pelo organismo através do conjunto de enzimas presentes no seu genoma anotado. A segunda fase consiste na inferência de vias metabólicas que contêm as reações identificadas na primeira fase. Nesta fase o módulo PathLogic do PathwayTools verifica cada elemento genético contido nos arquivos de entrada e reconstrói automaticamente as vias metabólicas. O banco de dados consultado pelo Pathway Tools para realizar essas inferências é o MetaCyc [44].

Os formatos dos arquivos de entrada do PathwayTools, exemplificados na Figura 4.4, são de três tipos, a saber:

1. O arquivo *genetic-elements.dat*, que contém um conjunto de registros, sendo um para cada elemento genético como cromossomos, genes, *contigs*, etc. Cada elemento

genético é descrito com uma ou mais anotações relacionadas. Atributos válidos para as anotações são:

- ID: Obrigatório, identifica o elemento genético;
 - NAME: Obrigatório, descreve o elemento genético;
 - TYPE: Opcional, pode ter os seguintes valores: CHRSM, PLASMID, CON-TIG, MT, PT;
 - CIRCULAR? Obrigatório (exceto para *contigs*), pode ter valores Y ou N, indicando se o elemento genético é circular ou não;
 - CODON-TABLE: Opcional, pode ser um número entre 1 e 15. A tabela padrão de códons é especificada no arquivo organism.dat;
 - ANNOT-FILE: Obrigatório (exceto para *contigs*). Provê caminho do arquivo que contém a anotação para o elemento genético. Este arquivo tem a extensão .pf. Pode ser também um arquivo no formato GenBank (.gbk, .gb);
 - SEQ-FILE: Opcional, caminho do arquivo que contém a sequência de nucleotídeos. Pode ter a extensão conforme o formato (.fna, .fsa, .fasta, .seq). No caso de se utilizar um arquivo no formato GenBank para o atributo ANNOT-FILE, este atributo pode ser omitido, pois as informações seriam redundantes;
 - CONTIG: Opcional, não permitido quando o atributo TYPE tiver valor CON-TIG.
2. Os arquivos com as sequências de cada elemento genético. Esse tipo de arquivo pode ter as extensões conforme o formato (.fna, .fsa, .fasta, .seq, .gbk, .gb). Nós utilizamos o formato *.fsa*.
3. Os arquivos com as anotações das sequências e seus ECs. Esse tipo de arquivo têm a extensão *.pf*.

A

```
ID PAAG_00001T0
NAME conserved hypothetical protein (1506 nt)
TYPE :CHRSM
CIRCULAR? N
SEQ-FILE /Users/admin/bioinfo/ptools-local/pg dbs/user/plutziicyc/2.0/input/PAAG_00001T0.fsa
ANNOT-FILE /Users/admin/bioinfo/ptools-local/pg dbs/user/plutziicyc/2.0/input/PAAG_00001T0.pf
//
```

ID PAAG_00002T0
NAME DNA polymerase epsilon subunit C (2976 nt)

B

```
>PAAG_00001T0 | Paracoccidioides brasiliensis Pb01 conserved hypothetical protein (1506 nt)
ATGTTGCCACCCCTTGCTTACCTCCCTACGTGTGATTGCCCTCTCGGGCGCA
CCGCCTTGATTCTGGCTGCTCGGTGATTTATCGACTCCGTTTACCCCTCACCCAG
TTTCAGGGGCCCTGGTATCGAGCTGCGACTTCGCTAACTTTGCCACCCCTCTTTGTT
AGACTTGTATCGTGTGGATAGAATACTTAGTGAAGAAGTATGGCGAACGGCACCAATC
CGAATCCAACCGAATCTCTTGTGCTCATCCATCGCACTAAAGGAAATATATTGC
GACCCAAATGCAACACCAAGTCCAGCATGTATGGAACAGGCCTTCGGTCTCCACAT
CTATTTGACCCCTTGACGGCGATAAGCATCGTGAGCTTCTGAAAGTCTCTAGGAGGTGCA
CCATCCACTATACCATCGCTCANCACACAGCTCCACCCACATTCGATCACCACCTCTC
```

C

```
ID PAAG_00001T0
NAME conserved hypothetical protein (1506 nt)
PRODUCT-TYPE P
STARTBASE 1
ENDBASE 1506
FUNCTION unspecific monooxygenase
EC 1.14.14.1
//
```

Figura 4.4: Trechos dos arquivos de entrada gerados a partir do banco de dados Plutzii: (A) genetic-elements.dat, (B) PAAG_0001T0.fsa, (C) PAAG_0001T0.pf.

Capítulo 5

Método para reconstrução *in silico* de redes metabólicas em fungos

Este capítulo apresenta o método criado para reconstruir *in silico* redes metabólicas de fungos a partir de seu genoma e/ou transcritoma, com destaque para o *pipeline* específico de identificação de vias de metabolismo secundário. O fungo *P. lutzii* foi utilizado como estudo de caso. Na Seção 5.1 o método proposto é descrito de forma geral. Na Seção 5.2 são detalhadas as etapas e *pipelines* do método.

5.1 Descrição geral do método

O método proposto é semi-automatizado e consiste de etapas com *pipelines* subsequentes apoiados por um banco de dados relacional, que armazena os arquivos de entrada e os resultados obtidos ao longo dos *pipelines* de predição de metabolismo secundário, identificação de vias metabólicas a partir dos dados do genoma do fungo e reanotação de enzimas. Opcionalmente, dados específicos de cada projeto podem ser agregados às anotações. A reconstrução da rede com os dados do banco colhidos é realizada pelo software PathwayTools. Uma visão geral do método é mostrada na Figura 5.1.

5.2 Detalhes das etapas

1. **Dados do genoma do fungo.** Um banco de dados relacional foi modelado e implementado utilizando o Sistema Gerenciador de Bancos de Dados MySQL. Este banco de dados tem dois propósitos: o primeiro é armazenar e recuperar os dados que são usados como entrada dos *pipelines*. O segundo é prover uma base de dados acessível multiplataforma para facilitar a geração de relatórios. O modelo de entidade-relacionamento pode ser visto na Figura 5.2.

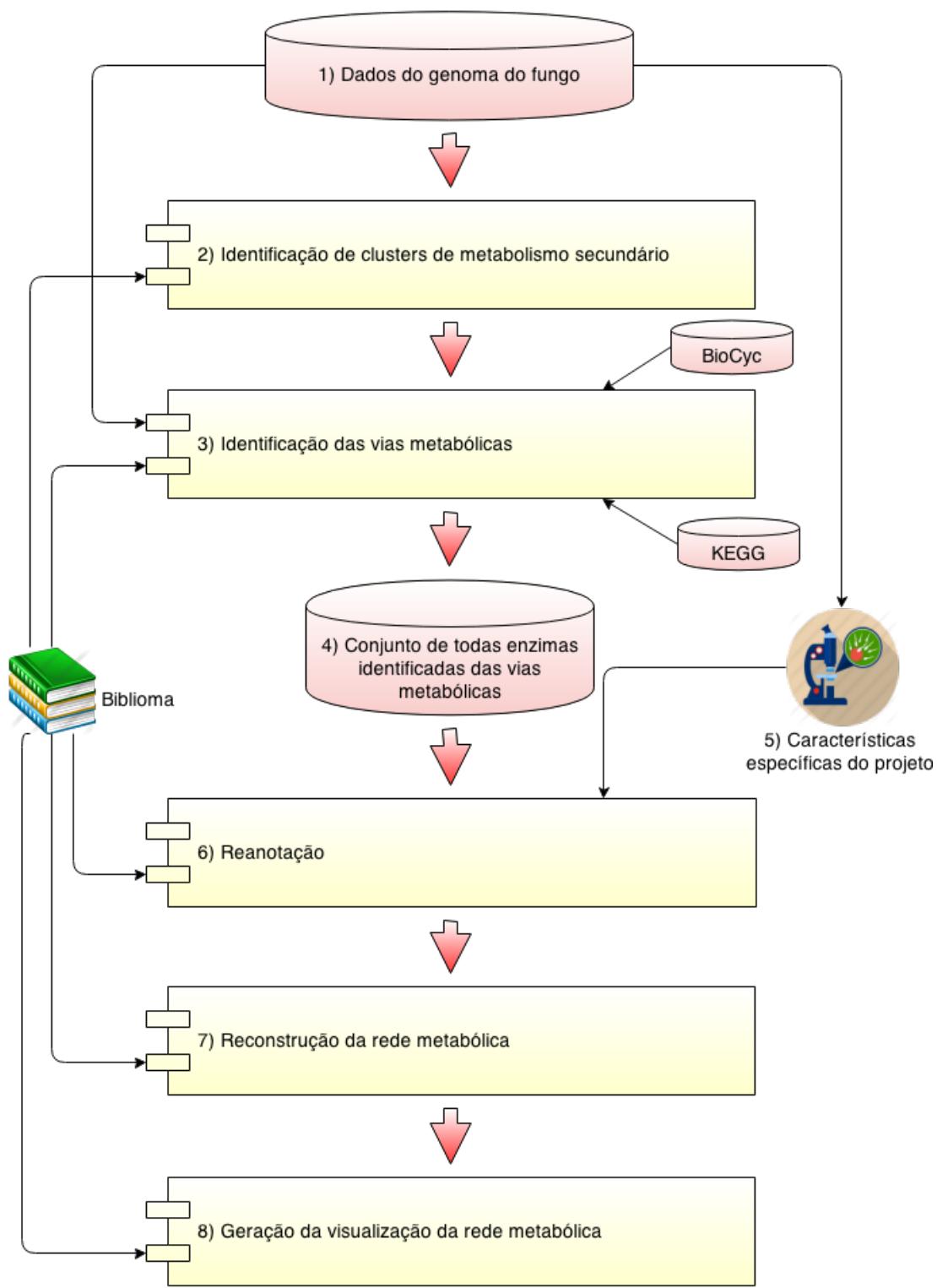


Figura 5.1: Método para reconstrução *in silico* de redes metabólicas de fungos.

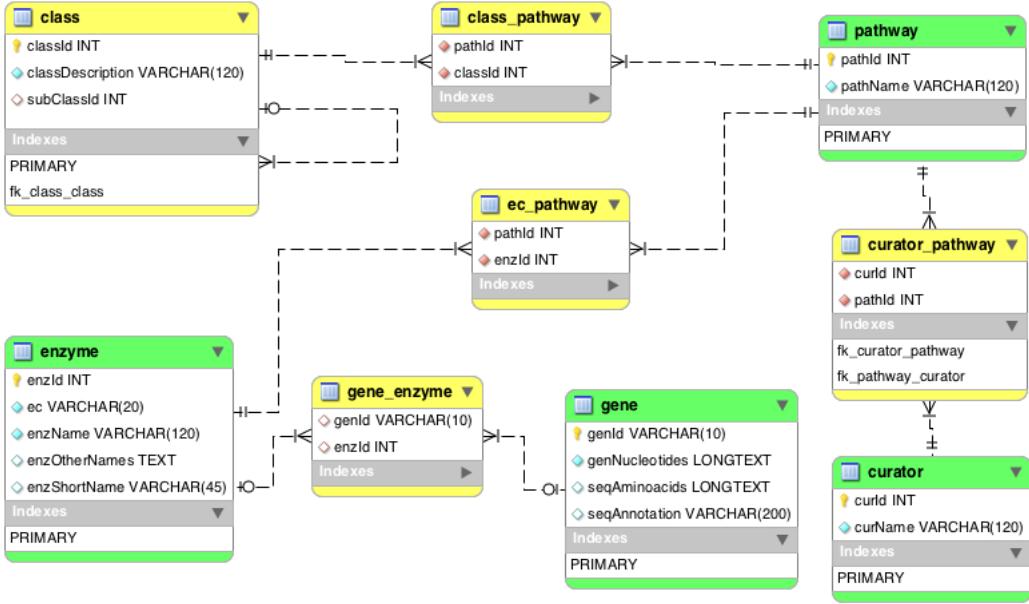


Figura 5.2: Modelo de Entidade-Relacionamento do banco de dados

2. **Identificação de clusters de metabolismo secundário.** Uma vez alimentado o banco de dados com os dados do genoma do fungo, esses dados são recuperados e manipulados através de *scripts*, para gerar os arquivos de entrada nos formatos das ferramentas de predição de genes de metabolismo secundário. Em particular SMURF [49] e antiSMASH [57] são softwares que têm como saída *clusters* de genes preditos para metabolismo secundário. A partir desses *clusters*, é possível identificar nos dados do genoma, as enzimas que os genes codificam.
3. **Identificação das vias metabólicas.** As enzimas do fungo¹ são buscadas nos bancos de dados KEGG [41] e BioCyc [13]. As vias metabólicas encontradas a partir dessas enzimas, são filtradas para separar as vias que ocorrem ou potencialmente ocorrem em fungos. Depois disso, as vias metabólicas com suas respectivas enzimas (enzimas preditas e demais enzimas das vias) são armazenadas no banco de dados com seus ECs. Adicionalmente todas as demais informações das reações são catalogadas.
4. **Conjunto de todas enzimas identificadas das vias metabólicas.** Neste ponto o banco de dados já conta com uma coleção de enzimas do genoma do fungo com seus ECs identificados.
5. **Características específicas do projeto.** Se o organismo alvo da reconstrução tiver alguma característica que se deseje inserir na anotação, essas características

¹Todas as enzimas, inclusive as preditas em *clusters* de genes para metabolismo secundário.

devem ser inseridas após o pipeline de anotação. Uma vez anotados, esses dados servirão de entrada para a reconstrução da rede e não é recomendável anotá-los após a reconstrução da rede metabólica devido ao risco de corromper a rede reconstruída.

6. **Reanotação.** As sequências de resíduos de aminoácidos destas enzimas na sua forma canônica e isoenzimas são então obtidas nos bancos de dados de enzimas. Em particular, sugerimos o UniProt [55] e BRENDA [70]. Essas enzimas passam a ser consideradas enzimas candidatas. É feito um alinhamento das sequências das enzimas candidatas contra as sequências das enzimas do fungo utilizando *Blastp*.

As enzimas do fungo que alcançam melhores *hits*, têm seu domínio conservado verificado utilizando CDD [56] e InterPro [35]. Este procedimento visa identificar se a enzima do fungo realmente catalisa a reação alvo para a qual ela pretende ser atribuída. Se a enzima do fungo for confirmada para a via e ainda não estiver anotada, ela tem sua anotação atualizada.

O conjunto de todas as enzimas do fungo é então submetido ao KAAS [61] para uma anotação automática. Este procedimento, visa confirmar ou atualizar a anotação existente e completar a anotação quando possível. Adicionalmente, informações exclusivas do projeto podem ser adicionadas na anotação.

7. **Reconstrução da rede metabólica.** Os dados de anotação são então armazenados no banco de dados. Os dados são recuperados via *script* e convertidos nos arquivos de entrada do PathwayTools [43]. O PathwayTools realiza primeiramente a inferência do reactoma a partir das enzimas e num segundo momento a inferência das vias metabólicas a partir do reactoma. Depois essas vias são classificadas e agrupadas. Esse processamento pode ser feito de forma automática ou manual pelo módulo PathoLogic do PathwayTools.
8. **Geração da visualização da rede metabólica.** A geração da visualização da rede reconstruída é realizada automaticamente pelo módulo Omics Viewer [42] do PathwayTools. O resultado é uma figura em formato *pdf* de alta resolução com todas as vias metabólicas com nível de detalhe customizável. Além disso é possível utilizar o PathwayTools em conjunto com um servidor web e um banco de dados MySQL e gerar uma versão para web da rede metabólica. Esta versão web provê a navegação entre os objetos da rede metabólica através de *links*.

Biblioma. O biblioma é o conjunto de informações que podem ser obtidas na literatura especializada, como livros e artigos científicos. O biblioma é consultado ao longo de todo o método nas suas diferentes etapas para que haja uma cura incremental dos dados gerados nos pipelines.

Capítulo 6

Estudo de caso: rede metabólica do *P. lutzii*

Este capítulo apresenta o estudo de caso para o fungo *P. lutzii*. A Seção 6.1 detalha a aplicação do método proposto na reconstrução *in silico* da rede metabólica do fungo *P. lutzii*. A Seção 6.2 apresenta os dados gerais da rede metabólica do *P. lutzii* além de comparar esses dados com outros resultados disponíveis. Na Seção 6.3 discutimos três vias de metabolismo identificadas.

6.1 Método aplicado ao *P. lutzii*

O transcritoma do *P. lutzii* identificou 6.022 ESTs, indicando um número estimado de aproximadamente 8.000 genes [22]. Posteriormente, o BROAD Institute sequenciou o genoma do *P. lutzii* e identificou 9.132 genes [17]. Notamos que no transcritoma, há informação disponível sobre a forma de ocorrência dos genes: micélio ou levedura. Este estudo de caso utilizou os dados do transcritoma, obtidos no Laboratório de Bioinformática do Departamento de Biologia Celular da UnB e dados do genoma, obtidos do BROAD Institute. Além desses dados, Alvarez [3] reconstruiu uma rede metabólica do *P. lutzii* com 295 vias metabólicas. Os dados deste trabalho foram revisitados e atualizados. O método definido para reconstrução *in silico* de redes de metabolismo em fungos foi aplicado ao *P. lutzii* e é mostrado na Figura 6.1. Os detalhes são descritos a seguir.

1. A rede metabólica do *P. lutzii*, previamente reconstruída por Alvarez [3], proporcionou um grande volume de informações armazenadas em arquivos, muitas vezes com redundâncias. O esquema do banco de dados relacional foi criado com nome de *Plutzii* (Figura 6.2) e com ele organizamos e armazenamos os dados disponíveis da rede metabólica previamente reconstruída por Alvarez [3] e dados de transcritoma

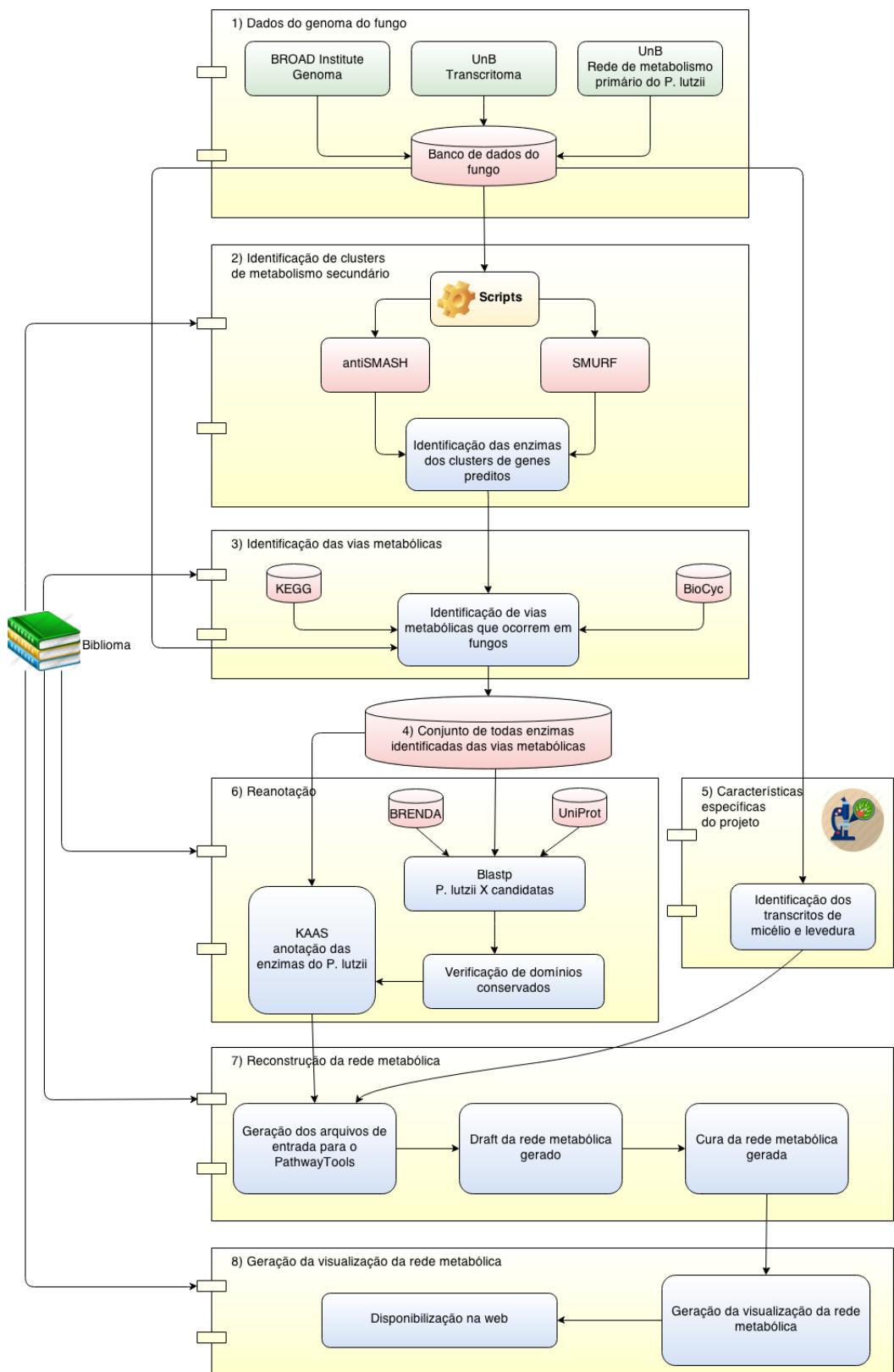


Figura 6.1: Método para reconstrução *in silico* de redes metabólicas em fungos aplicado ao *P. lutzii*.

e genoma. Além disso, armazenamos os dados resultantes da separação de forma de micélio/levedura, dados de enzimas e evidências de sua participação nas vias metabólicas.

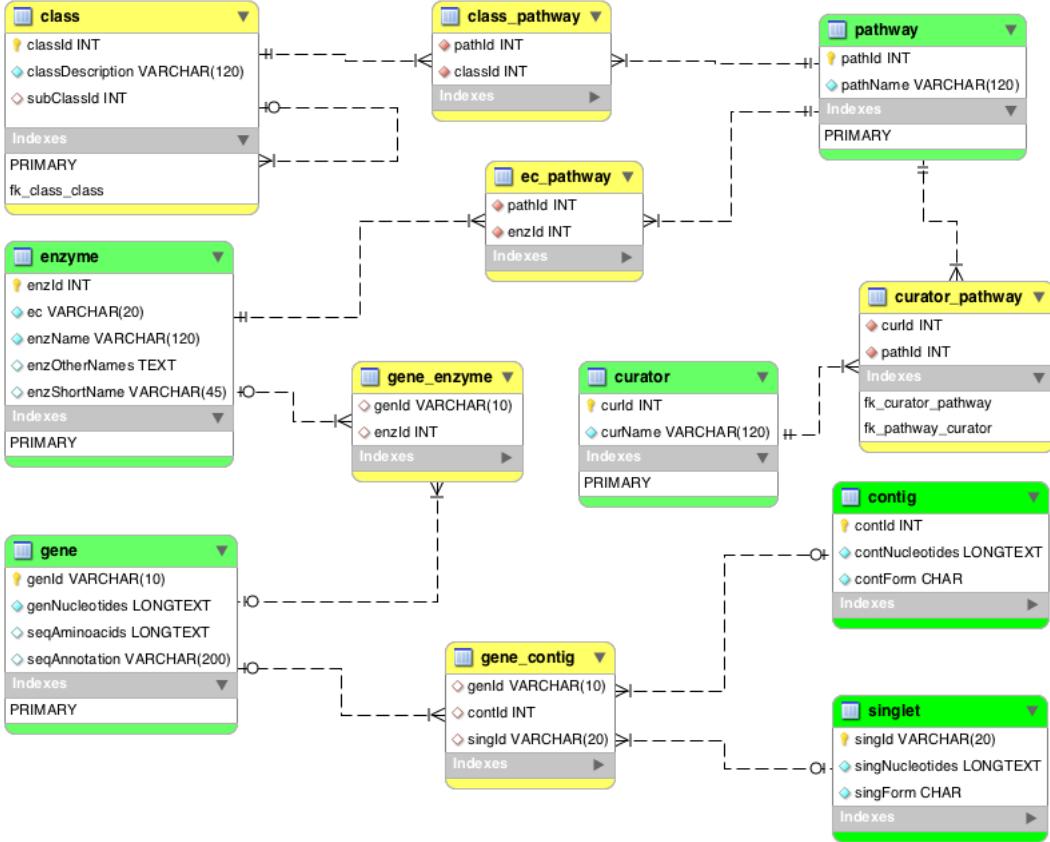


Figura 6.2: Modelo de Entidade-Relacionamento (MER) do banco de dados adaptado para o *P. lutzii*. Este MER foi chamado de Plutzii e contém as tabelas e relacionamentos com os dados do *P. lutzii* e os dados da rede previamente reconstruída. É utilizado como ponto de partida para o método e ao longo de sua execução, os dados são atualizados com os resultados intermediários.

