



UNIVERSIDADE FEDERAL DO MARANHÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM OCEANOGRAFIA

IGOR SANTANA RIBEIRO

**MODELAGEM METABÓLICA DA CIANOBACTÉRIA *Geminocystis* sp.
GBBB08, ISOLADA NO PARQUE NACIONAL DA CHAPADA DAS
MESAS, MARANHÃO**

SÃO LUÍS
2022

IGOR SANTANA RIBEIRO

**MODELAGEM METABÓLICA DA CIANOBACTÉRIA *Geminocystis* sp.
GBBB08, ISOLADA NO PARQUE NACIONAL DA CHAPADA DAS
MESAS, MARANHÃO**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência e Tecnologia / UFMA como parte dos pré-requisitos para obtenção do título de mestre.

Orientador(a): Dr. Leonardo Texeira Dall'Agnol

Coorientador(a): Dr. Alex Ranieri Jerônimo Lima

Linha de pesquisa: Biodiversidade e Ecologia de Ecossistemas Aquáticos de acordo com as linhas do PPGOceano/PPGC&TAmb

SÃO LUÍS
2022

IGOR SANTANA RIBEIRO

**MODELAGEM METABÓLICA DA CIANOBACTÉRIA *Geminocystis* sp.
GBBB08, ISOLADA NO PARQUE NACIONAL DA CHAPADA DAS
MESAS, MARANHÃO**

Dissertação apresentada ao Programa de
Pós-Graduação em Ciência e Tecnologia /
UFMA como parte dos pré-requisitos para
obtenção do título de mestre.

Aprovada em: 31 / 10 / 2022

BANCA EXAMINADORA

Prof. Dr. Leonardo Texeira Dall’Agnol (DEBIO/UFMA) – Orientador

Prof(a) Dr(a). Hivana Patricia Melo Barbosa Dall’ Agnol (DPAT/UFMA)

Dr. Lucas Miguel de Carvalho (UNICAMP) - Membro Externo

Dra.Tania Keiko Shishido (Universidade de Helsinque) - Membro Externo

Ficha gerada por meio do SIGAA/Biblioteca com dados fornecidos pelo(a) autor(a).
Diretoria Integrada de Bibliotecas/UFMA

Santana Ribeiro, Igor.

MODELAGEM METABÓLICA DA CIANOBACTÉRIA *Geminocystis* sp. GBBB08, ISOLADA NO PARQUE NACIONAL DA CHAPADA DAS MESAS, MARANHÃO / Igor Santana Ribeiro. - 2022.

66 p.

Coorientador(a): Alex Ranieri Jerônimo Lima.

Orientador(a): Leonardo Texeira Dallagnol.

Dissertação (Mestrado) - Programa de Pós-graduação em Oceanografia/ccbs, Universidade Federal do Maranhão, São Luís, 2022.

1. Engenharia metabólica. 2. Flux Balance Analysis. 3. Mineração genômica. 4. Modelo metabólico em escala genômica. 5. Reconstrução baseada em restrições. I. Jerônimo Lima, Alex Ranieri. II. Texeira Dallagnol, Leonardo. III. Título.

AGRADECIMENTOS

Com certeza a parte mais difícil e mais esperada de fazer deste trabalho, agradecer. Primeiramente a Deus que na sua bondade foi e é fiel mesmo em tempos tão difíceis. Ele, que não me desamparou e cuidou de mim mesmo na minha infidelidade.

Agradeço incomensuravelmente ao grande pai, João Ribeiro, que me acompanhou ao longo da minha formação e que hoje está na Graça de Deus (*in memoriam*). Meu grande amigo, mentor, parceiro, conselheiro, meu pai. A ti dedico esse trabalho que foi com muita luta para perseverar. Não tenho palavras para dizer o quanto o senhor foi marcante para mim.

À minha família pelo apoio nesses tempos, em especial, minhas mães Conceição Santana e Maria Rangel que souberam ser mães pra me consolar e me dar forças acompanhado da minha namorada Camila Vieira que soube auxiliar com paciência e carinho. O suporte família da minha namorada e em especial por Israel Vieira e Osvaldo Vieira, pessoas que inseriram em conversas familiares assuntos de TI e conseguiram me dar estímulo.

Ao meu orientador, Leonardo Dall' Agnol, que buscou além de me orientar, deu-me ânimo, estímulo numa dosagem firme e ao mesmo tempo muito respeitosa, juntamente com sua esposa, Hivana Dall' Agnol. Além disso, me proporcionaram um caminho na pesquisa e em especial na Bioinformática. Sou muito grato ao longo desses 6 anos de pesquisa no nosso grupo GB₃ que foram de experiências únicas.

Ao Alex Ranieiri que é meu co-orientador que é um grande parceiro, uma referência para mim que me estimulou na Biologia de Sistemas e na programação. A banca da qualificação que fizeram correções que me auxiliaram muito na pesquisa e em especial ao Lucas Miguel que trouxe grandes *insights* para a formação desse trabalho

Ao programa PPGOceano, em especial ao prof. Marco Valério, prof(a). Marianna Basso e prof. Ricardo Luvizotto que contribuíram em momentos nas disciplinas e também remotamente.

Gostaria de agradecer a minha formadora comunitária Fátima Régia e meu formador pessoal Márcio que foram voz de Deus para mim. A minha comunidade na qual amo que é a minha vocação Comunidade Católica Shalom.

Agradeço também aos profissionais de saúde que com muito profissionalismo e zelo também me auxiliaram que são a minha psicóloga Miranethy, meu médico psiquiatra Dr. Vitor e também a minha médica psicanalista Dra Janaína.

EPÍGRAFE

Lembre-se da Minha Ordem:

“Seja forte e corajoso! Não fique desanimado, nem tenha medo, porque Eu, o Senhor, seu Deus, estarei com você em qualquer lugar para onde você for!” Js 1, 9-11

“Se me amais, guardai meus mandamentos; e eu pedirei ao Pai que vos envie o Valedor que esteja convosco sempre: o Espírito da verdade, que o mundo não o vê nem o conhece. Vós o conheceis, pois permanece convosco e está em vós. Não vos deixo órfãos, mas voltarei a visitar-vos.” Jo 14, 15-

RESUMO

As ferramentas de bioinformática para a análise de predição de dados ômicos têm crescido exponencialmente aliadas ao avanço das novas tecnologias de sequenciamento e da computação. Modelos em escala genômica são instrumentos eficazes para a engenharia metabólica e investigação das redes metabólicas. Porém, essas reconstruções podem demandar muito tempo para abordagens mais amplas, e sua qualidade é dependente do conjunto de características bioquímicas e fenotípicas a serem avaliadas, além dos dados disponíveis sobre o grupo em estudo. Nesse trabalho utilizamos a integração da análise mineração de genoma (antiSMASH e MIBiG) para descoberta de vias biossintéticas de produtos naturais aliado a ferramentas de reconstrução metabólica em escala genômica (CarveMe) do genoma da cianobactéria *Geminocystis* sp. GBBB08. As análises comparativas levaram a reprodução *in silico* da via de produção de Terpeno (Não – Melavonato) com a análise de balanço de fluxo a fornecer características metabólicas para a biossíntese dos terpenóides envolvidos. Além disso, foi identificado agrupamentos gênicos biossintéticos de anabaenopeptina e de heptadeceno. De modo geral, o genoma da *Geminocystis* sp. GBBB08 fornece dados importantes ao potencial metabólico do gênero com uma abordagem *in silico* para uma via metabólica frequente e de relevância econômica como a dos terpenóides. As análises comparativas em abordagem de mineração genômica e biologia de sistemas podem favorecer a reconstrução de redes metabólicas e levar a uma compreensão melhor do metabolismo e seu potencial biotecnológico.

Palavras-chave: Mineração genômica, engenharia metabólica, modelo metabólico em escala genômica, reconstrução baseada em restrições, Flux Balance Analysis

ABSTRACT

Bioinformatics tools for predictive analysis of omic data have grown exponentially together with the advancement of new sequencing and computing technologies. Genomic scale models are effective tools for metabolic engineering and investigation of metabolic networks. However, these reconstructions can demand a lot of time for broader approaches, and their quality depends on the set of biochemical and phenotypic characteristics to be evaluated, in addition to the available data on the group under study. In this work, we used the integration of genome mining analysis (antiSMASH and MIBiG) for the discovery of biosynthetic natural products combined with genomic-scale metabolic reconstruction tools (CarveMe) of the genome of the cyanobacterium *Geminocystis* sp. GBBB08. Comparative analyzes led to the in silico reproduction of the Terpene (Non-Melavonate) production pathway with the flow balance analysis providing metabolic characteristics for the biosynthesis of the terpenoids involved. In addition, biosynthetic gene clusters of anabaenopeptin and heptadecene were identified. In general, the genome of *Geminocystis* sp. GBBB08 provides important data on the metabolic potential of the genus with an in silico approach to a frequent and economically relevant metabolic pathway such as terpenoids. Comparative analyzes in a genomic mining approach and systems biology can favor the reconstruction of metabolic networks and lead to a better understanding of metabolism and its biotechnological potential.

Keywords: Genome mining, metabolic engineering, genome-scale metabolic model, constraint-based reconstruction, FBA

SUMÁRIO

1. INTRODUÇÃO	11
2. FUNDAMENTAÇÃO TEÓRICA	12
2.1. O GÊNERO <i>GEMINOCYSTIS</i>	12
2.2. A ENGENHARIA GENÉTICA DE HIDROCARBONETOS PARA AS CIANOBACTÉRIAS.....	12
2.3. A VIA DE PRODUÇÃO DE TERPENOS E TERPENÓIDES DE CIANOBACTÉRIAS.....	13
2.4. RECONSTRUÇÃO DE MODELOS METABÓLICOS EM SUAS ABORDAGENS.....	16
2.5. ANÁLISE DE MODELOS METABÓLICOS.....	22
2.5.1. Análise de Balanço de Fluxo (FBA)	22
2.6. FLUXO DE TRABALHO DE RECONSTRUÇÃO DE MODELOS METABÓLICOS DE INTERESSE <i>IN SILICO</i>	25
2.6.1. Reconstrução do rascunho das vias (fase preliminar) - 1º Etapa	25
2.6.2. Curadoria manual do rascunho (refinamento) – 2º Etapa	25
2.6.3. Conversão da reconstrução curada das vias metabólicas em um modelo computacional – Estágio 3	26
2.6.4. Validação do modelo metabólico ou uso de reconstrução e integração de dados de alto desempenho – Estágio 4	26
3. JUSTIFICATIVA	27
4. OBJETIVOS	28
4.1. GERAL.....	28
4.2. ESPECÍFICOS.....	28
5. MATERIAL E MÉTODOS	29
5.1. AMOSTRA UTILIZADA E CONDIÇÕES DE MANUTENÇÃO DO CULTIVO.....	29
5.2. SEQUENCIAMENTO GENÔMICO, MONTAGEM E ANOTAÇÃO.....	29
5.3. MINERAÇÃO DO GENOMA E ANÁLISE COMPARATIVA.....	30
5.4. O PROCESSO DE RECONSTRUÇÃO DA REDE METABÓLICA.....	30
5.5. APLICAÇÕES DE ENGENHARIA METABÓLICA NO MODELO.....	32
5.5.1. Identificação das Reações da Via MEP	32
6. RESULTADOS E DISCUSSÃO	33
1.1. FILOGENÔMICA, MINERAÇÃO DO GENOMA E ANÁLISE COMPARATIVA	33
6.1. VERIFICAÇÃO DA RECONSTRUÇÃO DOS MODELOS METABÓLICOS DA <i>GEMINOCYSTIS</i> . SP. GBBB08 E SIMULAÇÃO FBA.....	39
6.2. ANÁLISE DE BALANÇO DE FLUXO PARCIMONIOSO (PFBA) PARA A VIA DE TERPENOS.....	41
7. CONCLUSÃO	48
8. REFERÊNCIAS BIBLIOGRÁFICAS	49
9. ANEXOS	59

1. INTRODUÇÃO

As cianobactérias são organismos procariotos do reino Bacteria, ou seja, sem núcleo definido e ausência de organelas membranares. Possuem parede celular de peptidoglicano, além de produzirem uma mucilagem que pode ser encontrada revestindo essa parede. Apresentam tamanhos e arranjos variados com espécies unicelulares, coloniais a filamentosas (TOMASELLI, 2004). São consideradas fotoautotóficas, capazes de converter a energia luminosa, principalmente pela ação da clorofila a, em íons de carbono orgânico. Possuem uma grande diversidade taxonômica de famílias, com importante papel na regulação do ciclo biogeoquímico de oxigênio e carbono a partir da fotossíntese, além do ciclo do nitrogênio a partir da fixação de N₂. Além de sua diversidade morfofisiológica, fruto do tempo de evolução a que estão submetidas, elas possuem distribuição cosmopolita e de ampla variedade de habitats, o que se reflete na expressão de uma gama de metabólitos secundários (PERCIVAL; WILLIAMS, 2014; SINGH, S. P.; MONTGOMERY, 2011).

Diante de habitats variados e sucessivas pressões seletivas, as cianobactérias exibem um potencial grande de produção de metabólitos secundários que favorecem seu sucesso evolutivo. Metabólitos foto protetores, como a citoemina e micosporina, são exemplos de compostos que tem a função de redução dos danos da radiação ultravioleta. Além disso, destacam-se devido à habilidade de produção de biocombustíveis como biodiesel ou biobutanol aliado ao crescimento fotoautotrófico de baixo custo. Esses biocombustíveis de 3ª geração, derivados de microalgas e cianobactérias (LEE, R. A.; LAVOIE, 2012; ZILLIGES *et al.*, 2013), têm atraído crescente interesse por aliar a produção de baixo custo à absorção de CO₂ gerado nas atividades antrópicas (RASTOGI; SINHA, RAJESHWAR P, 2009; SINGH, S. P. *et al.*, 2008; SINHA, R P; KLISCH; H, 1998).

Neste contexto ressalta-se a importância de se conhecer os mecanismos e fluxos metabólicos, com seus respectivos passos, de maneira a promover uma melhor compreensão das interconversões realizadas durante o crescimento fotoautotrófico (DUCAT; WAY; SILVER, 2011; HESS, 2011).

A genômica funcional, a mineração genômica e o estudo dos fluxos metabólicos, utilizando ferramentas *in silico*, favorecem uma compreensão mais profunda do funcionamento dos organismos e permitem modelar o comportamento do sistema apesar dos desafios acerca da complexidade do mesmo (BRUGGEMAN; WESTERHOFF, 2007).

2. FUNDAMENTAÇÃO TEÓRICA

2.1. O GÊNERO *GEMINOCYSTIS*

Este novo gênero de cianobactéria foi descrito por Korelusová em 2009, a partir de uma linhagem identificada como *Synechocystis*. No entanto, esta cepa, quando comparada com a linhagem de referência *Synechocystis* sp. PCC 6803, apresentou algumas diferenças, como taxas menores que 90% de similaridade gênica, com baixa quantidade de conteúdo GC (34,7%) em comparação com a alta porcentagem de GC da PCC 6803 (47,5%) e diferentes estruturas secundárias na região ITS (KORELUSOVA; KASTOVSKY; KOMAREK, 2009).

Segundo Korelusová (2009), o gênero *Geminocystis* apresenta células procarióticas, solitárias (unicelulares), esféricas ou levemente ovais e divisão hemisférica na divisão binária. As células possuem de 3 a 10 µm de diâmetro, sem mucilagem ou involucro, incolor ao redor das células solitárias. Células aeruginosas, verde acinzentada ou rosa pálido. O conteúdo celular é homogêneo, sem separação do centro cromatoplasma. As células do sistema tilacoide tem largura mais ou menos irregular. A divisão em duas células é equivalente em duas gerações subsequentes. As células se dividem em dois planos perpendiculares em gerações subsequentes.

2.2. A ENGENHARIA GENÉTICA DE HIDROCARBONETOS PARA AS CIANOBACTÉRIAS

As cianobactérias têm grande impacto na Terra e um dos potenciais econômicos que essas possuem é a capacidade de produzir hidrocarbonetos associados aos biocombustíveis. As primeiras análises de identificação de produção de cadeias longas de hidrocarbonetos foram em 1960 (COATES *et al.*, 2014; HAN *et al.*, 1968; HAN; CHAN; CALVIN, 1969; WINTERS; PARKER; BAALEN, VAN, 1969).

