UNIVERSIDADE ESTADUAL PAULISTA - UNESP

CÂMPUS DE JABOTICABAL

METAGENÔMICA E BIOINFORMÁTICA APLICADA À BIOENERGIA: EXPLORANDO UM CONSÓRCIO BACTERIANO DEGRADADOR DE BIOMASSA

Joana Gabriela Desiderato

Bióloga

2017

UNIVERSIDADE ESTADUAL PAULISTA - UNESP

CÂMPUS DE JABOTICABAL

METAGENÔMICA E BIOINFORMÁTICA APLICADA À BIOENERGIA: EXPLORANDO UM CONSÓRCIO BACTERIANO DEGRADADOR DE BIOMASSA

Joana Gabriela Desiderato

Orientador: Prof. Dr. Alessandro de Mello Varani
Coorientadora: Prof[a]. Dr[a]. Lucia Maria Carareto Alves

Dissertação apresentada à Faculdade de Ciências Agrárias e Veterinárias – Unesp, Campus de Jaboticabal, como parte das exigências para a obtenção do título de Mestre em Microbiologia Agropecuária

2017

Desiderato, Joana Gabriela
D458m Metagenômica e bioinformática aplicada à bioenergia: explorando um consórcio bacteriano degradador de biomassa / Joana Gabriela Desiderato. – – Jaboticabal, 2017
x, 78 p. : il. ; 29 cm

Dissertação (mestrado) - Universidade Estadual Paulista, Faculdade de Ciências Agrárias e Veterinárias, 2017
Orientador: Alessandro de Mello Varani
Banca examinadora: Jackson Antonio Marcondes de Souza, Victor Satler Pylro
Bibliografia

1. Bioetanol. 2. Biomassa. 3. *Brazilian Microbiome Project* (BMP). 4. *Dyella jiangningensis*. 5. Etanol de segunda geração. I. Título. II. Jaboticabal-Faculdade de Ciências Agrárias e Veterinárias.

CDU 576.8:661.722

Ficha catalográfica elaborada pela Seção Técnica de Aquisição e Tratamento da Informação – Diretoria Técnica de Biblioteca e Documentação - UNESP, Câmpus de Jaboticabal.

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Jaboticabal

CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: METAGENÔMICA E BIOINFORMÁTICA APLICADA À BIOENERGIA: EXPLORANDO UM CONSÓRCIO BACTERIANO DEGRADADOR DE BIOMASSA

AUTORA: JOANA GABRIELA DESIDERATO
ORIENTADOR: ALESSANDRO DE MELLO VARANI
COORIENTADORA: LUCIA MARIA CARARETO ALVES

Aprovada como parte das exigências para obtenção do Título de Mestra em MICROBIOLOGIA AGROPECUÁRIA, pela Comissão Examinadora:

Prof. Dr. ALESSANDRO DE MELLO VARANI
Departamento de Tecnologia / FCAV / UNESP - Jaboticabal

Prof. Dr. VICTOR SATLER PYLRO
Departamento de Ciências do Solo / ESALQ/USP / Piracicaba/SP

Prof. Dr. JACKSON ANTONIO MARCONDES DE SOUZA
Departamento de Biologia Aplicada à Agropecuária / FCAV / UNESP - Jaboticabal

Jaboticabal, 04 de julho de 2017

DADOS CURRICULARES DO AUTOR

Joana Gabriela Desiderato – Nasceu em 18 de abril de 1986, em São Paulo, capital, Brasil, filha de Renato Antônio Desiderato e Sonia Maria Desiderato. Graduou-se em Ciências Biológicas, Licenciatura e Bacharelado, em fevereiro de 2011 pela Universidade Estadual Paulista “Júlio de Mesquita Filho”, Faculdade de Ciências e Letras, Campus de Assis. Ingressou em março de 2015 no Curso de Mestrado, área de concentração em Microbiologia Agropecuária, na Faculdade de Ciências Agrárias e Veterinárias – UNESP, Campus de Jaboticabal. Em 04 de julho de 2017 defendeu sua dissertação de mestrado em Microbiologia Agropecuária.

DEDICATÓRIA

Aos meus antepassados que, exercendo seus papéis nas teias da vida, me permitiram vir ao mundo;

Aos meus familiares, especialmente meus pais, Renato e Sonia, e ao meu irmão, Renato (Tato), pelo apoio e amor incondicional, essenciais para que eu trilhasse esse caminho;

A vocês, dedico meu trabalho e o meu amor.

AGRADECIMENTOS

Ao prof. Dr. Alessandro de Mello Varani pela oportunidade e pela confiança depositada em mim, pela orientação desempenhada com excelência, por compartilhar seus conhecimentos e suas experiências, tão importantes para o meu aprimoramento e para o desenvolvimento desse trabalho.

À profa. Dra. Lucia Maria Carareto Alves pela coorientação, por compartilhar seus conhecimentos, pelas dicas e ideias, contribuindo enormemente para a realização desse trabalho e para o meu aprimoramento profissional.

À profa. Dra. Eliana G. de Macedo Lemos por ter me recebido e pela oportunidade de trabalhar no Laboratório de Bioquímica de Microrganismos e Plantas (LBMP).

Aos membros da banca do exame de qualificação, Prof. Dr. Manoel Victor F. Lemos e Dr. Luciano T. Kishi por compartilharem seus conhecimentos e pelas sugestões oferecidas que contribuíram para o melhoramento desse trabalho.

À banca examinadora, Prof. Dr. Jackson Antônio Marcondes de Souza e ao Dr. Victor Satler Pylro por compartilharem seus conhecimentos, pelas correções e sugestões oferecidas durante a defesa de mestrado, contribuindo para o enriquecimento desse trabalho.

Ao Dr. João Carlos Campanharo e à Dra. Camila Cesário Fernandes pelo auxílio no desenvolvimento desse trabalho e pelos conhecimentos compartilhados.

À Milena T. Lima Constancio e ao Dr. Danillo O. Alvarenga por todos os ensinamentos e por todo auxílio oferecido para o desenvolvimento desse trabalho.

Ao Ricardo Lopes da Silva (Batata) por ter deixado meu notebook em perfeito estado e pelos ensinamentos iniciais com o Linux.

A todos os companheiros do Laboratório de Bioquímica de Microrganismos e Plantas, pelos ensinamentos, pelas contribuições a esse trabalho e pelos bons momentos que passamos juntos.

A todos os funcionários do Departamento de Tecnologia pelo auxílio durante o desenvolvimento desse trabalho.

A todos os meus familiares, especialmente meus pais, Renato A. Desiderato e Sonia M. Desiderato, ao meu irmão Renato C. Desiderato e à minha cunhada Amanda Leite F. Desiderato, pelo amor, apoio e carinho compartilhados, pela compreensão e por acreditarem em mim. Vocês foram essenciais para a realização desse sonho.

Ao Guilherme F. Pitaluga pela compreensão, pelo carinho, apoio, amor, companheirismo, pelos finais de semanas que passamos estudando juntos e por fazer tudo o que estava em seu alcance para me ajudar, sempre que precisei.

A todos os meus amigos que me acompanharam e apoiaram durante essa caminhada, independente dos quilômetros que nos separam. A amizade e o apoio de vocês foram fundamentais durante essa etapa da minha vida.

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

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

Meus sinceros agradecimentos.