2. Após a execução, os resultados do SMURF [49] e do antiSMASH [57] apontaram um total de onze *clusters* de genes, sendo quatro deles preditos pelos dois softwares, cinco exclusivamente pelo SMURF e dois exclusivamente pelo antiSMASH. Foi realizada a identificação das enzimas dos *clusters* de genes de metabolismo secundário preditos para o *P. lutzii*.
3. As enzimas identificadas nos *clusters* de genes foram consultadas no KEGG [41] e BioCyc [13] e as vias metabólicas encontradas a partir dessas enzimas, foram filtradas para separar aquelas que ocorriam ou potencialmente ocorriam em fungos. Depois disso, as vias metabólicas identificadas com suas respectivas enzimas (enzimas preditas e demais enzimas), foram armazenadas em nosso banco de dados.

4. Cada uma das enzimas das vias metabólicas identificadas, teve sua sequência obtida do UniProt [55] e BRENDa [70] e foram chamadas de enzimas candidatas. Um alinhamento com *Blastp* das sequências das enzimas do *P. lutzii* contra as sequências das enzimas candidatas, revelou aquelas que tinham potencial para preencher os *gaps* da via.
5. Para confirmar este potencial indicado pelo alinhamento das sequências, verificamos também domínios conservados da enzima, no CDD [56] e InterPro [35]. As enzimas que preencheram os requisitos de afinidade por alinhamento e presença de domínios conservados para a reação alvo, foram mantidas para ocupar os *gaps* e tiveram seus ECs identificados e armazenados em nosso banco de dados.
6. O conjunto das enzimas do *P. lutzii* foi submetido ao KAAS para atualizar sua anotação anterior e anotar as enzimas hipotéticas, preditas ou putativas. O trabalho de Alvarez [3] forneceu ECs anotados de 1.802 enzimas. Essas enzimas foram novamente submetidas ao KAAS [61] e algumas anotações foram atualizadas e/ou expandidas. Após anotadas ou não, as enzimas que catalizam reações identificadas foram armazenados em nosso banco de dados.
7. Em paralelo, foi realizada a separação de micélio/levedura do transcritoma, obtido na UnB. *Scripts* escritos em Perl realizaram a análise dos arquivos de transcritoma para obter informação sobre a forma em que os genes ocorrem. Para cada *contig* foi verificada a sua origem que podia ser micélio (M) ou levedura (L) ou ambos (B). Na montagem do *contig*, o critério de separação foi a presença de 80% ou mais sequências de micélio ou levedura. Para cada *singlet* foi verificada a sua origem que podia ser micélio (M) ou levedura (L). Ao final deste processo, os *contigs* e *singlets* foram separados em: micélio, levedura, ambos ou nenhum¹. As enzimas oriundas dos 9.132 genes do *P. lutzii* obtidas no *BROAD Institute*, foram submetidas a um *Blastx* contra o transcritoma obtido na UnB, para identificar a forma de expressão dos genes de acordo com o transcritoma.

Foram identificados, com *evaluе* < 10^{-100} , 1.442 transcritos a partir dos contigs. Também com *evaluе* < 10^{-100} , foram identificados 715 transcritos a partir dos singlets. Para este conjunto de 2.157 transcritos, foi possível verificar a forma de ocorrência (M/L/B). Foi acrescentada à anotação dos genes essa nova informação. Os dados de forma dos transcritos foram armazenados em nosso banco de dados.

8. Com todas as informações armazenadas no banco de dados *Plutzii*, foram gerados os arquivos de entrada para o PathwayTools [43] através de um *script*.

¹Quando não foi possível identificar a forma de ocorrência.

9. Com os arquivos de entrada gerados, acessamos o módulo PathoLogic [42] do PathwayTools e criamos uma nova rede metabólica vazia. Em seguida a função *automated build* (Figura 6.3) foi acionada com a rede vazia selecionada. O *draft* da rede foi então reconstruído.

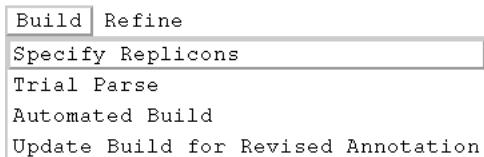


Figura 6.3: Funcionalidade do PathoLogic [42] para reconstrução automática da rede metabólica a partir dos arquivos com elementos genéticos gerados com informações do banco de dados Plutzii.

10. Com o *draft* da rede, foi feito o procedimento de cura, verificando de forma comparativa as vias geradas pelo PathwayTools com aquelas identificadas no KEGG [40, 41, 46] e no BioCyc [10, 13]. Novas reações e vias metabólicas puderam ser adicionadas manualmente. Onde houve divergências, a literatura foi consultada e a via foi ajustada com a função refine mostrada na Figura 6.4.

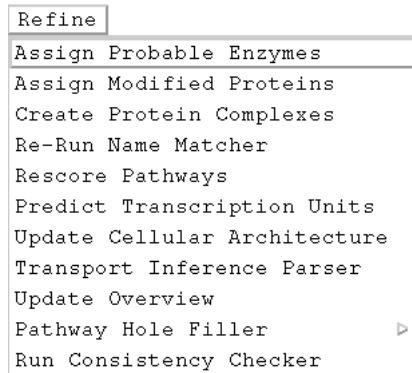


Figura 6.4: Funcionalidade do PathoLogic [42] para refinar dados de vias metabólicas reconstruídas.

11. A geração da visualização da rede reconstruída foi realizada automaticamente pelo módulo Omics Viewer [42] do PathwayTools. A rede metabólica está disponível no endereço <http://www.biomol.unb.br/metabolismo>.

6.2 Dados da rede metabólica *in silico* do *P. lutzii*

Após a execução do nosso método, obtivemos 2.087 reações enzimáticas sobre as quais foi possível identificar 335 vias metabólicas. Essas reações enzimáticas são catalizadas por

1.437 enzimas sobre 1.464 compostos. Esses resultados podem ser vistos na Tabela 6.1 juntamente com outros resultados existentes para redes metabólicas do *P. lutzii*.

Tabela 6.1: Tabela comparativa das informações das redes metabólicas reconstruídas *in silico* para o *P. lutzii*

Resultado	Alvarez [3]	BROAD [85]	Waldeyr
Reações enzimáticas	1.743	1.800	2.087
Enzimas	1413	1.463	1.437
Compostos	1.202	1.234	1.464
Vias metabólicas	295	290	335

A visualização da rede metabólica reconstruída pode ser vista na Figura 6.5. A distribuição das vias metabólicas de acordo com a classificação proposta pelo BioCyc, pode ser vista no gráfico da Figura 6.6, com destaque para as vias de metabolismo secundário.

A partir do alinhamento dos *contigs* e *singlets* transcritos do *P. lutzii* contra as sequências do genoma obtido do *BROAD Institute*, identificamos a forma (micélio/levedura/ambas) de 1.972 transcritos. O gráfico da Figura 6.7 mostra a distribuição das formas de ocorrência dos 1.972 transcritos do *P. lutzii*. A forma de cada transrito pode ser verificada no Anexo C.

Foram encontradas 20 vias de metabolismo secundário completas, sendo 15 vias de biossíntese de metabólitos secundários e 5 vias de degradação de metabólitos secundários. Estas vias representam cerca de 4.5% do total de vias que compõem a rede metabólica do *P. lutzii* e sua classificação é mostrada no Anexo A. A tabela 6.2 apresenta algumas aplicações dessas 20 vias metabólicas secundárias.

Após o alinhamento das sequências, a verificação de domínios conservados e uma revisão bibliográfica sobre as reações que as enzimas catalizam, foram propostas anotações para nove enzimas (ver Tabela 6.3), anteriormente anotadas como hipotéticas ou preditas.

Foram identificados 310 *gaps* em reações enzimáticas. Uma tabela com essas reações, ECs das enzimas e vias metabólicas das quais essas reações participam pode ser vista no Anexo B.

A rede metabólica reconstruída disponível na web, permite visualizar dados gerais da rede (Figura 6.8), navegar pelas vias através de sua classificação (Figura 6.9) e visualizar detalhes da via selecionada (Figura 6.10).

PlutziiCyc: *Paracoccidioides* sp. *lutzii* PbU1 Cellular Overview

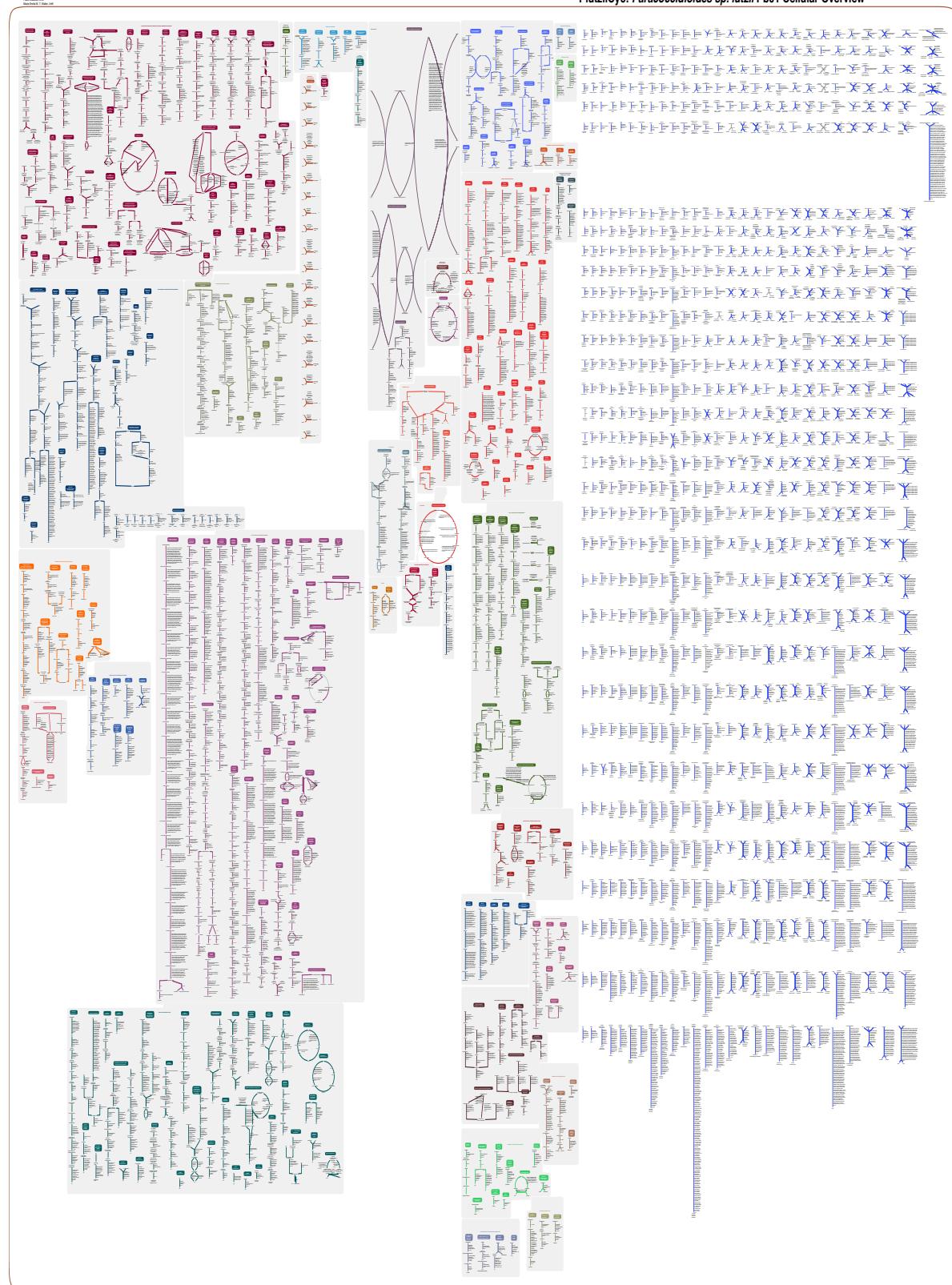


Figura 6.5: Rede metabólica reconstruída *in silico* para o fungo *P. lutzii*, incluindo vias metabólicas completas e reações não atreladas a vias.

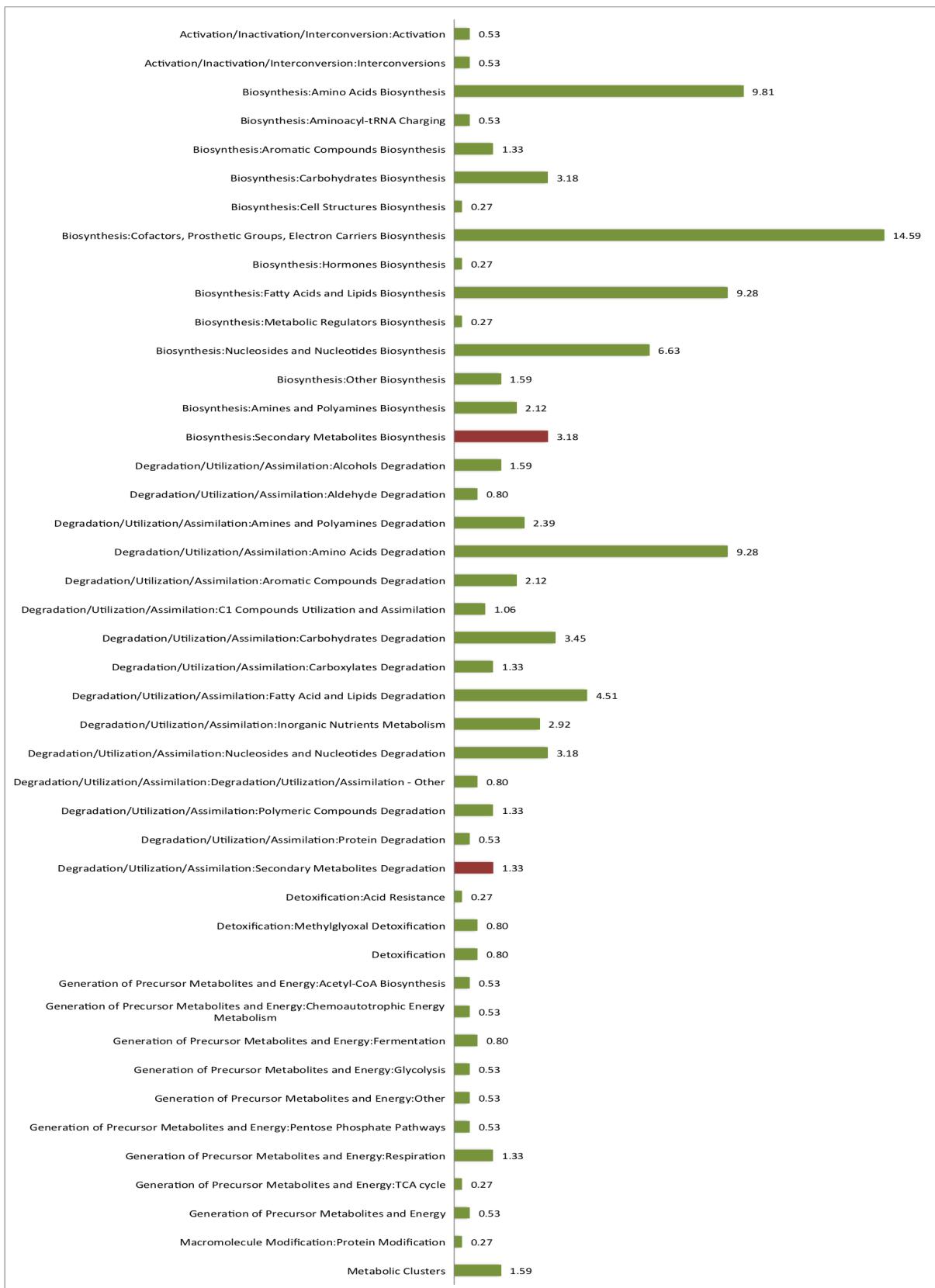


Figura 6.6: Distribuição (%) das vias metabólicas identificadas, de acordo com a classificação do BioCyc.

Tabela 6.2: Possíveis aplicações de vias metabólicas secundárias identificadas *in silico* para o *P. lutzii*.

Via	Possível aplicação
Betanidin degradation	A enzima peroxidase, que catalisa esta via, é usada no tratamento de contaminantes fenólicos na presença de H_2O_2 [31].
Taxiphyllin bioactivation	A partir de glicosídeos cianogênicos, ocorre a liberação de cianeto de hidrogênio, que desempenha um papel central no sistema de defesa química nas plantas em interações planta-inseto [26].
1D-myo-inositol hexakisphosphate biosynthesis I (from Ins(1,4,5)P3)	Em fungos, está relacionada à regulação da saída do mRNA do núcleo celular [84].
N-acetylglucosamine degradation I	A N-acetilglucosamina é relativamente abundante na natureza, sendo um componente da quitina e das paredes celulares dos fungos [2].
Mannitol degradation	O manitol é utilizado clinicamente em osmoterapia para reduzir a pressão intracraniana [72].
Betaxanthin biosynthesis	É um pigmento. As betaxantinas amarelo-laranja são restritas aos fungos [51].
Phenylethanol biosynthesis	Está presente em boa parte das essências naturais de flores e frutos. Pode ser usado como matéria-prima na produção de perfumes com aromas de rosa ou com outras combinações florais e na formação de sabores artificiais [18].
D-myo-inositol (1,4,5)-trisphosphate biosynthesis	Os mio-inositols contendo fosfolípidos estão envolvidos no controle de funções celulares de eucariotos, incluindo algumas vias de sinalização intracelular e de vários reguladores, como hormônios, neurotransmissores e fatores de crescimento [54].
zeaxanthin, antheraxanthin and violaxanthin interconversion	São carotenóides, que não podem ser sintetizados por animais e são uma fonte essencial de retinóides e vitamina A. A zeaxantina, por exemplo, está ligada à redução de catarratas. Nas plantas, através da dissipaçāo da luz solar, protege contra o estresse fotooxidativo [15].
Geranylgeranyl-diphosphate biosynthesis	Também conhecido como pirofosfato geranilgeranil, ou GGPP, é um composto essencial envolvido na biossíntese de vários terpenos e terpenóides, incluindo ubiquinonas e menaquinonas [78].
Mevalonate pathway	Em células eucarióticas, a via do mevalonato conduz a biossíntese de esterol vegetal, da biossíntese do ergosterol I e dolicol, através da formação de difosfato de farnesil (FPP). O ergosterol é um dos principais constituintes da membrana plasmática em fungos [1].
Neurosporaxanthin biosynthesis	bi- É um apocarotenóide, um pigmento derivado de terpenóides proveniente de vias de carotenóides. Geralmente o pigmento é amarelo brilhante ou laranja [69].
Epoxysqualene biosynthesis	Esqualeno é um composto que atua como intermediário na biossíntese de esteróis. Esqualeno é uma alusão ao nome científico para o tubarão (<i>Squalus spp.</i>), uma vez que o óleo de fígado de tubarão é considerada a mais rica fonte de esqualeno. O esqualeno é abundante em muitos óleos vegetais, como o azeite ⁴⁴ , óleo de palma, óleo de gérmen de trigo [48].

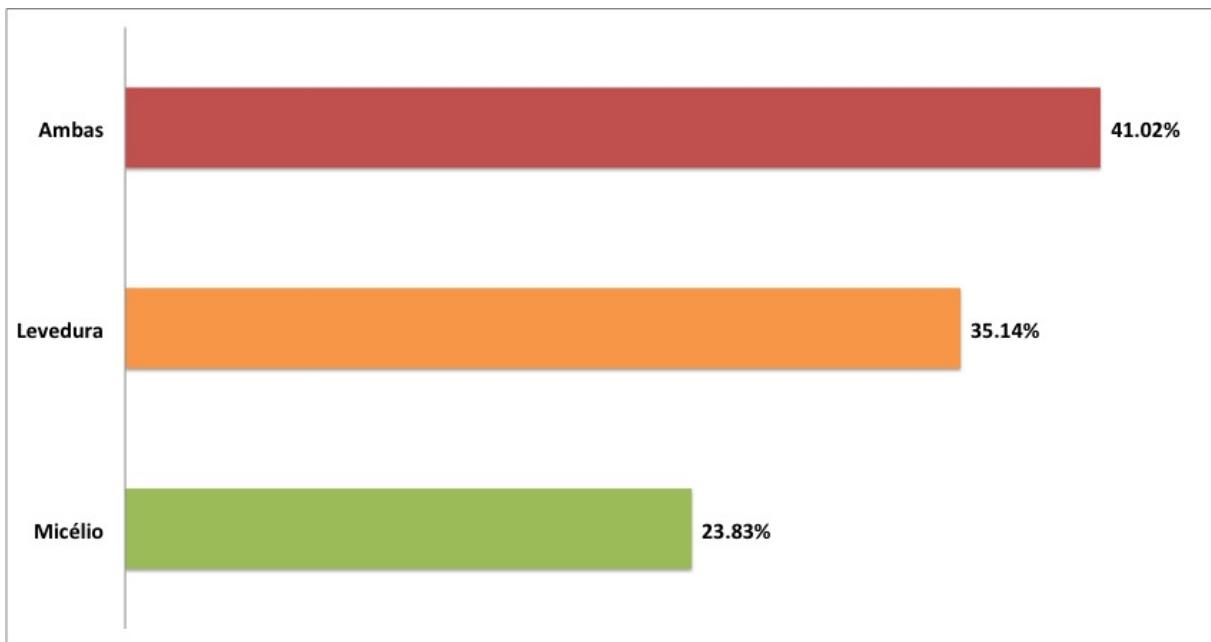


Figura 6.7: Distribuição de transcritos oriundos dos *contigs* e *singlets* do *P. lutzii*, sob as formas de levedura, micélio e ambas as formas.