Os hidrocarbonetos são basicamente compostos químicos que contém hidrogênio e carbono. Suas diferenças são descritas em: variação do tamanho da cadeia, os

rearranjos dessa cadeia e níveis de saturação (XIE *et al.*, 2017). A produção por cianobactérias e microalgas tem despertado interesse de pesquisadores e empreendedores por ser uma das soluções para a demanda de energia aplicadas para a diversificação no uso de matriz energética (HANKAMER *et al.*, 2007; PARMAR *et al.*, 2011).

Seja para a produção de biogás, biohidrogênio, bioetanol, biodiesel e biobutanol, as cianobactérias têm certas vantagens na produção em relação as plantas superiores como o milho, cana-de-açúcar, algas, dendê, como por exemplo: elas não estão associadas a inflação de alimentos, não envolvem grandes áreas agrícolas, utilizam água como doador de elétrons durante a fotossíntese, além de, ter um cultivo de massa renovável (PARMAR *et al.*, 2011; TAMAGNINI *et al.*, 2007).

Há diferenças em manipulação genética de cianobactérias em relação a plantas e microalgas. A conversão do carbono atmosférico fixado por cianobactérias em biomassa bruta é uma das vantagens na engenharia genética, contudo, diversos produtos metabólicos podem ser avaliados para a produção de interesse biotecnológico, como os compostos foto protetores ou outros metabólitos farmacêuticos como os antivirais ou antibióticos (QUINTANA *et al.*, 2011; SARSEKEYEVA *et al.*, 2015).

2.3. A VIA DE PRODUÇÃO DE TERPENOS E TERPENÓIDES DE CIANOACTÉRIAS

Os terpenóides é uma classe superior de produtos naturais que pode ser encontrado nas diversas formas de vida e que podem desempenhar tantas funções estruturais (a partir de colesterol nas membranas) ou características funcionais (quinonas na transferência de elétrons, carotenoides na fotossíntese, etc) (CHAPPELL, 1995; OLDFIELD; LIN, 2012).

A origem é a partir de um substrato de difosfato de dimetilalil (DMAPP) (C₅) e isopentil difosfato (IPP) que a cadeia tem suas variações com a condensação inicial de DMAPP seja com uma ou mais moléculas de IPP (OLDFIELD; LIN, F. Y., 2012). A produção de terpeno apresenta duas vias distintas de produção: a via do melavonato (MVA) e a via do Metileritritol-fosfato (MEP). A via MVA que usa acetil-CoA como substrato e converte-o em seis etapas, em que, o melavonato é um intermediário para IPP que

pode ser interconvertido em DMAPP por uma isomerase (Figura 1). Essa via é a via exclusiva em animais e fungos, podendo ocorrer em plantas, algas e bactérias.

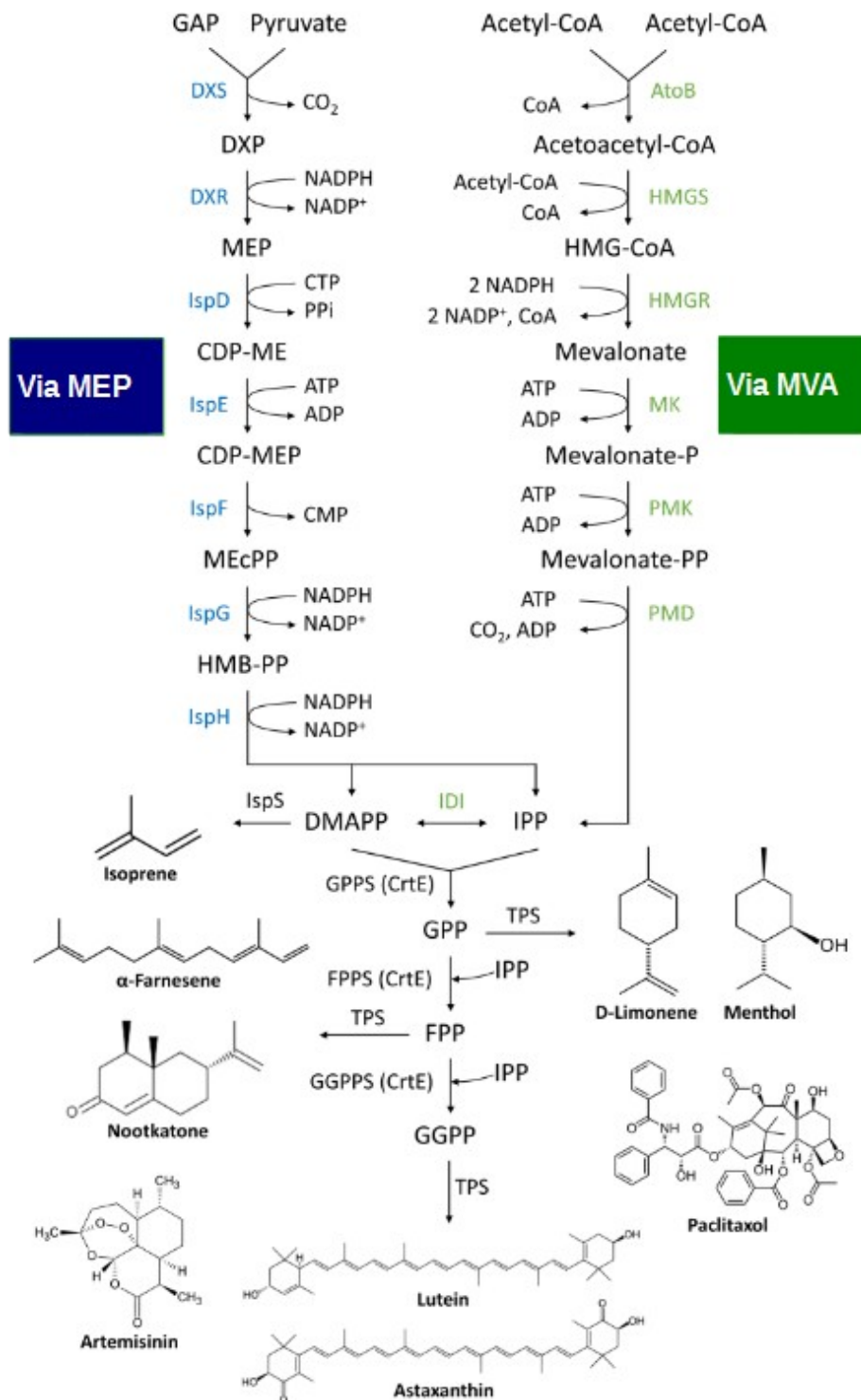


Figura 1. Vias metabólicas da biossíntese de terpenoides, adaptado de Lin e Pakrasi (2019).

A via MEP, por sua vez, é a via de produção de terpenóides em cianobactérias e contém 7 reações enzimáticas e onde seu início é a partir do:

1°. Gliceraldeído 3-fosfato (GAP) e o piruvato para formar o DXP;

2°. A DXP redutoisomerase (IspC ou DXR) reduz DXP a 2-C-metil-D-eritritol-4-fosfato (MEP);

3°. A MEP-citidiltransferase (IspD) que catalisa a citidilação de MEP em 4-difosfocitidil-2-C-metil-D-eritritol-2-fosfato (CDP-ME);

4°. CDP-ME quinase (ispE) que fosforila CDP-ME em difosfocitidil-2-C-metil-D-eritritol-2-fosfato (CDP-MEP);

5°. O 2-C-metil-D-eritritol-2,4-ciclodifosfato sintase (MecPP Sintase ou IspF) que catalisa a síntese de MecPP a partir de CDP-MEP;

6°. O 1-hidroxi-2-metil-2-(E)-butenil-4-difosfato sintase (HMBPP sintase ou IspG) que catalisa a redução de MecP para formar um difosfato linear HMBPP;

7°. HMBPP redutase (IspH) que catalisa a produção de IPP e DMAPP a partir de HMBPP para produzir a estrutura principal de isoprenóides (LIN, P. C.; PAKRASI, 2019; OLDFIELD; LIN, F. Y., 2012).

Os terpenóides são caracterizados em grupos de acordo com o número de unidades de isoprenos: hemiterpenos (C5), monoterpenos (C10), sesquiterpenos (C15), diterpenos (C20), triterpenos (C30), tetraterpenos (C40) e politerpenos (com mais de C40) (AJIKUMAR *et al.*, 2008; LIN, P. C.; PAKRASI, 2019).

Os terpenóides possuem função central na formação dos centros de reação fotossintética, além de estarem relacionados a resposta a estresses ou mecanismos de defesa, sendo, portanto, essenciais em organismos fotoautotróficos. Entre as aplicações biotecnológicas conhecidas podemos citar: (i) a artemisina, com ação antimalária (*Artemisia annua* - erva) ; (ii) a Paclitaxel com atividade anticâncer (*Taxus brevifolia* – casca dessa conífera); (iii) a avicina com atividade antitumoral (*Acacia victoriae* - leguminosa); e (iv) carotenóides como licopeno e astaxantina (microalgas, fungos e diversos organismos) com ação antioxidante (PATTANAİK; LINDBERG, 2015).

2.4. RECONSTRUÇÃO DE MODELOS METABÓLICOS EM SUAS ABORDAGENS

A microbiologia se desenvolveu a partir de estudos de características fenotípicas específicas a partir de organismos isolados, contudo, com o avanço das tecnologias de

sequenciamento de DNA e computação, ampliou-se para estudos sistêmicos como os microarranjos e posteriormente os sequenciamentos genômicos/transcriptômicos, os quais promovem uma melhor compreensão a nível gênico e molecular. Contudo, as moléculas não se encontram nas células sem interação com outras células e ou moléculas, o que leva a um estudo cada vez mais interativo que busca o entendimento no comportamento celular (BARABÁSI; OLTVAI, 2004).

A interdisciplinaridade de biologia de sistemas é proveniente de múltiplas abordagens ômicas geradas por técnicas de sequenciamento de DNA de nova geração, com compreensões mais sensíveis sobre a fisiologia da células, as interações ambientais, os organismos, a comunidade, o seu crescimento e progressão para doenças (PINU *et al.*, 2019).

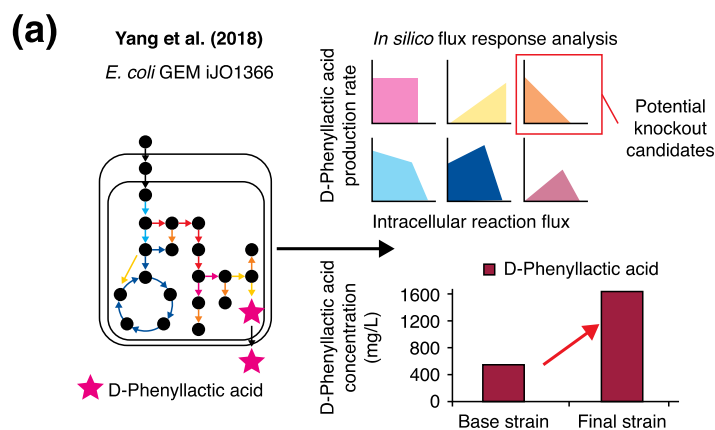
Na biologia de sistemas, existem duas abordagens para avaliar as redes de interação molecular: (i) top down (de cima para baixo), onde se parte do genoma completo e anotado, com base no comportamento molecular sobre a ótica ôhmica; (ii) ou bottom up (de baixo para cima) que aborda das propriedades funcionais, com as interações de componentes conhecidos e já estabelecidos. Com isso, tem-se um método para avaliar as interações moleculares, com seus respectivos sistemas, na busca de informações, de vias bioquímicas, produtos formados, consumo de substratos que leva a descrições de princípios realizáveis no ambiente celular (BRUGGEMAN; WESTERHOFF, 2007).

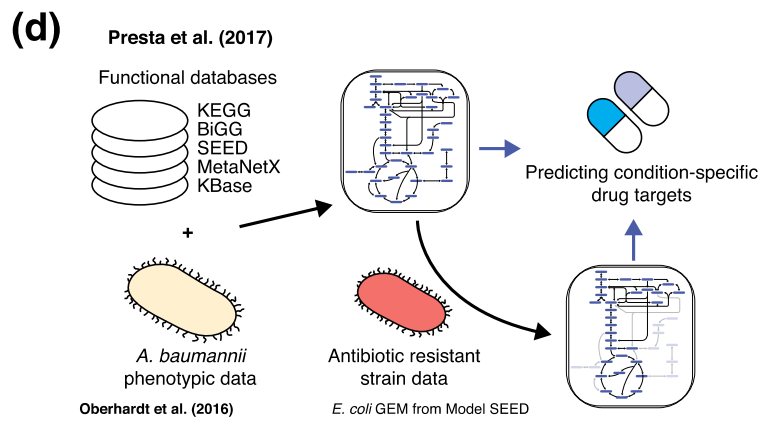
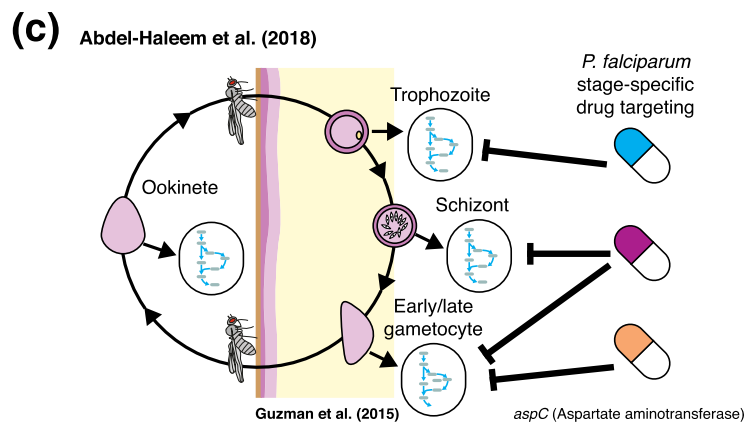
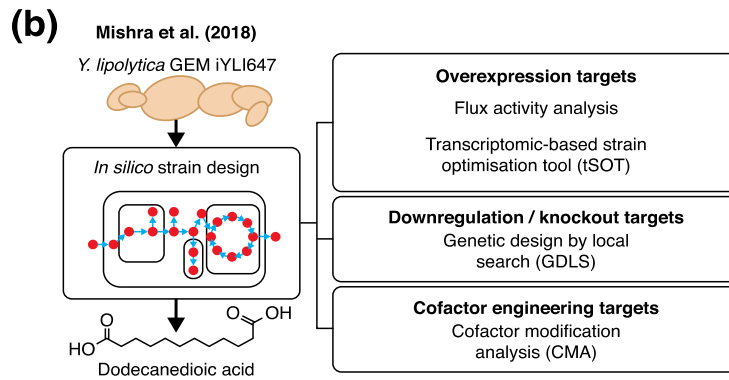
A abordagem Top-Down tem como base a criação de modelos reguladores para analisar dados, em que, parte de dados experimentais obtidos por dados ômicos (metabolômicos, transcriptômicos, proteômicos, fluxômicos e genômicos) com o intuito de buscar nestes dados, uma correlação nas concentrações dessas moléculas. A partir disso, pode-se formar questionamentos diante da regulação ou coregulação dos grupos de moléculas. As hipóteses desenvolvidas diante desses questionamentos condicionam a novas correlações nas quais podem ser ciclicamente testadas diante do experimento a ser repetido por análises bioquímicas (BRUGGEMAN; WESTERHOFF, 2007).