SUMÁRIO

Página

RESUMO.....IV
ABSTRACT.....V
LISTA DE FIGURAS.....VI
LISTA DE TABELAS.....IX
1. INTRODUÇÃO.....1
2. OBJETIVOS.....3
3. REVISÃO DE LITERATURA.....4
3.1 Breve História sobre o Etanol no Brasil.....4
3.2 Biomassa e a Conversão em Biocombustíveis.....5
3.3 Processo de Produção de Bioetanol de Segunda Geração.....7
3.4 Enzimas Degradadoras da Biomassa.....7
3.5 Parede Celular Vegetal e Suas Enzimas Degradadoras.....11
3.5.1 Estrutura da Celulose e Enzimas Degradadoras da Celulose.....12
3.5.2 Estrutura da Hemicelulose e Enzimas Degradadoras da Hemicelulose.....14
3.5.3 Estrutura da Pectina e Enzimas Degradadoras da Pectina.....16
3.5.4 Estrutura da Lignina e Enzimas Degradadoras da Lignina.....17
3.6 Consórcios Microbianos.....18
3.7 Metagenômica e Sequenciamento de Nova Geração.....19
3.8 *Brazilian Microbiome Project* (BMP).....21
4. MATERIAL E MÉTODOS.....22
4.1 Coleta do Material, Isolamento e Adaptação do Consórcio Bacteriano em Meio de Cultura Contendo Bagaço de Cana-de-Açúcar.....22
4.2 Extração e Quantificação do DNA Metagenômico do Consórcio.....23
4.3 Amplificação do Gene 16S rRNA através da Técnica de PCR (Reação em Cadeia da Polimerase) e Purificação dos Produtos da PCR.....24
4.4 Sequenciamento do 16S rRNA em Plataforma Ion PGM™ Torrent.....25
4.4.1 Análises das Sequências Obtidas pela Plataforma Ion PGM™ Torrent.....25
4.4.2 Procedimento de Análise do 16S rRNA de Acordo com as Recomendações do BMP.....26

Filtragem e Seleção dos *Reads*.....26
Preparação das Sequências, Formação de Matriz das Unidades Taxonômicas Operacionais (UTOs) e Classificação.....26
Estimativa de Riqueza de UTOs e Diversidade dos Micro-organismos Presentes no Consórcio Bacteriano a Partir da Análise do 16S rRNA.....27
4.5 Predição Metabólica Funcional do Consórcio a Partir do 16S rRNA.....27
4.6 Sequenciamento Total do Metagenoma em Plataforma Illumina® HiSeq 2500 e Análise da Qualidade dos *Reads*.....29
4.7 Identificação, Recuperação, Montagem e Anotação de um Genoma Completo Proveniente do Consórcio por Abordagens de Bioinformática.....30
5. RESULTADOS.....31
5.1 Diversidade da Comunidade Bacteriana Presente no Consórcio Através da Análise do Sequenciamento do 16S rRNA Realizado pela Plataforma Ion PGM™.....31
5.2 Estimativa de Riqueza de UTOs do Consórcio Bacteriano, Realizada Através do Programa QIIME de Acordo com o Índice de Chao-1.....32
5.2.1 Agrupamento de UTOs das Amostras das Semanas Avaliadas Baseado em Análise de Coordenadas Principais (PCoA) Utilizando as Distâncias Métricas UniFrac Ponderada e Não-Ponderada.....33
5.3 Diversidade Bacteriana Presente no Consórcio.....36
5.4 Predição do Potencial Metabólico Funcional do Consórcio Bacteriano a Partir do 16S rRNA.....43
5.5 Sequenciamento do DNA Total do Metagenoma em Plataforma Illumina® HiSeq 2500.....48
5.6 Identificação, Montagem e Anotação do Genoma da *Dyella jiangningensis* FCAV SCS01.....49
5.7 Características do Genoma de *Dyella jiangningensis* FCAV SCS01.....50
5.8 Número de Genes Associados aos Subsistemas Gerados pela Plataforma RAST.....51
5.9 Visualização Genômica Circular de *Dyella jiangningensis* FCAV SCS01.....53
5.10 Enzimas com Potencial Degradador de Biomassa Encontradas no Genoma de *Dyella jiangningensis* FCAV SCS01.....54

6. DISCUSSÃO.....56

6.1 Provável Papel do Gênero *Burkholderia* Presente no Consórcio.....56

6.2 O Genoma da *Dyella jiangningensis* FCAV SCS01 e seu Possível Papel na Degradação da Biomassa.....58

6.3 Considerações Sobre os Outros Gêneros Bacterianos Identificados no Consórcio.....60

6.4 Predição do Perfil Metabólico do Consórcio Através da Análise do 16S rRNA Sugere que as Bactérias são Capazes de Utilizar os Açúcares Disponíveis no Meio de Cultivo Contendo Bagaço de Cana-de-açúcar Triturado.....61

7. CONCLUSÃO.....62

8. REFERÊNCIAS.....63

METAGENÔMICA E BIOINFORMÁTICA APLICADA À BIOENERGIA: EXPLORANDO UM CONSÓRCIO BACTERIANO DEGRADADOR DE BIOMASSA

RESUMO – A indústria sucroalcooleira gera elevado número de resíduos de origem da biomassa lignocelulósica que apresentam grande potencial para produção de biocombustíveis, em particular o etanol de segunda geração. Uma das formas promissoras para a desconstrução da biomassa lignocelulósica é através da utilização de consórcios bacterianos que produzem enzimas altamente específicas com a capacidade de quebrar a estrutura da lignocelulose. Porém, para a otimização desse processo é importante entender as funções metabólicas presentes nesses consórcios degradadores de biomassa. Portanto, esse estudo objetivou identificar e classificar a composição de uma comunidade bacteriana proveniente de consórcio oriundo de solo contendo bagaço de cana-de-açúcar em decomposição e avaliar sua capacidade metabólica através de sequenciamento genômico de alto rendimento e análises de bioinformática. Esse consórcio foi cultivado durante vinte semanas, sendo que amostras de DNA foram extraídas para sequenciamento a cada sete dias. O sequenciamento da subunidade 16S do operon ribossomal (16S rRNA) foi realizado utilizando a plataforma Ion PGM™, enquanto que o sequenciamento do DNA total foi realizado na plataforma HiSeq 2500, Illumina®. Os resultados do sequenciamento do 16S rRNA indicam 5 diferentes famílias bacterianas ao longo das 20 semanas, sendo *Burkholderiaceae* (73%) e *Rhodanobacteraceae* (24%) as mais abundantes. Análises do potencial funcional do consórcio realizadas através dos programas PICRUSt e STAMP, indicam o enriquecimento da função de transportadores, incluindo transportadores do tipo ABC, principalmente na primeira semana, sugerindo um provável papel na metabolização do material lignocelulolítico presente no meio. Os dados gerados pelo sequenciamento total do metagenoma foram montados ("*de novo" assembly*) e permitiram recuperar a sequência genômica de um dos organismos mais abundantes presentes no consórcio (24%). Esse genoma foi finalizado e circularizado, apresentando 4.758.639 pb e GC% de 65,25, e similaridade de 99% da sequência do 16S rRNA e 90,77% de identidade de suas sequências codificadoras com o genoma da bactéria *Dyella jiangningensis* SBZ3-12 (*Rhodanobacteraceae*). A anotação realizada através da plataforma RAST, indicou 4.194 genes codificadores de proteínas, dentre os quais 36 Glicosil Hidrolases potencialmente envolvidas com a degradação da biomassa lignocelulósica. Portanto, esse estudo permitiu a elucidação da diversidade microbiana e perfil metabólico de um consórcio bacteriano, revelando um novo genoma pertencente ao gênero *Dyella*, potencialmente relacionado com o processo de degradação da biomassa lignocelulósica. Em linhas gerais, os resultados obtidos também corroboraram o potencial que ferramentas de sequenciamento de alto rendimento, metagenômica e bioinformática apresentam para o estudo de comunidades microbianas com potencial biotecnológico, mostrando-se como uma valiosa alternativa para a investigação de novos alvos para a pesquisa em bioenergia.