Tabela 6.3: Anotações propostas para enzimas antes anotadas como preditas ou hipotéticas

Gene	EC	Anotação proposta
PAAG_05405	1.1.1.35	3-hydroxyacyl-CoA dehydrogenase
PAAG_03986	1.11.1.7	peroxidase
PAAG_04175	1.11.1.7	peroxidase
PAAG_04128	1.14.13.7	phenol 2-monooxygenase
PAAG_00037	1.3.3.9	cytochrome P450 monooxygenase
PAAG_07667	2.3.1.135	phosphatidylcholine-retinol O-acyltransferase
PAAG_02729	4.2.1.28	aromatic-L-amino-acid decarboxylase
PAAG_06057	5.1.3.3	aldose 1-epimerase
PAAG_00112	3.1.1.1	carboxylesterase

Enter a gene, protein, metabolite or pathway... Quick Search Gene Search
Searching *Paracoccidioides sp. lutzii Pb01* change organism database

Home | **Search** | **Genome** | **Metabolism** | **Analysis** | **Help**

Summary of *Paracoccidioides sp. lutzii*, Strain Pb01, version 2.0

Authors: Waldeyr M. C. da Silva, IFG; Paulo Alvarez, UnB; Maria Emilia M. T. Walter, UnB

Replicon	Total Genes	Protein Genes	RNA Genes	Pseudogenes	Size (bp)	NCBI Link
9137 Undisplayed Contigs/Replicons	9137	9137	0	0	12,148,199	
Total:	9137	9137	0	0	12,148,199	

Pathways:	335
Enzymatic Reactions:	2087
Transport Reactions:	12
Polypeptides:	9140
Protein Complexes:	3
Enzymes:	1437
Transporters:	42
Compounds:	1464
Transcription Units:	0
tRNAs:	0

Taxonomic lineage: cellular organisms , Eukaryota , Opisthokonta , Fungi , Dikarya , Ascomycota , saccharomyceta , Pezizomycotina , leotiomyceta , Eurotiomycetes , Eurotiomycetidae , Orygenes , mitosporic Orygenes , Paracoccidioides , Paracoccidioides sp. lutzii , Paracoccidioides sp. lutzii Pb01

Unification Links: NCBI-Taxonomy:502779

Genetic Code Number:
1 -- Standard

Figura 6.8: Exemplo da página com dados gerais da rede metabólica do *P. lutzii* na web.



Figura 6.9: Exemplo de navegação pela rede metabólica através das classificações das vias metabólicas.

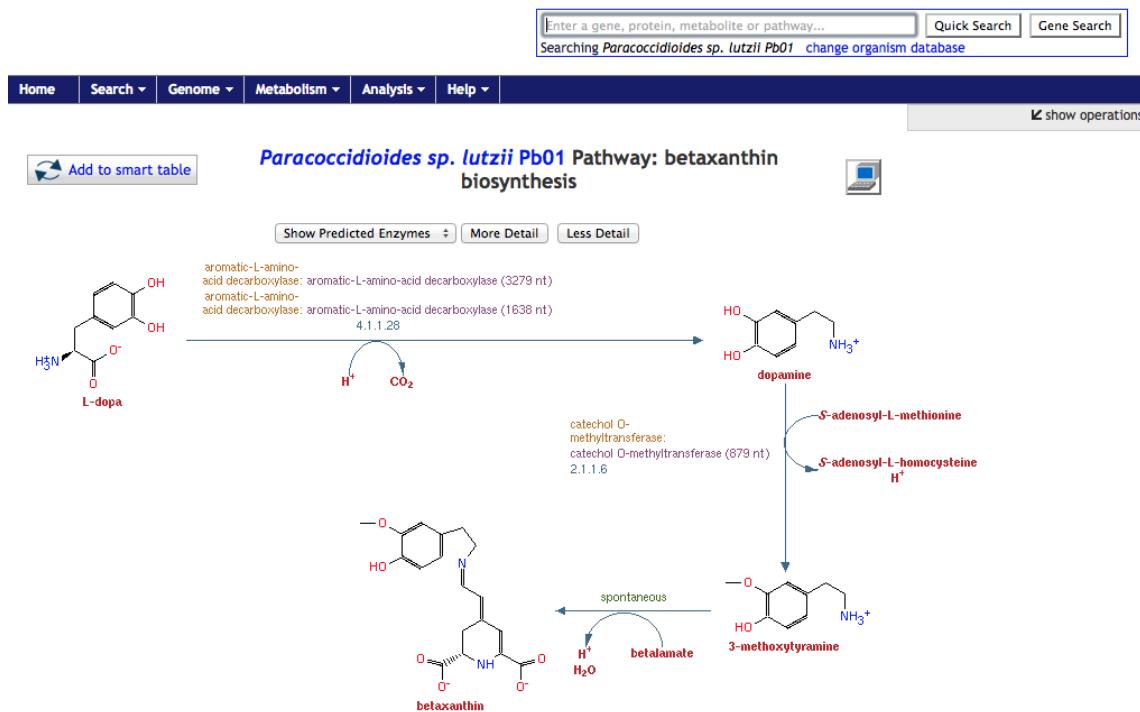


Figura 6.10: Exemplo de exibição da via metabólica de biossíntese de betaxantina.

6.3 Discussão

Ciclo do Glioxilato

A ocorrência do ciclo do glioxilato foi originalmente desboberta em bactéria [50]. Em plantas, este ciclo está envolvido no armazenamento de óleos durante a germinação das sementes [12]. O ciclo do glioxilato é uma via metabólica alternativa ao ciclo do ácido cítrico para oxidação do acetato (*actilCoA*) em ácidos dicarboxílicos (succinato, malato, oxaloacetato). A via é essencialmente, uma versão modificada do ciclo do TCA I (procariótico). Estudos apontam que o ciclo do glioxilato é necessário para a virulência em tanto da bactéria (*M. tuberculosis*) quanto do fungo (*C. albicans*) que podem sobreviver no interior de um macrófago [53].

O ciclo do glioxilato já foi descrito para o *P. lutzii* e pode ser inibido bloqueando a disponibilidade de nutrientes, o que impede a sobrevivência dos agentes patogênicos no interior do macrófago. Os compostos já desenvolvidos que inibem a disponibilidade de nutrientes, são herbicidas como o glifosato. Neste caso os alvos são enzimas produzidas por plantas, e não por animais [53]. Como as enzimas do ciclo do glioxilato não são encontradas em mamíferos, o que inclue o *H sapiens*, elas são os principais alvos para agentes anti-bacterianos e anti-fúngicos [6, 53].

No ciclo do gioxilato, uma enzima chave, isocitrato liase (EC 4.1.3.1²), converte D-threo-isocitrato para succinato e gioxilato. Uma segunda enzima, malato sintase (EC 2.3.3.9³), juntamente com uma segunda molécula de acetil-CoA obtida da via de decarboxilação do piruvato, converte o gioxilato para a forma (S)-malato. Subsequentemente, a oxidação do malato, catalisada pela dehidrogenase do malato (EC 1.1.1.37⁴) converte (S)-malato em oxaloacetato. Neste ponto, o succinato que foi convertido pela isocitrato liase, pode ser retirado do ciclo e usado para a síntese de carbono da célula. Todas as demais enzimas desta via são expressas por genes em ambas as formas (M/L), o que leva a crer que esta via ocorre em qualquer uma das formas do *P. lutzii*. A Figura 6.11, foi extraída de nossa rede metabólica e mostra esta via.

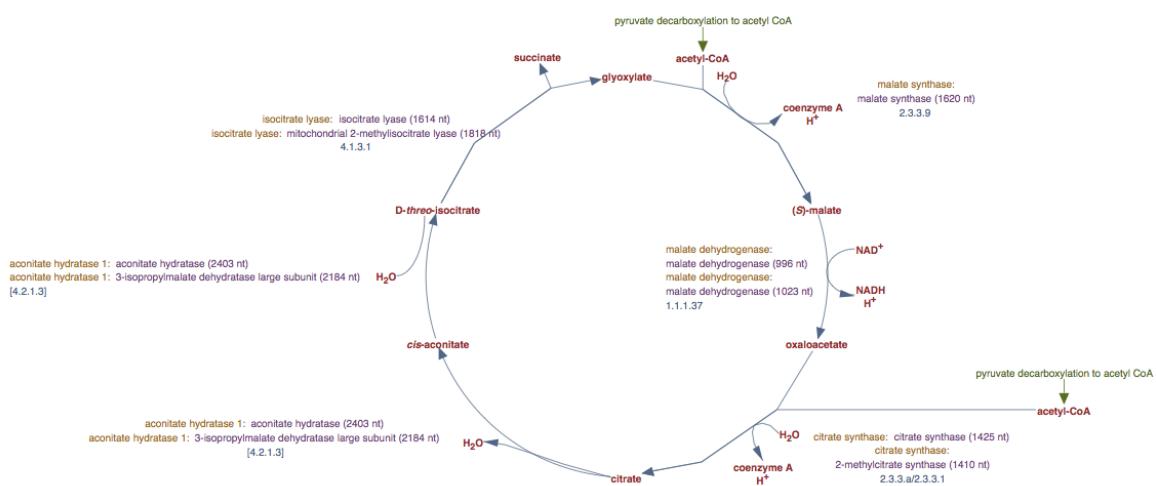


Figura 6.11: Via metabólica do ciclo do gioxilato.

Degradação de Mannitol I

Dos dez hexitóis existentes (*allitol*, *D-altritol*, *L-altritol*, *D-sorbitol*, *L-sorbitol*, *D-manitol*, *L-manitol*, *D-iditol*, *L-iditol* e *galactitol*) apenas três ocorrem naturalmente: *D-manitol*, *D-sorbitol* (também conhecido como *D-glucitol*), e galactitol (devido à simetria, *D-galactitol* e *L-galactitol* são idênticos). O manitol é o poliol mais comum no reino *fungi*, sendo encontrado em quantidades abundantes em esporos, micélios, escleródios e corpos de frutificação [80]. O manitol pode assumir diferentes papéis em diferentes fungos, por exemplo, em fungos como o *Agaricus bisporus* o manitol pode representar entre 10% e 50% do seu peso. Estudos de Niehaus [64] mostraram que as enzimas da via de biosíntese de manitol são expressas em resposta à presença de glucose.

²Transcritos PAAG_06951 e PAAG_04549

³Transcrito PAAG_04542

⁴Transcritos PAAG_08449 e PAAG_00053

A reação desta via⁵, ocorre nos dois sentidos. Quando a molécula de *D – manitol* é fosforilada, a molécula resultante é chamada de *D – manitol1 – fosfato*. A enzima *mannitol-1-phosphate 5-dehydrogenase* (EC 1.1.1.17⁶) catalisa a redução da *fructose – 6 – fosfato* para *1 – fosfatodemanitol* utilizando o cofator *NADH*. A Figura 6.12 foi extraída de nossa rede metabólica e mostra a via de degradação do manitol. O Transcrito que codifica a enzima desta foi expresso apenas sob a forma de micélio, o que indica que esta via ocorre também apenas nesta forma.

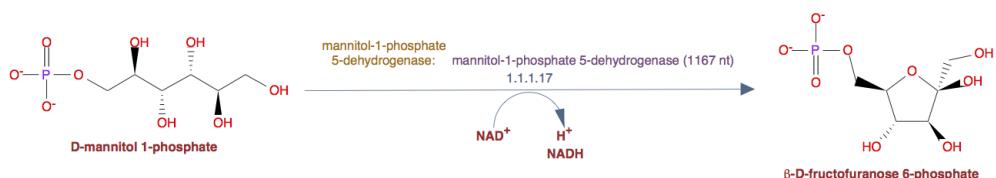


Figura 6.12: Via metabólica da degradação de manitol I.

Biossíntese de Geranilgeranil di-fosfato

Nos últimos anos, grande atenção tem sido dada aos compostos isoprenóides, que incluem não só moléculas essenciais para a sobrevivência da célula, mas vários compostos bioativos produzidos através do metabolismo secundário [74]. *All-trans-difosfato geranylgeranil*, também conhecido como pirofosfato geranylgeranil, ou GGPP) é um composto essencial envolvido na biossíntese de vários terpenos e terpenóides, incluindo ubiquinonas e menaquinonas [78].

Terpenóides são sintetizados pela condensação consecutiva total de unidades de isopreno C5, tais como a enzima difosfato dimetilalilo (DMAPP) e isopentenil a enzima difosfato (IPP) [79]. Os terpenóides são classificados pela quantidade de unidades de cinco carbonos presentes nas suas estruturas [79]. Durante o processo de alongamento da unidade de isopreno, a enzima difosfato geranylgeranil (GGPP) sintase catalisa a terceira reação de condensação do IPP para produzir GGPP (C20) [74]. GGPP é o precursor de di-terpenos e representa o ponto de ramificação para muitos isoprenóides importantes tanto para metabolitos primários quanto para os metabolitos secundários, como por exemplo carotenóides. Múltiplos genes têm sido descobertos para sintases de GGPP em fungos filamentosos [74].

No *P. lutzii* duas sintases (EC 2.5.1.29⁷) foram identificadas. Para o transcrito PAAG_00717, que expressa a *farnesyltranstransferase subunidade alfa*, não foi possível identificar a forma

⁵*D – manitol 1 – phosphate + NAD(P)⁺ ⇌ β – D – fructofuranose 6 – phosphate + NAD(P)H + H⁺*

⁶Transcrito PAAG_06473

⁷Genes PAAG_00717 e PAAG_00510

em que ocorre. O transcrito PAAG_00510, que expressa a enzima *geranylgeranyl pyrophosphate synthase*, ocorre exclusivamente sob a forma de levedura. Como esta via tem apenas uma reação catalisada por esta enzima, pode-se dizer que quando a reação é catalisada pela enzima *geranylgeranyl pyrophosphate synthase*, a via ocorre exclusivamente sob a forma de levedura no *P. lutzii*. A Figura 6.13 mostra a via extraída da rede metabólica do *P. lutzii*.

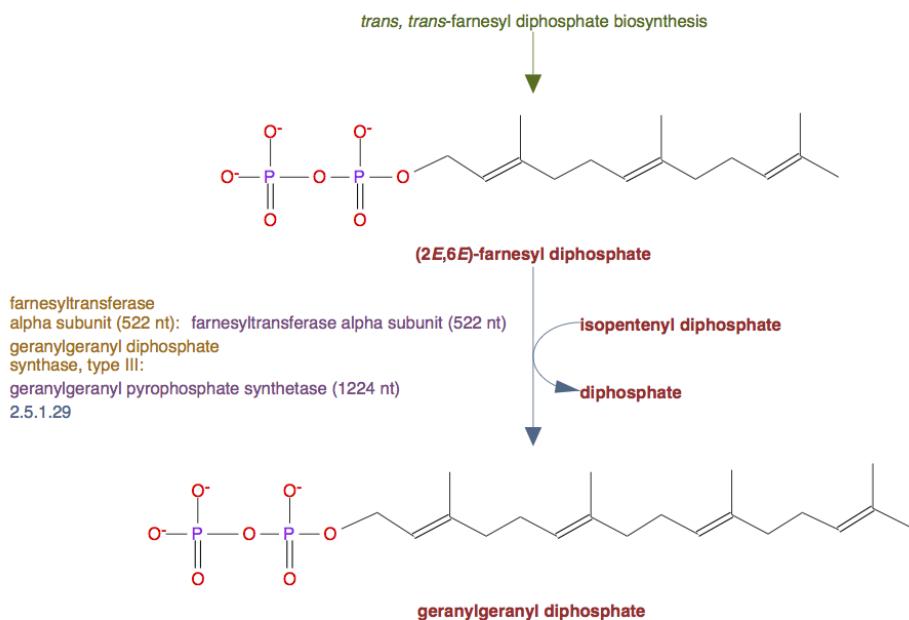


Figura 6.13: Via metabólica da biossíntese de Geranilgeranil di-fosfato.

Capítulo 7

Conclusão

A grande disponibilidade de dados oriundos de genomas, transcritomas, proteomas e metabolomas trouxe novos desafios e possibilidades. O número de genomas fúngicos completamente sequenciados está em franco crescimento e cada vez mais disponíveis. Além disso, dezenas de projetos de sequenciamento estão em andamento [30]. Um grande desafio da pesquisa pós-genoma é explorar o complexo biológico de vias metabólicas de um genoma, tal como sequências de DNA, sequências de proteínas e perfis de expressão gênica [82]. Um caminho para explorar esse complexo biológico é a reconstrução de redes metabólicas de organismos, especialmente de fungos, que são organismos produtores de metabólitos com diversas aplicações. A reconstrução de redes metabólicas é essencial para análises de dados experimentais e estudos computacionais em Biologia de Sistemas [20]. Integrar o conhecimento em vias metabólicas de uma rede em escala genômica permite uma melhor compreensão de como os componentes individuais no sistema interagem e influenciam as funções das células em geral [24].

As diversas ferramentas e bancos de dados disponíveis para anotação, análise, predição e organização de dados biológicos foram utilizadas e convergiram no método proposto, que se mostrou eficaz e eficiente na reconstrução *in silico* de redes metabólicas em fungos. O método proposto neste trabalho para reconstrução *in silico* de redes metabólicas de fungos, foi aplicado ao fungo *P. lutzii*, escolhido como estudo de caso. O método provou-se eficaz porque foi efetivamente reconstruída a rede metabólica para o fungo *P. lutzii* como estudo de caso e está disponível na web. O método, também é eficiente, porque uma vez alimentado o banco de dados com os dados de entrada, a reconstrução da rede torna-se mais rápida quando comparada ao tempo de um ano ou mais como indicado em Förster *et al.* [24]. Além disso, o método proposto abrange um *pipeline* específico para metabolismo secundário em fungos. A rede reconstruída avançou o conhecimento das vias em relação a Alvarez [3], parte devido à atualização constante dos dados provenientes dos bancos de dados consultados (KEGG e BioCyc), parte devido à inserção de um *pipeline*

específico para metabolismo secundário.

Em comparação às outras redes reconstruídas para o *P. lutzii* [3, 85], nossa rede difere em dois aspectos. O primeiro é que as demais redes reconstruídas não têm informação explícita sobre a forma dos genes que a compõem (Micélio/Levedura). É fato que essa informação não está disponível para todo o conjunto dos genes, mas é conhecida para um subconjunto muito expressivo. Outro aspecto em que nossa rede difere das demais conhecidas é na cura alcancada com um *pipeline* de anotação abrangente. As reações que compõem as vias podem ser encontradas, sem exceção, no MetaCyc, que é um banco de dados não-redundante de vias metabólicas com evidências experimentais.

Algumas vias do *P. lutzii* tem especial importância, conforme a literatura, e três delas foram selecionadas para um breve discussão, a fim de mostrar a importância de conhecer o metabolismo dos organismos.

7.1 Contribuições

As duas contribuições deste trabalho foram o método para reconstrução *in silico* de redes metabólicas em fungos com *pipeline* específico para metabolismo secundário e o estudo de caso da rede metabólica do *P. lutzii* reconstruída e disponível na web.

Quanto às limitações, o método exige um esforço de instalar e utilizar ferramentas e bancos de dados, o que demanda conhecimento computacional.

7.2 Trabalhos futuros

O método tem algumas lacunas, como o grande número de *gaps* em reações vinculadas ou não à uma via metabólica. No estudo de caso, 310 enzimas das vias metabólicas não foram identificadas no genoma. Uma abordagem mais aprofundada poderia revelar novas enzimas candidatas, por exemplo, o FBA (Flux Balance Analysis), amplamente utilizado para estudar redes bioquímicas. O FBA calcula o fluxo de metabólitos da rede metabólica, tornando possível prever a taxa de crescimento de um organismo ou a taxa de produção de um metabolito biotecnologicamente importante [67].

Outro projeto importante, seria a criação de uma interface amigável para alimentar o banco de dados relacional, bem como o ajuste do seu modelo para suportar mais de um organismo, além da aplicação de técnicas de *tuning* de desempenho no SGBD.

Referências

- [1] *The Molecular Biology of the Yeast Saccharomyces*, chapter 8 -Regulation and Compartmentalization of Lipid Synthesis in Yeast. 2009. 44
- [2] *Fungi: Biology and Applications.* •, 2nd edition, 2011. 44
- [3] P. A. Alvarez. Reconstrução in silico da rede metabólica do fungo Paracoccidioides lutzii. Master's thesis, Programa de Pós-Graduação em Informática - Departamento de Ciência da Computação - Instituto de Ciências Exatas - Universidade de Brasília, 2012. x, 2, 9, 15, 19, 20, 21, 36, 39, 41, 51, 52
- [4] Genoma Anaplasma. Projeto Genoma Anaplasma. <http://www.biomol.unb.br/anaplasma>, acessado em 07 de agosto de 2012. 2
- [5] R. V. Andrade. *Análise do Transcriptoma e da Expressão Diferencial de Genes de Micélio e Levedura de Paracoccidioides brasiliensis*. Tese de doutorado, Programa de Pós-Graduação em Biologia Molecular, Instituto de Ciências Biológicas - UnB, 2006. 15, 16, 18, 19
- [6] F. Arraes and et al. General metabolism of the dimorphic and pathogenic fungus Paracoccidioides brasiliensis. *Genetics and molecular research : GMR*, 4(2):290–308, January 2005. 47
- [7] R. Bailek, A. Ibricevic, A. Fothergill, and D. Begerow. Small subunit ribosomal DNA sequence shows *Paracoccidioides brasiliensis* closely related to blastomyces dermatidis. *J Clin Microbiol*, 38:3190–3193, 2000. 15
- [8] J. M. Berg, J. Tymoczko, and L. Stryer. *Biochemistry*. New York: Freeman, New York, 6nd edition, 2007. x, 4, 6, 7, 8
- [9] J. M. Bergamaschi. Terpenos. <http://www.terpenoil.com.br/tecnologia/terpenos.pdf> acessado em 01 de julho de 2014. 13
- [10] BIOCYC. Biocyc. <http://biocyc.org>, acessado em 26 de setembro de 2012. 19, 40
- [11] BIOFOCO. BIOFOCO. <http://www.biofoco.org/biofoco3/>, acessado em 07 de agosto de 2012. 2
- [12] M. D. Brownleader, J. B. Harborne, and P. M. Dey. Carbohydrate metabolism: primary metabolism of monosaccharides. *Plant Biochemistry*, 1997. 47

- [13] R. Caspi and et al. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic acids research*, 40:D742–53, jan 2012. 23, 24, 34, 38, 40
- [14] H. Chial. DNA sequencing technologies key to the human genome project. *Nature Education*, 2008. 1
- [15] F. X. Cunningham and E. Gantt. Genes and Enzymes of Carotenoid Biosynthesis in Plants. *Annual review of plant physiology and plant molecular biology*, 49:557–583, June 1998. 44
- [16] C. J. Date. *An Introduction to Database Systems*. Pearson Education, EUA, 8nd edition, 2004. 22, 23
- [17] C. A. Desjardins, M. D. Champion, J. W. Holder, A. Muszewska, J. Goldberg, and et al. Comparative genomic analysis of human fungal pathogens causing paracoccidioidomycosis. *PloS Genetics*, 7(10):e1002345, 2011. x, 16, 17, 36
- [18] M M W Etschmann, W Bluemke, D Sell, and J Schrader. Biotechnological production of 2-phenylethanol. *Applied microbiology and biotechnology*, 59(1):1–8, June 2002. 44
- [19] N. D. Fedorova, V. Moktali, and M. H. Medema. *Methods in Molecular Biology*. Springer, xii edition, 2010. 27
- [20] A. M. Feist, M. J. Herrgård, and I. Thiele. Reconstruction of biochemical networks in microorganisms. *Nature Reviews*, 7(2):129–143, 2008. 28, 51
- [21] M. S. S. Felipe and et al. Transcriptome characterization of the dimorphic and pathogenic fungus *Paracoccidioides brasiliensis* by EST analysis. *Yeast*, 20(3):263–271, 2003. 1, 16
- [22] M. S. S. Felipe and et al. Transcriptional profiles of the human pathogenic fungus *Paracoccidioides brasiliensis* in mycelium and yeast cells. *The Journal of Biological Chemistry*, 280:224706–24714, 2005. 1, 16, 36
- [23] W. Fontes and M. S. Castro. *Processos de manutenção da Vida: Estrutura e metabolismo de biomoléculas I e II*. Consórcio Setentrional, SEED, 2007. 4
- [24] J. Förster, I. Famili, and P. Fu. Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. *Genome research*, pages 244–253, 2003. 28, 51
- [25] M. R. P. Fortes, H. A. Miot, C. S. Kurokawa, M. E. A. Marques, and S. A. Marques. Imunologia da paracoccidioidomicose. *Anais Brasileiros de Dermatologia*, 86:516 – 524, 2011. x, 18, 19
- [26] D. Ganjewala and S. Kumar. Advances in cyanogenic glycosides biosynthesis and analyses in plants: A review. *Acta Biologica* . . . , 54(1):1–14, 2010. 44
- [27] GenomaGuaraná. Projeto Genoma Guaraná. <http://www.biomol.unb.br/GR>, acessado em 07 de agosto de 2012. 2