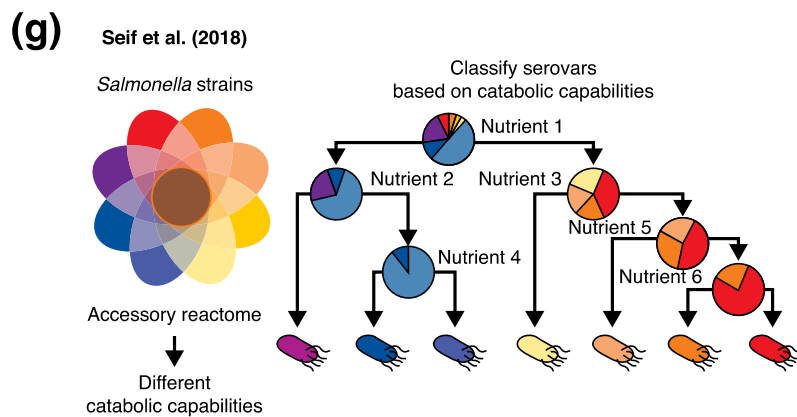
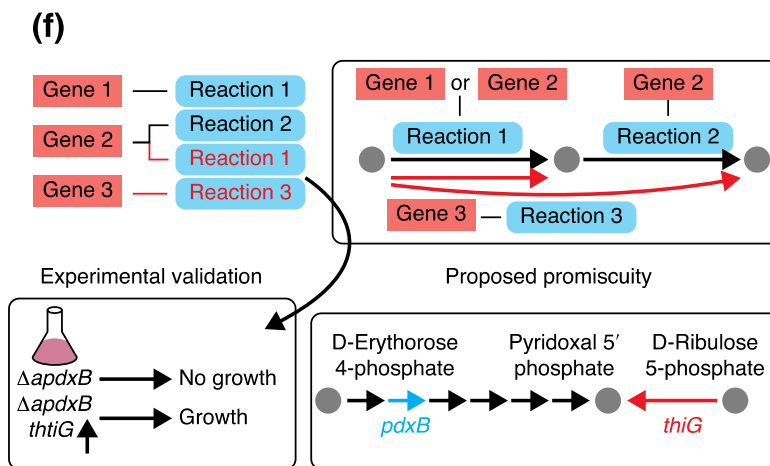
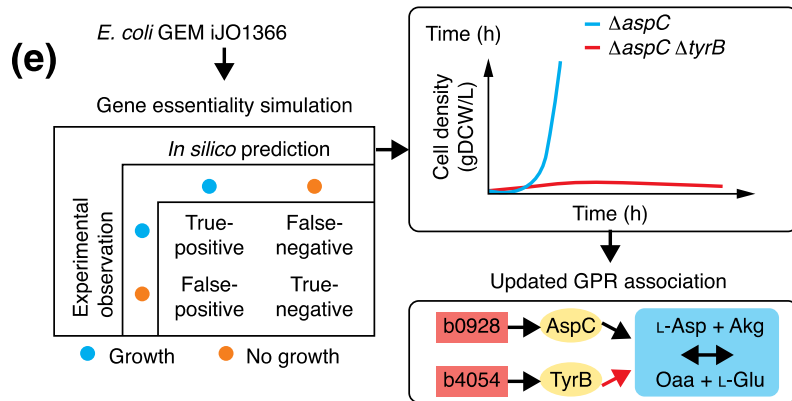
Enquanto na abordagem Bottom Up, o princípio é as partes constitutivas, a própria molécula em si, para com isso ir se formulando os comportamentos interativos em cada processo que levam a integração com formulações para predizer o comportamento do sistema. O cerne desta abordagem é combinar cada modelo de um trecho para todo o

modelo do sistema (BRUGGEMAN; WESTERHOFF, 2007). Nessa perspectiva, basicamente não se faz a busca de informações por induções como na abordagem em Top-Down, a Bottom-up tem como identidade a busca da funcionalidade deste metabólito derivado também de um subsistema bem caracterizado, com níveis de detalhes mais específicos diante dos métodos moleculares (BRUGGEMAN; WESTERHOFF, 2007; KHOLODENKO, BORIS N *et al.*, 1999; KIYATKIN, ANATOLY *et al.*, 2006; MARKEVICH, N I *et al.*, 2004; SUENAGA *et al.*, 2004).

A essência da reconstrução de modelos metabólicos de escala genômica é a montagem e a organização de todas as reações bioquímicas constituídas no organismo alvo (MONTAGUD *et al.*, 2010; OBERHARDT; PALSSON, B. Ø.; PAPIN, 2009). Diante dessa reconstrução de modo otimizada, bem avaliada e com o modelo bem formulado, se realiza a confecção de um organismo virtual *in silico* que se pode imprimir simulações ambientais, perturbações genéticas em distribuições de fluxo diante desses fatores em diferentes condições ambientais simuladas. A Figura 2 traz informações de quais aplicabilidades as reconstruções de redes modelos metabólicos em escala genômica (GSM – *Genome Scale Model*) com o uso de restrições podem ser aplicadas (GU *et al.*, 2019).







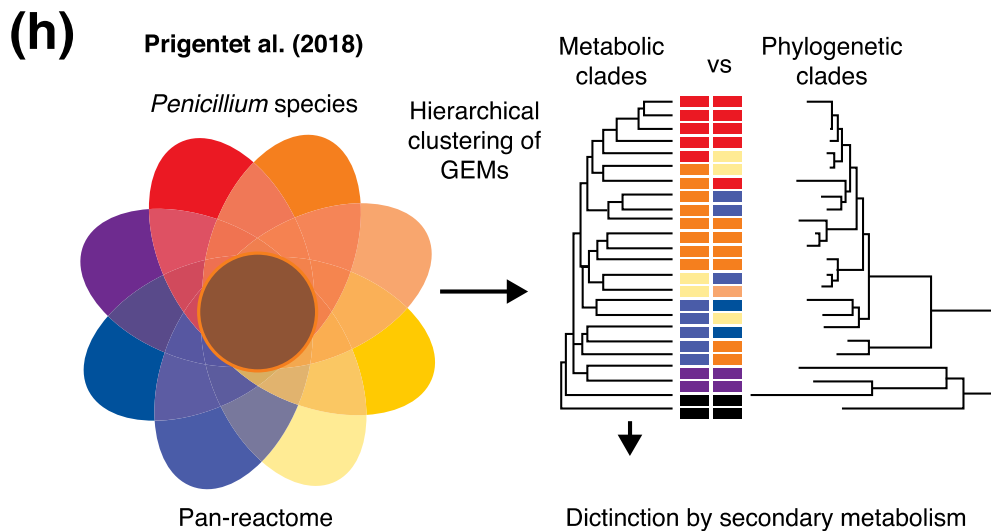


Figura 2. Aplicações de GSM para produção de produtos químicos e materiais, direcionamento de medicamentos a patógenos, previsão de funções enzimáticas e análise de pan-reactome (descrição no artigo publicado). a) Análise de Fluxo-resposta em *Escherichia coli* GSM iJO1366; b) Reconstrução do *Yarrowia lipolytica* GEM iYLI647 e sua aplicação na previsão de alvos de engenharia de reação usando quatro estratégias diferentes de projeto de deformação *in silico*; c) Identificação de alvos de medicamentos antimaláricos específicos para estágios para *Plasmodium falciparum* usando GSMs específicos para estágios que representam cinco estágios diferentes do ciclo de vida; d) Reconstrução de um GSM para *Acinetobacter baumannii* usando vários bancos de dados e sua aplicação para a previsão de alvos de drogas específicas para combater o *A. baumannii* resistente a antibióticos; e) Descoberta de novas funções de isozima para genes que demonstraram não ser essenciais em experimentos, mas que foram previstos como essenciais em uma simulação de essencialidade de genes do *E. coli* iJO1366 GEM (ou seja, previsão falso-negativa para o gene *aspC*); f) O método PROMiscuity PrEdictOR (PROPER) identifica enzimas promíscuas em escala genômica em um organismo-alvo; g) A simulação dos GEMs sob várias condições nutricionais revelou os diferentes fatores catabólicos capacidades das diferentes linhagens, bem como seus ambientes de crescimento preferidos e h) O agrupamento gênico hierárquico dos 24 GEMs revelou informações adicionais sobre via biosintética de metabólitos secundários, que diferenciaram com sucesso os clados metabólicos. Fonte (GU et al., 2019).

2.5. ANÁLISE DE MODELOS METABÓLICOS

2.5.1. Análise de Balanço de Fluxo (FBA)

A Análise de Balanço de Fluxo (Flux balance analysis – FBA) é o uso de uma abordagem matemática que permite analisar o fluxo de reações metabólicas em uma rede metabólica. As redes metabólicas que são alvo de estudo bioquímico em genomas, contém todas as reações metabólicas do organismo, referentes aos genes que codificam as possíveis enzimas do mesmo. Além disso, a FBA permite a localização das vias metabólicas com suas respectivas informações que tiveram o seu fluxo alterado diante de uma alteração genética e ou ambiental (DUARTE *et al.*, 2007; FEIST; PALSSON, B. Ø., 2008; OBERHARDT; PALSSON, B. Ø.; PAPIN, 2009; THIELE; PALSSON, B. Ø., 2010).

Orth e colaboradores (2010) demonstram como são realizadas as representações do metabolismo matematicamente. As reações metabólicas são representadas por matriz estequiométrica (S) de tamanho $m \times n$ (Figura 3).

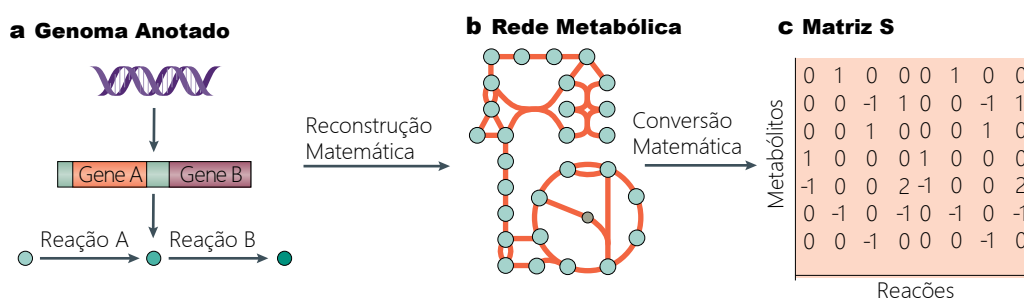


Figura 3. Princípios básicos da modelagem baseada em restrições. a) Os genes metabólicos são provenientes de genomas anotados e são associados a reações metabólicas conhecidas. b) Todas as reações metabólicas integradas que compartilham resultados de metabólitos para reconstrução da rede metabólica de interesse. c) A conversão dessa associação GPR (Gene-proteína-reação) em uma matriz estequiométrica, na qual, as linhas representam os metabólitos e as colunas as reações e que os coeficientes estequiométricos das reações são dispostos na matriz. Cada entrada representa esse coeficiente de reação de um metabólito particular na reação. Adaptado de (Fang et al.2020).

Cada linha dessa matriz representa apenas um único metabólito (para um sistema com m metabólitos) e cada coluna representa uma única reação (n reações). Os valores de entrada de cada coluna é correlato a cada coeficiente do metabólito que participa da reação. Os coeficientes positivos correspondem a metabólitos que são produzidos, enquanto os coeficientes negativos são referenciais de que um metabólito está sendo consumido. A presença do zero como coeficiente estequiométrico diz respeito ao metabólito que não faz parte da reação. A matriz estequiométrica dessa forma se torna

uma matriz esparsa por causa da grande quantidade de reações bioquímicas que superam os números de diferentes metabólitos. O fluxo de todas as reações em uma rede metabólica é representado pelo vetor v , que detém o tamanho de n . Para a concentração de metabólitos que é representada pelo vetor x , que apresenta o valor m .

A sua representação matemática da variação da concentração de metabólitos em função do tempo, é:

(1)

O fluxo de todas as reações que percorrem o sistema S é representado pelo vetor v , que possui o tamanho n . As concentrações de todos os metabólitos são expressas pelo vetor x , de tamanho m , em um determinado tempo t .

A estequiometria propõe restrições ao fluxo de metabólitos através da rede, nas quais, necessitam diferenciar de abordagens que exigem parâmetros cinéticos para medir e tendem a dificultar a análise (COVERT, M. W. *et al.*, 2001; EDWARDS; COVERT, M.; PALSSON, B., 2002). As restrições são aplicadas para duas realidades: uma é para que equações equilibrem entrada e saída de reação e a outra é para as desigualdades que se impõe limites ao sistema. A matriz estequiométrica promove restrições de equilíbrio de fluxo permitindo que a quantidade total de qualquer composto produzido seja igual à quantidade total consumida no estado estacionário, isso evita os comportamentos imprevisíveis e estocásticos. Diante disso, resumidamente, as concentrações de metabólitos intracelulares não variam com o tempo, sendo, portanto, sempre balanceadas. Nisso, se assume um estado estacionário que não promove um acúmulo de metabólitos internos.

A equação que com a restrição na nulidade da variação da concentração de metabólitos em função do tempo será:

(2)

$$S_{ij} \cdot v_j = 0, \forall i \in m, \forall j \in n$$

Os modelos de fluxo metabólico assumem essa restrição a ser analisada para um estado estacionário pois a taxa de metabólitos transitórios é mais rápida do que a taxa de crescimento celular, destacando que no metabolismo convencionalmente os metabólitos transitórios são mais curtos em alguns minutos, assumindo assim um estado estável em relação ao crescimento e aos transitórios típicos do processo (VANNA; PATSSON, 1994).

Além disso, cada reação pode receber limite superior e inferior que vão definir os fluxos máximo e mínimo de possíveis reações e com isso se promove um limite definido e um espaço permissível das distribuições de fluxo em um sistema (PRICE; REED, Jennifer L; PALSSON, B. Ø., 2004). A Figura 3 ilustra as restrições em modelo representativo de solução.

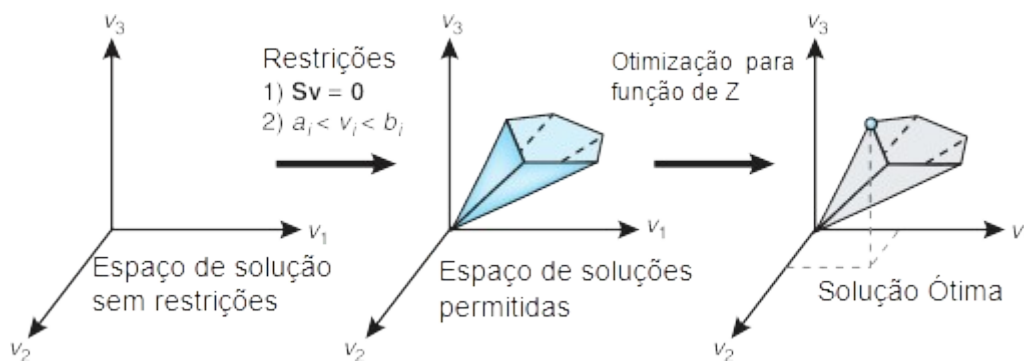


Figura 4. Base conceitual para a modelagem de sistemas baseada em restrições. Da esquerda para direita, o primeiro gráfico demonstra o espaço de solução sem restrições, no segundo aplica-se duas restrições para tornar-se um espaço de solução permissível (descrição no texto) e o terceiro faz a demonstração do uso de uma função objetivo Z . Fonte: (Orth et al, 2010).

Para se obter uma otimização máxima de um metabólito, ou a taxa máxima de crescimento, em uma FBA se faz o uso da função objetivo Z , expressa pela equação abaixo:

(3)

$$Z = c^T \cdot v$$

Nesta equação, c é um vetor de peso que referencia o quanto cada reação v contribui para o objetivo. A função objetivo é utilizada em ocasiões para se prever o máximo e mínimo de crescimento da taxa de biomassa, nisso se atribui o valor 1 no coeficiente linear de interesse (como exemplo na reação de biomassa) enquanto c é um vetor de zero no restante das reações (ORTH; THIELE; PALSSON, B. O., 2010). Com isso, reações de interesses podem ser maximizadas numa análise.

2.6. FLUXO DE TRABALHO DE RECONSTRUÇÃO DE MODELOS METABÓLICOS DE INTERESSE *IN SILICO*.

O processo de reconstrução de modelos metabólicos e curadoria aplicado para os principais representantes de todos os domínios da vida foi descrito em protocolo do trabalho de Thiele & Palsson (THIELE; PALSSON, B. Ø., 2010).

2.6.1. Reconstrução do rascunho das vias (fase preliminar) - 1º Etapa

A 1º etapa consiste na reconstrução do rascunho a partir do genoma anotado automaticamente, em que, se analisa no organismo alvo a presença de identificadores particulares dessa sequência genômica completa. Com isso, pode se deter informações das enzimas ativas que se fazem presentes no organismo, com a informação de genes que geram produtos interativos (isozimas, proteínas de complexo e subunidades) que podem fazer parte de determinada reação metabólica, ou seja, se tem uma associação GPR (gene, proteína e reação) (FEIST *et al.*, 2009).

As reações bioquímicas são analisadas para saber quais as enzimas deste organismo participam, em que, podem ser avaliadas automaticamente e manualmente utilizando algumas ferramentas. Bases metabólicas serão utilizadas, como: KEGG (Kyoto Encyclopedia for Genes and Genomes) (KANEHISA, M., 2000; KANEHISA, Minoru *et al.*, 2019), BRENDA (SCHOMBURG, 2004) e Uniprot (BAIROCH, Amos *et al.*, 2005; GASTEIGER, E *et al.*, 2005). Diante do uso dessas ferramentas, irá se reconstruir com maior facilidade.

Os softwares para reconstrução de genoma serão explorados diante da qualidade comparativa e diante do escopo do projeto (MENDOZA *et al.*, 2019). Serão alternados e ou complementados mediante a eficiência das reconstruções.