Palavras-chaves: bioetanol, biomassa, *Brazilian Microbiome Project* (BMP), *Dyella jiangningensis*, etanol de segunda geração, lignocelulose

METAGENOMIC AND BIOINFORMATIC APPLIED TO BIOENERGY: EXPLORING A BACTERIAL CONSORTIUM RELATED TO THE BIOMASS DEGRADATION

ABSTRACT – The sugarcane ethanol industry generates a number of residues related to the lignocellulosic biomass which exhibit potential for the production of biofuels, in particular second generation ethanol. One of the promising ways for the deconstruction of lignocellulosic biomass to valuable products is through the use of bacterial consortia. These consortia encode a specific set of enzymes capable to metabolize and decompose the lignocellulose structure. However, to optimize this process, the metabolic functions present in a biomass-degrading consortia must be well known. Therefore, this study aimed to unravel the taxonomic composition, and to evaluate the metabolic profile of a bacterial consortium from a sugar-cane derived soil containing decomposing straw, through high-throughput genomic sequencing and bioinformatics analyzes. The selected consortium was cultivated in laboratory for twenty weeks, and DNA samples were extracted for sequencing every seven days. The 16S subunit of the ribosomal operon (16S rRNA) and total DNA sequencing were performed through the Ion PGM™ (Thermo Fischer) and HiSeq 2500 (Illumina) platforms, respectively. The 16S rRNA sequencing indicate at least 5 different bacterial families during the 20 weeks of cultivation. The *Burkholderiaceae* (73%) and *Rhodanobacteraceae* (24%) are the most abundant. Functional analyses, indicate an enrichment of the transporter-related function, including ABC-transporters mainly in the first week of cultivation, suggesting a probable role related with the decomposing of the lignocellulolytic material. The whole metagenome sequencing uncover the genomic sequence of one of the most abundant organisms present in the consortium (abundance ~24% of the sequenced *reads*). This genome was finished and circularized, exhibiting 4,758,639 bp, GC% of 65.25, and 99% of similarity of the 16S rRNA and 90.77% identity of the coding sequences to the genome of the bacterium *Dyella jiangningensis* SBZ3- 12 (*Rhodanobacteraceae*). The annotation revealed 4,194 proteins coding genes, among which 36 Glycosyl Hydrolases potentially involved with the degradation of lignocellulosic biomass. Therefore, the microbial diversity and metabolic profile of the consortium was revealed. In addition a new genome belonging to the *Dyella* genus potentially involved with degradation of the lignocellulosic biomass were also uncovered. Overall, the results also corroborated the potential of high-throughput metagenomic sequencing and bioinformatic aproach for the elucidation of microbial communities with biotechnological potential, proving to be a valuable alternative for the bioenergy research.

Keywords: bioethanol, biomass, *Brazilian Microbiome Project* (BMP), *Dyella jiangningensis*, second generation ethanol, lignocellulose

LISTA DE FIGURAS

Página

Figura 1. Estrutura química da celulose e suas ligações de hidrogênio intramoleculares e intermoleculares (círculos em vermelho).....13

Figura 2. Estrutura química da hemicelulose.....16

Figura 3. Estrutura química da pectina constituída por ácidos galacturônicos unidos por ligações *α*-1,4.....17

Figura 4. Estrutura química da lignina.....18

Figura 5. Valores de riqueza de UTOs do consórcio bacteriano estimados pelo índice Chao-1 através do programa QIIME, demonstrando maior riqueza de UTOs presentes no consórcio na nona, décima primeira, décima terceira e décima nona semanas avaliadas.....33

Figura 6. Análise de Coordenadas Principais (PCoA) utilizando a distância métrica UniFrac Não-Ponderada gerada pelo programa QIIME.....34

Figura 7. Análise de Coordenadas Principais (PCoA) utilizando a distância métrica UniFrac Ponderada gerada pelo programa QIIME.....35

Figura 8. Diversidade dos principais gêneros presentes no consórcio bacteriano, obtida a partir de dados da matriz de UTOs, em cada semana analisada. Os gêneros *Burkholderia* e o Não Classificado, pertencente à família *Rhodanobacteraceae*, foram abundantemente representados. Além desses, foram identificados: gênero *Rhizobium*, *Cryocola* e um gênero assinalado como desconhecido (família *Isosphaeraceae*). Não classificados ou detectados em abundâncias pouco significativas foram agrupados e nomeados como Não Classificados e Outros Gêneros.....41

Figura 9. Análise comparativa entre a primeira (representada em azul) e terceira (representada em laranja) semanas de cultivo dos potenciais perfis metabólicos funcionais do consórcio bacteriano. Os dados foram analisados estatisticamente utilizando o software STAMP para diferenças entre proporções (%), utilizando G-teste + teste de Fisher para análise estatística, filtragem por q-value < 0,05 e correção de testes múltiplos usando o método FDR de *Benjamini-Hochber*. O banco de dados KEGG (nível 1) foi utilizado como referência.....43

Figura 10. Análise comparativa entre a primeira (representada em azul) e nona (representada em laranja) semanas de cultivo dos potenciais perfis metabólicos funcionais do consórcio bacteriano. Os dados foram analisados estatisticamente utilizando o software STAMP para diferenças entre proporções (%), utilizando G-teste + teste de Fisher para análise estatística, filtragem por q-value < 0,05 e correção de testes múltiplos usando o método FDR de *Benjamini-Hochber*. O banco de dados KEGG (nível 1) foi utilizado como referência.....44

Figura 11. Análise comparativa entre a primeira (representada em azul) e décima nona (representada em laranja) semanas de cultivo dos potenciais perfis metabólicos funcionais do consórcio bacteriano. Os dados foram analisados estatisticamente utilizando o software STAMP para diferenças entre proporções (%), utilizando G-teste + teste de Fisher para análise estatística, filtragem por q-value < 0,05 e correção de testes múltiplos usando o método FDR de *Benjamini-Hochber*. O banco de dados KEGG (nível 1) foi utilizado como referência.....45

Figura 12. Análise comparativa entre a primeira (representada em azul) e terceira (representada em laranja) semanas de cultivo dos potenciais perfis metabólicos funcionais do consórcio bacteriano. Os dados foram analisados estatisticamente utilizando o software STAMP para diferenças entre proporções (%), utilizando G-teste + teste de Fisher para análise estatística, filtragem por q-value < 0,05 e correção de testes múltiplos usando o método FDR de *Benjamini-Hochber*. O banco de dados KEGG (nível 3) foi utilizado como referência.....46

Figura 13. Análise comparativa entre a primeira (representada em azul) e décima primeira (representada em laranja) semanas de cultivo dos potenciais perfis metabólicos funcionais do consórcio bacteriano. Os dados foram analisados estatisticamente utilizando o software STAMP para diferenças entre proporções (%), utilizando G-teste + teste de Fisher para análise estatística, filtragem por q-value < 0,05 e correção de testes múltiplos usando o método FDR de *Benjamini-Hochber*. O banco de dados KEGG (nível 3) foi utilizado como referência.....47

Figura 14. Análise comparativa entre a primeira (representada em azul) e décima nona (representada em laranja) semanas de cultivo dos potenciais perfis metabólicos funcionais do consórcio bacteriano. Os dados foram analisados estatisticamente utilizando o software STAMP para diferenças entre proporções (%), utilizando G-teste + teste de Fisher para análise estatística, filtragem por q-value < 0,05 e correção de testes múltiplos usando o método FDR de *Benjamini-Hochber*. O banco de dados KEGG (nível 3) foi utilizado como referência.....48

Figura 15. Genoma da bactéria *Dyella jiangningensis* FCAV SCS01 gerado através da ferramenta DNA Plotter. A partir dos círculos presentes de dentro para fora da figura: GC Skew ((G-C) / (G+C)) (vermelho escuro), conteúdo GC % (cinza), Alien Hunter, Regiões Genômicas com Conteúdo Atípico (verde), regiões de Profago (amarelo), Ilhas Genômicas (rosa claro), Sequências de Inserção (vermelho), tRNA (cinza escuro), rRNA (rosa escuro), regiões codificadoras no sentido 3’ → 5’ (azul claro) e 5’ → 3’ (azul escuro).....53

LISTA DE TABELAS

Página

Tabela 1. Principais enzimas que participam da degradação de polissacarídeos presentes na parede celular vegetal. As enzimas estão separadas conforme o substrato em que são capazes de degradar, o número EC e a família de classificação: Glicosil Hidrolase (GH), Carboidrato Esterase (CE), Atividades Auxiliares (AA) e Polissacarídeo Liase (PL). Enzimas marcadas com asterisco (*) são degradadoras de quitina.....9

Tabela 2. Total de *reads* gerados após o sequenciamento do 16S rRNA do consórcio bacteriano na plataforma Ion PGM™, antes de realizar a filtragem seguindo o protocolo do BMP.....32

Tabela 3. Abundância relativa dos principais filos presentes no consórcio bacteriano em cada semana analisada. O filo *Proteobacteria* foi representado em todas as semanas. A classificação Outros Filos contêm filos que foram detectados ao longo das semanas, porém em abundância pouco significativa, sendo: *Bacteroidetes*, *Chlamydiae*, *Cyanobacteria*, *Firmicutes* e *Fusobacteria*. Filos Não Classificados não ultrapassaram a abundância de 0,03% ao longo das semanas.....37