- [28] GenomaPb. Genoma *Paracoccidioides brasiliensis*. <https://helix.biomol.unb.br/Pb>, acessado em 08 de agosto de 2012. 15, 16
- [29] GOLD. Genomes Online Database. <http://www.genomesonline.org>, acessado em 07 de agosto de 2012. 1
- [30] S. Grossetête, B. Labedan, and O. Lespinet. FUNGIpath: a tool to assess fungal metabolic pathways predicted by orthology. *BMC genomics*, 11:81, 2010. 51
- [31] M. Hamid. Potential applications of peroxidases. *Food Chemistry*, 115(4):1177–1186, August 2009. 44
- [32] B. Heavner and et al. Yeast 5 - an expanded reconstruction of the *Saccharomyces cerevisiae* metabolic network. *BMC systems biology*, 6(1):55, 2012. 28
- [33] HGP. Human Genome Project. http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml, acessado em 07 de agosto de 2012. 1
- [34] D. Hoffmeister and N. P. Keller. Natural products of filamentous fungi: enzymes, genes, and their regulation. *The Royal Society of Chemistry*, 24:393–416, 2007. 12, 26
- [35] S. Hunter and et al. InterPro in 2011: new developments in the family and domain prediction database. *Nucleic acids research*, 40:D306–12, jan 2012. 26, 35, 39
- [36] BROAD Institute. Broad Institute of MIT and Harvard. <http://www.broadinstitute.org/>, acessado em 25 de setembro de 2012. 16
- [37] International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature*, 409:860–921, 2001. 1
- [38] IUBMB. International Union of Biochemistry and Molecular Biology. <http://www.iubmb.org>, acessado em 21 de setembro de 2012. 5
- [39] S. G. Júnior and et al. Analysis of membrane protein genes in a Brazilian isolate of *Anaplasma marginale*. *Memórias do Instituto Oswaldo Cruz*, 105:843–849, 2010. 2
- [40] M. Kanehisa and et al. Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic acids research*, 42:D199–205, January 2014. 24, 40
- [41] M. Kanehisa and S. Goto. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research*, 28(1):27–30, jan 2000. 24, 34, 38, 40
- [42] P. D. Karp and et al. Pathway Tools version 13.0: integrated software for pathway/-genome informatics and systems biology. *Briefings in Bioinformatics*, II(I):40–79, 2009. xi, 29, 35, 40
- [43] P. D. Karp and et al. Pathway Tools version 13.0: integrated software for pathway/-genome informatics and systems biology. *Briefings in bioinformatics*, 11(1):40–79, 2010. 19, 29, 35, 39

- [44] P. D. Karp, M. Latendresse, and R. Caspi. The pathway tools pathway prediction algorithm. *Standards in genomic sciences*, 5(3):424–9, 2011. 29
- [45] K. Kazuharu, Y. Yohei, K. Kosaku, Y. Nakayama, and M. Tomita. GEM System: automatic prototyping of cell-wide metabolic pathway models from genomes. *BMC bioinformatics*, 7:168, January 2006. 28
- [46] KEGG. Kyoto Encyclopedia of Genes and Genomes. <http://www.genome.jp/kegg/>, acessado em 12 de setembro de 2012. x, 19, 25, 40
- [47] N. P. Keller, G. Turner, and J. W. Bennett. Fungal secondary metabolism - from biochemistry to genomics. *Nature Reviews - Microbiology*, 3:937–947, 2005. x, 12, 13, 14
- [48] G S Kelly. Squalene and its potential clinical uses. *Alternative medicine review : a journal of clinical therapeutic*, 4(1):29–36, February 1999. 44
- [49] N. Khaldi and et al. SMURF: genomic mapping of fungal secondary metabolite clusters. *Fungal Genet Biol*, 47(9):736–741, 2010. 11, 12, 27, 34, 38
- [50] H. L. Kornberg and H. A. Krebs. Synthesis of cell constituents from C2-units by a modified tricarboxylic acid cycle. *Nature*, (179):988–991, 1957. 47
- [51] Florian Kugler, Stephan Graneis, Florian C Stintzing, and Reinhold Carle. Studies on betaxanthin profiles of vegetables and fruits from the Chenopodiaceae and Cactaceae. *Zeitschrift für Naturforschung. C, Journal of biosciences*, 62(5-6):311–8, 2007. 44
- [52] C. S. Lacaz. *Paracoccidioides brasiliensis*, morphology, evolutionary cycle, maintenance during saprophytic life, biology, virulence, taxonomy. *Boca Raton, CRC Press*, pages 13–25, 1994. 15
- [53] M. C. Lorenz and G. R. Fink. The glyoxylate cycle is required for fungal virulence. *Nature*, 412(6842):83–6, July 2001. 47
- [54] Robert H. M. Inositol phospholipids and cell surface receptor function. *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes*, 415(1):81 – 147, 1975. 44
- [55] M. Magrane and Uniprot Consortium. UniProt Knowledgebase: a hub of integrated protein data. *Database : the journal of biological databases and curation*, 2011:bar009, jan 2011. 23, 35, 39
- [56] A. Marchler-Bauer and et al. CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic acids research*, 39:D225–9, jan 2011. 26, 35, 39
- [57] M. H. Medema and et al. antiSMASH: Rapid identifications, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Research*, 39:339–346, 2011. x, 12, 26, 27, 34, 38
- [58] METACYC. Encyclopedia of Metabolic Pathways. <http://www.metacyc.org>, acessado em 12 de setembro de 2012. x, 10, 29

- [59] G. Michal and D. Schomburg. *Biochemical Pathways*. Jhon Wiley Sons Inc., Hoboken, New jersey, 2nd edition, 2012. 7, 8, 9, 11
- [60] P. V. Milreu. Análise de nutrientes utilizando redes metabólicas. Dissertação de Mestrado, Departamento de Computação e Estatística da Universidade Federal de Mato Grosso do Sul., 2008. 4, 5
- [61] Y. Moriya, M. Itoh, S. Okuda, A. C. Yoshizawa, and M. Kanehisa. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic acids research*, 35:W182–5, jul 2007. 26, 35, 39
- [62] MS. Portal da Saúde - Ministério da Saúde, Brasil. http://portal.saude.gov.br/portal/saude/profissional/area.cfm?id_area=1662, acessado em 19 de agosto de 2012. 18, 19
- [63] D. L. Nelson and M. C. Michael. *Princípios de Bioquímica de Lehninger - Ed. Comemorativa 25 Anos*. Artmed, São Paulo - SP, 5nd edition, 2010. 4, 5, 6, 7, 8, 11, 13
- [64] W. G. Niehaus and T. Flynn. Regulation of mannitol biosynthesis and degradation by Cryptococcus neoformans. *Journal of bacteriology*, 176(3):651–5, February 1994. 48
- [65] L. R. Nunes and et al. Transcriptome analysis of *Paracoccidioides brasiliensis* cells undergoing mycelium-to-yeast transition. *Eukariotic Cell*, 4(2):2115–2128, 2003. 16
- [66] University of Sydney. Fungal biology. <http://bugs.bio.usyd.edu.au/learning/resources/Mycology/>, acessado em 01 de julho de 2013. x, 11, 12, 26
- [67] J. Orth, I. Thiele, and B. ØPalsson. What is flux balance analysis? *Nature biotechnology*, 28(3):245–8, March 2010. 52
- [68] T. Pfau, N. Nils, and O. Ebenhöh. Systems approaches to modelling pathways and networks. *Briefings in functional genomics*, 10(5):266–79, September 2011. 28
- [69] Gerhard Sandmann, Shinishi Takaichi, and Paul D Fraser. C(35)-apocarotenoids in the yellow mutant Neurospora crassa YLO. *Phytochemistry*, 69(17):2886–90, December 2008. 44
- [70] M. Scheer and et al. BRENDA, the enzyme information system in 2011. *Nucleic acids research*, 39:D670–6, jan 2011. 23, 35, 39
- [71] J. Setubal and J. Meidanis. *Introduction to Computacional Molecular Biology*. PWS Publishing Company, Pacific Groove, CA, 1997. 1
- [72] H. Shawkat, M. M. Westwood, and A. Mortimer. Mannitol: a review of its clinical uses. *Continuing Education in Anaesthesia, Critical Care & Pain*, 12(2):82–85, January 2012. 44
- [73] A. J. G. Simpson and et al. The genome sequence of the plant pathogen *Xylella fastidiosa*. *Nature*, 406:151–157, 2000. 1

- [74] S. Singkaravanit, H. Kinoshita, F. Ihara, and T. Nihira. Geranylgeranyl diphosphate synthase genes in entomopathogenic fungi. *Applied microbiology and biotechnology*, 85(5):1463–72, February 2010. 49
- [75] A. Splendore. Un affezione micotica con localizzazione nella mucosa della bocca, osservata in brazile, determinata da fungi appartenenti alla tribu degli exoascei (*Zygomonema brasiliensi*, n. sp.). *Roma, G, Bertero*, in onore del prof. Angelo Celli nel 25º anno di insegnamento(3421-458), 1912. 15
- [76] M. Teixeira and et al. Phylogenetic analysis reveals a high level of speciation in the Paracoccidioides genus. *Molecular phylogenetics and evolution*, 52(2):273–83, August 2009. 2, 16
- [77] R. Theodoro and et al. Genus paracoccidioides: Species recognition and biogeographic aspects. *PloS one*, 7(5):e37694, January 2012. x, 17
- [78] J. S. Tkacz and L. Lange. *Advances in Fungal Biotechnology for Industry, Agriculture, and Medicine*. Springer, 1nd edition, 2004. 44, 49
- [79] D. Umeno and F. H. Arnold. Evolution of a pathway to novel long-chain carotenoids. *Journal of bacteriology*, 186(5):1531–1536, 2004. 49
- [80] H. Véllez, N. J. Glassbrook, and M. E. Daub. Mannitol metabolism in the phytopathogenic fungus *Alternaria alternata*. *Fungal genetics and biology : FG & B*, 44(4):258–68, April 2007. 48
- [81] J. C. Venter and et al. The Sequence of the Human Genome. *Science*, 291(5507), 2011. 1
- [82] H. Wang, H. H. Lu, and T. Chueh. Constructing Biological Pathways by a Two-Step Counting Approach. *PLoS ONE*, 6(6):10, 2011. 51
- [83] J. D. Watson and F. H. C. Crick. A structure for deoxyribose nucleic acid. *Nature*, 7:737–738, 1953. 1
- [84] J. D. York. A Phospholipase C-Dependent Inositol Polyphosphate Kinase Pathway Required for Efficient Messenger RNA Export. *Science*, 285(5424):96–100, July 1999. 44
- [85] J. Zucker. Fungycyc. <http://fungycyc.broadinstitute.org/PAAG/organism-summary>, acessado em 12 de janeiro de 2014. 41, 52
- [86] P. C. S. Ângelo and et al. Guarana (*Paulliniacupanavar sorbilis*), ananciently consumed stimulant from the Amazon rain forest: the seeded-fruit transcriptome. *Plant Cell Reports*, 27:117–124, 2005. 2

Anexo A

Tabela com as vias de metabolismo secundário identificadas

Vias de degradação de metabólitos secundários

Classificação	Via metabólica
Degradation-Utilization-Assimilation → Secondary Metabolites Degradation → Nitrogen Containing Secondary Compounds Degradation → Alkaloids Degradation	betanidin degradation
Degradation-Utilization-Assimilation → Secondary Metabolites Degradation → Nitrogen Containing Secondary Compounds Degradation → Nitrogen Containing	taxiphyllin bioactivation
Degradation-Utilization-Assimilation → Secondary Metabolites Degradation → Sugar Derivatives Degradation → Cyclitols Degradation	ID-myo-inositol hexakisphosphate biosynthesis I (from Ins(1,4,5)P3)
Degradation-Utilization-Assimilation → Secondary Metabolites Degradation → Sugar Derivatives Degradation → N-acetylglucosamine degradation	N-acetylglucosamine degradation I
Degradation-Utilization-Assimilation → Secondary Metabolites Degradation → Sugar Derivatives Degradation → Sugar Alcohols Degradation	mannitol degradation I

Vias de síntese de metabólitos secundários

Classificação	Via metabólica
Biosynthesis → Secondary Metabolites Biosynthesis → Nitrogen-Containing Secondary Compounds Biosynthesis → Alkaloids Biosynthesis → Betalaine Al-	betaxanthin biosynthesis
Biosynthesis → Secondary Metabolites Biosynthesis → Nitrogen-Containing Secondary Compounds Biosynthesis → Alkaloids Biosynthesis → Betalaine Al-	betaxanthin biosynthesis (via dopa- mine)
Biosynthesis → Secondary Metabolites Biosynthesis → Phenylpropanoid Derivatives Biosynthesis	phenylethanol biosynthesis
Biosynthesis → Secondary Metabolites Biosynthesis → Sugar Derivatives Biosynthesis → Cyclitols Biosynthesis	D-myo-inositol (1,4,5)- trisphosphate biosynthesis
Biosynthesis → Secondary Metabolites Biosynthesis → Sugar Derivatives Biosynthesis → Cyclitols Biosynthesis	D-myo-inositol (1,4,5)- trisphosphate degradation
Biosynthesis → Secondary Metabolites Biosynthesis → Sugar Derivatives Biosynthesis → Cyclitols Biosynthesis → myo-Inositol Biosynthesis	myo-inositol biosynthesis
Biosynthesis → Secondary Metabolites Biosynthesis → Sugar Derivatives Biosynthesis → Cyclitols Biosynthesis → Phytate Biosynthesis	ID-myo-inositol hexakisphosphate biosynthesis I (from Ins(1,4,5)P3)
Biosynthesis → Secondary Metabolites Biosynthesis → Terpenoids Biosynthesis → Carotenoids Biosynthesis	zeaxanthin, antheraxanthin and violaxanthin interconversion
Biosynthesis → Secondary Metabolites Biosynthesis → Terpenoids Biosynthesis → Diterpenoids Biosynthesis	superpathway of geranylgeranyl-diphosphate biosynthesis I (via mevalonate)
Biosynthesis → Secondary Metabolites Biosynthesis → Terpenoids Biosynthesis → Hemiterpenes Biosynthesis → Isopentenyl Diphosphate Biosynthesis	mevalonate pathway I
Biosynthesis → Secondary Metabolites Biosynthesis → Terpenoids Biosynthesis → Monoterpeneoids Biosynthesis → Linalool Biosynthesis	superpathway of linalool biosynthesis
Biosynthesis → Secondary Metabolites Biosynthesis → Terpenoids Biosynthesis → Tetraterpenoids Biosynthesis	zeaxanthin, antheraxanthin and violaxanthin interconversion
Biosynthesis → Secondary Metabolites Biosynthesis → Terpenoids Biosynthesis → Tetraterpenoids Biosynthesis → Apocarotenoids Biosynthesis	neurosporaxanthin biosynthesis
Biosynthesis → Secondary Metabolites Biosynthesis → Terpenoids Biosynthesis → Triterpenoids Biosynthesis	epoxysqualene biosynthesis
Biosynthesis → Secondary Metabolites Biosynthesis → Terpenoids Biosynthesis → Triterpenoids Biosynthesis	plant sterol biosynthesis

Anexo B

Tabela de reações com gaps e as vias
metabólicas onde ocorrem

Tabela de reações com gaps e as vias metabólicas onde ocorrem

Reação	Gap (EC)	Via metabólica que precisa da reação
all-trans-lycopene + an oxidized electron acceptor → all-trans-3,4-didehydrolycopene + a reduced electron acceptor	1.3.99.30	neurosporaxanthin biosynthesis
all-trans-3,4-didehydrolycopene → apo-4'-lycopenal	[none]	neurosporaxanthin biosynthesis
apo-4'-lycopenate → neurosporaxanthin	2.5.1.-	neurosporaxanthin biosynthesis
all-trans-3,4-didehydrolycopene → torulene	2.5.1.-	neurosporaxanthin biosynthesis
torulene + oxygen → 4'-apo-β-carotenal + 3-methyl-2-butenal	1.13.11.59	neurosporaxanthin biosynthesis
4'-apo-β-carotenal + NAD+ + H2O → neurosporaxanthin + NADH + 2 H+	1.2.1.82	neurosporaxanthin biosynthesis
all-trans-hexaprenyl diphosphate + 4-aminobenzoate → 3-hexaprenyl-4-aminobenzoate + diphosphate	2.5.1.-	ubiquinol-6 biosynthesis from 4-aminobenzoate (eukaryotic)
3-hexaprenyl-4-aminobenzoate + NADPH + oxygen + H+ → 3-hexaprenyl-4-amino-5-hydroxybenzoate + NADP+ + H2O	1.14.13.-	ubiquinol-6 biosynthesis from 4-aminobenzoate (eukaryotic)
3-hexaprenyl-4-amino-5-hydroxybenzoate + S-adenosyl-L-methionine → 3-hexaprenyl-4-amino-5-methoxybenzoate + S-adenosyl-L-homocysteine + H+	2.1.1.-	ubiquinol-6 biosynthesis from 4-aminobenzoate (eukaryotic)
3-hexaprenyl-4-amino-5-methoxybenzoate + CO2 → 2-hexaprenyl-6-methoxyaminobenzene + CO2	4.1.1.-	ubiquinol-6 biosynthesis from 4-aminobenzoate (eukaryotic)
2-hexaprenyl-6-methoxyaminobenzene + NADPH + oxygen + H+ → 2-hexaprenyl-6-methoxyaminophenol + NADP+ + H2O	1.14.13.-	ubiquinol-6 biosynthesis from 4-aminobenzoate (eukaryotic)
2-hexaprenyl-6-methoxyaminophenol + S-adenosyl-L-methionine → 2-hexaprenyl-3-methyl-6-methoxyaminophenol + S-adenosyl-L-homocysteine + H+	2.1.1.-	ubiquinol-6 biosynthesis from 4-aminobenzoate (eukaryotic)
2-hexaprenyl-6-methoxyaminophenol + H+ + H2O → 2-methoxy-6-all trans-hexaprenyl-2-methoxy-1,4-benzoquinol + ammonium	3.5.99.-	ubiquinol-6 biosynthesis from 4-aminobenzoate (eukaryotic)
2-hexaprenyl-3-methyl-6-methoxyaminophenol + H+ + H2O → 6-methoxy-3-methyl-2-all-trans-hexaprenyl-1,4-benzoquinol + ammonium	3.5.99.-	ubiquinol-6 biosynthesis from 4-aminobenzoate (eukaryotic)
di-homo-γ-linolenate + a reduced electron acceptor + oxygen → arachidonate + an oxidized electron acceptor + 2 H2O	1.14.19.-	docosahexanoate biosynthesis I
malonyl-CoA + γ-linolenoyl-CoA + H+ → 3-oxo-dihomo γ-linolenoyl-CoA + CO2 + coenzyme A	2.3.1.-	docosahexanoate biosynthesis I
3-hydroxy-dihomo γ-linolenoyl-CoA → dihomo γ-linolenoyl-2-enoyl-CoA + H2O	4.2.1.-	docosahexanoate biosynthesis I
dihomo γ-linolenoyl-2-enoyl-CoA + NADH + H+ → dihomo γ-linolenoyl-CoA + NAD+	1.3.1.-	docosahexanoate biosynthesis I
3-oxo-dihomo γ-linolenoyl-CoA + NAD(P)H + H+ → 3-hydroxy-dihomo γ-linolenoyl-CoA + NAD(P)+	1.1.1.-	docosahexanoate biosynthesis I
dihomo γ-linolenoyl-CoA + H2O → di-homo-γ-linolenate + coenzyme A + H+	3.1.2.2	docosahexanoate biosynthesis I
arachidonate + a reduced electron acceptor + oxygen → eicosapentaenoate + an oxidized electron acceptor + 2 H2O	1.14.19.-	docosahexanoate biosynthesis I
L-saccharopine + NADP+ + H2O ← L-lysine + 2-oxoglutarate + NADPH + H+	1.5.1.8	lysine degradation II
L-saccharopine + NAD+ + H2O → L-glutamate + (S)-2-amino-6-oxohexanoate + NADH + H+	1.5.1.9	lysine degradation II
2-oxoadipate + coenzyme A + NAD+ → CO2 + glutaryl-CoA + NADH	[none]	lysine degradation II
selenite + AMP + glutathione disulfide + 3 H+ ← adenosine 5'-phosphoselenate + 2 glutathione	1.8.4.9	selenate reduction
glutathioselenol + NADPH + 2 H+ → hydrogen selenide + glutathione + NADP+	[none]	selenate reduction
selenodiglutathione + NADPH → glutathioselenol + glutathione + NADP+	[none]	selenate reduction
sphingosine + ATP → sphingosine 1-phosphate + ADP + H+	[none]	sphingosine and sphingosine-1-phosphate metabolism
sphingosine 1-phosphate → 2-trans-hexadecenal + phosphoryl-ethanolamine	4.3.2.-	sphingosine and sphingosine-1-phosphate metabolism
sphingosine 1-phosphate + H2O → sphingosine + phosphate	[none]	sphingosine and sphingosine-1-phosphate metabolism
palmitaldehyde + NADP+ ← 2-trans-hexadecenal + NADPH + H+	1.3.1.27	sphingosine and sphingosine-1-phosphate metabolism
UDP-α-D-galactose + an N-acetyl-β-D-glucosaminyl-R → β-D-galactosyl-(1→4)-N-acetyl-β-D-glucosaminyl-R + UDP + H+	2.4.1.38	terminal O-glycans residues modification
GDP-L-fucose + β-D-galactosyl-(1→4)-N-acetyl-β-D-glucosaminyl-R → Lewis x epitope + GDP + H+	2.4.1.152	terminal O-glycans residues modification
UDP-α-D-galactose + β-D-galactosyl-(1→4)-N-acetyl-β-D-glucosaminyl-R → α-D-galactosyl-(1,3)-β-D-galactosyl-(1,4)-N-acetyl-D-glucosaminyl-R + UDP + H+	2.4.1.87	terminal O-glycans residues modification
N-acetylneuraminyl-(2→3)-β-D-galactosyl-(1→4)-N-acetyl-β-D-glucosaminyl-R + CMP + H+	2.4.99.6	terminal O-glycans residues modification
N-acetyl-β-D-glucosaminyl-1,3-β-D-galactosyl-1,4-N-acetyl-β-D-glucosaminyl-R + UDP + H+	2.4.1.149	terminal O-glycans residues modification
CMP-N-acetyl-β-neuraminate + β-D-galactosyl-(1→4)-N-acetyl-β-D-glucosaminyl-R → α-N-acetylneuraminyl-2,6-β-D-galactosyl-(1→4)-N-acetyl-β-D-glucosaminyl-R + CMP + H+	2.4.99.1	terminal O-glycans residues modification