2.6.2. Curadoria manual do rascunho (refinamento) – 2º Etapa

Esta etapa conta com a curadoria manual da reconstrução automática serve para corrigir possíveis erros que podem ocorrer, decorrente de percas ou reações que podem ser equivocadamente incluídas ou lacunas deixadas (FEIST *et al.*, 2009).

As funções metabólicas e as reações metabólicas são analisadas individualmente, e feitas buscas ativas na literatura específica além de consultar em bancos de dados (THIELE; PALSSON, B. Ø., 2010).

Como controle de qualidade da reconstrução, independente do tipo de fonte de reconstrução, o controle básico para a análise do modelo pode ser submetido ao Memote (v.0.13) que apresenta como teste a proposta de pontuações comparáveis, de modo independente da amostra em linguagem de marcação de biologia de sistemas (cepa única ou multi-cepa em SBML) e uma visão compreensiva com implementações para possíveis erros (LIEVEN *et al.*, 2020). Suas métricas de teste são bem delimitadas previamente (HEAVNER; PRICE, 2015; MONK, J.; NOGALES; PALSSON, B. O., 2014; THIELE; PALSSON, B. Ø., 2010) com duas medidas importantes: 1° ter todas as métricas na seção de consistência e a 2° Identificador de metabólito uniforme Namespace (identificador principal para metabólitos) (NORSIGIAN; FANG; *et al.*, 2020).

2.6.3. Conversão da reconstrução curada das vias metabólicas em um modelo computacional – Estágio 3

Este passo é que a rede reconstruída é convertida em uma representação matemática (BORODINA; NIELSEN, 2005; NOTEBAART *et al.*, 2006). Essa conversão auxilia na forma de melhor questionar a consistência dos dados e se a rede reconstruída pode ou não ser executada. Pode se implantar o uso de mais ferramentas computacionais para analisar as propriedades da rede e avaliar as funções da rede diante das restrições físico-químicas ofertadas a célula (FEIST *et al.*, 2009).

2.6.4. Validação do modelo metabólico ou uso de reconstrução e integração de dados de alto desempenho – Estágio 4

Neste último passo, o processo é de reconstrução, verificação e validação para verificar possíveis erros na reconstrução. A etapa 3 é responsável por identificar pelo modelo metabólico criado a capacidade de sintetizar os precursores de biomassa (aminoácidos, nucleotídeos trifosfato e lipídio, por exemplo). Porém lacunas na rede que são falhas na identificação de funções metabólicas no processo de reconstrução serão avaliadas pelo processo de repetição dos passos 2 e 3 (THIELE; PALSSON, B. Ø., 2010).

3. JUSTIFICATIVA

Os modelos metabólicos são ferramentas essenciais para a descoberta de princípios do metabolismo celular, evolução microbiana e subsidiar a engenharia genética (BORDBAR *et al.*, 2014; MACHADO *et al.*, 2018).

Em decorrência da diminuição dos custos de sequenciamento genômicos uma grande quantidade de dados tem sido gerada e mais pesquisas ômicas voltadas a um estudo integrativo para o funcionamento das células e interação dos componentes tem sido aprimorado (BORDBAR *et al.*, 2014).

Os modelos metabólicos, contemporaneamente, tem sido aplicado para uma variedade de finalidades, como: micróbios patogênicos (KIM, T. Y.; KIM, H. U.; LEE, S. Y., 2010), inclusão racional de linhagem pra design de um produto bioquímico de interesse biotecnológico (BROCHADO *et al.*, 2010; MACHADO; HERRGÅRD, 2015) e estudo de doenças com metabólicos associados (SHLOMI *et al.*, 2011; VÄREMO; NOOKAEW; NIELSEN, 2013).

Esse estudo tem o objetivo de promover a descoberta de metabólitos secundários de interesse biotecnológico e compreender a dinâmica metabólica de maneira a subsidiar condições ótimas para produção de um composto de interesse como os terpenóides. A cianobactéria *Geminocystis sp. GBBB08* obtida do Parque Nacional da Chapada das Mesas foi um dos primeiros isolados sequenciados do Estado do Maranhão. Em vista disso, há necessidade de se realizar reconstrução de modelos metabólicos baseados em restrição para investigar o potencial genômico e bioquímico dessa linhagem com vistas a valorização da biodiversidade local.

4. OBJETIVOS

4.1. Geral

Reconstruir um modelo metabólico para a *Geminocystis* sp. GBBB08 com foco na produção de terpenóides.

4.2. ESPECÍFICOS

- Realizar a mineração de agrupamentos gênicos biossintéticos;
- Reconstruir modelo metabólico da via dos terpenóides;
- Testar e validar o modelo computacional reconstruído;
- Analisar a simulação do modelo computacional construído aplicando métodos de FBA;
- Analisar a análise de balanço de fluxos do modelo computacional construído;
- Avaliar estratégias para o aumento da produção de metabólicos almejados.

5. MATERIAL E MÉTODOS

5.1. AMOSTRA UTILIZADA E CONDIÇÕES DE MANUTENÇÃO DO CULTIVO

O projeto foi executado com o uso da linhagem de uma cianobactéria isolada no Parque Nacional da Chapada das Mesas, cultivada pelo grupo de pesquisa em Biodiversidade, Bioprospecção e Biotecnologia (GB3) que se integra ao Laboratório de Genética e Biologia Molecular (LabGeM) da UFMA (BUTARELLI et al., 2022; FERREIRA et al., 2021)(BUTARELLI et al., 2022; FERREIRA et al., 2021)(BUTARELLI et al., 2022; FERREIRA et al., 2021).

A linhagem obtida em espécime, *in vivo*, está sendo cultivada usando o meio de cultivo BG-11, com a transferência de inóculos de 1.0 mL do cultivo para 49 mL de meio BG-11 em ciclos de 4 semanas como forma de manutenção da cultura (FERRIS; HIRSCH, 1991).

5.2. SEQUENCIAMENTO GENÔMICO, MONTAGEM E ANOTAÇÃO

Bibliotecas metagenômicas para sequenciamento shotgun foram preparadas e indexadas com o Nextera DNA Sample Prep kit (Illumina, Inc., San Diego, EUA), de acordo com o protocolo do fabricante. Após quantificação e normalização, a biblioteca foi sequenciada na plataforma MiSeq (Illumina). O sequenciamento foi feito utilizando o MiSeq Reagent kit v2 (Illumina) de 500 ciclos com leituras Paired-end de 250 nucleotídeos cada. Ao final de cada corrida foi gerado um arquivo de saída com as sequências (read1 e read2) contendo as leituras e métricas de qualidade de cada nucleotídeo sequenciado (arquivo FastQ). Em média as corridas tiveram boa performance, gerando >75% de sequências com pontuação de qualidade acima de 30 (Q30).

Para análise e controle de qualidade, foi utilizado o módulo Read_qc do Metawrap (URITSKIY; DIRUGGIERO; TAYLOR, 2018) onde as sequências foram trimadas e

filtradas, baseadas na qualidade PHRED como parâmetro, garantindo que apenas as sequências de alta qualidade (Phred score ≥ 30) fossem mantidas. Após a obtenção das reads filtradas, os contigs foram montados através do módulo Assembly do Metawrap utilizando o metaSPAdes v. 3.13 (NURK *et al.*, 2017).

O isolamento do genoma foi realizado através do módulo Binning, onde foram realizadas três rodadas de binning diferentes, uma com o CONCOCT (ALNEBERG *et al.*, 2013), outra utilizando o MaxBin2 (WU, Y.; SIMMONS; SINGER, 2015) e outra utilizando o MetaBAT2 (KANG *et al.*, 2019). Os três resultados obtidos foram comparados através do módulo Bin_refinement, dessa forma foram selecionados os bins de maior qualidade. Para anotar e classificar taxonomicamente os bins obtidos, as sequências foram submetidas ao Prokaryotic Genome Annotation Pipeline (PGAP) (TATUSOVA *et al.*, 2016) do NCBI. Todo fluxo de trabalho de sequenciamento, montagem e anotação é proveniente de trabalhos anteriores de nosso grupo (FERREIRA, L. S. De S. *et al.*, 2021).

5.3. MINERAÇÃO DO GENOMA E ANÁLISE COMPARATIVA

A mineração do genoma completo foi realizada a partir de um pipeline adaptado baseado em genômica comparativa para microrganismos (LIMA, A. R. J. *et al.*, 2018). Foi utilizado o servidor antiSMASH 6.0 (BLIN *et al.*, 2019) para predizer agrupamentos gênicos bem estabelecidos.

5.4. O PROCESSO DE RECONSTRUÇÃO DA REDE METABÓLICA

Para a reconstrução do modelo em escala Genômica (GSM) foram utilizados três ferramentas, o CarveMe (versão 1.4.1) com os parâmetros (carve genome/GCA_016743215.1_ASM1674321v1_protein.faa -o models/iGbb08.xml --universe cyanobacteria --fbc2 --mediadb) (MACHADO *et al.*, 2018). Essa é uma ferramenta de reconstrução de espécies em nível metabólico com uma abordagem automática que é manualmente curada e de pronta simulação. A sua utilização é por linha de comando baseado em linguagem de interpretação de programação em python que é desenvolvido

para criar GSM em formato SBML versão 3 (linguagem de marcação de biologia de sistemas). Essa ferramenta, constrói um modelo universal de metabolismo anotado de rascunho (carving), que é obtido do banco de dados do BiGG (KING, Z. A. *et al.*, 2016) para se identificar com predições por homologia e ortologia quais associações gene-proteína-reação são presentes na *Geminocystis* sp. GBBB08 a partir do uso do genoma anotado em formato FASTA.

Para uma abordagem de reconstrução automática de GSM também foi utilizado o Kbase (ARKIN *et al.*, 2018) com a ferramenta e banco de dados bioquímico Model SEED (SEEVER *et al.*, 2021).

Para prever as taxas capazes de absorção e secreção do organismo apenas em evidências genéticas e para a produção de simulação-real do modelo do meio específico, se utilizou o meio BG-11 (FERRIS; HIRSCH, 1991) em formato tsv (Valores separados por tabulações) para análise com preenchimento de lacunas restrito apenas a esse meio.

O COBRApy (EBRAHIM *et al.*, 2013), é um pacote python que serviu para executar a reconstrução de modelos por restrição e para fazer refino no modelo em rascunho. Além disso, foi utilizado para simular a Análise de Balanço de Fluxo (FBA) (ORTH; THIELE; PALSSON, B. O., 2010) e pFBA (LEWIS *et al.*, 2010a). O FBA é uma abordagem em redes metabólicas de escala genômica que calcula o fluxo de metabólitos através de uma rede metabólica, o que possibilita prever a taxa de crescimento de um organismo ou de metabólito importante (ORTH; THIELE; PALSSON, B. O., 2010). O pFBA adiciona a minimização de todos os fluxos ao objetivo do modelo. Essa proposição é motivada pela ideia de que fluxos elevados têm um turn-over enzimático mais elevado e que a produção de enzimas é custosa biologicamente, a célula tentará minimizar o fluxo global enquanto ainda maximiza a função objetivo original e a taxa de crescimento (LEWIS *et al.*, 2010). Ambos dentro do ambiente web gráfico Jupyter Notebook (KLUYVER *et al.*, 2016) para utilizar código de programação linear e texto em interface gráfica.

A visualização de vias bioquímicas foi feita utilizando o Escher (KING, Z. A. *et al.*, 2015) e o Fluxer (HARI; LOBO, 2020).

Para a curadoria manual na foi utilizado o banco de dados BiGG (NORSIGIAN; PUSARLA; *et al.*, 2020) e o KEGG (KANEHISA, M., 2000).

5.5. APLICAÇÕES DE ENGENHARIA METABÓLICA NO MODELO

5.5.1. Identificação das Reações da Via MEP

Para avaliar possíveis otimizações na produção de terpeno, foram realizadas alterações na função objetivo. Compostos construtores como o Isopentil difosfato (IPP), Geranil difosfato e Farnesil difosfato que compõe a via MEP (LIN, P. C.; PAKRASI, 2019; PATTANAİK; LINDBERG, 2015) nas quais foram aplicados com função objetivo, respectivamente. Além disso, análise de variabilidade de fluxo para as reações (GUDMUNDSSON; THIELE, 2010) e a análise de balanço de fluxo parsimonioso (LEWIS *et al.*, 2010a).

6. RESULTADOS E DISCUSSÃO

1.1. FILOGENÔMICA, MINERAÇÃO DO GENOMA E ANÁLISE COMPARATIVA

O diagrama do genoma da *Geminocystis* sp. GBBB08 circular está representado na Figura 4.

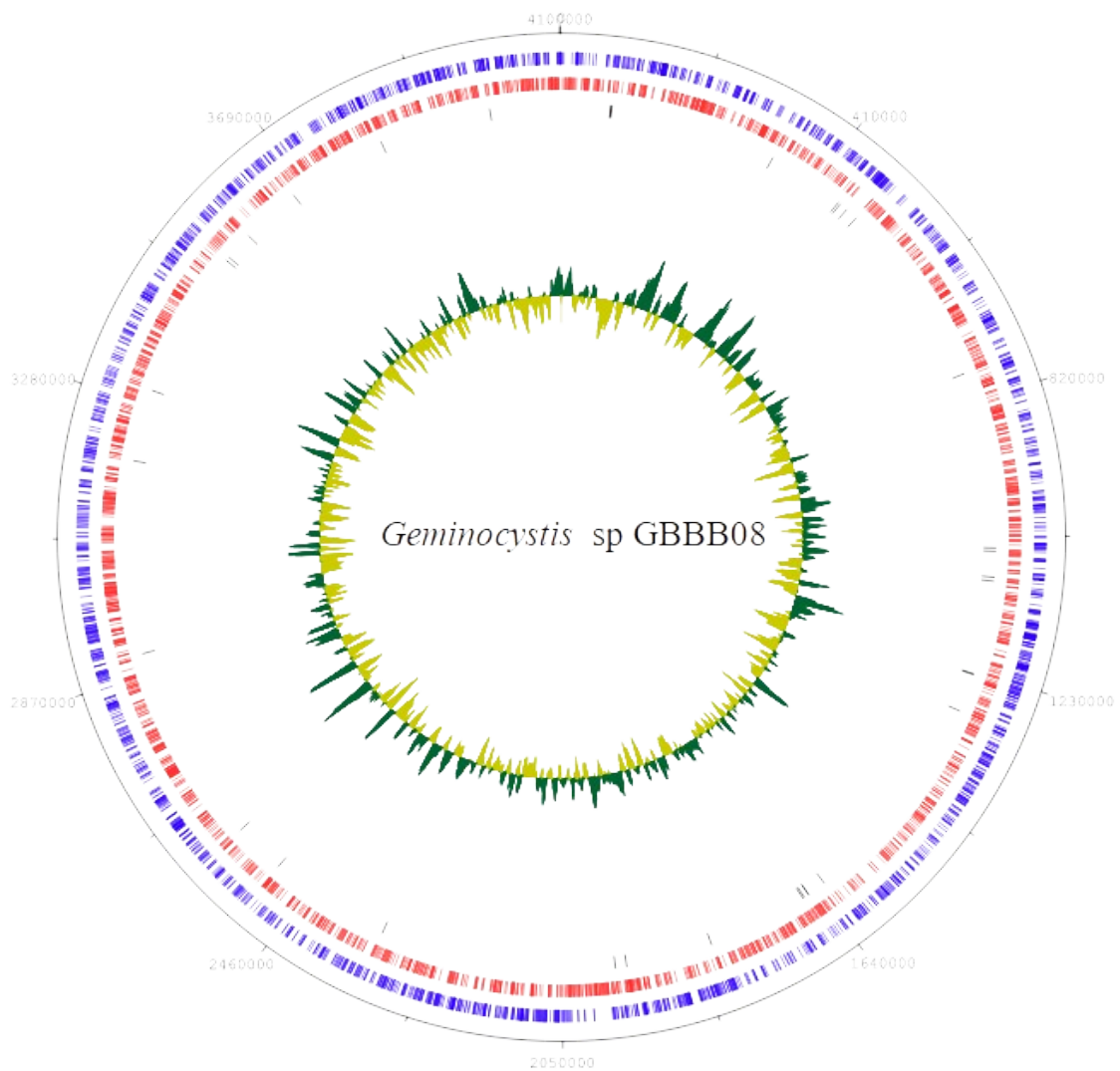


Figura 5. Apresentação circular do genoma da *Geminocystis* sp. GBBB08 realizado na ferramenta Artemis. As faixas do lado de fora representam: (1) CDS Sentido horário - Azul, (2) CDS Sentido anti-horário - Vermelho, (3) tRNA e (4) %GC, respectivamente

A linhagem apresenta 4.101.793 pares de base. As características e os dados de qualidade do genoma podem ser visualizados na Tabela 1.