Tabela 4. Abundância relativa das principais classes presentes no consórcio bacteriano em cada semana analisada. A classe representada em todas as semanas foi *Betaproteobacteria*, pertencentes ao filo *Proteobacteria*. Não Classificados não ultrapassaram a abundância de 0,03% ao longo das semanas e a classificação Outras Classes incluem as que foram detectadas presentes, porém em abundâncias pouco significativas, portanto foram agrupadas.....39

Tabela 5. Abundâncias relativas dos principais gêneros presentes no consórcio bacteriano em cada semana analisada. Os gêneros abundantemente representados foram *Burkholderia* e um gênero Não Classificado, pertencente à família *Rhodanobacteraceae*. Além desses, os gêneros identificados foram: *Rhizobium*, um gênero assinalado como desconhecido, pertencente à família *Isosphaeraceae* e o gênero *Cryocola*. Não classificados e os detectados em abundâncias pouco significativas (< 1%), anotados como Outros Gêneros foram agrupados (Não Classificados e Outros Gêneros).....42

Tabela 6. Características Genômicas da *Dyella jiangningensis* FCAV SCS01.....50

Tabela 7. Descrição dos subsistemas classificados através da plataforma RAST contendo o número de atribuições e porcentagens para cada subsistema.....52

Tabela 8. Enzimas com potencial degradador de biomassa encontradas no genoma de *Dyella jiangningensis* FCAV SCS01 através da classificação de subsistemas realizadas pela plataforma RAST. A tabela apresenta o nome, número EC da enzima, as famílias GHn e AA onde são encontradas e a quantidade de enzimas identificadas no genoma da bactéria.....55

1. INTRODUÇÃO

A busca por fontes de energias renováveis e sustentáveis como alternativa a dependência mundial de combustíveis fósseis é de grande importância no cenário econômico mundial. Essa busca é impulsionada por fatores relevantes como o crescimento populacional global e a preocupação ambiental com as emissões de gases na atmosfera que provocam o aquecimento global, como o caso da liberação de dióxido de carbono (CO_2). O processo de queima de combustíveis fósseis é um dos processos que mais contribuem para a liberação desse gás para a atmosfera (MUNIR; LEVIN, 2016; TYE et al., 2016).

Os combustíveis fósseis (petróleo, carvão mineral e gás natural) são reconhecidamente recursos finitos e não renováveis, que suprem atualmente 84% do consumo energético mundial, sendo que 75% são voltados para o setor de transportes (EIA, 2016; GU; HELD; FAIK, 2013). O setor de transportes rodoviários é responsável pelo aumento de mais de 80% de emissões de CO_2 para a atmosfera (IEA, 2015).

Segundo a uma publicação no “website” da ONU-BR (Nações Unidas no Brasil), em 2015 a quantidade mundial de veículos, considerando carros, vans, caminhões e ônibus, ultrapassou um bilhão e a estimativa é de que ultrapasse dois bilhões em 2050 (ONU, 2015).

Desse modo, uma alternativa para geração de energia são as fontes energéticas renováveis, incluindo a eólica, a solar, as hidroelétricas e a energia gerada a partir da biomassa (MATATIELE; GULUMIAN, 2016). A biomassa lignocelulósica, como resíduos agrícolas e florestais, é constituída de lignocelulose, o principal componente da parede celular das plantas, fornecendo estrutura, resistência e proteção às células vegetais (BAYER; SHOVAM; LAMED, 2013). A lignocelulose é constituída por polímeros de carboidratos, principalmente celulose (35 a 50%), hemicelulose (20 a 35%) e pectina, e por polímeros fenólicos, como a lignina (15 a 20%) (BRETHAUER; STUDER, 2015; HAGHIGHI MOOD et al., 2013). A degradação dessa estrutura lignocelulolítica na natureza ocorre através da ação de diversas enzimas produzidas por vários micro-organismos presentes no solo e

que atuam sinergicamente, utilizando os carbonos liberados como fonte de energia (BAYER; SHOVAM; LAMED, 2013; MHUANTONG et al., 2015).

A celulose é um homopolímero linear formado por unidades de glicose unidas por ligações glicosídicas *β*-1,4 em feixes de microfibrilas (BRETHAUER; STUDER, 2015; KANG et al., 2014). A estrutura química da hemicelulose é de um heteropolímero linear ou ramificado, constituído de diferentes pentoses (como xilose e arabinose), de hexoses (como manose, glicose e galactose) e por ácidos urônicos, que possuem uma estrutura amorfa (BRETHAUER; STUDER, 2015; KANG et al., 2014). Já a lignina é um composto fenólico que possui estrutura heterogênea, amorfa, formada por ligações entre unidades de fenilpropano, *p*-coumaril, álcool coniferil e álcool sinapil, conferindo rigidez e resistência à parede celular vegetal (BRETHAUER; STUDER, 2015; KANG et al., 2014).

Assim, os polímeros formadores da lignocelulose estão associados entre si formando uma estrutura resistente, principalmente devido à presença de lignina, que limita a ação de micro-organismos, tornando a biomassa lignocelulósica um material resistente à degradação biológica (TAHERZADEH; KARIMI, 2008). Devido a essa característica químico-física da lignocelulose são necessários pré-tratamentos físicos, químicos ou biológicos do material lignocelulolítico, permitindo que ocorra a quebra da parede celular vegetal (BRETHAUER; STUDER, 2015; CASTRO; PEREIRA JR, 2010). Assim, após o pré-tratamento, os carboidratos que compõem a biomassa lignocelulósica ficam mais acessíveis para a ação de enzimas hidrolíticas que degradam os oligo e polissacarídeos presentes na parede celular vegetal, liberando os monômeros de açúcares que serão fermentados para a produção de etanol (BRETHAUER; STUDER, 2015). Portanto, uma das formas promissoras para a desconstrução da biomassa é através do estudo, otimização e utilização de consórcios microbianos que produzem enzimas altamente específicas com a capacidade de quebrar a estrutura da lignocelulose (WONGWILAIWALIN et al., 2013).

Dessa maneira, um consórcio bacteriano proveniente de solo contendo bagaço de cana-de-açúcar em decomposição foi alvo desse estudo, onde os resultados das análises de bioinformática sugerem que esse consórcio apresenta capacidade de degradação da biomassa lignocelulósica.

Esse estudo também avaliou o perfil taxonômico e a estrutura desse consórcio ao longo de vinte semanas, onde os principais gêneros identificados foram *Burkholderia* e *Dyella*. Os resultados de predição funcional indicam que esse consórcio apresenta possível enriquecimento de genes relacionados com a função de transporte, incluindo transportadores do tipo ABC, na primeira semana de cultivo.

Através das análises realizadas a partir do sequenciamento do metagenoma total, foi possível recuperar o genoma de uma bactéria pertencente ao gênero *Dyella* (*Rhodanobacteraceae*). As sequências genômicas foram recuperadas e montadas. Esse genoma foi finalizado e circularizado através de análises comparativas. Em seguida, a anotação genômica foi realizada utilizando a plataforma RAST. Análises comparativas permitiram classificar esse micro-organismo como *Dyella jiangningensis*, denominada linhagem FCAV SCS01, apresentando pelo menos 36 enzimas da família das Glicosil Hidrolases, com possível potencial capacidade de degradar a biomassa lignocelulósica.

2. OBJETIVOS

Esse trabalho apresenta como objetivo principal estudar e caracterizar, através de técnicas de sequenciamento de alto rendimento e abordagens de bioinformática, um consórcio bacteriano isolado do solo contendo bagaço de cana-de-açúcar em decomposição, com capacidade de degradar a biomassa lignocelulósica.

Objetivos específicos:

i) Cultivar um consórcio bacteriano proveniente do solo de região de colheita mecanizada de cana-de-açúcar contendo bagaço;

ii) Identificar a comunidade bacteriana presente no consórcio através da análise das sequências geradas, revelar sua estrutura e predizer suas funções através de ferramentas de metagenômica e bioinformática, utilizando o protocolo do *Brazilian Microbiome Project* (BMP);

iii) Recuperar, identificar, montar e anotar um genoma presente no consórcio e avaliar seu potencial metabólico para enzimas que possuem capacidade de degradar a biomassa, através de ferramentas de metagenômica e bioinformática.