UDP- α -D-galactose + N-acetyl- β -D-glucosaminyl-1,3- β -D-galactosyl-1,4-N-acetyl- β -D-glucosaminyl-R → i antigen + UDP + H+	2.4.1.38	terminal O-glycans residues modification
acetyl- β -D-glucosaminyl-(1,3)- β -D-galactosyl-(1,4)-N-acetyl- β -D-glucosaminyl-R → I antigen + UDP + H+	2.4.1.38	terminal O-glycans residues modification
succinyl-CoA + acetyl-CoA ← 3-oxoadipyl-CoA + coenzyme A	2.3.1.174	phenylacetate degradation I (aerobic)
phenylacetyl-CoA + NADPH + oxygen + H+ → 2-(1,2-epoxy-1,2-dihydrophenyl)acetyl-CoA + NADP+ + H2O	1.14.13.149	phenylacetate degradation I (aerobic)
2,3-didehydroadipyl-CoA + acetyl-CoA ← 3-oxo-5,6-didehydrosuberyl-CoA + coenzyme A	2.3.1.223	phenylacetate degradation I (aerobic)
2-(1,2-epoxy-1,2-dihydrophenyl)acetyl-CoA → 2-oxepin-2(3H)-ylideneacetyl-CoA	5.3.3.18	phenylacetate degradation I (aerobic)
2-oxepin-2(3H)-ylideneacetyl-CoA + H2O → 3-oxo-5,6-didehydrosuberyl-CoA semialdehyde	3.3.2.12	phenylacetate degradation I (aerobic)
3-oxo-5,6-didehydrosuberyl-CoA semialdehyde + NADP+ + H2O → 3-oxo-5,6-didehydrosuberyl-CoA + NADPH + 2 H+	1.17.1.7	phenylacetate degradation I (aerobic)
phenylacetate + ATP + coenzyme A → phenylacetyl-CoA + AMP + diphosphate	6.2.1.30	phenylacetate degradation I (aerobic)
galactitol + NAD+ ↔ L-xylo-3-hexulose + NADH + H+	1.1.1.-	galactose degradation IV
L-xylo-3-hexulose + a reduced electron acceptor ↔ D-sorbitol + an oxidized electron acceptor	[none]	galactose degradation IV
galactitol + an oxidized electron acceptor ↔ keto-L-sorbose + a reduced electron acceptor	[none]	galactose degradation IV
D-sorbitol + NADP+ ← keto-L-sorbose + NADPH + H+	1.1.1.289	galactose degradation IV
L-xylo-3-hexulose ↔ keto-L-sorbose	[none]	galactose degradation IV
β -D-fructofuranose + ATP → β -D-fructofuranose 6-phosphate + ADP + H+	2.7.1.4	galactose degradation IV
a 1-phosphatidyl-1D-myo-inositol 4,5-bisphosphate + H2O → a 1-phosphatidyl-1D-myo-inositol 5-phosphate + phosphate	3.1.3.78	3-phosphoinositide degradation
a 1-phosphatidyl-1D-myo-inositol 4,5-bisphosphate + H2O → a 1-phosphatidyl-1D-myo-inositol 4-phosphate + phosphate	3.1.3.36	3-phosphoinositide degradation
1-phosphatidyl-1D-myo-inositol 3,4,5-trisphosphate + H2O → a 1-phosphatidyl-1D-myo-inositol 3,4-bisphosphate + phosphate	3.1.3.86	3-phosphoinositide degradation
a 1-phosphatidyl-1D-myo-inositol 3,4-bisphosphate + H2O → a 1-phosphatidyl-1D-myo-inositol 3-phosphate + phosphate	3.1.3.66	3-phosphoinositide degradation
a 1-phosphatidyl-1D-myo-inositol 3-phosphate + H2O → an L-1-phosphatidyl-inositol + phosphate	3.1.3.64	3-phosphoinositide degradation
a 1-phosphatidyl-1D-myo-inositol 4-phosphate + H2O → an L-1-phosphatidyl-inositol + phosphate	3.1.3.-	3-phosphoinositide degradation
a 1-phosphatidyl-1D-myo-inositol 5-phosphate + H2O → an L-1-phosphatidyl-inositol + phosphate	3.1.3.-	3-phosphoinositide degradation
4-hydroxy-2-nonenal-glutathione conjugate + H2O → 4-hydroxy-2-nonenal-[Cys-Gly] conjugate + L-glutamate	3.4.19.9	4-hydroxy-2-nonenal detoxification
4-hydroxy-2-nonenal-[Cys-Gly] conjugate + H2O → 4-hydroxy-2-nonenal-[L-Cys] conjugate + glycine	3.4.11.2	4-hydroxy-2-nonenal detoxification
4-hydroxy-2-nonenal-[L-Cys] conjugate + acetyl-CoA → 4-hydroxy-2-nonenal-N-acetyl-L-cysteine + coenzyme A + H+	2.3.1.80	4-hydroxy-2-nonenal detoxification
2-methoxy-6-(all-trans-nonaprenyl)phenol + 2 H+ + oxygen → 2-methoxy-6-all-trans-nonaprenyl-2-methoxy-1,4-benzoquinol + H2O	1.14.3.-	ubiquinol-9 biosynthesis (eukaryotic)
S-adenosyl-L-methionine + 3-demethylubiquinol-9 → S-adenosyl-L-homocysteine + ubiquinol-9 + H+	2.1.1.64	ubiquinol-9 biosynthesis (eukaryotic)
6-methoxy-3-methyl-2-all-trans-nonaprenyl-1,4-benzoquinol + a reduced electron acceptor + oxygen → an oxidized electron acceptor + 3-demethylubiquinol-9 + H2O	1.14.-.-	ubiquinol-9 biosynthesis (eukaryotic)
3-nonaprenyl-4-hydroxybenzoate + NADPH + oxygen + H+ → 3,4-dihydroxy-5-all-trans-nonaprenylbenzoate + NADP+ + H2O	1.14.13.-	ubiquinol-9 biosynthesis (eukaryotic)
3-methoxy-4-hydroxy-5-all-trans-nonaprenylbenzoate + H+ → 2-methoxy-6-(all-trans-nonaprenyl)phenol + CO2	4.1.1.-	ubiquinol-9 biosynthesis (eukaryotic)
all-trans-nonaprenyl diphosphate + 4-hydroxybenzoate → 3-nonaprenyl-4-hydroxybenzoate + diphosphate	2.5.1.39	ubiquinol-9 biosynthesis (eukaryotic)
2-methoxy-6-(all-trans-heptaprenyl)phenol + NADPH + oxygen + H+ → 2-methoxy-6-all-trans-heptaprenyl-2-methoxy-1,4-benzoquinol + NADP+ + H2O	1.14.13.-	ubiquinol-7 biosynthesis (eukaryotic)
6-methoxy-3-methyl-2-all-trans-heptaprenyl-1,4-benzoquinol + NADPH + oxygen + H+ → 3-demethylubiquinol-7 + NADP+ + H2O	1.14.13.-	ubiquinol-7 biosynthesis (eukaryotic)
all-trans-heptaprenyl diphosphate + 4-hydroxybenzoate → 3-heptaprenyl-4-hydroxybenzoate + diphosphate	2.5.1.39	ubiquinol-7 biosynthesis (eukaryotic)
S-adenosyl-L-methionine + 3-demethylubiquinol-7 → S-adenosyl-L-homocysteine + ubiquinol-7 + H+	2.1.1.64	ubiquinol-7 biosynthesis (eukaryotic)
3-methoxy-4-hydroxy-5-all-trans-heptaprenylbenzoate + H+ → 2-methoxy-6-(all-trans-heptaprenyl)phenol + CO2	4.1.1.-	ubiquinol-7 biosynthesis (eukaryotic)
3-heptaprenyl-4-hydroxybenzoate + NADPH + oxygen + H+ → 3,4-dihydroxy-5-all-trans-heptaprenylbenzoate + NADP+ + H2O	1.14.13.-	ubiquinol-7 biosynthesis (eukaryotic)
2-methoxy-6-(all-trans-decaprenyl)phenol + NADPH + oxygen + H+ → 2-methoxy-6-all-trans-decaprenyl-2-methoxy-1,4-benzoquinol + NADP+ + H2O	1.14.13.-	ubiquinol-10 biosynthesis (eukaryotic)
6-methoxy-3-methyl-2-all-trans-decaprenyl-1,4-benzoquinol + a reduced electron acceptor + oxygen → an oxidized electron acceptor + 3-demethylubiquinol-10 + H2O	1.14.-.-	ubiquinol-10 biosynthesis (eukaryotic)
S-adenosyl-L-methionine + 3-demethylubiquinol-10 → S-adenosyl-L-homocysteine + ubiquinol-10 + H+	2.1.1.64	ubiquinol-10 biosynthesis (eukaryotic)
all-trans-decaprenyl diphosphate + 4-hydroxybenzoate → 3-decaprenyl-4-hydroxybenzoate + diphosphate	2.5.1.39	ubiquinol-10 biosynthesis (eukaryotic)

3-decaprenyl-4-hydroxybenzoate + NADPH + oxygen + H+ → 3,4-dihydroxy-5-all-trans-decaprenylbenzoate + NADP+ + H2O	1.14.13.-	ubiquinol-10 biosynthesis (eukaryotic)
3-methoxy-4-hydroxy-5-all-trans-decaprenylbenzoate + H+ → 2-methoxy-6-(all-trans-decaprenyl)phenol + CO2	4.1.1.-	ubiquinol-10 biosynthesis (eukaryotic)
eicosatrienoate + a reduced electron acceptor + oxygen → eicosatetraenoate + an oxidized electron acceptor + 2 H2O	[none]	docosahexanoate biosynthesis II
eicosatetraenoate + a reduced electron acceptor + oxygen → eicosapentaenoate + an oxidized electron acceptor + 2 H2O	1.14.19.-	docosahexanoate biosynthesis II
all-cis-7,10,13,16,19-docosapentaenoate + a reduced electron acceptor + oxygen → docosahexanoate + an oxidized electron acceptor + 2 H2O	[none]	docosahexanoate biosynthesis I, docosahexanoate biosynthesis II
eicosatrienoyl-2-enoyl CoA + NADH + H+ → eicosatrienoyl-CoA + NAD+	1.3.1.-	docosahexanoate biosynthesis II
3-hydroxy-eicosatrienoyl-CoA → eicosatrienoyl-2-enoyl CoA + H2O	4.2.1.-	docosahexanoate biosynthesis II
eicosatrienoyl-CoA + H2O → eicosatrienoate + coenzyme A + H+	3.1.2.-	docosahexanoate biosynthesis II
α-linolenoyl-CoA + malonyl-CoA + H+ → 3-oxo-eicosatrienoyl-CoA + CO2 + coenzyme A	[none]	docosahexanoate biosynthesis II
eicosapentaenoyl-CoA + malonyl-CoA + H+ → 3-oxo-docosapentaenoyl-CoA + CO2 + coenzyme A	[none]	docosahexanoate biosynthesis I, docosahexanoate biosynthesis II
3-hydroxy-docosapentaenoyl-CoA → (5Z,8Z,11Z,14Z,17Z)-eicosapenta-5,8,11,14,17-enoyl-CoA + H2O	4.2.1.-	docosahexanoate biosynthesis I, docosahexanoate biosynthesis II
(5Z,8Z,11Z,14Z,17Z)-eicosapenta-5,8,11,14,17-enoyl-CoA + 2 H+ → docosapenta-7,10,13,16,19-enoyl-CoA	1.3.1.-	docosahexanoate biosynthesis I, docosahexanoate biosynthesis II
docosapenta-7,10,13,16,19-enoyl-CoA + H2O → all-cis-7,10,13,16,19-docosapentaenoate + coenzyme A + H+	3.1.2.-	docosahexanoate biosynthesis I, docosahexanoate biosynthesis II
9-cis, 11-trans-octadecadienoyl-CoA + 2 coenzyme A + 2 NAD+ + 2 oxygen + 2 H2O → 5-cis, 7-trans-tetradecadienoyl-CoA + 2 acetyl-CoA + 2 hydrogen peroxide + 2 NADH + 2 H+	[none]	9-cis, 11-trans-octadecadienoyl-CoA degradation (isomerase-dependent, yeast)
2-trans, 5-cis, 7-trans-tetradecatrienoyl-CoA + H2O → 3-hydroxy- 5-cis, 7-trans-tetradecadienoyl-CoA	4.2.1.119	9-cis, 11-trans-octadecadienoyl-CoA degradation (isomerase-dependent, yeast)
3-hydroxy- 5-cis, 7-trans-tetradecadienoyl-CoA + NAD+ → 5-cis, 7-trans-3-oxo-tetradecadienoyl-CoA + NADH + H+	1.1.1.-	9-cis, 11-trans-octadecadienoyl-CoA degradation (isomerase-dependent, yeast)
3-cis, 5-trans-dodecadienoyl-CoA → 2-trans, 5-trans-dodecadienoyl-CoA	5.3.3.-	9-cis, 11-trans-octadecadienoyl-CoA degradation (isomerase-dependent, yeast)
2-trans, 5-trans-dodecadienoyl-CoA + H2O → 3-hydroxy-5-trans-dodecenoyl-CoA	4.2.1.119	9-cis, 11-trans-octadecadienoyl-CoA degradation (isomerase-dependent, yeast)
3-hydroxy-5-trans-dodecenoyl-CoA + NAD+ → 5-trans-3-oxo-dodecenoyl-CoA + NADH + H+	1.1.1.-	9-cis, 11-trans-octadecadienoyl-CoA degradation (isomerase-dependent, yeast)
3-trans-decenoyl-CoA → trans-Δ2-decenoyl-CoA	5.3.3.-	9-cis, 11-trans-octadecadienoyl-CoA degradation (isomerase-dependent, yeast)
L-lactaldehyde + NADP+ ↔ methylglyoxal + NADPH + H+	1.1.1.283	methylglyoxal degradation V
L-lactaldehyde + NAD+ + H2O → (S)-lactate + NADH + 2 H+	1.2.1.22	methylglyoxal degradation V
D-myo-inositol (1,4)-bisphosphate + H2O → 1D-myo-inositol 4-monophosphate + phosphate	3.1.3.57	D-myo-inositol (1,4,5)-trisphosphate degradation
D-myo-inositol (1,4,5)-trisphosphate + H2O → D-myo-inositol (1,4)-bisphosphate + phosphate	3.1.3.56	D-myo-inositol (1,4,5)-trisphosphate degradation
2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline + H+ → (S)-(+) -allantoin + CO2	4.1.1.h	urate degradation to allantoin I
5-hydroxyisourate + H2O → 2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline + H+	3.5.2.17	urate degradation to allantoin I
L-arabitol + NAD+ ↔ L-xylulose + NADH + H+	1.1.1.12	L-arabinose degradation II
xylitol + NADP+ ↔ L-xylulose + NADPH + H+	1.1.1.10	L-arabinose degradation II
cellulose + n H2O → n a cellulose	3.2.1.4	cellulose degradation II (fungi)
a cellulose + n H2O → n β-D-cellulobiose	3.2.1.91	cellulose degradation II (fungi)
a phosphatidyl-N-methylethanolamine + S-adenosyl-L-methionine → a phosphatidyl-N-dimethylethanolamine + S-adenosyl-L-homocysteine + H+	2.1.1.71	phosphatidylcholine biosynthesis V
a phosphatidyl-N-dimethylethanolamine + S-adenosyl-L-methionine → a phosphatidylcholine + S-adenosyl-L-homocysteine + H+	2.1.1.71	phosphatidylcholine biosynthesis V
26,27-dehydrozymosterol + methanol + S-adenosyl-L-methionine + H+ → 24-alkyl sterol 2 + S-adenosyl-L-homocysteine + methyl-group	[none]	26,27-dehydrozymosterol metabolism
26,27-dehydrozymosterol + methanol + S-adenosyl-L-methionine + H+ → 24-alkyl sterol 3 + S-adenosyl-L-homocysteine + methyl-group	[none]	26,27-dehydrozymosterol metabolism
4-guanidinobutyrate + H2O → urea + 4-aminobutanoate	3.5.3.7	arginine degradation X (arginine monooxygenase pathway)
L-arginine + oxygen → 4-guanidinobutyramide + CO2 + H2O	1.13.12.1	arginine degradation X (arginine monooxygenase pathway)
2-oxo-4-methylthiobutanoate + H+ → 3-methylthiopropanal + CO2	[none]	methionine degradation III
L-methionine + 2-oxo-3-phenylpropanoate ↔ 2-oxo-4-methylthiobutanoate + L-phenylalanine	2.6.1.28	methionine degradation III
a sphingoid base + ATP → a sphingoid 1-phosphate + ADP + H+	[none]	ceramide degradation
a sphingoid 1-phosphate → phosphoryl-ethanolamine + an aldehyde	[none]	ceramide degradation
α-D-glucose 1-phosphate + H2O → D-glucopyranose + phosphate	3.1.3.10	glucose and glucose-1-phosphate degradation

an ubiquinone + D-glucopyranose[out] → an ubiquinol + D-glucono-1,5-lactone[out]	1.1.5.2	glucose and glucose-1-phosphate degradation
6-methoxy-3-methyl-2-all-trans-octaprenyl-1,4-benzoquinol + a reduced electron acceptor + oxygen → an oxidized electron acceptor + 3-demethylubiquinol-8 + H2O	1.14.-.-	ubiquinol-8 biosynthesis (eukaryotic)
2-methoxy-6-(all-trans-octaprenyl)phenol + NADPH + oxygen + H+ → 2-methoxy-6-all-trans-octaprenyl-2-methoxy-1,4-benzoquinol + NADP+ + H2O	1.14.13.-	ubiquinol-8 biosynthesis (eukaryotic)
all-trans-octaprenyl diphosphate + 4-hydroxybenzoate → 3-octaprenyl-4-hydroxybenzoate + diphosphate	2.5.1.39	ubiquinol-8 biosynthesis (eukaryotic)
chorismate → 4-hydroxybenzoate + pyruvate	4.1.3.40	ubiquinol-8 biosynthesis (eukaryotic)
3-octaprenyl-4-hydroxybenzoate + NADPH + oxygen + H+ → 3,4-dihydroxy-5-all-trans-octaprenylbenzoate + NADP+ + H2O	1.14.13.-	ubiquinol-8 biosynthesis (eukaryotic)
3-methoxy-4-hydroxy-5-all-trans-octaprenylbenzoate + H+ → 2-methoxy-6-(all-trans-octaprenyl)phenol + CO2	4.1.1.-	ubiquinol-8 biosynthesis (eukaryotic)
sphinganine + NADPH + H+ + oxygen → phytosphingosine + NADP+ + H2O	1.14.13.169	sphingolipid biosynthesis (yeast)
an α hydroxypyrophoceramide + an L-1-phosphatidyl-inositol → an inositol-phospho-α hydroxypyrophoceramide + a 1,2-diacyl-sn-glycerol	2.7.1.-	sphingolipid biosynthesis (yeast)
an inositol-phospho-α hydroxypyrophoceramide + GDP-α-D-mannose → a mannosyl-inositol-phospho-α hydroxypyrophoceramide + GDP + H+	2.4.1.-	sphingolipid biosynthesis (yeast)
a mannosyl-inositol-phospho-α hydroxypyrophoceramide + an L-1-phosphatidyl-inositol → a mannosyl-diphosphoinositol-α hydroxypyrophoceramide + a 1,2-diacyl-sn-glycerol	2.7.1.-	sphingolipid biosynthesis (yeast)
a phytoceramide + a ferrocyclochrome b5 + oxygen + 2 H+ → an α hydroxypyrophoceramide + a ferricytochrome b5 + H2O	1.14.19.-	sphingolipid biosynthesis (yeast)
3-hexaprenyl-4-hydroxybenzoate + NADPH + oxygen + H+ → 3,4-dihydroxy-5-all-trans-hexaprenylbenzoate + NADP+ + H2O	1.14.13.-	ubiquinol-6 biosynthesis from 4-hydroxybenzoate (eukaryotic)
3-methoxy-4-hydroxy-5-all-trans-hexaprenylbenzoate + H+ → 2-hexaprenyl-6-methoxyphenol + CO2	4.1.1.-	ubiquinol-6 biosynthesis from 4-hydroxybenzoate (eukaryotic)
2-hexaprenyl-6-methoxyphenol + NADPH + oxygen + H+ → 2-methoxy-6-all-trans-hexaprenyl-2-methoxy-1,4-benzoquinol + NADP+ + H2O	1.14.13.-	ubiquinol-6 biosynthesis from 4-hydroxybenzoate (eukaryotic)
6-methoxy-3-methyl-2-all-trans-hexaprenyl-1,4-benzoquinol + a reduced electron acceptor + oxygen → 3-demethylubiquinol-6 + an oxidized electron acceptor + H2O	1.14.-.-	ubiquinol-6 biosynthesis from 4-aminobenzoate (eukaryotic), ubiquinol-6 biosynthesis from 4-hydroxybenzoate (eukaryotic)
3-demethylubiquinol-6 + S-adenosyl-L-methionine → S-adenosyl-L-homocysteine + ubiquinol-6 + H+	2.1.1.64	ubiquinol-6 biosynthesis from 4-aminobenzoate (eukaryotic), ubiquinol-6 biosynthesis from 4-hydroxybenzoate (eukaryotic)
L-phenylalanine + H+ + oxygen + H2O → phenylacetaldehyde + ammonium + hydrogen peroxide + CO2	[none]	phenylethanol biosynthesis
2-phenylethanol + NADP+ ← phenylacetaldehyde + NADPH + H+	1.1.1.-	phenylethanol biosynthesis
L-phenylalanine + 2 NADPH + 2 oxygen + 2 H+ → (E)-phenylacetaldehyde oxime + CO2 + 2 NADP+ + 3 H2O	1.14.13.124	phenylethanol biosynthesis
(Z)-phenylacetaldehyde oxime + H+ + 2 H2O → phenylacetaldehyde + ammonium + hydrogen peroxide	[none]	phenylethanol biosynthesis
N1-acetylspermine + oxygen + H2O → spermidine + 3-acetamidopropanal + hydrogen peroxide	1.5.3.13	spermine and spermidine degradation I
spermine + oxygen + H2O → spermidine + 3-aminopropanal + hydrogen peroxide	1.5.3.16/1.5.3.17	spermine and spermidine degradation I
N1-acetylspermidine + oxygen + H2O → putrescine + 3-acetamidopropanal + hydrogen peroxide	1.5.3.13/1.5.3.17	spermine and spermidine degradation I
oxaloacetate + propionyl-CoA + H2O → (2S,3S)-2-methylcitrate + coenzyme A + H+	2.3.3.5	2-methylcitrate cycle I
(2R,3S)-2-methylisocitrate → succinate + pyruvate	4.1.3.30	2-methylcitrate cycle I
(2R,3S)-2-methylisocitrate ← cis-2-methylaconitate + H2O	4.2.1.99	2-methylcitrate cycle I
ADP ribose 1"-phosphate + H2O → ADP-D-ribose + phosphate	3.1.3.84	tRNA splicing
ADP ribose 1",2"-cyclic phosphate + H2O → ADP ribose 1"-phosphate + H+	3.1.4.-	tRNA splicing
molecule with a hydroxyl on its 5' end + ATP → a 2'-phospho-[ligated tRNA] + AMP + diphosphate	6.5.1.-	tRNA splicing
octanoyl-ACP + coenzyme A → octanoyl-CoA + a holo-[acyl-carrier protein]	2.8.3.-	lipoate biosynthesis and incorporation (pyruvate dehydrogenase and oxoglutarate dehydrogenase, yeast)
octanoyl-CoA + [pyruvate dehydrogenase E2 lipoyl-carrier protein]-L-lysine → a [pyruvate dehydrogenase E2 protein] N6-octanoyl-L-lysine + coenzyme A + H+	2.3.1.-	lipoate biosynthesis and incorporation (pyruvate dehydrogenase and oxoglutarate dehydrogenase, yeast)
octanoyl-CoA + [2-oxoglutarate dehydrogenase E2 lipoyl-carrier protein]-L-lysine → a [2-oxoglutarate-dehydrogenase E2 protein] N6-octanoyl-L-lysine + coenzyme A + H+	2.3.1.-	lipoate biosynthesis and incorporation (pyruvate dehydrogenase and oxoglutarate dehydrogenase, yeast)
an S-sulfanyl-[L-cysteine desulfurase] + a [disordered-form [Fe-S] cluster scaffold protein] → an S-sulfanyl-[cysteine desulfurase]-[disordered-form scaffold protein] complex	[none]	[2Fe-2S] iron-sulfur cluster biosynthesis
a [cysteine desulfurase]-[S-sulfanyl]2-[disordered-form scaffold protein] complex + 2 Fe3+ → a [cysteine desulfurase]-[scaffold protein-(2Fe-2S)] complex	[none]	[2Fe-2S] iron-sulfur cluster biosynthesis
co-chaperone] → a [co-chaperone]-[scaffold protein-(2Fe-2S)] complex + an [L-cysteine desulfurase]	[none]	[2Fe-2S] iron-sulfur cluster biosynthesis
a [co-chaperone]-[scaffold protein-(2Fe-2S)] complex + an [Fe-S cluster biosynthesis chaperone]-ATP → a [chaperone-ATP]-[co-chaperone]-[scaffold protein-(2Fe-2S)] complex	[none]	[2Fe-2S] iron-sulfur cluster biosynthesis
sulfur protein → a [chaperone-ADP]-[disordered-form scaffold protein] complex + an [Fe-S cluster biosynthesis chaperone] + an [2Fe-2S] cluster protein + phosphate	[none]	[2Fe-2S] iron-sulfur cluster biosynthesis
a [chaperone-ADP]-[disordered-form scaffold protein] complex + ATP → an [Fe-S cluster biosynthesis chaperone]-ATP + a [disordered-form [Fe-S] cluster scaffold protein] + ADP	[none]	[2Fe-2S] iron-sulfur cluster biosynthesis
GMP + diphosphate ← guanine + 5-phospho-α-D-ribose 1-diphosphate	2.4.2.8	guanine and guanosine salvage II
ammonium + CO2 + ATP ← carbamoyl-phosphate + ADP + 2 H+	2.7.2.2	citrulline degradation