Tabela 1. Características genômicas da *Geminocystis* sp. GBBB08.

Características	Cromossomo
Cepa	<i>Geminocystis</i> GBBB08
Número de contigs	327
L50 value	60
N50 value (bp)	20.754
Completeness	98.94%
Contaminação	0.4%
GC content	32.70%
Tamanho estimado do cromossomo (bp)	4,101,793
Genes codificadores de proteínas (CDS)¹	4.153
Total de genes de RNAs	36

A identificação da *Geminocystis* sp. GBBB08 e sua respectiva árvore filogenômica é apresentada na Figura 5, em que, foi comparado 20 genomas públicos de referência depositados no Kbase com a aplicação SpeciesTree (v2.2.0), de diferentes gêneros e famílias. A árvore foi construída a partir de um conjunto de 49 marcadores, genes universais definido por família de genes – COG (Cluster Orthologous Groups). A relação é determinada pela semelhança de alinhamento com 49 domínios COG, com alinhamento de múltiplas sequências (MSA) para cada família COG. Além disso, foi trimado utilizando o GBLOCKS para remover seções mal alinhadas de MSA. A análise confirmou a identificação da *Geminocystis* sp. GBBB08 por similaridade gênica utilizando o método Neighbor public genome count, com a linhagem *Geminocystis. herdmanii* PCC 6308.

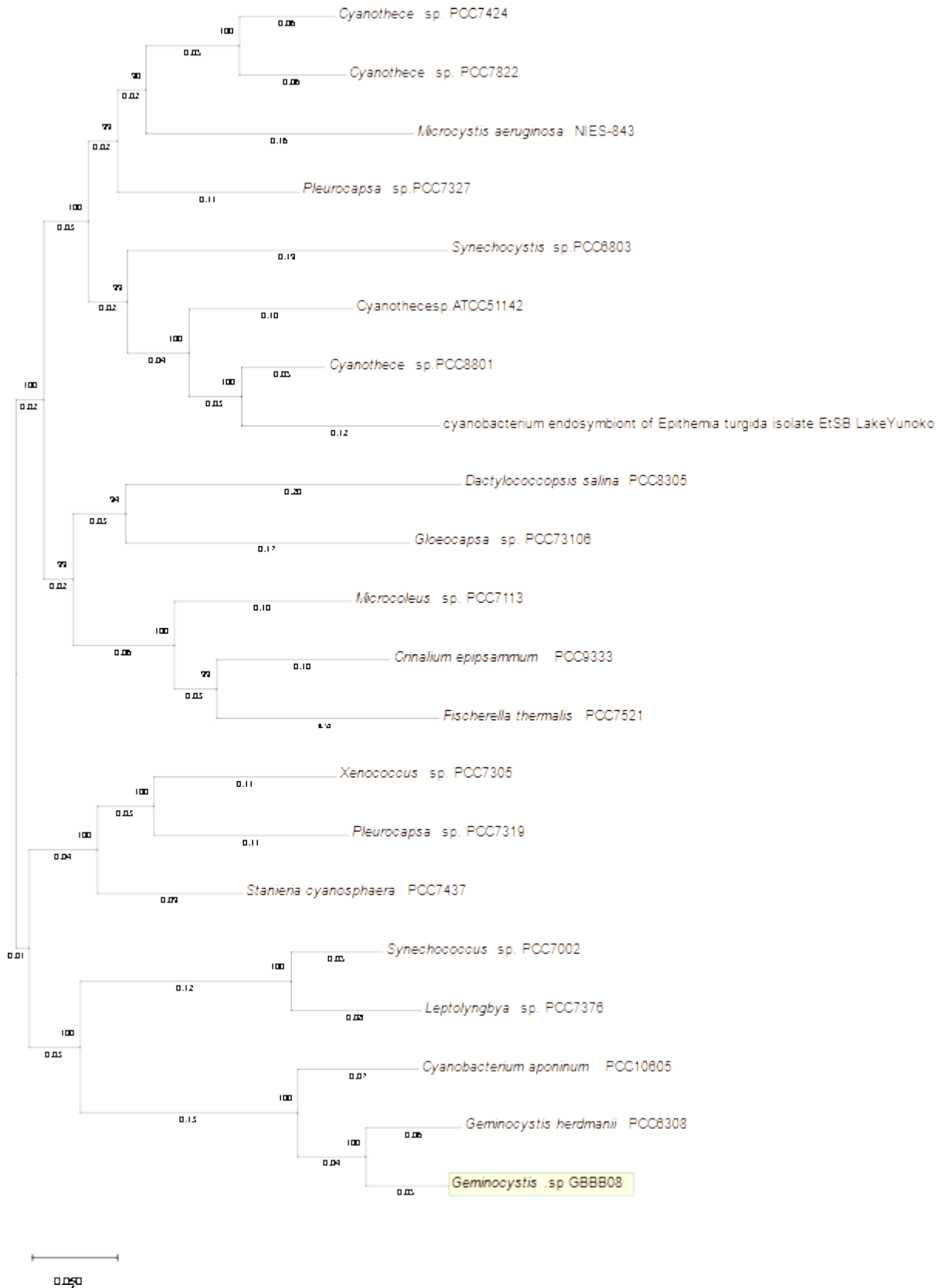


Figura 6: Árvore filogenômica da *Geminocystis .sp GBBB08* em comparação de diversidade e ancestralidade com 20 genomas núcleos de COG depositados no KBase. A relação filogenética foi inferida usando o método neighbor-joining (agrupamento de vizinhos) usando 49 marcadores de genes.

A submissão do genoma da *Geminocystis* sp. GBBB08 para o servidor online gratuito antiSMASH (MEDEMA *et al.*, 2011) definiu o genoma em agrupamentos específicos biosintéticos (BLIN *et al.*, 2021), que são demonstrados na Tabela 2. Foram identificados 12 agrupamentos gênicos de vias biossintéticas diversas. Houve uma classificação de acordo com o tipo de agrupamento mais próximo, decorrente de uma análise a partir do algoritmo de perfil HMM (enzimas-núcleo) que especifica o tipo de agrupamento gênico biossintético (BGC type) e além de regras definidas individualmente para a identificação de protoclusters codificados na região (BLIN *et al.*, 2021).

A Região 1 identificou os agrupamentos do tipo Policetídeo Sintase Tipo I, o 1-Heptadeceno (olefina), um alceno ($C_{17}H_{34}$) com 100% de similaridade ao núcleo gênico biossintético identificado no servidor. Sendo esse um dos tipos de hidrocarbonetos produzidos através de ácidos graxos pelas cianobactérias (COATES *et al.*, 2014).

As α -olefinas, alcenos com ligação dupla, são importantes para a fabricação de uma variedade de produtos químicos de valor agregado, como: detergentes, poliolefinas, detergentes, lubrificantes, plastificantes e óleos sintéticos (ELIASSON; JENSEN, V. R., 2019; GUO *et al.*, 2018; KILLIAN; JOHNSON; BROOKHART, 1997; WANG, J. *et al.*, 2018; ZHANG, Z. *et al.*, 2020).

Coates *et al.*, (2014), investigou a produção de hidrocarbonetos a partir de ácidos graxos com 128 linhagens de genoma completo e apesar das cianobactérias realizarem a produção de hidrocarbonetos nem todos são produzidos pela mesma via metabólica. Duas vias metabólicas são bem caracterizadas na produção de hidrocarbonetos a partir de ácidos graxos que envolvem a retirada do átomo de carbono oxidado terminal, uma via realiza a conversão de acil-ACPs para acil-aldeídos graxos através da acil-ACP redutase (FAAR), seguido por uma reação única dependente de oxigênio do aldeído (ADO) para produzir um alceno saturado de comprimento de cadeia ímpar (LI, N. *et al.*, 2012; SCHIRMER, A. *et al.*, 2010). A outra via, é a Olefina sintetase (OLS) que realiza o alongamento da cadeia acil seguida de uma descarboxilação para produzir um alceno terminal (CHANG, Z. *et al.*, 2004; MCCARTHY, J. G. *et al.*, 2012).

O 1-heptadeceno é um dos hidrocarbonetos biossintéticos produzidos por linhagens unicelulares (Chroococcales), baecóstios (Peurocapsales) e filamentosas (Oscillatoriales) apenas de via Olefina Sintetase (OLS) (COATES *et al.*, 2014).

A utilização desse alceno derivado de biomassa pode ser uma solução de baixo custo na produção de α -olefinas, que podem gerar hidrocarbonetos eficientes através de desoxigenação. A desoxigenação de ácidos graxos é referida como estratégia complementar à oligomerização de etileno que envolve altos custos de produção, principalmente por manutenção no controle de crescimento da cadeia de carbono longa para obter produtos com um único comprimento de cadeia de carbono, a própria matéria-prima que é derivada de fontes de carbono e as α -olefinas obtidas são de comprimento de cadeia uniforme (HOU *et al.*, 2018; WANG, J. *et al.*, 2018; ZHANG, Z. *et al.*, 2018, 2020).

Tabela 2. Resultados da mineração genômica para a predição de compostos e produtos presentes na *Geminocystis* sp. GBBB08 no antiSMASH.

Região	De	Para	Tipo do agrupamento	Agrupamento mais conhecido similarmente		Similaridade
Região 1	656,136	704,898	Tipo T1PKS,NRPS	1-heptadeceno	Polyketide: Modular type I	100%
Região 2	865,418	894,134	Betalactona			
Região 3	1,103,493	1,143,944	Tipo NRPS	minutissamida A/C / D	NRPS	23%
Região 4	1,391,326	1,411,584	Contendo elemento de reconhecimento RiPP	Glicano da camada S	Sacarídeo	8%
Região 5	1,887,181	1,907,320	Terpeno			
Região 6	2,415,237	2,430,776	Contendo elemento de reconhecimento RiPP			

Continuação da Tabela 2. Resultados da mineração genômica para a predição de compostos e produtos presentes na *Geminocystis* sp. GBBB08 no antiSMASH.

Região	De	Para	Tipo do agrupamento	Agrupamento mais conhecido similarmente	Similaridade	Região
--------	----	------	---------------------	---	--------------	--------

Região 7	2,526,150	2,544,904	Terpeno			
Região 8	2,595,981	2,617,954	Terpeno			
Região 9	2,729,090	2,804,139	Tipo NRPS	Anabaenopeptina	NRPS	57%
Região 10	3,006,643	3,048,958	NRPS	GE2270	RiPP: Thiopeptido	19%
Região 11	3,290,398	3,333,259	Tipo transAT-PKS			
Região 12	3,362,392	3,382,847	Contendo elemento de reconhecimento RIPP			

6.1. Verificação da reconstrução dos modelos metabólicos da *Geminocystis*. sp. GBBB08 e simulação FBA

A GSM realizada pelo CarveMe (iGbb08) em condições padrão de reconstrução e de inspeção do modelo foram obtidos 1509 reações, 1127 metabólitos e que representam a anotação de 758 genes. Os compartimentos preditos foram o citosol, periplasma, o espaço extracelular e membrana tilacóide (Arquivo Adicional 1 'ipynb'). Para o GSM rascunho reconstruído pelo Kbase (iGbbb08_kb) com meio de cultivo BG-11 para preenchimentos de espaços de lacuna (gap fill) foram obtidos 1042 reações, 1056 metabólitos para 743 anotados. Os compartimentos preditos foram apenas o citosol e o espaço extracelular (Arquivo Adicional 2 'ipynb'). Enquanto para o GSM rascunho sem o meio de cultivo BG-11 utilizando o Kbase foram 1026 reações, com 1056 metabólitos e 758 genes anotados (Arquivo Adicional 3 'ipynb'). Os dados estão distribuídos na Tabela 3.

Tabela 3. Extrato de reconstruções de *Geminocystis*. sp GBBB08 e suas principais propriedades que foram construídas com o CarveMe e o Kbase.

Org	Fer	MC	Versão	GR	Mets	Rxn	Comp
<i>Geminocystis</i> sp. GBBB 08	Kbase	BG-11	iGbbb08_kb	743	1056	1042	2 (c,e)
<i>Geminocystis</i> sp. GBBB 08	Kbase	Sem meio	IGbbb08_BG 11_FBA_with Media	758	1056	1026	2 (c,e)
<i>Geminocystis</i> sp. GBBB 08	CarveMe	BG-11	iGbb08	758	1127	1509	4 (c, p, e, t)

Abreviações: Org, Organismo; Fer, ferramenta de Reconstrução; MC, meio de cultivo; GR, genes em reconstrução; Mets, metabólitos; Rxn, reações; Comp, compartimentos (c - Citosol, e - Espaço extracelular, p - Periplasma, t - Tilacóide)

A utilização de diversas ferramentas de reconstrução de modelos metabólicos em rascunho permite uma melhor curadoria, preenchimento de espaços de lacuna e uma melhor performance individual de cada reconstrução do organismo escolhido (KIM, W. J.; KIM, H. U.; LEE, S. Y., 2017; MENDOZA *et al.*, 2019).

O CarveMe (MACHADO *et al.*, 2018) é uma ferramenta baseada em linha de comando em python que utiliza a abordagem Top down para a recriação do modelo, baseado em um modelo universal curado e depositado no BIGG models. Para esta análise, foi utilizado o gap-filling (meio BG-11), como uma abordagem Bottom - up que poderia ser implementado de modo opcional pois o CarveMe já apresenta o gap-filling ativo para diversos fenótipos.

O Kbase (ARKIN *et al.*, 2018) é um software de código aberto que apresenta o uso interativo de ferramentas de anotação automática, reconstrução, preenchimento de gap fills, análise em uma plataforma robusta e de fácil aprendizagem, sem a necessidade prévia de programação linear. Apresenta também um grande número de aplicações e de tutoriais demonstrativos para o uso desse software. Para a reconstrução automática como ferramenta que integra o Kbase, no qual pipeline padrão do model SEED foi utilizado (HENRY *et al.*, 2010).

Ambos os reconstrutores têm como premissa serem de uso amigável, apresentam uma variedade de aplicações presentes em seus respectivos programas, com até a possibilidade de realizar uma abordagem GEM com múltiplas cepas. A maior vantagem para esses programas de reconstrução é que ambos são de reconstrução automática. Contudo, ambos necessitam realizar curadoria em seus modelos reconstruídos (NORSIGIAN; FANG; *et al.*, 2020).

A sua função objetivo que é a própria reação de biomassa *in silico* foi de 0.02189 h⁻¹ com o CarveMe. O comportamento de entrada e saída do modelo para a análise de solução de balanço de fluxo (FBA) é demonstrada na Tabela Suplementar 1. Destaca-se o comportamento para a formação de biomassa do modelo, detém a absorção de 15 metabólitos provenientes do meio de crescimento através do espaço extracelular simulado e como produção residual há 3 metabólitos referentes a biomassa.

Em rede metabólica da linhagem de *Synechocystis* sp. PCC 6803 com condições de luz constante o FBA para a taxa de biomassa é de 0.029 h⁻¹.

A Tabela Suplementar 2 representa o comportamento do rascunho do modelo executado pelo CarveMe em análise de variação do fluxo (FVA), em que, exibe informações sobre o comportamento de entrada e saída do modelo associado com a função objetiva padrão, com intervalos de fluxo mínimo e máximo alcançáveis.

A pontuação geral do MEMOTE para o modelo do CarveMe (iGbb08) é de 23% com a consistência de 49% (Figura Suplementar 1) enquanto para o modelo reconstruído pelo Kbase (iGbbb08_kb) é de 75% com a consistência de 53%. Apesar de melhor nota de validação pelo MEMOTE (iGbbb08_kb) o modelo não apresenta compartimento de periplasma e nem de membrana tilacoide o que nos fez utilizar as simulações apenas para o modelo reconstruído no CarveME (iGbb08)

6.2. ANÁLISE DE BALANÇO DE FLUXO PARCIMONIOSO (PFBA) PARA A VIA DE TERPENOS

Nós refinamos a investigação da via de reação da produção de terpeno *in silico* a partir do modelo em escala genômica da *Gemynocystis* sp. GBBB 08 (CarveMe), utilizando informações do mapa da via no banco de dados como o KEGG (M00096)

(KANEHISA, M., 2000) e literatura específica (LIN, P. C.; PAKRASI, 2019; PATTANAIK; LINDBERG, 2015). Nenhuma reação metabólica foi adicionada ou retirada mas os valores de limite superior para as reações de: Dimetilaliltransferase (DMATT), Geranil Difosfatase (GRTT) e Decaprenil sintase (DPPS), foram definidos (Figura 6).