3. REVISÃO DE LITERATURA

3.1 Breve História sobre o Etanol no Brasil

O etanol é um biocombustível que pode ser obtido a partir de processos fermentativos de matérias-primas que são constituídas de açúcar, como a cana-de-açúcar, ou de amido, como o milho (TANAKA, S; LIN, Y, 2006). A cana-de-açúcar foi introduzida no Brasil no período colonial, tornando-se importante para o desenvolvimento econômico do país, passando a ser utilizada como a principal matéria-prima para a produção de açúcar e álcool (VIEIRA; LIMA; BRAGA, 2007).

Até 1973, quando ocorreu a crise mundial do petróleo, o Brasil importava aproximadamente 80% de combustível para o consumo energético, porém, apesar de não regulamentada, a adição de etanol à gasolina (etanol anidro) já era realizada na década de 30, em proporções variadas, conforme as práticas regionais do país (VIEIRA; LIMA; BRAGA, 2007). Assim, em 1975, foi criado o Programa Proálcool (Programa Nacional do Álcool) com o intuito de diminuir a dependência das importações de petróleo e para apoiar a diversificação na produção de açúcar pelas indústrias, como a transformação do excesso de açúcar em etanol anidro, que passou a ser regulamentada (KOHLHEPP, 2010; VIEIRA; LIMA; BRAGA, 2007).

A produção de cana-de-açúcar passou de aproximadamente 100 milhões de toneladas (anos que antecederam o Proálcool), para 220 milhões de toneladas nos anos de 1986 e 87, chegando a 520 milhões de toneladas em 2007, produzindo 31,3 Mt de açúcar e 22,5 Gl de etanol no mesmo ano (CGEE, 2009).

Já para as safras de 2015/16, segundo a Companhia Nacional de Abastecimento (Conab), o país produziu 665,6 milhões de toneladas de cana-de-açúcar, 33,49 milhões de toneladas de açúcar e 30,5 bilhões de litros em etanol (CONAB, 2016).

Atualmente, os interesses na produção e do consumo de biocombustíveis são voltados para além dos propostos inicialmente pelo Programa Proálcool, como a busca por fontes renováveis de energia devido às preocupações relacionadas ao meio ambiente, à aceitação e procura de biocombustíveis por outros países, à transformação do etanol em comódite, entre outros (IPEA, 2010).

Dentre esses interesses, a busca por fontes renováveis de energia trouxe uma nova visão com relação ao bagaço, subproduto constituído de lignocelulose, gerado após o processo de moagem da cana-de-açúcar e extração dos açúcares para a produção de etanol (etanol de primeira geração). O processo de moagem fornece 140 Kg de bagaço (peso seco) para cada tonelada de cana moída. O bagaço apresenta potencial para a produção de etanol de segunda geração, possibilitando o aumento da produção de biocombustíveis sem necessariamente expandir a área de plantio de cana-de-açúcar (CGEE, 2009; LEAL; LARSON; WILLIAMS, 2001).

A biomassa lignocelulósica é composta basicamente por celulose, hemicelulose e lignina. O processo de conversão desse material lignocelulolítico em etanol de segunda geração necessita de algumas etapas essenciais, como o pré-tratamento, a hidrólise dos componentes da lignocelulose para a liberação dos açúcares e a fermentação desses açúcares liberados (CORTEZ, 2010). Porém, ainda existem obstáculos nesses processos, impedindo a desconstrução eficiente da lignocelulose e o melhor aproveitamento dos açúcares liberados, proporcionando aumento do custo de produção (CORTEZ, 2010).

3.2 Biomassa e a Conversão em Biocombustíveis

A biomassa é uma fonte de energia renovável proveniente de matéria orgânica de origem animal ou vegetal, que pode ser convertida em bioenergia, como em biocombustíveis, energia elétrica ou térmica (AQUINO et al., 2014).

Os biocombustíveis podem ser classificados em duas vertentes: em primários, quando são utilizadas matérias-primas não processadas, como lenhas e restos de madeiras, para a produção de calor, e podem ser classificados em secundários, quando são provenientes de biomassa processada como, por exemplo, o bioetanol

produzido a partir da fermentação de açúcares provenientes de plantas (cana-de-açúcar, beterraba) e o biodiesel, produzido a partir de óleos vegetais de sementes (CHAKRABORTY et al., 2012). Dentre essas classificações, os biocombustíveis secundários podem ser categorizados conforme o tipo de matéria-prima e da tecnologia utilizados para a produção de biocombustíveis, sendo assim, a classificação dos biocombustíveis secundários também pode ser de primeira, segunda ou terceira geração (CHAKRABORTY et al., 2012).

Os principais biocombustíveis de primeira geração são: o etanol, que é produzido através da fermentação de açúcares presentes em fontes agrícolas, como a cana-de-açúcar e beterraba ou através de processos de hidrólise e fermentação de amido, derivado de milho ou trigo (HO; NGO; GUO, 2014; NAIK et al., 2010); o biodiesel, obtido através do processo de transesterificação de óleos extraídos de sementes, como a soja, ou de palma (NAIK et al., 2010; UNCTAD, 2016); e o biogás, gerado a partir de processos anaeróbicos de dejetos orgânicos (NAIK et al., 2010).

Já os biocombustíveis de segunda geração são produzidos a partir de materiais lignocelulolíticos, como os resíduos agrícolas (palha, bagaço), florestais e resíduos urbanos (EMBRAPA, 2013). Segundo Mosier et al. (2005), o processo de conversão do material lignocelulolítico em bioetanol de segunda geração ocorre através de alguns passos importantes: a) Pré-tratamento, para que a estrutura da biomassa lignocelulósica seja alterada, melhorando, assim, o processo de hidrólise dos polímeros de carboidratos e o rendimento na obtenção de açúcares; b) Hidrólise, que quebra os polímeros de carboidratos, transformando-os em monômeros; c) Fermentação dos açúcares monoméricos; d) Separação ou purificação do produto final (MOSIER et al., 2005).

Por último, os denominados biocombustíveis de terceira geração são produzidos a partir de biomassa de algas, através da extração de lipídios, que são produzidos e armazenados por determinadas espécies de algas ou através da fermentação de seus carboidratos, convertendo-os em biodiesel e etanol, respectivamente (COSTA; DE MORAIS, 2011; NAIK et al., 2010).

3.3 Processo de Produção de Bioetanol de Segunda Geração

A produção de bioetanol de segunda geração requer a degradação da biomassa lignocelulósica para que os polissacarídeos que compõem a celulose e a hemicelulose se tornem acessíveis para a hidrólise enzimática (RUBIN, 2008). Assim, para que os componentes da parede celular vegetal possam ser hidrolisados em monômeros de açúcares, a biomassa necessita passar por um pré-tratamento, ocorrendo, assim, a remoção da lignina e facilitando o acesso à hemicelulose e celulose. Isso aumenta a eficácia da hidrólise enzimática dos polissacarídeos que a compõe para, em seguida, passarem pelo processo de fermentação (RAVINDRAN; JAISWAL, 2016). Existe uma variedade de pré-tratamentos, incluindo métodos físicos, químicos (ácido ou alcalino), a combinação de ambos ou pré-tratamento biológico, que utiliza de micro-organismos capazes de degradar a lignina através de suas enzimas potenciais, como por exemplo, lacases e peroxidases (SARITHA; ARORA; LATA, 2012).

3.4 Enzimas Degradadoras da Biomassa

A estrutura heterogênea e rígida da biomassa lignocelulósica torna desafiante o processo em escala industrial da sua conversão ou degradação em carboidratos simples (SWEENEY; XU, 2012). Já na natureza, a desconstrução da lignocelulose e utilização de seus produtos ocorrem através de enzimas altamente específicas, que agem sinergicamente, provindas de organismos lignocelulolíticos, como as bactérias, fungos e arquéias (SWEENEY; XU, 2012). Assim, a conversão da biomassa através de enzimas específicas tem sido de grande interesse na área biotecnológica e industrial para a produção de combustíveis e outros produtos que necessitam dos açúcares simples liberados, e que são utilizados em processos fermentativos (SWEENEY; XU, 2012). Além disso, também há interesse em relação a esse processo de conversão da biomassa utilizando enzimas por ser de baixo consumo energético e químico, e por não proporcionar altos índices de poluição ambiental (SWEENEY; XU, 2012).