glycine + acetyl-CoA \leftarrow 2-amino-3-oxobutanoate + coenzyme A + H+	2.3.1.29	threonine degradation II
cytosine + H+ + H2O \rightarrow ammonium + uracil	3.5.4.1	pyrimidine ribonucleosides salvage III
tRNA _{Gln} + L-glutamate + ATP + H+ \rightarrow L-glutamyl-tRNA _{Gln} + AMP + diphosphate	[none]	L-glutamine biosynthesis II (tRNA-dependent)
adenosine + H2O \rightarrow D-ribofuranose + adenine	3.2.2.7	adenine and adenosine salvage II
a D-xylopyranose \rightarrow D-xylulose	5.3.1.5	xylose degradation I
XMP + diphosphate \leftarrow xanthine + 5-phospho-α-D-ribose 1-diphosphate	2.4.2.22	xanthine and xanthosine salvage
linoleoyl-CoA + a reduced electron acceptor + oxygen \rightarrow γ-linolenoyl-CoA + an oxidized electron acceptor + 2 H2O	1.14.19.3	docosahexanoate biosynthesis I, gamma-linolenate biosynthesis II (animals)
thiamin + H2O \rightarrow 4-methyl-5-(β-hydroxyethyl)thiazole + hydroxymethylpyrimidine + H+	3.5.99.2	thiamin salvage IV (yeast)
D-ribose 5-phosphate + uracil \leftarrow pseudouridine 5'-phosphate + H2O	4.2.1.70	pseudouridine degradation
methyl-5-phosphomethylpyrimidine + 5'-deoxyadenosine + L-methionine + formate + carbon monoxide + 3 H+	4.1.99.17	4-amino-2-methyl-5-diphosphomethylpyrimidine biosynthesis
4-amino-4-deoxychorismate \rightarrow 4-aminobenzoate + pyruvate + H+	4.1.3.38	4-aminobenzoate biosynthesis
α-D-glucose 6-phosphate + ADP-α-D-glucose \rightarrow α,α-trehalose 6-phosphate + ADP + H+	2.4.1.-	trehalose biosynthesis II
an oxidized thioredoxin + 2 e- \leftarrow a reduced thioredoxin	[none]	thioredoxin pathway
a [lipoyl-carrier protein]-L-lysine + octanoate + ATP \rightarrow a [lipoyl-carrier protein] N6-octanoyl-L-lysine + AMP + diphosphate + H+	[none]	lipoate biosynthesis and incorporation II
oleoyl-CoA + H2O \rightarrow oleate + coenzyme A + H+	3.1.2.-	oleate biosynthesis II (animals and fungi)
4-hydroxyphenylpyruvate + H+ \rightarrow CO2 + (4-hydroxyphenyl)acetaldehyde	4.1.1.80	tyrosine degradation III
L-tyrosine + pyruvate \leftrightarrow 4-hydroxyphenylpyruvate + L-alanine	2.6.1.58	tyrosine degradation III
3-hydroxy-N6,N6,N6-trimethyl-L-lysine \rightarrow 4-trimethylammoniobutanal + glycine	[none]	L-carnitine biosynthesis
4-trimethylammoniobutanal + NAD+ + H2O \rightarrow γ-butyrobetaine + NADH + 2 H+	1.2.1.47	L-carnitine biosynthesis
coproporphyrinogen III + 2 S-adenosyl-L-methionine \rightarrow protoporphyrinogen IX + 2 CO2 + 2 L-methionine + 2 5'-deoxyadenosine	1.3.99.22	heme biosynthesis from uroporphyrinogen-III II (anaerobic)
protoporphyrinogen IX + 3 a menaquinone \rightarrow protoporphyrin IX + 3 a menaquinol	1.3.5.3	heme biosynthesis from uroporphyrinogen-III II (anaerobic)
a long-chain aldehyde + coenzyme A + NAD+ \leftarrow a long-chain acyl-CoA + NADH + H+	1.2.1.-	wax esters biosynthesis II
a long-chain aldehyde + NADH + H+ \rightarrow a long-chain alcohol + NAD+	1.2.1.-	wax esters biosynthesis II
oleoyl-CoA + 2 coenzyme A + 2 NAD+ + 2 oxygen + 2 H2O \rightarrow 5-cis-tetradecenoyl-CoA + 2 hydrogen peroxide + 2 acetyl-CoA + 2 NADH + 2 H+	[none]	oleate beta-oxidation (isomerase-dependent, yeast)
2-trans,5-cis-tetradecadienoyl-CoA + coenzyme A + NAD+ + oxygen + H2O + H+ \rightarrow 3-cis-dodecenoyl-CoA + hydrogen peroxide + acetyl-CoA + NADH	[none]	oleate beta-oxidation (isomerase-dependent, yeast)
antheraxanthin + L-ascorbate + H+ \rightarrow zeaxanthin + L-dehydro-ascorbate + H2O	[none]	zeaxanthin, antheraxanthin and violaxanthin interconversion
violaxanthin + L-ascorbate + H+ \rightarrow antheraxanthin + L-dehydro-ascorbate + H2O	[none]	zeaxanthin, antheraxanthin and violaxanthin interconversion
UDP-α-D-glucose + a glycogenin \rightarrow a glucosyl-glycogenin + UDP + H+	2.4.1.186	glycogen biosynthesis II (from UDP-D-Glucose)
UDP-α-D-glucose + a glucosyl-glycogenin[cytosol] \rightarrow (1,4-α-D-glucosyl)n-glucosyl glucogenin + UDP + H+[out]	2.4.1.186	glycogen biosynthesis II (from UDP-D-Glucose)
di-trans, poly-cis-polypropenyl diphosphate (C80) + H2O \rightarrow di-trans, poly-cis-polyprenol (C80) + diphosphate	[none]	dolichol and dolichyl phosphate biosynthesis
di-trans, poly-cis-polyprenol (C80) + NADPH + H+ \rightarrow a dolichol + NADP+	[none]	dolichol and dolichyl phosphate biosynthesis
2-oxo-3-phenylpropanoate + H+ \rightarrow CO2 + phenylacetaldehyde	4.1.1.43	phenylalanine degradation III
L-phenylalanine + pyruvate \leftrightarrow 2-oxo-3-phenylpropanoate + L-alanine	2.6.1.58	phenylalanine degradation III
NADPH + H2O \rightarrow NADH + phosphate	3.1.3.-	NAD/NADH phosphorylation and dephosphorylation
NADP+ + H2O \rightarrow NAD+ + phosphate	3.1.3.-	NAD/NADH phosphorylation and dephosphorylation
ammonium + ATP + nicotinate adenine dinucleotide \rightarrow AMP + NAD+ + diphosphate + H+	6.3.1.5	NAD biosynthesis I (from aspartate)
α-iminosuccinate + dihydroxyacetone phosphate \rightarrow quinolinate + phosphate + 2 H2O	2.5.1.72	NAD biosynthesis I (from aspartate)
L-aspartate + oxygen \rightarrow α-iminosuccinate + hydrogen peroxide + H+	1.4.3.16	NAD biosynthesis I (from aspartate)
L-ornithine \rightarrow L-Nδ-acetylmithine	2.3.1.-	L-Nδ-acetylmithine biosynthesis
L-glutamate + ATP + NADPH + H+ \rightarrow ADP + L-glutamate-5-semialdehyde + NADP+ + phosphate	[none]	L-Nδ-acetylmithine biosynthesis
L-arginine + H2O \rightarrow ammonium + L-citrulline	3.5.3.6	proline biosynthesis II (from arginine)

L-citrulline + 2 H+ + H2O → L-ornithine + CO2 + ammonium	3.5.1.20	proline biosynthesis II (from arginine)
2-methyl-3-hydroxybutyryl-CoA + NAD+ → 2-methylacetoacetyl-CoA + NADH + H+	1.1.1.178	isoleucine degradation I
(S)-3-methyl-2-oxopentanoate + coenzyme A + NAD+ → 2-methylbutanoyl-CoA + CO2 + NADH	1.2.1.-	isoleucine degradation I
2-methylbutanoyl-CoA + an oxidized electron-transfer flavoprotein → (E)-2-methylcrotonyl-CoA + a reduced electron-transfer flavoprotein	1.3.8.5	isoleucine degradation I
3-hydroxy, 4-trans-undecenoyl-CoA + NAD+ → 4-trans-3-oxo-undecenoyl-CoA + NADH + H+	1.1.1.-	10-trans-heptadecenoyl-CoA degradation (MFE-dependent, yeast)
10-cis-heptadecenoyl-CoA + 2 coenzyme A + 2 NAD+ + 2 H2O + 2 oxygen → 6-cis-tridecenoyl-CoA + 2 acetyl-CoA + 2 hydrogen peroxide + 2 NADH + 2 H+	[none]	10-cis-heptadecenoyl-CoA degradation (yeast)
3-hydroxy, 6-cis-tridecenoyl-CoA + NAD+ → 6-cis, 3-oxo-tridecenoyl-CoA + NADH + H+	1.1.1.-	10-cis-heptadecenoyl-CoA degradation (yeast)
10-trans-heptadecenoyl-CoA + 2 coenzyme A + 2 NAD+ + 2 H2O + 2 oxygen → 6-trans-tridecenoyl-CoA + 2 acetyl-CoA + 2 hydrogen peroxide + 2 NADH + 2 H+	[none]	10-trans-heptadecenoyl-CoA degradation (reductase-dependent, yeast)
3-hydroxy, 6-trans-tridecenoyl-CoA + NAD+ → 6-trans-3-oxo-tridecenoyl-CoA + NADH + H+	1.1.1.-	10-trans-heptadecenoyl-CoA degradation (reductase-dependent, yeast)
3-trans-undecenoyl-CoA → 2-trans-undecenoyl-CoA	5.3.3.-	10-cis-heptadecenoyl-CoA degradation (yeast), 10-trans-heptadecenoyl-CoA degradation (reductase-dependent, yeast)
3-hydroxy-undecanoyl-CoA + coenzyme A + NAD+ + H2O + oxygen → 3-hydroxy-nonanoyl-CoA + acetyl-CoA + hydrogen peroxide + NADH + H+	[none]	10-cis-heptadecenoyl-CoA degradation (yeast), 10-trans-heptadecenoyl-CoA degradation (reductase-dependent, yeast)
3-hydroxy-nonanoyl-CoA + coenzyme A + NAD+ + H2O + oxygen → 3-hydroxy-heptanoyl-CoA + acetyl-CoA + hydrogen peroxide + NADH + H+	[none]	heptadecenoyl-CoA degradation (MFE-dependent, yeast), 10-trans-heptadecenoyl-CoA degradation (reductase-dependent, yeast)
3-hydroxy-heptanoyl-CoA + coenzyme A + NAD+ + H2O + oxygen → 3-hydroxy-pentanoyl-CoA + acetyl-CoA + hydrogen peroxide + NADH + H+	[none]	heptadecenoyl-CoA degradation (MFE-dependent, yeast), 10-trans-heptadecenoyl-CoA degradation (reductase-dependent, yeast)
S-adenosyl-L-methionine + a cytosine49 in tRNA = S-adenosyl-L-homocysteine + a 5-methylcytosine49 in tRNA + H+	2.1.1.202	tRNA methylation (yeast)
S-adenosyl-L-methionine + a cytosine48 in tRNA = S-adenosyl-L-homocysteine + a 5-methylcytosine48 in tRNA + H+	2.1.1.202	tRNA methylation (yeast)
S-adenosyl-L-methionine + a cytosine40 in tRNA precursor = S-adenosyl-L-homocysteine + a 5 methylcytosine40 in tRNA precursor	2.1.1.202	tRNA methylation (yeast)
S-adenosyl-L-methionine + a cytosine34 in tRNA precursor = S-adenosyl-L-homocysteine + a 5 methylcytosine34 in tRNA precursor	2.1.1.202	tRNA methylation (yeast)
S-adenosyl-L-methionine + a cytidine4 in tRNAPro = S-adenosyl-L-homocysteine + a 2'-O-methylcytidine4 in tRNAPro + H+	2.1.1.225	tRNA methylation (yeast)
S-adenosyl-L-methionine + a cytidine4 in tRNAGly(GCC) = S-adenosyl-L-homocysteine + a 2'-O-methylcytidine4 in tRNAGly(GCC)	2.1.1.225	tRNA methylation (yeast)
S-adenosyl-L-methionine + an adenosine4 in tRNAHis = S-adenosyl-L-homocysteine + a 2'-O-methyladenosine4 in tRNAHis + H+	2.1.1.225	tRNA methylation (yeast)
sphinganine 1-phosphate + H2O → sphinganine + phosphate	3.1.3.-	sphingolipid recycling and degradation (yeast)
phytosphingosine 1-phosphate + H2O → phytosphingosine + phosphate	3.1.3.-	sphingolipid recycling and degradation (yeast)
a mannosyl-diphosphoinositol- α hydroxypyrophoceramide + H2O → an α hydroxypyrophoceramide + phosphoinositol- β -mannoside-phosphoinositol	3.1.4.-	sphingolipid recycling and degradation (yeast)
a mannosyl-inositol-phospho- α hydroxypyrophoceramide + H2O → an α hydroxypyrophoceramide + phosphoinositol- β -mannoside + H+	3.1.4.-	sphingolipid recycling and degradation (yeast)
an inositol-phospho- α hydroxypyrophoceramide + H2O → an α hydroxypyrophoceramide + 1D-myo-inositol 1-monophosphate + H+	3.1.4.-	sphingolipid recycling and degradation (yeast)
GTP → cyclic pyranopterin phosphate + diphosphate	4.1.99.18	molybdenum cofactor biosynthesis
a small subunit of molybdopterin synthase + ATP → a carboxy-adenylated small subunit of molybdopterin synthase + diphosphate	2.7.7.80	molybdenum cofactor biosynthesis
desulfurase] → a thiocarboxylated small subunit of molybdopterin synthase + an [L-cysteine desulfurase] + AMP	2.8.1.11	molybdenum cofactor biosynthesis
trimethyl sulfonium + tetrahydrofolate ← dimethyl sulfide + 5-methyl-tetrahydrofolate + H+	2.1.1.19	folate transformations I
2 an oxidized ferredoxin + 5-methyl-tetrahydrofolate ← 2 a reduced ferredoxin + 5,10-methylenetetrahydrofolate + 2 H+	1.5.7.1	folate transformations I
acetaldehyde + NADP+ + H2O → acetate + NADPH + 2 H+	1.2.1.4	NAD/NADP-NADH/NADPH cytosolic interconversion (yeast)
a diphthine-[translation elongation factor 2] + ammonium + ATP → a diphthamide-[translation elongation factor 2] + AMP + diphosphate + H+	6.3.1.14	diphthamide biosynthesis
amino-3-carboxypropyl]-L-histidine-[translation elongation factor 2] + S-methyl-5'-thioadenosine + H+	2.5.1.108	diphthamide biosynthesis
D-glucosamine 1-phosphate + acetyl-CoA → N-acetyl- α -D-glucosamine 1-phosphate + coenzyme A + H+	2.3.1.157	UDP-N-acetyl-D-glucosamine biosynthesis I
D-glucosamine 1-phosphate ← D-glucosamine 6-phosphate	5.4.2.10	UDP-N-acetyl-D-glucosamine biosynthesis I
7,8-dihydronopterin 3'-phosphate + H2O → 7,8-dihydronopterin + phosphate	3.6.1.-	6-hydroxymethyl-dihydronopterin diphosphate biosynthesis I
7,8-dihydronopterin triphosphate + H2O → 7,8-dihydronopterin 3'-phosphate + diphosphate + H+	3.6.1.-	6-hydroxymethyl-dihydronopterin diphosphate biosynthesis I
pyruvate + coenzyme A + NAD+ → acetyl-CoA + CO2 + NADH	1.2.1.-	respiration (anaerobic)
β -nicotinamide D-ribonucleotide + H2O → ammonium + β -nicotinate D-ribonucleotide	3.5.1.42	NAD salvage pathway I
NAD+ + H2O → ADP-D-ribose + nicotinamide + H+	3.2.2.6	NAD salvage pathway I
β -nicotinamide D-ribonucleotide + H2O → D-ribose 5-phosphate + nicotinamide + H+	3.2.2.14	NAD salvage pathway I

a 1-phosphatidyl-1D-myo-inositol 4-phosphate + ATP → a 1-phosphatidyl-1D-myo-inositol 3,4-bisphosphate + ADP + H+	2.7.1.154	3-phosphoinositide biosynthesis
a 1-phosphatidyl-1D-myo-inositol 4,5-bisphosphate + ATP → 1-phosphatidyl-1D-myo-inositol 3,4,5-trisphosphate + ADP + H+	2.7.1.153	3-phosphoinositide biosynthesis
1-phosphatidyl-1D-myo-inositol 3,5-bisphosphate + H2O → a 1-phosphatidyl-1D-myo-inositol 3-phosphate + phosphate	3.1.3.-	3-phosphoinositide biosynthesis
10-formyl-tetrahydrofolate + H2O → formate + tetrahydrofolate + H+	3.5.1.10	formylTHF biosynthesis, folate transformations I
glycine + tetrahydrofolate + NAD+ → ammonium + CO2 + 5,10-methylenetetrahydrofolate + NADH	[none]	formylTHF biosynthesis, folate transformations I
L-homocysteine + 5-methyl-tetrahydrofolate → L-methionine + tetrahydrofolate	2.1.1.13	formylTHF biosynthesis, folate transformations I, methionine biosynthesis III
L-quinate + NAD+ → 3-dehydroquinate + NADH + H+	1.1.1.24	quinate degradation II
an L-1-phosphatidylglycerol-phosphate + H2O → an L-1-phosphatidyl-glycerol + phosphate	3.1.3.27	cardiolipin biosynthesis II, cardiolipin biosynthesis III
L-1-glycero-3-phosphocholine + H2O → sn-glycerol 3-phosphate + choline + H+	3.1.4.2	phosphatidylcholine resynthesis via glycerophosphocholine
D-arabinofuranose + NAD(P)+ → D-arabinono-1,4-lactone + NAD(P)H + H+	1.1.1.117	dehydro-D-arabinono-1,4-lactone biosynthesis
a lipid hydroperoxide + 2 glutathione + H+ → a lipid + glutathione disulfide + 2 H2O	1.11.1.12	glutathione redox reactions I
3-methyl-2-oxobutanoate + H+ → isobutanal + CO2	4.1.1.72	valine degradation II
(S)-3-methyl-2-oxopentanoate + H+ → 2-methylbutanal + CO2	4.1.1.72	isoleucine degradation II
3-trans-tetradecenoyl-CoA → 2-trans-tetradecenoyl-CoA	5.3.3.-	oleate beta-oxidation (reductase-dependent, yeast)
2 acetaldehyde → acetoin	[none]	acetoin biosynthesis III
CO2 + 7,8-diaminopelargonate + ATP → dethiobiotin + ADP + phosphate + 3 H+	6.3.3.3	biotin biosynthesis from 8-amino-7-oxononanoate I
L-glutamate + 7,8-dihydropteroate + ATP → ADP + 7,8-dihydrofolate + phosphate + H+	6.3.2.12	formylTHF biosynthesis, tetrahydrofolate biosynthesis
thiamin + ATP → thiamin diphosphate + AMP + H+	2.7.6.2	thiamin diphosphate biosynthesis IV (eukaryotes)
pyritthiamine + H2O → hydroxymethylpyrimidine + 2-(2-methylpyridin-3-yl)ethanol + 2 H+	3.5.99.-	thiamin formation from pyritthiamine and oxythiamine (yeast)
oxythiamine + H2O → 4-methyl-5-(β-hydroxyethyl)thiazole + 5-(hydroxymethyl)-2-methyl-4(1H)-pyrimidinone + H+	3.5.99.-	thiamin formation from pyritthiamine and oxythiamine (yeast)
β-D-glucose 1-phosphate → β-D-glucose 6-phosphate	5.4.2.6	galactose degradation I (Leloir pathway)
a butyryl-[acp] + a malonyl-[acp] + H+ → a 3-oxo-hexanoyl-[acp] + CO2 + a holo-[acyl-carrier protein]	2.3.1.41	octanoyl-ACP biosynthesis (mitochondria, yeast)
a 3-oxo-hexanoyl-[acp] + a malonyl-[acp] + 2 NADPH + 3 H+ → a 3-oxo-octanoyl-[acp] + CO2 + a holo-[acyl-carrier protein] + 2 NADP+ + H2O	[none]	octanoyl-ACP biosynthesis (mitochondria, yeast)
a 3-oxo-octanoyl-[acp] + 2 NADPH + 2 H+ → octanoyl-ACP + 2 NADP+ + H2O	[none]	octanoyl-ACP biosynthesis (mitochondria, yeast)
5-phospho-β-D-ribosylamine → 2'-hydroxypyridoxol	[none]	4-amino-2-methyl-5-phosphomethylpyrimidine biosynthesis (yeast)
2'-hydroxypyridoxol + 2 H+ → pyridoxine + H2O	4.2.1.-	4-amino-2-methyl-5-phosphomethylpyrimidine biosynthesis (yeast)
[an HMP-P synthase]-L-histidine + pyridoxal 5'-phosphate → 4-amino-2-methyl-5-phosphomethylpyrimidine + an HMP-P synthase	[none]	4-amino-2-methyl-5-phosphomethylpyrimidine biosynthesis (yeast)
L-2-amino adipate + 2-oxoglutarate ↔ L-glutamate + 2-oxoadipate	2.6.1.39	lysine degradation II, lysine biosynthesis IV
acetyl-CoA + phosphate → acetyl phosphate + coenzyme A	2.3.1.8	mixed acid fermentation
formate + acetyl-CoA ← pyruvate + coenzyme A	2.3.1.54	mixed acid fermentation, respiration (anaerobic)
oxaloacetate + phosphate ← phosphoenolpyruvate + bicarbonate	4.1.1.31	mixed acid fermentation, respiration (anaerobic)
formate + H+ → CO2 + H2	[none]	mixed acid fermentation, respiration (anaerobic)
a menaquinol + fumarate[in] → a menaquinone + succinate[in]	1.3.5.4	mixed acid fermentation, respiration (anaerobic)
maltotriose + H2O → maltose + D-glucopyranose	3.2.1.20	glycogen degradation I
β-D-glucose + ATP → β-D-glucose 6-phosphate + ADP + H+	2.7.1.2	glycogen degradation I, glucose and glucose-1-phosphate degradation
2 a 1,2-diacyl-sn-glycerol ↔ a triacyl-sn-glycerol + a 2-monoglyceride	2.3.1.-	triacylglycerol biosynthesis
CTP + a 1,2-diacyl-sn-glycerol → CDP + a 1,2-diacyl-sn-glycerol 3-phosphate + H+	2.7.1.174	triacylglycerol biosynthesis
sphinganine + a long-chain acyl-CoA → a dihydroceramide + coenzyme A + H+	[none]	ceramide de novo biosynthesis
S-adenosyl-L-methionine + a demethylated methyl acceptor → S-adenosyl-L-homocysteine + a methylated methyl acceptor	2.1.1.-	S-adenosyl-L-methionine cycle II, methionine degradation I (to homocysteine)
IMP + diphosphate ← hypoxanthine + 5-phospho-α-D-ribose 1-diphosphate	2.4.2.8	adenine and adenosine salvage III
sn-glycerol 3-phosphate + NAD(P)+ ← dihydroxyacetone phosphate + NAD(P)H + H+	1.1.1.94	CDP-diacylglycerol biosynthesis I