Os terpenóides são um grupo diverso e bem descrito de produtos naturais de grande diversidade, com participação efetiva na respiração a partir do transporte de elétrons em cadeia (ubiquinona), em plantas são importantes para a integridade e termo tolerância. Seus tamanhos são decorrente a partir do número de isoprenos (BRENNER *et al.*, 2006; LIN, P. C.; PAKRASI, 2019; SIWKO *et al.*, 2007).

A via para produção de terpenóides de compostos construtores como o Isopentenil difosfato (IPP) que servem para construção de outros terpenos, como geranil difosfato (GDP) e farnesil difosfato (FDP) foram encontradas neste modelo. A via canônica da produção de terpenóides em cianobactérias (via MEP – Não Melavonato) tem sete reações enzimáticas até a formação do IPP, com descrição bem estabelecida na literatura (LIN, P. C.; PAKRASI, 2019; PATTANAIK; LINDBERG, 2015).

MEP

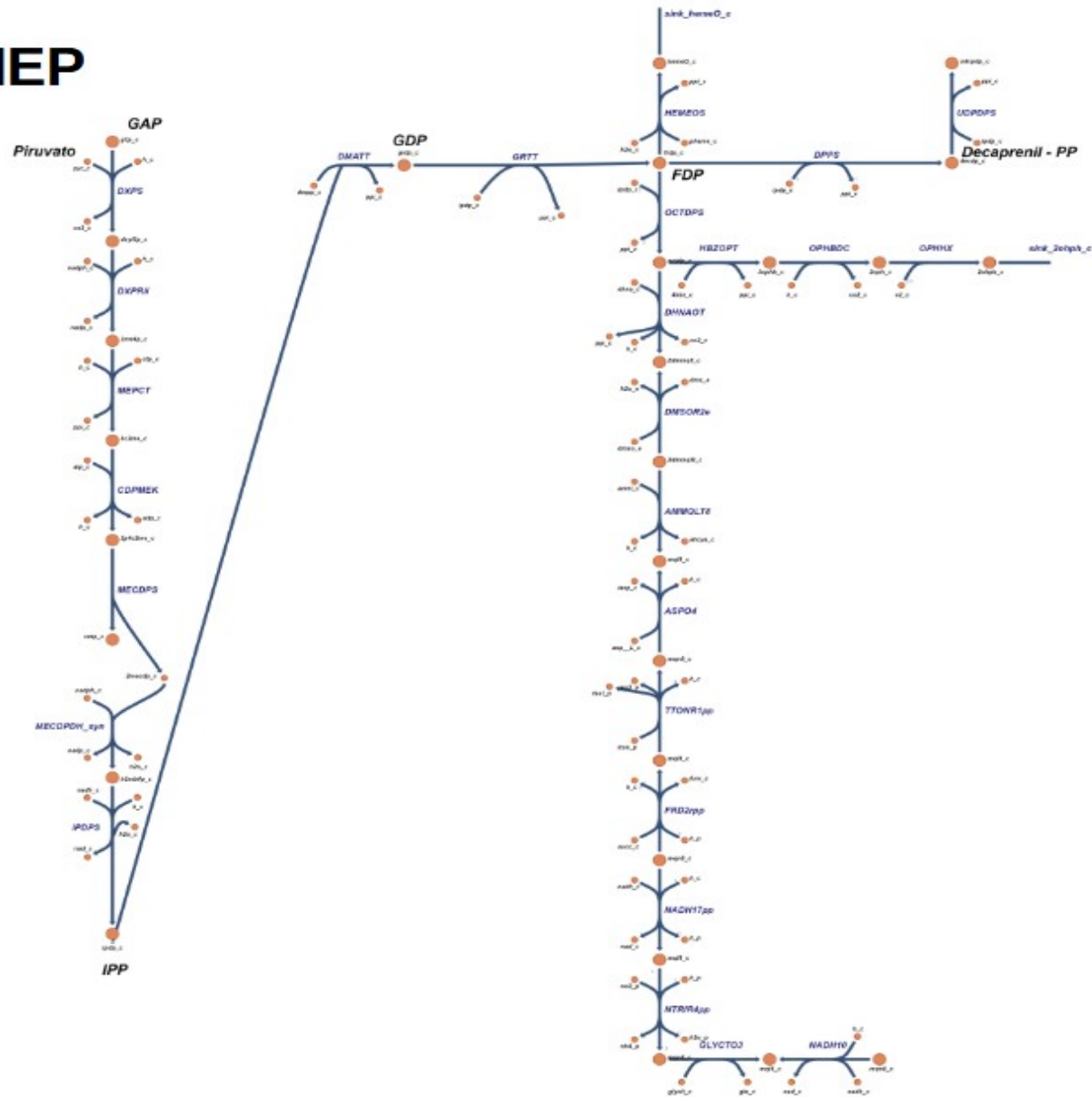


Figura 7: Via metabólica da biosíntese de Terpeno em cianobactérias (MEP) na *Geminocystis* sp. GBBB08. A produção de terpeno é catalisada por múltiplos passos de reações. Abreviações: setas em azul indicam reações e círculos em amarelo indicam metabólitos.

As análises de balanço de fluxo parsimonioso (pFBA) com mudanças nas reações “objetivo” para os terpenóides GDP, FDP e Decaprenil - PP são dispostas abaixo (Tabelas 5 e 6). Nós aplicamos a função objetivo para avaliar de acordo com restrições de limite máximo de crescimento para esses isoprenos. Com essa definição de objetivo se simulou uma análise pFBA também para observar as propriedades do modelo, se ocorreria aumento ou diminuição do fluxo e também para a produção de biomassa.

O pFBA para todas as três reações resultou em um ligeiro aumento (em valores absoluto) de fluxo mas em ambos ocorreu um decréscimo do fluxo de biomassa. Isso indica que essas restrições de fenótipo para o aumento do limite de reação (0.9 em função objetivo) não favorece a cepa simulada (Tabela 5) diante do uso nos parâmetros padrão para o limite superior em cada um dos terpenóides como função objetivo (DMATT, GRTT e DPPS).

Outras restrições como aumento e redução de fótons, FVA e pFBA podem ser aplicadas para a investigação do fluxo desses metabólitos.

Apesar de termos alterado o valor de fluxo, como solução otimizada (função objetivo conseqüentemente) e tornado cada uma dessas três reações de formação de isoprenos utilizando o valor limite (0.9 da fração do ótimo) dado pela biblioteca COBRApy, essa modificação, não é uma alteração favorável ao crescimento biológico simulado ao modelo quando não se altera possíveis parâmetros do método da função pFBA no COBRApy pois em padrões normais para essa função do método inviabilizaria o crescimento da cianobactéria comparado aos valores padrão de reconstrução do modelo. Contudo, quando se promove a alteração em parâmetros (Tabelas 5 e 6), como a fração do ótimo (parâmetro do módulo de pFBA), em que, a reação objetiva original tem a restrição para ser maior que o valor máximo vezes a fração do ótimo que por padrão é 0.9 para 0.2 (que leva uma interferência menor da solução otimizada na produtividade teórica) há um aumento nas taxas dos terpenóides sem que haja ausência de crescimento de biomassa.

Em contraste com a alteração padrão anterior que apesar de levar a um estímulo na taxa de produção de terpenos mas com a produção nula de biomassa, a

alteração na solução objetivo com o ajuste do parâmetro da fração de solução pode representar um rendimento melhor na produção para este modelo reconstruído.

A alteração sugere que a medida que se intensifica o crescimento de terpenóides o fluxo por pFBA aumenta gradualmente e menos carbono fica disponível para a formação de biomassa o que promove que o fluxo das reações relacionadas a biomassa diminuam.

Em ensaios para testar a superprodução de isoprenos (terpenóides) em *Synechocystis* sp. PCC 6803 testando a capacidade da co-expressão das enzimas (Ipi, Dxs e IspD), com a avaliação do ciclo Calvin-Benson-Bassham com estresse de baixa de glicose e síntese de co-fatores já foram descritos (ENGLUND *et al.*, 2018). Nessa análise também houve um estímulo de limite superior com redução de carbono para biomassa.

Para resultados de FBA, há um crescimento na taxa de crescimento de biomassa se for analisado pelas taxas correspondentes de alteração da função objetivo. Porém, o pFBA é um método mais apurado pois é um módulo da função que adiciona a minimização de todos os fluxos ao objetivo do modelo. Essa é uma abordagem que é motivada pela ideia de que a alta taxa de fluxos tem uma maior rotatividade de enzimas e que como a produção de enzimas é cara a célula tentará minimizar o fluxo total enquanto ainda maximizava a função objetivo geral (taxa de crescimento) (LEWIS *et al.*, 2010).

Tabela 4. Análise de Balanço de Fluxo (FBA) e Análise de Variabilidade de Fluxo (FVA) com alteração da função objetivo para cada uma das reações DMATT, GRTT e DPPS, respectivamente.

S. No	Reação	Análise de balanço de Fluxo para Função Objetivo (Biomassa - mmol/gDW ⁻¹ /h ⁻¹)	Análise de variação de fluxo para Função Objetivo (mmol/gDW ⁻¹ /h ⁻¹)		Análise de variação de fluxo para Função Objetivo (mmol/gDW ⁻¹ /h ⁻¹)	
			90% de otimização		0.2 de fração do ótimo	
		Biomassa	FVA mínimo	FVA máximo	FVA mínimo	FVA máximo
1	Biomassa (Growth)	0.021897155983588913	0.021897	0.021897	0.004379	0.021897
2	Dimetilaliltransferase (DMATT)	0.0	0.019321	0.019321	0.0	0.01968
3	Geranyl Difosfatase (GRTT)	0.0	0.019321	0.019321	0.0	0.01968
4	Decaprenil sintase (DPPS)	0.0	0.011138	0.011138	0.0	0.02165

A visualização do mapa da via de biosíntese de terpeno tem seguimento para a formação de Menaquinona – 8 (mqn8_c) a partir da catalise da enzima Octaprenil pirofosfatase sintase (OCTDPS), no qual, esse isoprenóide é integrante do grupo das quinonas que participa da cadeia transportadora de elétrons em procariotos e eucariotos. Além disso, a via endógena para a produção desse isoprenóide foi previamente descrita para *Escherichia coli* (KONG; LEE, 2011).

Tabela 5. Análise de Enzima Parsimônia de Fluxo (pFBA) com alteração da função objetivo para cada uma das reações DMATT, GRTT e DPPS, respectivamente em 90% de otimização.

S. No	Reação	Análise de Enzima Parsimônia balanço de Fluxo para Função Objetivo (mmol/gDW ⁻¹ /h ⁻¹) pFBA da biomassa (objetivo alterado)	Solução ótima com valor objetivo
1	Biomassa	0.021897155983588913	65.320
2	Dimetilaliltransferase (DMATT)	-1.8174125185489555e-20	63.097
3	Geranil Difosfatase (GRTT)	-1.8174125185489555e-20	63.097
4	Decaprenil syntase (DPPS)	-1.2961106951006021e-30	63.819

Tabela 6. Análise de Enzima Parsimônia de Fluxo (pFBA) com alteração da função objetivo para cada uma das reações DMATT, GRTT e DPPS, respectivamente em 20% de otimização.

S. No	Reação	Análise de Enzima Parsimônia balanço de Fluxo para Função Objetivo (mmol/gDW ⁻¹ /h ⁻¹) pFBA da biomassa (objetivo alterado)	Solução ótima com valor objetivo
1	Biomassa	0.004379431196717786	12.623092492934301
2	Dimetilaliltransferase (DMATT)	0.0	12.554683840749437
3	Geranil Difosfatase (GRTT)	0.0	12.554683840749437
4	Decaprenil syntase (DPPS)	0.0	12.685185115734908

7. CONCLUSÃO

O modelo iGbb08 representa o primeiro modelo metabólico em escala genômica de uma cianobactéria do gênero *Geminocystis* aliado a uma abordagem de mineração funcional que permite subsidiar futuros trabalhos para exploração biotecnológica desse grupo. Tivemos sucesso em reconstruir o modelo metabólico em abordagem top down para este organismo adaptando possíveis mutações e alterações de fenótipo otimizando a produção de terpenóides. Futuramente outras aplicações ao modelo podem ser implantadas in silico para otimizar a via de terpeno em termos de engenharia metabólica.

8. REFERÊNCIAS BIBLIOGRÁFICAS

AJIKUMAR, P. K. *et al.* Terpenoids: opportunities for biosynthesis of natural product drugs using engineered microorganisms. **Molecular pharmaceutics**, 2008. v. 5, n. 2, p. 167–190.

ALNEBERG, J. *et al.* CONCOCT: clustering contigs on coverage and composition. **arXiv preprint arXiv:1312.4038**, 2013.

ARKIN, A. P. *et al.* KBase: The United States department of energy systems biology knowledgebase. **Nature Biotechnology**, 2018. v. 36, n. 7, p. 566–569.

BAIROCH, Amos *et al.* The Universal Protein Resource (UniProt). **Nucleic Acids Research**, 2005. v. 33, n. DATABASE ISS., p. 154–159.

BARABÁSI, A.; OLTVAI, Z. N. NETWORK BIOLOGY : UNDERSTANDING THE CELL ' S FUNCTIONAL ORGANIZATION. 2004. v. 5, n. February.

BLIN, K. *et al.* AntiSMASH 5.0: Updates to the secondary metabolite genome mining pipeline. **Nucleic Acids Research**, 2019. v. 47, n. W1, p. W81–W87.

BLIN, K *et al.* AntiSMASH 6.0: Improving cluster detection and comparison capabilities. **Nucleic Acids Research**, 2021. v. 49, n. W1, p. W29–W35.

BORDBAR, A. *et al.* Constraint-based models predict metabolic and associated cellular functions. **Nature Reviews Genetics**, 2014. v. 15, n. 2, p. 107–120.

BORODINA, I.; NIELSEN, J. From genomes to in silico cells via metabolic networks. **Current opinion in biotechnology**, 2005. v. 16, n. 3, p. 350–355.

BRENNER, M. *et al.* **Engineering Microorganisms for Energy Production. DTIC Document.**

BROCHADO, A. R. *et al.* Improved vanillin production in baker's yeast through in silico design. **Microbial Cell Factories**, 2010. v. 9, p. 1–15.

BRUGGEMAN, F. J.; WESTERHOFF, H. V. **The nature of systems biology. Trends in Microbiology.**

CHANG, Z. *et al.* Biosynthetic Pathway and Gene Cluster Analysis of Curacin A, an Antitubulin Natural Product from the Tropical Marine Cyanobacterium *Lyngbya majuscula*. **Journal of natural products**, 2004. v. 67, n. 8, p. 1356–1367.

CHAPPELL, J. Biochemistry and molecular biology of the isoprenoid biosynthetic pathway in plants. **Annual review of plant biology**, 1995. v. 46, n. 1, p. 521–547.

COATES, R. C. *et al.* Characterization of Cyanobacterial Hydrocarbon Composition and Distribution of Biosynthetic Pathways. **PLoS ONE**, 27 jan. 2014. v. 9, n. 1, p. e85140. Disponível em: <<https://dx.plos.org/10.1371/journal.pone.0085140>>.

COVERT, M. W. *et al.* Metabolic modeling of microbial strains in silico. **Trends in biochemical sciences**, 2001. v. 26, n. 3, p. 179–186.

DUARTE, N. C. *et al.* Global reconstruction of the human metabolic network based on genomic and bibliomic data. **Proceedings of the National Academy of Sciences**, 2007. v. 104, n. 6, p. 1777–1782.

DUCAT, D. C.; WAY, J. C.; SILVER, P. A. Engineering cyanobacteria to generate high-value products. **Trends in Biotechnology**, 2011. v. 29, n. 2, p. 95–103. Disponível em: <<http://dx.doi.org/10.1016/j.tibtech.2010.12.003>>.

EBRAHIM, A. *et al.* COBRApy: COstraints-Based Reconstruction and Analysis for Python. **BMC Systems Biology**, 2013. v. 7.

EDWARDS, J. S.; COVERT, M.; PALSSON, B. Metabolic modelling of microbes: the flux-balance approach. **Environmental microbiology**, 2002. v. 4, n. 3, p. 133–140.

ELIASSON, S. H. H.; JENSEN, V. R. Benefit of a hemilabile ligand in deoxygenation of fatty acids to 1-alkenes. **Faraday Discussions**, 2019. v. 220, p. 231–248.