Apesar de apresentar vantagens, a conversão da biomassa lignocelulósica em biocombustíveis e outros produtos químicos através da hidrólise enzimática possui alguns impasses, como a falta de conhecimento sobre a ação e função das enzimas, e de como se relacionam com as estruturas do substrato (YANG et al., 2011). Além desse impasse, a estrutura rígida da parede celular vegetal impede o acesso das enzimas, reduzindo o rendimento da hidrólise e aumentando os custos desse processo (PERRONE et al., 2016; YANG et al., 2011). Portanto, estudos relacionados a esse tema são de suma importância para que ocorram avanços significativos na área para a obtenção de altos rendimentos de etanol celulósico, com custos reduzidos, tornando-o competitivo no mercado (PERRONE et al., 2016; YANG et al., 2011).

As enzimas que possuem capacidade de degradar os polissacarídeos da parede celular vegetal, catalisando a hidrólise de oligo e polissacarídeos, foram agrupadas primariamente em uma família denominada Glicosil Hidrolases (GHn, onde “n” representa o número do grupo pertencente à família GH) segundo a determinação da *Enzyme Commission* (EC) (BAYER; SHOVAM; LAMED, 2013). Além disso, as enzimas são nomeadas e agrupadas em determinada família GH conforme as similaridades de suas sequências de aminoácidos e suas características estruturais (*Carbohydrate Actives Enzymes* – http://www.cazy.org/) (LOMBARD et al., 2014). Também são consideradas as interações com o substrato e reações que catalisam, por exemplo: se a clivagem de polissacarídeos acontece no meio da cadeia são nomeadas como “endo”, enquanto se a clivagem ocorre nas regiões terminais da cadeia são nomeadas como “exo” e agrupadas em diferentes categorias, como no caso da enzima celulase (endoglucanase), EC 3.2.1.4 e a celobiohidrolase (exoglucanase) EC 3.2.1.91 (BAYER; SHOHAM; LAMED, 2013). Informações completas sobre essas enzimas podem ser consultadas no banco de dados CAZy (*Carbohydrate Actives Enzymes* – http://www.cazy.org/) (LOMBARD et al., 2014).

As principais enzimas envolvidas na degradação de polissacarídeos da parede celular vegetal são apresentadas na Tabela 1.

Tabela 1. Principais enzimas que participam da degradação de polissacarídeos presentes na parede celular vegetal. As enzimas estão separadas conforme o substrato em que são capazes de degradar, o número EC e a família de classificação: Glicosil Hidrolase (GH), Carboidrato Esterase (CE), Atividades Auxiliares (AA) e Polissacarídeo Liase (PL). Enzimas marcadas com asterisco (*) são degradadoras de quitina.

Enzimas Degradadoras de (Hemi)Celulose	EC	**Família GH**
α-amilase	3.2.1.1	13, 14, 57, 119
α-galactosidase	3.2.1.22	4, 27, 31, 36, 57, 97, 110
α-glicosidase	3.2.1.20	4, 13, 31, 63, 97, 122
α-glucuronidase	3.2.1.139	67
α-L-arabinofuranosidase	3.2.1.55	51, 54, 62
α-L-fucosidase	3.2.1.51	29, 59
α-N-acetilgalactosaminidase	3.2.1.49	36
α-N-acetilglicosaminidase	3.2.1.50	89
α-1,2-manosidase	3.2.1.24	31, 89, 92
α-xilosidase	3.2.1.177	31
β-galactosidase	3.2.1.23	1, 2, 3, 35, 42, 50
β-glicosidase	3.2.1.21	1, 3
β-glucuronidase	3.2.1.31	1, 2
β-manosidase	3.2.1.25	1, 2, 5
β-xilosidase	3.2.1.37	3, 30, 39, 43, 52, 54
Arabinase	3.2.1.99	43
Celobiohidrolase	3.2.1.91	6, 9
Endo-1,4-*β*-galactanase	3.2.1.89	53
Endo-1,6-*β*-galactanase	3.2.1.164	5
Endo-1,3(4)-*β*-glucanase	3.2.1.6	16
Endo-1,3-*β*-xilanase	3.2.1.32	10, 11, 26
Endo-1,4-*β*-xilanase	3.2.1.8	5, 8, 10, 11, 43
Endoglucanase	3.2.1.4	5, 6, 7, 9, 12, 44, 45, 51, 74
Exo-1,4-*β*-glucanase	3.2.1.74	9
Glicano 1,3-*β*-glicosidase	3.2.1.58	3, 5, 17, 55
Glicano 1,4-*β*-glicosidase	3.2.1.74	3

Tabela 1. (Continuação)

Enzimas Degradadoras de (Hemi)Celulose	EC	**Família GH**
Galactano 1,3-*β*-galactosidase	3.2.1.145	43
Glicoamilase	3.2.1.3	15, 97
Manana endo-1,4-*β*-D-manosidase	3.2.1.78	5, 26, 113
Maltodextrina glicosidase	3.2.1.20	4, 13
Malto-oligosiltrealose trealohidrolase	3.2.1.141	13
Trealase	3.2.1.28	3, 15, 37, 65
Xilana *α*-1,2-glucuronidase	3.2.1.131	67, 115
Xilanase	3.2.1.8	5, 8, 10, 11, 43
β-hexosaminidase*	3.2.1.52	3, 5, 18, 20, 84, 116
Quitinase*	3.2.1.14	18, 19
Enzimas Degradadoras de (Hemi)Celulose	EC	**Família CE**
Acetil xilana esterase	3.1.1.72	1, 2, 3, 4, 5
Feruloil esterase	3.1.1.73	1
Pectina acetil esterase	3.1.1.72	12
Pectina metilesterase	3.1.1.11	8
Enzimas Degradadoras de Lignina	EC	**Família AA**
Catalase	1.11.1.6	2
Lignina-peroxidase	1.11.1.14	2
Manganês-peroxidase	1.11.1.13	2
Peroxidase	1.11.1.7	2
Enzimas Degradadoras de Pectina	EC	**Família PL**
Pectato liase	4.2.2.2	1, 3, 9, 10
Rhamnogalacturonana liase	4.2.2.-	11

Fonte: adaptação de BAYER; SHOHAM; LAMED, 2013, *Lignocellulose-Decomposing Bacteria and Their Enzyme Systems*; *Carbohydrate Actives Enzymes* (http://www.cazy.org/) (LOMBARD et al., 2014).

3.5 Parede Celular Vegetal e Suas Enzimas Degradadoras

A parede celular vegetal é formada por uma estrutura química complexa composta por polissacarídeos, principalmente a celulose e hemicelulose, e pectinas, por lignina e glicoproteínas, que estão intimamente associados, formando uma estrutura rígida, oferecendo, assim, sustentação e proteção às células vegetais de agressões que podem ocorrer no ambiente, como ataques mecânicos, químicos ou microbianos (BAYER; SHOVAM; LAMED, 2013; SANT'ANNA et al., 2013). Além disso, os polissacarídeos presentes nas células vegetais são fontes de carbono para diversos micro-organismos que são capazes de degradar a parede celular vegetal, convertendo os polissacarídeos ali presentes em monossacarídeos e outros produtos da degradação. Esse processo ocorre através da participação em conjunto de micro-organismos satélites (BAYER; SHOVAM; LAMED, 2013; SANT'ANNA et al., 2013).