ammonium + 3 NAD(P)+ + 2 H2O \leftarrow nitrite + 3 NAD(P)H + 5 H+	1.7.1.4	nitrate reduction V (assimilatory), alkylnitronates degradation
2-oleoylglycerol + CTP \rightarrow 1-oleyl-2-lyso-phosphatidate + CDP + H+	2.7.1.174	monoacylglycerol metabolism (yeast)
isobutanoyl-CoA + FAD + H+ \rightarrow methylacrylyl-CoA + FADH2	1.3.8.-	valine degradation I
3-methyl-2-oxobutanoate + coenzyme A + NAD+ \rightarrow isobutanoyl-CoA + CO2 + NADH	1.2.1.25	valine degradation I
5-amino-6-(5-phospho-D-ribitylamino)uracil + H2O \rightarrow 5-amino-6-(D-ribitylamino)uracil + phosphate	[none]	flavin biosynthesis III (fungi)
2,5-diamino-6-(5-phospho-D-ribitylamino)pyrimidin-4(3H)-one + H+ + H2O \rightarrow 5-amino-6-(5-phospho-D-ribitylamino)uracil + ammonium	3.5.4.-	flavin biosynthesis III (fungi)
D-myo-inositol (1,4,5)-trisphosphate + ATP \rightarrow D-myo-inositol (1,3,4,5)-tetrakisphosphate + ADP + H+	2.7.1.127	1D-myo-inositol hexakisphosphate biosynthesis I (from Ins(1,4,5)P3)
a phosphatidylcholine + H2O \rightarrow a 1-lysophosphatidylcholine + a fatty acid + H+	3.1.1.32	phospholipases
ATP + NADH \rightarrow ADP + NADPH + H+	2.7.1.86	(yeast), NAD/NADH phosphorylation and dephosphorylation, NAD/NADP-NADH/NADPH cytosolic interconversion (yeast)
a (3R)-3-hydroxyacyl-CoA + NAD+ \rightarrow a 3-oxoacyl-CoA + NADH + H+	1.1.1.-	fatty acid beta-oxidation (peroxisome, yeast)
episterol + NADPH + H+ + oxygen \rightarrow 5-dehydro episterol + NADP+ + 2 H2O	1.14.19.-	ergosterol biosynthesis I
lanosterol \rightarrow ergosterol	[none]	sterol:steryl ester interconversion (yeast)
a 1-phosphatidyl-1D-myo-inositol 5-phosphate + ATP \rightarrow a 1-phosphatidyl-1D-myo-inositol 4,5-bisphosphate + ADP + H+	2.7.1.149	D-myo-inositol (1,4,5)-trisphosphate biosynthesis
shikimate + NAD(P)+ \leftarrow 3-dehydroshikimate + NAD(P)H + H+	1.1.1.282	chorismate biosynthesis from 3-dehydroquinate
stearoyl-CoA + H2O \rightarrow stearate + coenzyme A + H+	3.1.2.2	stearate biosynthesis III (fungi)
an ubiquinone + (S)-dihydroorotate[in] \rightarrow an ubiquinol + orotate[in]	1.3.5.2	UMP biosynthesis
4-methyl-2-oxopentanoate + coenzyme A + NAD+ \rightarrow isovaleryl-CoA + CO2 + NADH	1.2.1.-	leucine degradation I
4 α -carboxy-4 β -methyl-5 α -cholest-8,24-dien-3 β -ol + NAD(P)+ \rightarrow 3-dehydro-4-methylzymosterol + NAD(P)H + CO2	1.1.1.170	zymosterol biosynthesis
4 α -carboxy-5 α -cholest-8,24-dien-3 β -ol + NAD(P)+ \rightarrow 5 α -cholest-8,24-dien-3-one + NAD(P)H + CO2	1.1.1.170	zymosterol biosynthesis
(S)-malate + NADP+ \rightarrow CO2 + pyruvate + NADPH	1.1.1.40	gluconeogenesis I
2-oxoglutarate + coenzyme A + NAD+ \rightarrow succinyl-CoA + CO2 + NADH	[none]	TCA cycle II (plants and fungi)
imidazole-glycerol-phosphate + 5-amino-1-(5-phospho-D-ribosyl)imidazole-4-carboxamide + H+	2.4.2.-	histidine biosynthesis
a dodecanoyl-[acp] + H2O \rightarrow a holo-[acyl-carrier protein] + laurate + H+	3.1.2.21	palmitate biosynthesis I (animals and fungi)
a palmitoyl-[acp] + coenzyme A \rightarrow palmitoyl-CoA + a holo-[acyl-carrier protein]	2.3.1.-	palmitate biosynthesis I (animals and fungi)
a palmitoyl-[acp] + H2O \rightarrow palmitate + a holo-[acyl-carrier protein] + H+	3.1.2.14	palmitate biosynthesis I (animals and fungi)
pyruvate + ATP + H2O \rightarrow phosphoenolpyruvate + AMP + phosphate + 2 H+	2.7.9.2	glycolysis I (from glucose-6P), gluconeogenesis I
a (mannosyl)3-(N-acetylglucosaminyl)2-diphosphodolichol + 2 GDP- α -D-mannose \rightarrow a (mannosyl)5-(N-acetylglucosaminyl)2-diphosphodolichol + 2 GDP + 2 H+	2.4.1.131	dolichyl-diphosphooligosaccharide biosynthesis

Anexo C

Tabela de transcritos em suas
respectivas formas de ocorrência

Micélio	Levedura	Ambas
PAAG_00017	PAAG_00014	PAAG_00004
PAAG_00054	PAAG_00023	PAAG_00019
PAAG_00058	PAAG_00024	PAAG_00028
PAAG_00064	PAAG_00030	PAAG_00046
PAAG_00070	PAAG_00035	PAAG_00047
PAAG_00072	PAAG_00063	PAAG_00048
PAAG_00090	PAAG_00065	PAAG_00050
PAAG_00100	PAAG_00068	PAAG_00053
PAAG_00106	PAAG_00088	PAAG_00071
PAAG_00117	PAAG_00128	PAAG_00102
PAAG_00132	PAAG_00144	PAAG_00103
PAAG_00149	PAAG_00151	PAAG_00116
PAAG_00153	PAAG_00165	PAAG_00123
PAAG_00154	PAAG_00173	PAAG_00142
PAAG_00169	PAAG_00177	PAAG_00143
PAAG_00175	PAAG_00201	PAAG_00174
PAAG_00192	PAAG_00212	PAAG_00176
PAAG_00203	PAAG_00224	PAAG_00202
PAAG_00204	PAAG_00225	PAAG_00221
PAAG_00215	PAAG_00271	PAAG_00236
PAAG_00228	PAAG_00276	PAAG_00242
PAAG_00249	PAAG_00281	PAAG_00244
PAAG_00287	PAAG_00288	PAAG_00257
PAAG_00289	PAAG_00306	PAAG_00264
PAAG_00293	PAAG_00330	PAAG_00273
PAAG_00320	PAAG_00352	PAAG_00299
PAAG_00365	PAAG_00358	PAAG_00302
PAAG_00373	PAAG_00391	PAAG_00303
PAAG_00377	PAAG_00396	PAAG_00315
PAAG_00386	PAAG_00403	PAAG_00316
PAAG_00394	PAAG_00416	PAAG_00331
PAAG_00422	PAAG_00433	PAAG_00335
PAAG_00440	PAAG_00438	PAAG_00339
PAAG_00459	PAAG_00455	PAAG_00340
PAAG_00473	PAAG_00460	PAAG_00342
PAAG_00476	PAAG_00477	PAAG_00347
PAAG_00514	PAAG_00481	PAAG_00356

PAAG_00547	PAAG_00485	PAAG_00364
PAAG_00550	PAAG_00487	PAAG_00376
PAAG_00558	PAAG_00492	PAAG_00397
PAAG_00562	PAAG_00494	PAAG_00417
PAAG_00581	PAAG_00495	PAAG_00427
PAAG_00617	PAAG_00510	PAAG_00430
PAAG_00664	PAAG_00524	PAAG_00435
PAAG_00670	PAAG_00535	PAAG_00446
PAAG_00695	PAAG_00546	PAAG_00450
PAAG_00707	PAAG_00559	PAAG_00462
PAAG_00719	PAAG_00621	PAAG_00468
PAAG_00731	PAAG_00626	PAAG_00479
PAAG_00763	PAAG_00639	PAAG_00493
PAAG_00764	PAAG_00642	PAAG_00503
PAAG_00788	PAAG_00657	PAAG_00512
PAAG_00789	PAAG_00689	PAAG_00529
PAAG_00803	PAAG_00706	PAAG_00548
PAAG_00805	PAAG_00726	PAAG_00555
PAAG_00825	PAAG_00734	PAAG_00566
PAAG_00827	PAAG_00739	PAAG_00594
PAAG_00857	PAAG_00774	PAAG_00598
PAAG_00866	PAAG_00791	PAAG_00606
PAAG_00986	PAAG_00795	PAAG_00623
PAAG_00994	PAAG_00802	PAAG_00627
PAAG_01013	PAAG_00813	PAAG_00629
PAAG_01067	PAAG_00818	PAAG_00647
PAAG_01093	PAAG_00850	PAAG_00659
PAAG_01096	PAAG_00851	PAAG_00661
PAAG_01110	PAAG_00865	PAAG_00685
PAAG_01127	PAAG_00867	PAAG_00697
PAAG_01136	PAAG_00868	PAAG_00698
PAAG_01143	PAAG_00871	PAAG_00702
PAAG_01151	PAAG_00889	PAAG_00703
PAAG_01159	PAAG_00890	PAAG_00716
PAAG_01165	PAAG_00893	PAAG_00724
PAAG_01176	PAAG_00914	PAAG_00732
PAAG_01189	PAAG_00934	PAAG_00742
PAAG_01191	PAAG_00935	PAAG_00744

PAAG_01192	PAAG_00953	PAAG_00748
PAAG_01207	PAAG_00968	PAAG_00752
PAAG_01232	PAAG_00971	PAAG_00769
PAAG_01237	PAAG_00997	PAAG_00771
PAAG_01269	PAAG_01023	PAAG_00773
PAAG_01278	PAAG_01034	PAAG_00781
PAAG_01280	PAAG_01082	PAAG_00799
PAAG_01293	PAAG_01085	PAAG_00801
PAAG_01299	PAAG_01087	PAAG_00808
PAAG_01300	PAAG_01095	PAAG_00833
PAAG_01307	PAAG_01123	PAAG_00836
PAAG_01320	PAAG_01141	PAAG_00843
PAAG_01321	PAAG_01146	PAAG_00856
PAAG_01377	PAAG_01147	PAAG_00869
PAAG_01437	PAAG_01150	PAAG_00875
PAAG_01466	PAAG_01177	PAAG_00878
PAAG_01469	PAAG_01209	PAAG_00948
PAAG_01481	PAAG_01226	PAAG_00952
PAAG_01539	PAAG_01227	PAAG_00954
PAAG_01583	PAAG_01243	PAAG_00962
PAAG_01601	PAAG_01244	PAAG_00966
PAAG_01602	PAAG_01251	PAAG_00969
PAAG_01606	PAAG_01252	PAAG_00976
PAAG_01653	PAAG_01256	PAAG_00978
PAAG_01656	PAAG_01262	PAAG_00981
PAAG_01665	PAAG_01273	PAAG_00982
PAAG_01681	PAAG_01289	PAAG_00988
PAAG_01704	PAAG_01291	PAAG_01002
PAAG_01706	PAAG_01302	PAAG_01015
PAAG_01715	PAAG_01305	PAAG_01037
PAAG_01741	PAAG_01306	PAAG_01041
PAAG_01755	PAAG_01317	PAAG_01044
PAAG_01775	PAAG_01339	PAAG_01052
PAAG_01786	PAAG_01381	PAAG_01053
PAAG_01790	PAAG_01390	PAAG_01078
PAAG_01807	PAAG_01412	PAAG_01079
PAAG_01877	PAAG_01415	PAAG_01097
PAAG_01933	PAAG_01423	PAAG_01098

PAAG_01941	PAAG_01425	PAAG_01111
PAAG_01943	PAAG_01462	PAAG_01125
PAAG_01959	PAAG_01479	PAAG_01131
PAAG_01963	PAAG_01494	PAAG_01134
PAAG_02004	PAAG_01495	PAAG_01168
PAAG_02056	PAAG_01515	PAAG_01170
PAAG_02070	PAAG_01517	PAAG_01171
PAAG_02071	PAAG_01527	PAAG_01178
PAAG_02087	PAAG_01541	PAAG_01179
PAAG_02095	PAAG_01563	PAAG_01190
PAAG_02112	PAAG_01570	PAAG_01229
PAAG_02152	PAAG_01574	PAAG_01233
PAAG_02189	PAAG_01581	PAAG_01234
PAAG_02192	PAAG_01591	PAAG_01255
PAAG_02194	PAAG_01647	PAAG_01257
PAAG_02208	PAAG_01682	PAAG_01258
PAAG_02210	PAAG_01696	PAAG_01264
PAAG_02255	PAAG_01713	PAAG_01279
PAAG_02274	PAAG_01718	PAAG_01282
PAAG_02278	PAAG_01722	PAAG_01292
PAAG_02282	PAAG_01725	PAAG_01310
PAAG_02361	PAAG_01743	PAAG_01324
PAAG_02381	PAAG_01750	PAAG_01327
PAAG_02398	PAAG_01751	PAAG_01330
PAAG_02409	PAAG_01757	PAAG_01363
PAAG_02467	PAAG_01763	PAAG_01369
PAAG_02501	PAAG_01776	PAAG_01389
PAAG_02520	PAAG_01791	PAAG_01396
PAAG_02573	PAAG_01821	PAAG_01399
PAAG_02577	PAAG_01825	PAAG_01432
PAAG_02582	PAAG_01831	PAAG_01436
PAAG_02589	PAAG_01834	PAAG_01445
PAAG_02594	PAAG_01847	PAAG_01452
PAAG_02604	PAAG_01864	PAAG_01454
PAAG_02612	PAAG_01874	PAAG_01461
PAAG_02636	PAAG_01876	PAAG_01463
PAAG_02700	PAAG_01898	PAAG_01468
PAAG_02732	PAAG_01899	PAAG_01472

PAAG_02762	PAAG_01906	PAAG_01475
PAAG_02766	PAAG_01907	PAAG_01485
PAAG_02782	PAAG_01917	PAAG_01499
PAAG_02803	PAAG_01921	PAAG_01501
PAAG_02821	PAAG_01922	PAAG_01513
PAAG_02838	PAAG_01924	PAAG_01521
PAAG_02854	PAAG_01928	PAAG_01524
PAAG_02866	PAAG_01932	PAAG_01526
PAAG_02937	PAAG_01935	PAAG_01534
PAAG_03024	PAAG_01938	PAAG_01536
PAAG_03030	PAAG_01970	PAAG_01538
PAAG_03032	PAAG_01973	PAAG_01569
PAAG_03057	PAAG_02011	PAAG_01578
PAAG_03069	PAAG_02012	PAAG_01587
PAAG_03075	PAAG_02016	PAAG_01594
PAAG_03096	PAAG_02050	PAAG_01599
PAAG_03127	PAAG_02068	PAAG_01603
PAAG_03165	PAAG_02096	PAAG_01605
PAAG_03188	PAAG_02102	PAAG_01631
PAAG_03189	PAAG_02110	PAAG_01633
PAAG_03209	PAAG_02111	PAAG_01650
PAAG_03224	PAAG_02129	PAAG_01660
PAAG_03226	PAAG_02150	PAAG_01699
PAAG_03227	PAAG_02164	PAAG_01727
PAAG_03231	PAAG_02169	PAAG_01733
PAAG_03259	PAAG_02173	PAAG_01752
PAAG_03297	PAAG_02180	PAAG_01756
PAAG_03312	PAAG_02232	PAAG_01777
PAAG_03317	PAAG_02289	PAAG_01778
PAAG_03320	PAAG_02293	PAAG_01785
PAAG_03333	PAAG_02294	PAAG_01808
PAAG_03336	PAAG_02297	PAAG_01823
PAAG_03342	PAAG_02303	PAAG_01845
PAAG_03359	PAAG_02327	PAAG_01850
PAAG_03381	PAAG_02334	PAAG_01861
PAAG_03407	PAAG_02343	PAAG_01882
PAAG_03413	PAAG_02369	PAAG_01889
PAAG_03416	PAAG_02391	PAAG_01893

PAAG_03453	PAAG_02399	PAAG_01905
PAAG_03491	PAAG_02410	PAAG_01926
PAAG_03494	PAAG_02416	PAAG_01931
PAAG_03511	PAAG_02417	PAAG_01934
PAAG_03513	PAAG_02439	PAAG_01958
PAAG_03527	PAAG_02443	PAAG_01969
PAAG_03532	PAAG_02446	PAAG_01974
PAAG_03544	PAAG_02452	PAAG_01989
PAAG_03556	PAAG_02460	PAAG_01995
PAAG_03574	PAAG_02486	PAAG_01996
PAAG_03613	PAAG_02521	PAAG_02024
PAAG_03641	PAAG_02526	PAAG_02051
PAAG_03662	PAAG_02546	PAAG_02055
PAAG_03667	PAAG_02548	PAAG_02078
PAAG_03698	PAAG_02551	PAAG_02084
PAAG_03701	PAAG_02553	PAAG_02107
PAAG_03718	PAAG_02558	PAAG_02130
PAAG_03750	PAAG_02561	PAAG_02133
PAAG_03823	PAAG_02567	PAAG_02157
PAAG_03843	PAAG_02569	PAAG_02159
PAAG_03878	PAAG_02580	PAAG_02162
PAAG_03894	PAAG_02592	PAAG_02163
PAAG_03924	PAAG_02602	PAAG_02179
PAAG_03932	PAAG_02603	PAAG_02220
PAAG_03941	PAAG_02608	PAAG_02226
PAAG_03969	PAAG_02610	PAAG_02227
PAAG_03976	PAAG_02613	PAAG_02228
PAAG_04051	PAAG_02614	PAAG_02242
PAAG_04058	PAAG_02616	PAAG_02254
PAAG_04061	PAAG_02627	PAAG_02262
PAAG_04161	PAAG_02628	PAAG_02266
PAAG_04166	PAAG_02633	PAAG_02272
PAAG_04195	PAAG_02644	PAAG_02331
PAAG_04233	PAAG_02648	PAAG_02333
PAAG_04235	PAAG_02650	PAAG_02352
PAAG_04238	PAAG_02653	PAAG_02363
PAAG_04258	PAAG_02656	PAAG_02365
PAAG_04259	PAAG_02671	PAAG_02367

PAAG_04281	PAAG_02686	PAAG_02371
PAAG_04282	PAAG_02698	PAAG_02379
PAAG_04285	PAAG_02723	PAAG_02382
PAAG_04293	PAAG_02743	PAAG_02389
PAAG_04294	PAAG_02747	PAAG_02401
PAAG_04308	PAAG_02748	PAAG_02415
PAAG_04327	PAAG_02772	PAAG_02418
PAAG_04380	PAAG_02776	PAAG_02419
PAAG_04530	PAAG_02789	PAAG_02427
PAAG_04548	PAAG_02801	PAAG_02436
PAAG_04555	PAAG_02817	PAAG_02437
PAAG_04577	PAAG_02859	PAAG_02456
PAAG_04581	PAAG_02865	PAAG_02458
PAAG_04586	PAAG_02869	PAAG_02466
PAAG_04594	PAAG_02876	PAAG_02482
PAAG_04611	PAAG_02883	PAAG_02512
PAAG_04693	PAAG_02907	PAAG_02530
PAAG_04699	PAAG_02920	PAAG_02536
PAAG_04744	PAAG_02925	PAAG_02554
PAAG_04748	PAAG_02960	PAAG_02576
PAAG_04752	PAAG_02965	PAAG_02583
PAAG_04754	PAAG_02971	PAAG_02584
PAAG_04755	PAAG_03009	PAAG_02585
PAAG_04757	PAAG_03021	PAAG_02588
PAAG_04770	PAAG_03022	PAAG_02593
PAAG_04811	PAAG_03027	PAAG_02632
PAAG_04814	PAAG_03050	PAAG_02634
PAAG_04825	PAAG_03058	PAAG_02638
PAAG_04835	PAAG_03098	PAAG_02642
PAAG_04845	PAAG_03105	PAAG_02643
PAAG_04869	PAAG_03112	PAAG_02651
PAAG_04881	PAAG_03116	PAAG_02655
PAAG_04882	PAAG_03139	PAAG_02659
PAAG_04919	PAAG_03141	PAAG_02664
PAAG_04930	PAAG_03142	PAAG_02717
PAAG_04940	PAAG_03151	PAAG_02718
PAAG_04949	PAAG_03152	PAAG_02725
PAAG_04958	PAAG_03164	PAAG_02760