ENGLUND, E. *et al.* Systematic overexpression study to find target enzymes enhancing production of terpenes in *Synechocystis* PCC 6803, using isoprene as a model compound. **Metabolic Engineering**, 2018. v. 49, n. July, p. 164–177. Disponível em: <<https://doi.org/10.1016/j.ymben.2018.07.004>>.

FEIST, A. M. *et al.* **Reconstruction of biochemical networks in microorganisms. Nature Reviews Microbiology.**

FEIST, A. M.; PALSSON, B. Ø. The growing scope of applications of genome-scale metabolic reconstructions using *Escherichia coli*. **Nature biotechnology**, 2008. v. 26, n. 6, p. 659–667.

FERREIRA, L. S. De S. *et al.* High-Quality Draft Genome Sequence of *Pantanalinema* sp. GBBB05, a Cyanobacterium From Cerrado Biome. **Frontiers in Ecology and Evolution**, 2021. v. 9, n. June, p. 1–5.

FERRIS, M. J.; HIRSCH, C. F. Method for isolation and purification of cyanobacteria. **Applied and Environmental Microbiology**, 1991. v. 57, n. 5, p. 1448–1452.

FINN, R. D. *et al.* Pfam: The protein families database. **Nucleic Acids Research**, 2014. v. 42, n. D1, p. 222–230.

GASTEIGER, E *et al.* Protein Identification and Analysis Tools on the ExPASy Server. **The Proteomics Protocols Handbook**, 2005. p. 571–607. Disponível em: <<http://dx.doi.org/10.1385/1-59259-890-0:571>>.

GU, C. *et al.* Current status and applications of genome-scale metabolic models. **Genome Biology**, 2019. v. 20, n. 1, p. 1–18.

GUDMUNDSSON, S.; THIELE, I. Computationally efficient flux variability analysis. **BMC Bioinformatics**, 2010. v. 11, n. 2, p. 2–4.

GUO, L. *et al.* Directly converting carbon dioxide to linear α -olefins on bio-promoted catalysts. **Communications Chemistry**, 2018. v. 1, n. 1, p. 1–8.

HAN, J. *et al.* Organic geochemical studies, II. A preliminary report on the distribution of aliphatic hydrocarbons in algae, in bacteria, and in a recent lake sediment. **Proceedings of the National Academy of Sciences**, 1968. v. 59, n. 1, p. 29–33.

HAN, J.; CHAN, H. W. S.; CALVIN, M. Biosynthesis of alkanes in *Nostoc muscorum*. **Journal of the American Chemical Society**, 1969. v. 91, n. 18, p. 5156–5159.

HANKAMER, B. *et al.* Photosynthetic biomass and H₂ production by green algae: from bioengineering to bioreactor scale-up. **Physiologia plantarum**, 2007. v. 131, n. 1, p. 10–21.

HARI, A.; LOBO, D. Fluxer: a web application to compute, analyze and visualize genome-scale metabolic flux networks. **Nucleic acids research**, 2020. v. 48, n. W1, p. W427–W435.

HEAVNER, B. D.; PRICE, N. D. Transparency in metabolic network reconstruction enables scalable biological discovery. **Current Opinion in Biotechnology**, 2015. v. 34, p. 105–109. Disponível em: <<http://dx.doi.org/10.1016/j.copbio.2014.12.010>>.

HENRY, C. S. *et al.* High-throughput generation, optimization and analysis of genome-scale metabolic models. **Nature Biotechnology**, 2010. v. 28, n. 9, p. 977–982.

HESS, W. R. Cyanobacterial genomics for ecology and biotechnology. **Current Opinion in Microbiology**, 2011. v. 14, n. 5, p. 608–614. Disponível em: <<http://dx.doi.org/10.1016/j.mib.2011.07.024>>.

HOU, Y. *et al.* Trace mono-atomically dispersed rhodium on zeolite-supported cobalt catalyst for the efficient methane oxidation. **Communications Chemistry**, 2018. v. 1, n. 1, p. 1–7.

KANEHISA, M. KEGG: Kyoto Encyclopedia of Genes and Genomes. **Nucleic Acids Research**, 1 jan. 2000. v. 28, n. 1, p. 27–30. Disponível em: <<http://www.spandidos-publications.com/10.3892/ol.2020.11439>>.

KANEHISA, Minoru *et al.* New approach for understanding genome variations in KEGG. **Nucleic Acids Research**, 2019. v. 47, n. D1, p. D590–D595.

KANG, D. D. *et al.* MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. 2019. p. 1–13.

KHOLODENKO, Boris N *et al.* Quantification of short term signaling by the epidermal growth factor receptor. **Journal of Biological Chemistry**, 1999. v. 274, n. 42, p. 30169–30181.

KILLIAN, C. M.; JOHNSON, L. K.; BROOKHART, M. Preparation of linear α -olefins using cationic nickel (II) α -diimine catalysts. **Organometallics**, 1997. v. 16, n. 10, p. 2005–2007.

KIM, T. Y.; KIM, H. U.; LEE, S. Y. Metabolite-centric approaches for the discovery of antibacterials using genome-scale metabolic networks. **Metabolic engineering**, 2010. v. 12, n. 2, p. 105–111.

KIM, W. J.; KIM, H. U.; LEE, S. Y. Current state and applications of microbial genome-scale metabolic models. **Current Opinion in Systems Biology**, 2017. v. 2, p. 10–18.

KING, Z. A. *et al.* Escher: A Web Application for Building, Sharing, and Embedding Data-Rich Visualizations of Biological Pathways. **PLoS Computational Biology**, 2015. v. 11, n. 8, p. 1–13.

KING, Z. A. *et al.* BiGG Models: A platform for integrating, standardizing and sharing genome-scale models. **Nucleic Acids Research**, 2016. v. 44, n. D1, p. D515–D522.

KIYATKIN, Anatoly *et al.* Scaffolding protein Grb2-associated binder 1 sustains epidermal growth factor-induced mitogenic and survival signaling by multiple positive feedback loops. **Journal of biological chemistry**, 2006. v. 281, n. 29, p. 19925–19938.

KLUYVER, T. *et al.* Jupyter Notebooks-a publishing format for reproducible computational workflows. [S.l.]: [s.n.], 2016. p. 87–90.

KORELUSOVA, J.; KASTOVSKY, J.; KOMAREK, J. Heterogeneity of the cyanobacterial genus *synechocystis* and description of a new genus, *geminocystis*. **Journal of Phycology**, 2009. v. 45, n. 4, p. 928–937.

LEE, R. A.; LAVOIE, J. From first- to third-generation biofuels : Challenges of producing a commodity from a biomass of increasing complexity. 2012. n. August, p. 6–11.

LEWIS, N. E. *et al.* Omic data from evolved *E. coli* are consistent with computed optimal growth from genome-scale models. **Molecular Systems Biology**, 2010a. v. 6, n. 390.

LI, N. *et al.* Evidence for only oxygenative cleavage of aldehydes to alcohols and formate by cyanobacterial aldehyde decarbonylases. **Biochemistry**, 2012. v. 51, n. 40, p. 7908–7916.

LIEVEN, C. *et al.* MEMOTE for standardized genome-scale metabolic model testing. **Nature Biotechnology**, 2020. v. 38, n. 3, p. 272–276.

LIMA, A. R. J. *et al.* Insights into *Limnothrix* sp. metabolism based on comparative genomics. **Frontiers in Microbiology**, 2018. v. 9, n. NOV, p. 1–16.

LIN, P. C.; PAKRASI, H. B. Engineering cyanobacteria for production of terpenoids. **Planta**, 2019. v. 249, n. 1, p. 145–154. Disponível em: <<https://doi.org/10.1007/s00425-018-3047-y>>.

MACHADO, D. *et al.* Fast automated reconstruction of genome-scale metabolic models for microbial species and communities. **Nucleic Acids Research**, 2018. v. 46, n. 15, p. 7542–7553.

MACHADO, D.; HERRGÅRD, M. J. Co-evolution of strain design methods based on flux balance and elementary mode analysis. **Metabolic Engineering Communications**, 2015. v. 2, p. 85–92. Disponível em: <<http://dx.doi.org/10.1016/j.meteno.2015.04.001>>.

MARKEVICH, N I *et al.* Signal processing at the Ras circuit: what shapes Ras activation patterns? **Systems Biology**, 2004. v. 1, n. 1, p. 104–113.

MCCARTHY, J. G. *et al.* Structural basis of functional group activation by sulfotransferases in complex metabolic pathways. **ACS chemical biology**, 2012. v. 7, n. 12, p. 1994–2003.

MEDEMA, M. H. *et al.* AntiSMASH: Rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. **Nucleic Acids Research**, 2011. v. 39, n. SUPPL. 2, p. 339–346.

MENDOZA, S. N. *et al.* A systematic assessment of current genome-scale metabolic reconstruction tools. **Genome Biology**, 2019. v. 20, n. 1, p. 1–20.

MONK, J.; NOGALES, J.; PALSSON, B. O. Optimizing genome-scale network reconstructions. **Nature Biotechnology**, 2014. v. 32, n. 5, p. 447–452.

MONTAGUD, A. *et al.* Reconstruction and analysis of genome-scale metabolic model of a photosynthetic bacterium. **BMC Systems Biology**, 2010. v. 4.

NORSIGIAN, C. J.; PUSARLA, N.; *et al.* BiGG Models 2020: Multi-strain genome-scale models and expansion across the phylogenetic tree. **Nucleic Acids Research**, 2020. v. 48, n. D1, p. D402–D406.

NORSIGIAN, C. J.; FANG, X.; *et al.* A workflow for generating multi-strain genome-scale metabolic models of prokaryotes (S). **Nature Protocols**, 20 jan. 2020. v. 15, n. 1, p. 1–14. Disponível em: <<http://www.nature.com/articles/s41596-019-0254-3>>.

NOTEBAART, R. A. *et al.* Accelerating the reconstruction of genome-scale metabolic networks. **BMC bioinformatics**, 2006. v. 7, n. 1, p. 1–10.

NURK, S. *et al.* metaSPAdes: a new versatile metagenomic assembler. 2017. p. 824–834.

OBERHARDT, M. A.; PALSSON, B. Ø.; PAPIN, J. A. Applications of genome-scale metabolic reconstructions. **Molecular systems biology**, 2009. v. 5, n. 1, p. 320.

OLDFIELD, E.; LIN, F. Y. Terpene biosynthesis: Modularity rules. **Angewandte Chemie - International Edition**, 2012. v. 51, n. 5, p. 1124–1137.

ORTH, J. D.; THIELE, I.; PALSSON, B. O. What is flux balance analysis? **Nature Biotechnology**, 2010. v. 28, n. 3, p. 245–248. Disponível em: <<http://dx.doi.org/10.1038/nbt.1614>>.

PARMAR, A. *et al.* Cyanobacteria and microalgae: A positive prospect for biofuels. **Bioresource Technology**, 2011. v. 102, n. 22, p. 10163–10172. Disponível em: <<http://dx.doi.org/10.1016/j.biortech.2011.08.030>>.

PATTANAIK, B.; LINDBERG, P. Terpenoids and their biosynthesis in cyanobacteria. **Life**, 2015. v. 5, n. 1, p. 269–293.

PERCIVAL, S. L.; WILLIAMS, D. W. **Cyanobacteria**. Second Edition. [S.l.]: Elsevier, 2014.

PINU, F. R. *et al.* Systems Biology and Multi-Omics Integration: Viewpoints from the Metabolomics Research Community. 2019. p. 1–31.

PRICE, N. D.; REED, Jennifer L; PALSSON, B. Ø. Genome-scale models of microbial cells: evaluating the consequences of constraints. **Nature Reviews Microbiology**, 2004. v. 2, n. 11, p. 886–897.

QUINTANA, N. *et al.* Renewable energy from Cyanobacteria: energy production optimization by metabolic pathway engineering. **Applied Microbiology and Biotechnology**, 2011. v. 91, n. 3, p. 471–490. Disponível em: <<https://doi.org/10.1007/s00253-011-3394-0>>.

RASTOGI, R. P.; SINHA, Rajeshwar P. Biotechnological and industrial significance of cyanobacterial secondary metabolites. **Biotechnology Advances**, 2009. v. 27, n. 4, p. 521–539. Disponível em: <<http://dx.doi.org/10.1016/j.biotechadv.2009.04.009>>.

RIPPKA, R. *et al.* Generic Assignments, Strain Histories and Properties of Pure Cultures of Cyanobacteria. **Microbiology**, 1979. v. 111, n. 1, p. 1–61. Disponível em: <<https://www.microbiologyresearch.org/content/journal/micro/10.1099/00221287-111-1-1>>.

SARSEKEYEVA, F. *et al.* Cyanofuels: Biofuels from cyanobacteria. Reality and perspectives. **Photosynthesis Research**, 2015. v. 125, n. 1–2, p. 329–340. Disponível em: <<http://dx.doi.org/10.1007/s11120-015-0103-3>>.

SCHIRMER, A. *et al.* Microbial biosynthesis of alkanes. **Science**, 2010. v. 329, n. 5991, p. 559–562.

SCHOMBURG, I. BRENDA, the enzyme database: updates and major new developments. **Nucleic Acids Research**, 2004. v. 32, n. 90001, p. 431D – 433.

SEAVER, S. M. D. *et al.* The ModelSEED Biochemistry Database for the integration of metabolic annotations and the reconstruction, comparison and analysis of metabolic models for plants, fungi and microbes. **Nucleic Acids Research**, 2021. v. 49, n. D1, p. D575–D588.

SHLOMI, T. *et al.* Genome-scale metabolic modeling elucidates the role of proliferative adaptation in causing the Warburg effect. **PLoS Comput Biol**, 2011. v. 7, n. 3, p. e1002018.

SINGH, S. P. *et al.* Mycosporine-like amino acids (MAAs): chemical structure, biosynthesis and significance as UV-absorbing/screening compounds. **Indian journal of experimental biology**, 2008. v. 46, n. 1, p. 7–17. Disponível em: <<http://europepmc.org/abstract/MED/18697565>>.

SINGH, S. P.; MONTGOMERY, B. L. Determining cell shape: adaptive regulation of cyanobacterial cellular differentiation and morphology. **Trends in microbiology**, 2011. v. 19, n. 6, p. 278–285.

SINHA, R P; KLISCH, M.; H, D. Ultraviolet-absorbing / screening substances in cyanobacteria , phytoplankton and macroalgae. 1998. v. 1344, n. 98.

SIWKO, M. E. *et al.* Does isoprene protect plant membranes from thermal shock? A molecular dynamics study. Laboratory of Lipids and Liposomes, Institute of Biochemistry and Molecular Biology, University of Wrocław, Przybyszewskiego 63/77, 51-148 Wrocław, Poland.: **Biochimica et biophysica acta**, 2007. v. 1768, n. 2, p. 198–206. Disponível em: <<http://europepmc.org/abstract/MED/17125733>>.

SUENAGA, A. *et al.* Tyr-317 phosphorylation increases Shc structural rigidity and reduces coupling of domain motions remote from the phosphorylation site as revealed by molecular dynamics simulations. **Journal of Biological Chemistry**, 2004. v. 279, n. 6, p. 4657–4662.

TAMAGNINI, P. *et al.* Cyanobacterial hydrogenases: diversity, regulation and applications. **FEMS microbiology reviews**, 2007. v. 31, n. 6, p. 692–720.

TATUSOVA, T. *et al.* NCBI prokaryotic genome annotation pipeline. 2016. v. 44, n. 14, p. 6614–6624.

THIELE, I.; PALSSON, B. Ø. A protocol for generating a high-quality genome-scale metabolic reconstruction. **Nature Protocols**, 7 jan. 2010. v. 5, n. 1, p. 93–121. Disponível em: <<http://www.nature.com/articles/nprot.2009.203>>.

URITSKIY, G. V.; DIRUGGIERO, J.; TAYLOR, J. MetaWRAP - A flexible pipeline for genome-resolved metagenomic data analysis 08 Information and Computing Sciences 0803 Computer Software 08 Information and Computing Sciences 0806 Information Systems. **Microbiome**, 2018. v. 6, n. 1, p. 1–13.

VANNA, A.; PATSSON, B. Metabolic Flux Balancing: Basic Concepts, Scientific and Practical Use. 1994. v. 12, n. October, p. 994–998.

VÄREMO, L.; NOOKAEW, I.; NIELSEN, J. Novel insights into obesity and diabetes through genome-scale metabolic modeling. **Frontiers in physiology**, 2013. v. 4, p. 92.

WANG, J. *et al.* Directly converting syngas to linear α -olefins over core-shell Fe₃O₄@ MnO₂ catalysts. **ACS applied materials & interfaces**, 2018. v. 10, n. 50, p. 43578–43587.