A celulose é formada por resíduos de glicose que são unidos por ligações glicosídicas, gerando longas cadeias lineares. Essas cadeias lineares se associam paralelamente entre si formando as microfibrilas de celulose que, por sua vez, estão inseridas dentro da parede celular em uma matriz que pode conter proteínas e os polissacarídeos hemicelulose e pectinas (COOPER; HAUSMAN, 2016). As hemiceluloses se interligam a superfície das microfibrilas de celulose através de ligações de hidrogênio, formando uma rede de moléculas fibrosas e robustas, responsáveis pela resistência mecânica da parede celular vegetal (COOPER; HAUSMAN, 2016). Já a pectina, um polissacarídeo ramificado, é formada por resíduos de ácido galacturônico e possui a capacidade de ligar-se a íons carregados positivamente (pois o ácido galacturônico é carregado negativamente) atraindo, assim, moléculas de água que formam uma matriz gelatinosa onde as microfibrilas de celulose estão imersas (COOPER; HAUSMAN, 2016).

Além disso, a parede celular vegetal é dividida em primária e secundária, sendo suas características: a) A parede celular vegetal primária é denominada quando as células vegetais estão em crescimento e são relativamente finas e flexíveis para que as células possam expandir de tamanho. São compostas por quantidades equivalentes de celulose, hemicelulose e pectina e possuem

microfibrilas de celulose arranjadas de forma aleatória; b) Já a parede celular secundária é formada entre a parede celular primária e a membrana plasmática, quando as células vegetais deixam de crescer, apresentando maior espessura e rigidez do que a parede celular primária. Exibe de 50 a 80% de celulose em sua estrutura, é desprovida de pectina e é organizada em três camadas que diferem entre si pela orientação das microfibrilas de celulose formando, assim, estruturas laminares que proporcionam maior rigidez da parede celular vegetal (COOPER; HAUSMAN, 2016). Além disso, a parede celular secundária conta com a presença de lignina, um polímero composto por resíduos fenólicos que reforçam a parede celular (COOPER; HAUSMAN, 2016).

3.5.1 Estrutura da Celulose e Enzimas Degradadoras da Celulose

A parede celular vegetal é composta, em sua maioria, por celulose, um polímero linear formado por monômeros de D-glicose unidos por ligações glicosídicas do tipo *β*-1,4 e ligações de hidrogênio intra e intermoleculares (ligações entre unidades de glicoses da mesma molécula e ligações entre unidades de glicoses de moléculas adjacentes, respectivamente) (D'ALMEIDA et al., 2013). As interações intramoleculares são responsáveis pela rigidez do polímero, enquanto as intermoleculares pela formação das microfibrilas através do alinhamento das moléculas de celulose (D'ALMEIDA et al., 2013). Quimicamente, a celulose é composta por unidades de glicose, formadas pela unidade de repetição do dissacarídeo celobiose, onde cada resíduo de glicose se mantém em uma posição de 180° com relação ao resíduo adjacente, formando uma cadeia que pode conter de cem a 10 mil unidades de glicose que se dispõem paralelamente, formando microfibrilas de celulose (BAYER; SHOHAM; LAMED, 2013).

Esse arranjo de microfibrilas da celulose geram regiões altamente ordenadas, conhecidas como cristalinas, que são resistentes ao ataque químico e à degradação, devido à sua estrutura compacta que também impede a entrada de pequenas moléculas e o acesso de enzimas (D'ALMEIDA et al., 2013). Também formam regiões desordenadas dispostas entre as regiões cristalinas, onde os polímeros não se apresentam tão próximos uns dos outros, formando uma estrutura menos

compacta denominada de regiões amorfas (D'ALMEIDA et al., 2013; LIMA, 2014). Além disso, há regiões irregulares ou que apresentam micro poros entre as fibras de celulose, proporcionando aumento na superfície de contato e a entrada de moléculas variadas, incluindo enzimas com capacidade de degradar a celulose (D'ALMEIDA et al., 2013; LIMA, 2014).

Como visto anteriormente, as moléculas de celulose estão ligadas aos polímeros da hemicelulose e lignina e essa união torna o material resistente à tensão, à hidrólise, degradação microbiana e insolúvel a vários solventes (DING; HIMMEL, 2006). A Figura 1 apresenta a estrutura da celulose e suas ligações.

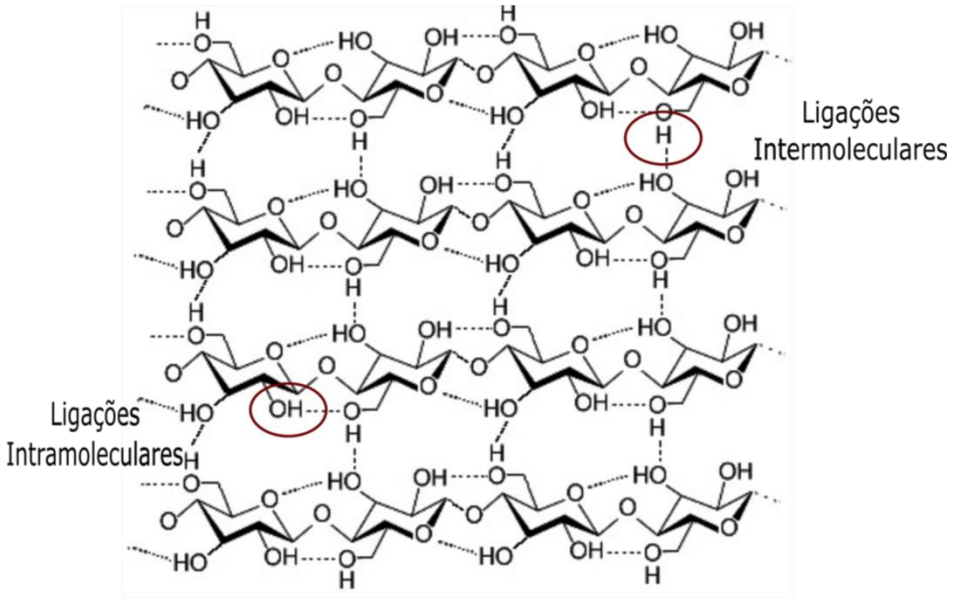

Figura 1. Estrutura química da celulose e suas ligações de hidrogênio intramoleculares e intermoleculares (círculos em vermelho).

Na natureza, a completa degradação da celulose ocorre através de diversas enzimas produzidas por micro-organismos celulolíticos e não-celulolíticos que podem atuar em sinergia (PÉREZ et al., 2002). Assim, as celulases (EC 3.2.1.x) são enzimas capazes de hidrolisar as ligações glicosídicas $\beta(1\rightarrow4)$ da celulose e são divididas em três grupos principais, endoglucanases, exoglucanases e β-glicosidades, conforme a região em que hidrolisam as cadeias de celulose (CASTRO; PEREIRA JR, 2010; PÉREZ et al., 2002).

As endoglucanases ou endo-1,4-*β*-glucanases (EC 3.2.1.4), hidrolisam as ligações internas de forma aleatória na estrutura amorfa da celulose, liberando oligossacarídeos que, por fim, liberam novas cadeias terminais onde as enzimas celobio-hidrolases atuam (CASTRO; PEREIRA JR, 2010). Já o grupo das celobio-hidrolases (EC 3.2.1.91) faz parte das exoglucanases ou exo-1,4-*β*-glucanases, que possuem capacidade de hidrolisar tanto regiões amorfas quanto cristalinas da celulose (CASTRO; PEREIRA JR, 2010; PÉREZ et al., 2002). Porém, esse é o único grupo capaz de hidrolisar as regiões cristalinas de forma eficiente, onde seu produto de hidrólise é a celobiose (dissacarídeo de glicose) e essas enzimas podem ser divididas em celobio-hidrolases do tipo I e II, que hidrolisam terminais redutores e não-redutores, respectivamente (PÉREZ et al., 2002; SWEENEY; XU, 2012). Por sua vez, as enzimas *β*-glicosidases (EC 3.2.1.21) hidrolisam as celobioses liberando moléculas de glicose, assim, a ação sinérgica dessas enzimas degrada a celulose de modo eficiente (CASTRO; PEREIRA JR, 2010; PÉREZ et al., 2002; SWEENEY; XU, 2012).

Devido às características da celulose, as enzimas celulases podem apresentar Módulos de Ligação a Carboidratos (CBM, do inglês, "*Carbohydrate-binding Modules"*), que se ligam a superfície do substrato e aumentam a proximidade entre as enzimas com a celulose, possibilitando uma hidrólise mais eficiente (LYND et al., 2002). Portanto, os CBM são importantes para que as exoglucanases iniciem sua ação (LYND et al., 2002).