PAAG_05005	PAAG_03166	PAAG_02769
PAAG_05009	PAAG_03171	PAAG_02793
PAAG_05032	PAAG_03182	PAAG_02805
PAAG_05044	PAAG_03201	PAAG_02823
PAAG_05045	PAAG_03215	PAAG_02824
PAAG_05066	PAAG_03237	PAAG_02836
PAAG_05108	PAAG_03246	PAAG_02843
PAAG_05112	PAAG_03276	PAAG_02901
PAAG_05113	PAAG_03279	PAAG_02903
PAAG_05122	PAAG_03295	PAAG_02913
PAAG_05191	PAAG_03378	PAAG_02921
PAAG_05241	PAAG_03385	PAAG_02926
PAAG_05245	PAAG_03417	PAAG_02935
PAAG_05254	PAAG_03419	PAAG_02951
PAAG_05272	PAAG_03448	PAAG_02961
PAAG_05273	PAAG_03472	PAAG_02973
PAAG_05292	PAAG_03482	PAAG_02985
PAAG_05328	PAAG_03496	PAAG_02990
PAAG_05334	PAAG_03503	PAAG_02994
PAAG_05360	PAAG_03508	PAAG_03004
PAAG_05385	PAAG_03512	PAAG_03019
PAAG_05415	PAAG_03514	PAAG_03025
PAAG_05424	PAAG_03555	PAAG_03028
PAAG_05452	PAAG_03558	PAAG_03031
PAAG_05459	PAAG_03561	PAAG_03042
PAAG_05461	PAAG_03568	PAAG_03043
PAAG_05507	PAAG_03572	PAAG_03045
PAAG_05532	PAAG_03573	PAAG_03092
PAAG_05586	PAAG_03586	PAAG_03106
PAAG_05603	PAAG_03590	PAAG_03125
PAAG_05648	PAAG_03596	PAAG_03137
PAAG_05672	PAAG_03599	PAAG_03138
PAAG_05712	PAAG_03612	PAAG_03153
PAAG_05748	PAAG_03621	PAAG_03163
PAAG_05754	PAAG_03622	PAAG_03191
PAAG_05756	PAAG_03624	PAAG_03216
PAAG_05826	PAAG_03626	PAAG_03243
PAAG_05837	PAAG_03629	PAAG_03290

PAAG_05845	PAAG_03639	PAAG_03299
PAAG_05861	PAAG_03640	PAAG_03305
PAAG_05864	PAAG_03697	PAAG_03309
PAAG_05866	PAAG_03711	PAAG_03330
PAAG_05873	PAAG_03736	PAAG_03334
PAAG_05880	PAAG_03737	PAAG_03386
PAAG_05887	PAAG_03746	PAAG_03388
PAAG_05940	PAAG_03748	PAAG_03418
PAAG_05954	PAAG_03768	PAAG_03442
PAAG_05967	PAAG_03770	PAAG_03447
PAAG_06065	PAAG_03777	PAAG_03464
PAAG_06086	PAAG_03792	PAAG_03468
PAAG_06087	PAAG_03816	PAAG_03530
PAAG_06096	PAAG_03828	PAAG_03536
PAAG_06139	PAAG_03835	PAAG_03537
PAAG_06158	PAAG_03851	PAAG_03548
PAAG_06163	PAAG_03874	PAAG_03550
PAAG_06170	PAAG_03890	PAAG_03557
PAAG_06298	PAAG_03909	PAAG_03563
PAAG_06317	PAAG_03931	PAAG_03564
PAAG_06392	PAAG_03944	PAAG_03595
PAAG_06395	PAAG_03945	PAAG_03618
PAAG_06441	PAAG_03954	PAAG_03623
PAAG_06446	PAAG_03978	PAAG_03625
PAAG_06466	PAAG_03984	PAAG_03631
PAAG_06473	PAAG_04005	PAAG_03652
PAAG_06500	PAAG_04006	PAAG_03687
PAAG_06522	PAAG_04030	PAAG_03689
PAAG_06535	PAAG_04040	PAAG_03693
PAAG_06568	PAAG_04043	PAAG_03722
PAAG_06585	PAAG_04050	PAAG_03726
PAAG_06587	PAAG_04060	PAAG_03735
PAAG_06588	PAAG_04064	PAAG_03742
PAAG_06590	PAAG_04082	PAAG_03743
PAAG_06595	PAAG_04083	PAAG_03744
PAAG_06601	PAAG_04117	PAAG_03760
PAAG_06616	PAAG_04130	PAAG_03804
PAAG_06631	PAAG_04137	PAAG_03827

PAAG_06669	PAAG_04142	PAAG_03839
PAAG_06684	PAAG_04148	PAAG_03857
PAAG_06696	PAAG_04165	PAAG_03873
PAAG_06713	PAAG_04190	PAAG_03881
PAAG_06720	PAAG_04191	PAAG_03898
PAAG_06725	PAAG_04225	PAAG_03903
PAAG_06764	PAAG_04232	PAAG_03904
PAAG_06788	PAAG_04251	PAAG_03935
PAAG_06811	PAAG_04262	PAAG_03938
PAAG_06820	PAAG_04264	PAAG_03946
PAAG_06821	PAAG_04276	PAAG_03952
PAAG_06865	PAAG_04286	PAAG_03953
PAAG_06897	PAAG_04287	PAAG_03956
PAAG_06937	PAAG_04290	PAAG_03960
PAAG_06941	PAAG_04291	PAAG_03965
PAAG_06954	PAAG_04313	PAAG_03970
PAAG_07037	PAAG_04321	PAAG_03974
PAAG_07059	PAAG_04328	PAAG_03979
PAAG_07111	PAAG_04333	PAAG_03987
PAAG_07118	PAAG_04336	PAAG_04004
PAAG_07152	PAAG_04339	PAAG_04014
PAAG_07157	PAAG_04340	PAAG_04028
PAAG_07191	PAAG_04347	PAAG_04041
PAAG_07196	PAAG_04374	PAAG_04042
PAAG_07200	PAAG_04378	PAAG_04054
PAAG_07236	PAAG_04394	PAAG_04078
PAAG_07245	PAAG_04401	PAAG_04103
PAAG_07269	PAAG_04415	PAAG_04114
PAAG_07320	PAAG_04423	PAAG_04124
PAAG_07347	PAAG_04433	PAAG_04146
PAAG_07390	PAAG_04435	PAAG_04147
PAAG_07410	PAAG_04443	PAAG_04162
PAAG_07412	PAAG_04464	PAAG_04164
PAAG_07419	PAAG_04468	PAAG_04168
PAAG_07425	PAAG_04478	PAAG_04196
PAAG_07442	PAAG_04550	PAAG_04201
PAAG_07444	PAAG_04596	PAAG_04205
PAAG_07469	PAAG_04601	PAAG_04211

PAAG_07470	PAAG_04622	PAAG_04249
PAAG_07474	PAAG_04661	PAAG_04253
PAAG_07480	PAAG_04678	PAAG_04274
PAAG_07485	PAAG_04686	PAAG_04277
PAAG_07498	PAAG_04711	PAAG_04310
PAAG_07500	PAAG_04712	PAAG_04319
PAAG_07506	PAAG_04724	PAAG_04348
PAAG_07526	PAAG_04725	PAAG_04371
PAAG_07545	PAAG_04742	PAAG_04372
PAAG_07553	PAAG_04766	PAAG_04379
PAAG_07566	PAAG_04847	PAAG_04414
PAAG_07589	PAAG_04860	PAAG_04458
PAAG_07595	PAAG_04925	PAAG_04466
PAAG_07598	PAAG_04931	PAAG_04489
PAAG_07625	PAAG_04953	PAAG_04504
PAAG_07629	PAAG_04961	PAAG_04525
PAAG_07670	PAAG_04979	PAAG_04540
PAAG_07676	PAAG_04999	PAAG_04541
PAAG_07680	PAAG_05028	PAAG_04542
PAAG_07694	PAAG_05042	PAAG_04545
PAAG_07714	PAAG_05046	PAAG_04549
PAAG_07728	PAAG_05049	PAAG_04559
PAAG_07730	PAAG_05050	PAAG_04563
PAAG_07740	PAAG_05067	PAAG_04570
PAAG_07764	PAAG_05071	PAAG_04572
PAAG_07824	PAAG_05077	PAAG_04599
PAAG_07826	PAAG_05078	PAAG_04605
PAAG_07830	PAAG_05079	PAAG_04614
PAAG_07851	PAAG_05103	PAAG_04620
PAAG_07852	PAAG_05107	PAAG_04636
PAAG_07869	PAAG_05114	PAAG_04637
PAAG_07875	PAAG_05150	PAAG_04642
PAAG_07879	PAAG_05151	PAAG_04651
PAAG_08016	PAAG_05160	PAAG_04669
PAAG_08017	PAAG_05162	PAAG_04692
PAAG_08023	PAAG_05163	PAAG_04698
PAAG_08028	PAAG_05251	PAAG_04701
PAAG_08030	PAAG_05253	PAAG_04726

PAAG_08063	PAAG_05300	PAAG_04732
PAAG_08106	PAAG_05338	PAAG_04810
PAAG_08116	PAAG_05363	PAAG_04819
PAAG_08139	PAAG_05379	PAAG_04821
PAAG_08178	PAAG_05387	PAAG_04824
PAAG_08198	PAAG_05395	PAAG_04838
PAAG_08199	PAAG_05427	PAAG_04846
PAAG_08231	PAAG_05436	PAAG_04851
PAAG_08237	PAAG_05437	PAAG_04866
PAAG_08260	PAAG_05438	PAAG_04886
PAAG_08261	PAAG_05439	PAAG_04888
PAAG_08265	PAAG_05442	PAAG_04891
PAAG_08268	PAAG_05465	PAAG_04892
PAAG_08356	PAAG_05474	PAAG_04896
PAAG_08391	PAAG_05481	PAAG_04899
PAAG_08421	PAAG_05482	PAAG_04903
PAAG_08429	PAAG_05485	PAAG_04904
PAAG_08432	PAAG_05502	PAAG_04913
PAAG_08441	PAAG_05510	PAAG_04914
PAAG_08449	PAAG_05513	PAAG_04917
PAAG_08458	PAAG_05516	PAAG_04959
PAAG_08483	PAAG_05527	PAAG_04966
PAAG_08524	PAAG_05534	PAAG_04976
PAAG_08532	PAAG_05551	PAAG_04996
PAAG_08535	PAAG_05575	PAAG_04998
PAAG_08612	PAAG_05576	PAAG_05007
PAAG_08637	PAAG_05617	PAAG_05018
PAAG_08662	PAAG_05619	PAAG_05026
PAAG_08664	PAAG_05657	PAAG_05029
PAAG_08694	PAAG_05660	PAAG_05036
PAAG_08701	PAAG_05674	PAAG_05037
PAAG_08711	PAAG_05676	PAAG_05048
PAAG_08733	PAAG_05690	PAAG_05064
PAAG_08766	PAAG_05692	PAAG_05068
PAAG_08792	PAAG_05696	PAAG_05091
PAAG_08804	PAAG_05704	PAAG_05128
PAAG_08823	PAAG_05731	PAAG_05140
PAAG_08836	PAAG_05732	PAAG_05147

PAAG_08839	PAAG_05734	PAAG_05148
PAAG_08862	PAAG_05735	PAAG_05171
PAAG_08895	PAAG_05743	PAAG_05179
PAAG_08904	PAAG_05770	PAAG_05182
PAAG_08931	PAAG_05772	PAAG_05194
PAAG_08941	PAAG_05776	PAAG_05226
PAAG_08973	PAAG_05780	PAAG_05232
PAAG_08986	PAAG_05786	PAAG_05248
PAAG_08992	PAAG_05803	PAAG_05249
PAAG_09011	PAAG_05806	PAAG_05250
PAAG_09049	PAAG_05810	PAAG_05256
PAAG_09072	PAAG_05818	PAAG_05265
PAAG_09076	PAAG_05824	PAAG_05297
PAAG_09082	PAAG_05833	PAAG_05331
	PAAG_05859	PAAG_05336
	PAAG_05877	PAAG_05356
	PAAG_05896	PAAG_05367
	PAAG_05932	PAAG_05369
	PAAG_05935	PAAG_05375
	PAAG_05949	PAAG_05376
	PAAG_05972	PAAG_05392
	PAAG_05999	PAAG_05396
	PAAG_06017	PAAG_05404
	PAAG_06027	PAAG_05417
	PAAG_06089	PAAG_05425
	PAAG_06093	PAAG_05426
	PAAG_06103	PAAG_05454
	PAAG_06129	PAAG_05460
	PAAG_06144	PAAG_05484
	PAAG_06150	PAAG_05486
	PAAG_06153	PAAG_05492
	PAAG_06173	PAAG_05497
	PAAG_06178	PAAG_05504
	PAAG_06189	PAAG_05518
	PAAG_06214	PAAG_05543
	PAAG_06224	PAAG_05550
	PAAG_06229	PAAG_05588
	PAAG_06230	PAAG_05590

PAAG_06238	PAAG_05610
PAAG_06279	PAAG_05621
PAAG_06286	PAAG_05628
PAAG_06301	PAAG_05640
PAAG_06322	PAAG_05642
PAAG_06333	PAAG_05643
PAAG_06334	PAAG_05644
PAAG_06343	PAAG_05664
PAAG_06365	PAAG_05679
PAAG_06368	PAAG_05683
PAAG_06380	PAAG_05716
PAAG_06387	PAAG_05737
PAAG_06403	PAAG_05740
PAAG_06404	PAAG_05764
PAAG_06420	PAAG_05781
PAAG_06439	PAAG_05798
PAAG_06450	PAAG_05825
PAAG_06454	PAAG_05847
PAAG_06455	PAAG_05850
PAAG_06464	PAAG_05857
PAAG_06489	PAAG_05860
PAAG_06493	PAAG_05863
PAAG_06494	PAAG_05871
PAAG_06505	PAAG_05879
PAAG_06515	PAAG_05904
PAAG_06529	PAAG_05911
PAAG_06533	PAAG_05928
PAAG_06538	PAAG_05929
PAAG_06542	PAAG_05939
PAAG_06550	PAAG_05944
PAAG_06558	PAAG_05960
PAAG_06561	PAAG_05961
PAAG_06564	PAAG_05963
PAAG_06582	PAAG_05971
PAAG_06583	PAAG_06023
PAAG_06589	PAAG_06031
PAAG_06597	PAAG_06032
PAAG_06606	PAAG_06057

PAAG_06628	PAAG_06068
PAAG_06652	PAAG_06079
PAAG_06668	PAAG_06082
PAAG_06721	PAAG_06083
PAAG_06738	PAAG_06084
PAAG_06765	PAAG_06102
PAAG_06766	PAAG_06132
PAAG_06783	PAAG_06136
PAAG_06858	PAAG_06155
PAAG_06862	PAAG_06156
PAAG_06891	PAAG_06168
PAAG_06947	PAAG_06172
PAAG_06966	PAAG_06175
PAAG_06970	PAAG_06192
PAAG_06984	PAAG_06206
PAAG_07003	PAAG_06211
PAAG_07016	PAAG_06222
PAAG_07019	PAAG_06228
PAAG_07030	PAAG_06237
PAAG_07036	PAAG_06250
PAAG_07047	PAAG_06254
PAAG_07067	PAAG_06261
PAAG_07070	PAAG_06269
PAAG_07080	PAAG_06281
PAAG_07086	PAAG_06299
PAAG_07114	PAAG_06309
PAAG_07131	PAAG_06320
PAAG_07133	PAAG_06329
PAAG_07154	PAAG_06344
PAAG_07159	PAAG_06347
PAAG_07180	PAAG_06363
PAAG_07208	PAAG_06367
PAAG_07243	PAAG_06377
PAAG_07266	PAAG_06414
PAAG_07276	PAAG_06416
PAAG_07313	PAAG_06444
PAAG_07340	PAAG_06484
PAAG_07349	PAAG_06487

PAAG_07359	PAAG_06501
PAAG_07375	PAAG_06503
PAAG_07393	PAAG_06506
PAAG_07402	PAAG_06511
PAAG_07414	PAAG_06513
PAAG_07432	PAAG_06524
PAAG_07472	PAAG_06541
PAAG_07476	PAAG_06574
PAAG_07484	PAAG_06581
PAAG_07497	PAAG_06617
PAAG_07516	PAAG_06643
PAAG_07518	PAAG_06671
PAAG_07528	PAAG_06673
PAAG_07532	PAAG_06693
PAAG_07564	PAAG_06701
PAAG_07567	PAAG_06703
PAAG_07572	PAAG_06751
PAAG_07583	PAAG_06769
PAAG_07586	PAAG_06776
PAAG_07610	PAAG_06777
PAAG_07631	PAAG_06796
PAAG_07682	PAAG_06807
PAAG_07722	PAAG_06809
PAAG_07760	PAAG_06816
PAAG_07769	PAAG_06823
PAAG_07775	PAAG_06835
PAAG_07777	PAAG_06843
PAAG_07791	PAAG_06851
PAAG_07804	PAAG_06869
PAAG_07831	PAAG_06873
PAAG_07833	PAAG_06876
PAAG_07845	PAAG_06886
PAAG_07853	PAAG_06906
PAAG_07856	PAAG_06932
PAAG_07864	PAAG_06945
PAAG_07878	PAAG_06949
PAAG_07887	PAAG_06951
PAAG_07900	PAAG_06977

PAAG_07902	PAAG_06986
PAAG_07948	PAAG_06996
PAAG_07955	PAAG_07013
PAAG_07959	PAAG_07017
PAAG_07998	PAAG_07028
PAAG_08019	PAAG_07035
PAAG_08020	PAAG_07038
PAAG_08059	PAAG_07039
PAAG_08064	PAAG_07051
PAAG_08082	PAAG_07054
PAAG_08086	PAAG_07060
PAAG_08088	PAAG_07089
PAAG_08151	PAAG_07103
PAAG_08152	PAAG_07108
PAAG_08153	PAAG_07150
PAAG_08155	PAAG_07160
PAAG_08163	PAAG_07165
PAAG_08170	PAAG_07170
PAAG_08184	PAAG_07182
PAAG_08203	PAAG_07188
PAAG_08217	PAAG_07192
PAAG_08220	PAAG_07210
PAAG_08224	PAAG_07264
PAAG_08230	PAAG_07267
PAAG_08251	PAAG_07279
PAAG_08279	PAAG_07285
PAAG_08295	PAAG_07307
PAAG_08324	PAAG_07312
PAAG_08337	PAAG_07321
PAAG_08338	PAAG_07335
PAAG_08341	PAAG_07339
PAAG_08349	PAAG_07348
PAAG_08359	PAAG_07353
PAAG_08360	PAAG_07354
PAAG_08365	PAAG_07368
PAAG_08382	PAAG_07374
PAAG_08399	PAAG_07377
PAAG_08420	PAAG_07378

PAAG_08426	PAAG_07440
PAAG_08439	PAAG_07447
PAAG_08447	PAAG_07477
PAAG_08463	PAAG_07482
PAAG_08473	PAAG_07492
PAAG_08479	PAAG_07502
PAAG_08495	PAAG_07505
PAAG_08558	PAAG_07509
PAAG_08569	PAAG_07529
PAAG_08583	PAAG_07541
PAAG_08591	PAAG_07577
PAAG_08620	PAAG_07592
PAAG_08633	PAAG_07605
PAAG_08655	PAAG_07633
PAAG_08658	PAAG_07634
PAAG_08666	PAAG_07638
PAAG_08674	PAAG_07675
PAAG_08685	PAAG_07702
PAAG_08700	PAAG_07707
PAAG_08725	PAAG_07708
PAAG_08729	PAAG_07717
PAAG_08743	PAAG_07734
PAAG_08747	PAAG_07735
PAAG_08758	PAAG_07742
PAAG_08770	PAAG_07746
PAAG_08795	PAAG_07750
PAAG_08876	PAAG_07772
PAAG_08878	PAAG_07786
PAAG_08881	PAAG_07788
PAAG_08903	PAAG_07811
PAAG_08915	PAAG_07813
PAAG_08920	PAAG_07821
PAAG_08926	PAAG_07832
PAAG_08944	PAAG_07834
PAAG_08951	PAAG_07840
PAAG_08966	PAAG_07842
PAAG_09003	PAAG_07866
PAAG_09013	PAAG_07868

PAAG_09017	PAAG_07873
PAAG_09028	PAAG_07883
PAAG_09032	PAAG_07886
PAAG_09053	PAAG_07903
PAAG_09080	PAAG_07909
PAAG_09085	PAAG_07917
PAAG_09086	PAAG_07925
PAAG_09094	PAAG_07956
PAAG_09118	PAAG_07961
PAAG_09123	PAAG_07987
	PAAG_07989
	PAAG_07992
	PAAG_08003
	PAAG_08006
	PAAG_08013
	PAAG_08032
	PAAG_08037
	PAAG_08050
	PAAG_08051
	PAAG_08052
	PAAG_08057
	PAAG_08058
	PAAG_08071
	PAAG_08074
	PAAG_08075
	PAAG_08095
	PAAG_08097
	PAAG_08103
	PAAG_08110
	PAAG_08118
	PAAG_08137
	PAAG_08141
	PAAG_08145
	PAAG_08148
	PAAG_08157
	PAAG_08164
	PAAG_08166
	PAAG_08174

PAAG_08196
PAAG_08227
PAAG_08247
PAAG_08252
PAAG_08280
PAAG_08297
PAAG_08301
PAAG_08303
PAAG_08308
PAAG_08313
PAAG_08315
PAAG_08323
PAAG_08325
PAAG_08336
PAAG_08340
PAAG_08351
PAAG_08357
PAAG_08369
PAAG_08370
PAAG_08375
PAAG_08386
PAAG_08389
PAAG_08400
PAAG_08405
PAAG_08414
PAAG_08433
PAAG_08450
PAAG_08451
PAAG_08468
PAAG_08469
PAAG_08488
PAAG_08505
PAAG_08517
PAAG_08529
PAAG_08539
PAAG_08545
PAAG_08548
PAAG_08587

PAAG_08593
PAAG_08596
PAAG_08597
PAAG_08606
PAAG_08623
PAAG_08634
PAAG_08680
PAAG_08682
PAAG_08693
PAAG_08695
PAAG_08721
PAAG_08745
PAAG_08748
PAAG_08759
PAAG_08797
PAAG_08801
PAAG_08805
PAAG_08820
PAAG_08834
PAAG_08835
PAAG_08837
PAAG_08847
PAAG_08856
PAAG_08859
PAAG_08875
PAAG_08880
PAAG_08882
PAAG_08888
PAAG_08908
PAAG_08913
PAAG_08919
PAAG_08937
PAAG_08955
PAAG_08969
PAAG_08981
PAAG_08987
PAAG_08989
PAAG_08990

PAAG_08994
PAAG_09004
PAAG_09006
PAAG_09015
PAAG_09043
PAAG_09046
PAAG_09050
PAAG_09056
PAAG_09062
PAAG_09074
PAAG_09083
PAAG_09100