WINTERS, K.; PARKER, P. L.; BAALEN, C. VAN. Hydrocarbons of Blue-Green Algae: Geochemical Significance. **Science**, 31 jan. 1969. v. 163, n. 3866, p. 467–468. Disponível em: <<https://www.science.org/doi/10.1126/science.163.3866.467>>.

WU, Y.; SIMMONS, B. A.; SINGER, S. W. MaxBin 2 . 0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. 2015. p. 1–2.

XIE, M. *et al.* Versatility of hydrocarbon production in cyanobacteria. **Applied Microbiology and Biotechnology**, 2017. v. 101, n. 3, p. 905–919. Disponível em: <<http://dx.doi.org/10.1007/s00253-016-8064-9>>.

ZHANG, Z. *et al.* In situ hydrogenation and decarboxylation of oleic acid into heptadecane over a Cu–Ni alloy catalyst using methanol as a hydrogen carrier. **Green Chemistry**, 2018. v. 20, n. 1, p. 197–205.

ZHANG, Z. *et al.* Highly selective production of linear 1-heptadecene from stearic acid over a partially reduced MoO_x catalyst. **Chemical Communications**, 2020. v. 56, n. 32, p. 4456–4459.

ZIEMERT, N. *et al.* The natural product domain seeker NaPDoS: A phylogeny based bioinformatic tool to classify secondary metabolite gene diversity. **PLoS ONE**, 2012. v. 7, n. 3, p. 1–9.

ZILLIGES, Y. *et al.* Flux Balance Analysis of Cyanobacterial Metabolism : The Metabolic Network of Synechocystis sp . PCC 6803. 2013. v. 9, n. 6.

9. ANEXOS

As imagens suplementares e os arquivos de programação estão hospedados no link: <https://github.com/igorshribeiro/dissertacao-ufma/tree/master>

https://drive.google.com/drive/folders/1-t2nUCqpa_LxkfATD3w109YX_LxEILaf?usp=share_link

A descrição dos materiais suplementares segue no link onde está hospedados os mesmos.

Tabela suplementar 1. Quadro de análise de balanço de fluxo do GSM G.sp GBBB08 pelo CarveMe

Absorção				
Metabólito	Reação	Fluxo	C-Número	C-Fluxo
ca2_e	EX_ca2_e	0.0001104	0	0.00%
cit_e	EX_cit_e	0.0003311	6	0.22%
cl_e	EX_cl_e	0.0001104	0	0.00%
co2_e	EX_co2_e	0.9028	1	99.78%
cobalt2_e	EX_cobalt2_e	2,12E-03	0	0.00%
cu2_e	EX_cu2_e	1,50E-02	0	0.00%
fe2_e	EX_fe2_e	0.0001424	0	0.00%
fe3_e	EX_fe3_e	0.0001656	0	0.00%
h2o_e	EX_h2o_e	0.3049	0	0.00%
k_e	EX_k_e	0.004139	0	0.00%
mg2_e	EX_mg2_e	0.000184	0	0.00%
mn2_e	EX_mn2_e	1,47E-02	0	0.00%
nh4_e	EX_nh4_e	0.2374	0	0.00%
photon_e	EX_photon_e	10	0	0.00%
pi_e	EX_pi_e	0.02091	0	0.00%
so4_e	EX_so4_e	0.005315	0	0.00%
zn2_e	EX_zn2_e	7,23E-03	0	0.00%
Sercreção				
Metabólito	Reação	Fluxo	C-Número	C-Fluxo
h_e	EX_h_e	-0.2066	0	0.00%
o2_e	EX_o2_e	-943	0	0.00%
4hba_c	sink_4hba_c	-4,73E-03	7	100.00%

Tabela Suplementar 2. Quadro de análise de variação de Fluxo do GSM G.sp GBBB08 pelo CarveMe

Absorção					
Metabólito	Reação	Fluxo	Intervalo	C-Número	C-Fluxo
ca2_e	EX_ca2_e	0.0001104	[0.0001049; 0.0001104]	0	0.00%
cit_e	EX_cit_e	0.0003311	[0; 0.0003311]	6	0.22%
cl_e	EX_cl_e	0.0001104	[0.0001049; 0.0001104]	0	0.00%
co2_e	EX_co2_e	0.9028	[0.65; 0.9389]	1	99.78%
cobalt2_e	EX_cobalt2_e	2,12E-03	[2.015E-06; 2.121E-06]	0	0.00%
cu2_e	EX_cu2_e	1,50E-02	[1.428E-05; 1.503E-05]	0	0.00%
fe2_e	EX_fe2_e	0.0001424	[0.0001353; 0.001455]	0	0.00%
fe3_e	EX_fe3_e	0.0001656	[0; 0.0001656]	0	0.00%
h2o_e	EX_h2o_e	0.3049	[0.04921; 0.3552]	0	0.00%
k_e	EX_k_e	0.004139	[0.003932; 0.004139]	0	0.00%
mg2_e	EX_mg2_e	0.000184	[0.0001748; 0.000184]	0	0.00%
mn2_e	EX_mn2_e	1,47E-02	[1.392E-05; 1.465E-05]	0	0.00%
nh4_e	EX_nh4_e	0.2374	[0.179; 0.2571]	0	0.00%
photon_e	EX_photon_e	10	[9.5; 10]	0	0.00%
pi_e	EX_pi_e	0.02091	[0.01987; 0.04253]	0	0.00%
so4_e	EX_so4_e	0.005315	[0.005049; 0.04188]	0	0.00%
zn2_e	EX_zn2_e	7,23E-03	[6.87E-06; 7.231E-06]	0	0.00%
Secreção					
Metabólito	Reação	Fluxo	Intervalo	C-Número	C-Fluxo
12ppd__S_e	EX_12ppd__S_e	0	[-0.01637; 0]	3	0.00%
LalaDgluMdap_e	EX_LalaDgluMdap_e	0	[-0.001701; 0]	15	0.00%
acald_e	EX_acald_e	0	[-0.007792; 0]	2	0.00%
ala__D_e	EX_ala__D_e	0	[-0.013; 0]	3	0.00%
alaala_e	EX_alaala_e	0	[-0.008072; 0]	6	0.00%
alltn_e	EX_alltn_e	0	[-0.007892; 0]	8	0.00%
anhgm_e	EX_anhgm_e	0	[-0.002884; 0]	19	0.00%
asp__L_e	EX_asp__L_e	0	[-0.02032; 0]	4	0.00%
dha_e	EX_dha_e	0	[-0.01414; 0]	3	0.00%

Continuação da Tabela Suplementar 2. Quadro de análise de variação de Fluxo do GSM G.sp GBBB08 pelo CarveMe;

Metabólito	Reação	Secreção		C-Número	C-Fluxo
		Fluxo	Intervalo		
enlipa_e	EX_enlipa_e	0	[-0.0003987; 0]	112	0.00%
etha_e	EX_etha_e	0	[-0.008499; 0]	2	0.00%
g3pe_e	EX_g3pe_e	0	[-0.005835; 0]	5	0.00%
g3pg_e	EX_g3pg_e	0	[-0.00772; 0]	6	0.00%
glu__L_e	EX_glu__L_e	0	[-0.01245; 0]	5	0.00%
glyc_e	EX_glyc_e	0	[-0.01319; 0]	3	0.00%
gm1lipa_e	EX_gm1lipa_e	0	[-0.0003371; 0]	142	0.00%
h2s_e	EX_h2s_e	0	[-0.03683; 0]	0	0.00%
h_e	EX_h_e	-0.2066	[-0.248; 0]	0	0.00%
hco3_e	EX_hco3_e	0	[0; 0.232]	2	0.00%
hxan_e	EX_hxan_e	0	[-0.007608; 0]	5	0.00%
inost_e	EX_inost_e	0	[-0.009876; 0]	6	0.00%
kdo2lipid4_e	EX_kdo2lipid4_e	0	[-0.000558; 0]	84	0.00%
lipa_cold_e	EX_lipa_cold_e	0	[-0.0004013; 0]	114	0.00%
lipa_e	EX_lipa_e	0	[-0.0004168; 0]	110	0.00%
mobd_e	EX_mobd_e	0	[0; 0.442]	0	0.00%
no3_e	EX_no3_e	0	[0; 0.04652]	0	0.00%
o2_e	EX_o2_e	-943	[-0.9889; -0.8959]	0	0.00%
pser__L_e	EX_pser__L_e	0	[-0.02266; 0]	3	0.00%
pydxn_e	EX_pydxn_e	0	[-0.006646; 0]	8	0.00%
s_e	EX_s_e	0	[-0.01179; 0]	0	0.00%
spm�_e	EX_spm�_e	0	[-0.004994; 0]	7	0.00%
uaagmda_e	EX_uaagmda_e	0	[-0.0005159; 0]	95	0.00%
udcpo5_e	EX_udcpo5_e	0	[-0.00058; 0]	81	0.00%
xtsn_e	EX_xtsn_e	0	[-0.006003; 0]	10	0.00%
2ohph_c	sink_2ohph_c	0	[-0.000966; 0]	46	0.00%
4hba_c	sink_4hba_c	-4,73E-03	[-4.729E-06; -4.492E-06]	7	100.00%
5drib_c	sink_5drib_c	0	[-0.008261; 0]	5	0.00%
hemeO_c	sink_hemeO_c	0	[-0.0009394; 0]	49	0.00%
lipopb_c	sink_lipopb_c	0	[-0.001333; 0]	8	0.00%
mobd_c	sink_mobd_c	0	[-0.442; 0]	0	0.00%
sheme_c	sink_sheme_c	0	[-0.001194; 0]	42	0.00%

Figura Suplementar 1. Visualização geral do teste de modelo metabólico do Igbg 08 realizado no Memote.

iGbb08 Expand All Readme

2022-06-20 18:39

Independent Section

Contains tests that are independent of the class of modeled organism, a model's

Consistency

Stoichiometric Consistency	0.0%	^{x3} ↓
Mass Balance	54.0%	↓
Charge Balance	99.9%	↓
Metabolite Connectivity	100.0%	↓
Unbounded Flux In Default Medium	86.2%	↓
Sub Total	49%	^{x3} ↓

Annotation - Metabolites

Presence of Metabolite Annotation	0.0%	↓
Metabolite Annotations Per Database	Info	↓
pubchem.compound	0.0%	↓
kegg.compound	0.0%	↓
seed.compound	0.0%	↓
inchikey	0.0%	↓
inchi	0.0%	↓
chebi	0.0%	↓
hmdb	0.0%	↓
reactome	0.0%	↓
metanetx.chemical	0.0%	↓
bigg.metabolite	0.0%	↓
biocyc	0.0%	↓
Metabolite Annotation Conformity Per Database	Info	↓
pubchem.compound	0.0%	↓
kegg.compound	0.0%	↓
seed.compound	0.0%	↓
inchikey	0.0%	↓
inchi	0.0%	↓
chebi	0.0%	↓

Specific Section

Covers general statistics and specific aspects of a metabolic network that are not

SBML

SBML Level and Version	Errored	↓
FBC enabled	Errored	↓

Basic Information

Model Identifier	iGbb08	↓
Total Metabolites	1,127	↓
Total Reactions	1,509	↓
Total Genes	758	↓
Total Compartments	4	↓
Metabolic Coverage	1.99	↓
Uncoserved Metabolites	7	↓
Minimal Inconsistent Net Stoichiometries	Skipped	↓

Metabolite Information

Unique Metabolites	1,127	↓
Duplicate Metabolites in Identical Compartments	0	↓
Metabolites without Charge	0	↓
Metabolites without Formula	0	↓
Medium Components	21	↓

Reaction Information

Purely Metabolic Reactions	994	↓
Purely Metabolic Reactions with Constraints	2	↓
Transport Reactions	376	↓
Transport Reactions with Constraints	1	↓
Reactions With Partially Identical Annotations	0.00	↓
Duplicate Reactions	0.00	↓

iGbb08

2022-06-20 18:39

metanetx.chemical	0.0%	▼
bigg.metabolite	0.0%	▼
biocyc	0.0%	▼
Uniform Metabolite Identifier Namespace	100.0%	▼
Sub Total	25%	▼

Annotation - Reactions

Presence of Reaction Annotation	0.0%	▼
Reaction Annotations Per Database	Info	▼
rhea	0.0%	▼
kegg.reaction	0.0%	▼
seed.reaction	0.0%	▼
metanetx.reaction	0.0%	▼
bigg.reaction	0.0%	▼
reactome	0.0%	▼
ec-code	0.0%	▼
brenda	0.0%	▼
biocyc	0.0%	▼
Reaction Annotation Conformity Per Database	Info	▼
rhea	0.0%	▼
kegg.reaction	0.0%	▼
seed.reaction	0.0%	▼
metanetx.reaction	0.0%	▼
bigg.reaction	0.0%	▼
reactome	0.0%	▼
ec-code	0.0%	▼
brenda	0.0%	▼
biocyc	0.0%	▼
Uniform Reaction Identifier Namespace	100.0%	▼
Sub Total	25%	▼

Annotation - Genes

Presence of Gene Annotation	0.0%	▼
-----------------------------	------	---

Gene-Protein-Reaction (GPR) Associations

Reactions without GPR	401	▼
Fraction of Transport Reactions without GPR	0.35	▼
Enzyme Complexes	134	▼

Biomass

Biomass Reactions Identified	1	▼
Biomass Consistency	-4.14	▼
Biomass Production In Default Medium	0.02	▼
Unrealistic Growth Rate In Default Medium	false	▼
Biomass Production In Complete Medium	87.63	▼
Blocked Biomass Precursors In Default Medium	0	▼
Blocked Biomass Precursors In Complete Medium	0	▼
Ratio of Direct Metabolites in Biomass Reaction	0.18	▼
Number of Missing Essential Biomass Precursors	1	▼

Energy Metabolism

Non-Growth Associated Maintenance Reaction	1	▼
Growth-associated Maintenance in Biomass Reaction	true	▼
Number of Reversible Oxygen-Containing Reactions	4	▼
Erroneous Energy-generating Cycles	Info	▼
MNXM3	Skipped	▼
MNXM63	Skipped	▼
MNXM51	Skipped	▼
MNXM121	Skipped	▼
MNXM423	Skipped	▼

iGbb08

2022-06-20 18:39

uniprot	0.0%	▼
ecogene	0.0%	▼
kegg_genes	0.0%	▼
ncbigi	0.0%	▼
ncbigene	0.0%	▼
ncbiprotein	0.0%	▼
ccds	0.0%	▼
hprd	0.0%	▼
asap	0.0%	▼

Gene Annotation Conformity
Per Database

Info

refseq	0.0%	▼
uniprot	0.0%	▼
ecogene	0.0%	▼
kegg_genes	0.0%	▼
ncbigi	0.0%	▼
ncbigene	0.0%	▼
ncbiprotein	0.0%	▼
ccds	0.0%	▼
hprd	0.0%	▼
asap	0.0%	▼

Sub Total 0% ▼

Annotation - SBO Terms

Metabolite General SBO Presence	0.0%	▼
Metabolite SBO:0000247 Presence	0.0%	▼
Reaction General SBO Presence	0.0%	▼
Metabolic Reaction SBO:0000176 Presence	0.0%	▼
Transport Reaction SBO:0000185 Presence	0.0%	▼
Exchange Reaction SBO:0000627 Presence	0.0%	▼
Demand Reaction SBO:0000628 Presence	Skipped	▼
Sink Reactions SBO:0000632 Presence	Skipped	▼
Gene General SBO Presence	0.0%	▼

MNXM38	Skipped	▼
MNXM208	Skipped	▼
MNXM191	Skipped	▼
MNXM223	Skipped	▼
MNXM7517	Skipped	▼
MNXM12233	Skipped	▼
MNXM558	Skipped	▼
MNXM21	Skipped	▼
MNXM89557	Skipped	▼

Network Topology

Universally Blocked Reactions	16	▼
Orphan Metabolites	0	▼
Dead-end Metabolites	0	▼
Stoichiometrically Balanced Cycles	62	▼
Metabolite Production In Complete Medium	171	▼
Metabolite Consumption In Complete Medium	171	▼

Matrix Conditioning

Ratio Min/Max Non-Zero Coefficients	0.00	▼
Independent Conservation Relations	50	▼
Rank	1077	▼
Degrees Of Freedom	432	▼

Experimental Data Comparison

Growth Prediction	Skipped	▼
Gene Essentiality Prediction	Skipped	▼

Misc. Tests

Environment

iGbb08

2022-06-20 18:39

SBO:0000629 Presence

Sub Total 0% x2 ▼

Total Score 23% ▼

Total Score

23%

Score per Category
Export