As enzimas celulases pertencem às principais famílias Glicosil Hidrolases: GH5, GH6, GH7 (celulases encontradas em fungos), GH8, GH9, GH12, GH44, GH45, GH48, GH74 e GH124 (BAYER; SHOHAM; LAMED, 2013).

3.5.2 Estrutura da Hemicelulose e Enzimas Degradadoras da Hemicelulose

A hemicelulose é um polímero heterogêneo, altamente ramificado e possui uma estrutura aleatória e amorfa, composta por um conjunto de polissacarídeos que estão fortemente ligados à celulose, sendo pentoses como D-xilose e L-arabinose, hexoses como D-galactose, D-manose e D-glicose e ácidos urônicos. Além disso, outros açúcares podem estar presentes em pequenas quantidades como L-ramnose

e L-fucose (GÍRIO et al., 2010). As xilanas são compostas por unidades de D-xilose que estão unidas por ligações glicosídicas *β*-1,4 e constituem os principais componentes da hemicelulose (GÍRIO et al., 2010).

A hemicelulose apresenta diferenças estruturais entre espécies e tipos celulares e é o segundo maior constituinte, representando de 20-35% da parede celular vegetal, onde seu principal papel é o de reforçar a estrutura da parede e dar flexibilidade através de suas ramificações que se ligam às microfibrilas de celulose (HAGHIGHI MOOD et al., 2013; LE FLOCH; JOURDES; TEISSEDRE, 2015). A classificação da hemicelulose é derivada de um ou dois açúcares que compõem sua estrutura principal, por exemplo, xilana é composta por unidades de xilopiranose unidas por ligações *β*-1,4, já galactoglucomanana é formada por ligações lineares entre unidade *β*-1,4-D-glicopiranose e *β*-D-manopiranose, ligados a resíduos de galactose (BAYER; SHOVAM; LAMED, 2013). A xilana está presente na parede celular vegetal, formando ligações cruzadas entre a celulose e a lignina e, ao contrário da região cristalina da celulose que apresenta uma estrutura ordenada e compacta, a hemicelulose apresenta uma consistência semelhante a um gel, gerando uma matriz amorfa, onde as microfibrilas de celulose estão inseridas (BAYER; SHOHAM; LAMED, 2013). Assim, a hemicelulose é facilmente hidrolisada, já que não possui regiões cristalinas (MENDES, 2013).

Associada à celulose, a hemicelulose é um composto heterogêneo, onde sua composição pode variar entre as diferentes espécies de plantas (CASTRO; PEREIRA JR, 2010). Apesar da sua estrutura, a hemicelulose é mais facilmente degradada quando comparada à celulose (BAYER; SHOVAM; LAMED, 2013). Para isso, enzimas específicas denominadas hemicelulases (que são classificadas conforme suas atividades) são responsáveis pela degradação (BAYER; SHOVAM; LAMED, 2013). Essas enzimas são classificadas em endoxilanases ou endo-*β*-xilanases (EC 3.2.1.8), que hidrolisam as ligações glicosídicas internas da cadeia da xilana, liberando xilobioses e que são degradadas pelas enzimas *β*-xilosidases (EC 3.2.1.37), liberando monossacarídeos de xilose (ALVES-PRADO et al., 2010; POLIZELI et al., 2005). Com relação ao agrupamento nas famílias Glicosil Hidrolases, as enzimas xilanases são reunidas nas famílias GH10, GH11 e uma pequena porção de enzimas que apresentam essa atividade são encontradas nas

GH5 e GH8 (BAYER; SHOVAM; LAMED, 2013). Já as *β*-xilosidases são agrupadas nas famílias GH3, GH30, GH39, GH43, GH52, GH54, GH116 e GH120 (BAYER; SHOHAM; LAMED, 2013).

As enzimas *α*-L-arabinofuranosidases (EC 3.2.1.55) atuam nas cadeias laterais da xilana removendo as ramificações de L-arabinose, enquanto as enzimas *α*-glucoronidases (EC 3.2.1.131) hidrolisam as ligações glicosídicas *α*(1→2) entre ácidos glucurônicos ou de ésteres metílicos (POLIZELI et al., 2005). Por fim, as acetilxilana esterases (EC 3.1.1.72) e ácido ferúlico esterases (EC 3.1.1.73), removem grupos acetil e ácidos fenólicos, respectivamente (SANTOS, 2014). Essas enzimas atuam sinergicamente entre si ou com as enzimas celulases para que o processo de degradação da parede celular vegetal ocorra eficientemente (SANTOS, 2014). As enzimas *α*-L-arabinofuranosidases são encontradas nas famílias GH43, GH51, GH54, GH62 e GH127 (BAYER; SHOHAM; LAMED, 2013).

A Figura 2 apresenta a estrutura química da hemicelulose.

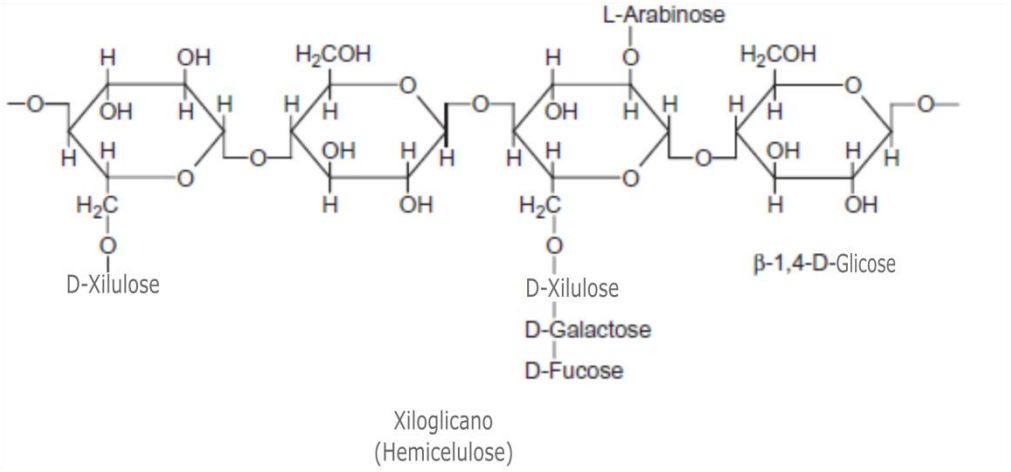

Figura 2. Estrutura química da hemicelulose.

3.5.3 Estrutura da Pectina e Enzimas Degradadoras da Pectina

A pectina é um heteropolissacarídeo estrutural constituída por ácidos galacturônicos unidos por ligações do tipo *α*-1,4, presente na parede celular vegetal primária de plantas terrestres, com a função de regular o desenvolvimento e as respostas de defesa das plantas, apresentando variações de quantidade,

distribuição, composição química e de sua estrutura conforme os processos de desenvolvimento vegetal como crescimento, maturação e envelhecimento (BAYER; SHOHAM; LAMED, 2013).

A estrutura química da pectina é apresentada na Figura 3.

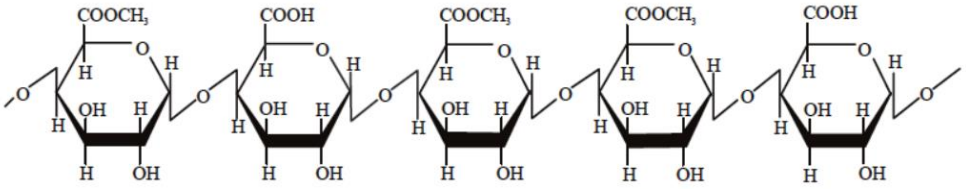

Figura 3. Estrutura química da pectina constituída por ácidos galacturônicos unidos por ligações *α*-1,4.

As enzimas polissacarídeo-liases (PLs) são proteínas modulares que clivam cadeias de polissacarídeos compostas por ácido urônico, gerando um resíduo de ácido hexaurônico (um novo terminal redutor), complementando, assim, as atividades das outras enzimas que são capazes de degradar os componentes da parede celular vegetal para que esse processo ocorra de modo mais eficiente (BAYER; SHOHAM; LAMED, 2013).

3.5.4 Estrutura da Lignina e Enzimas Degradadoras da Lignina

A lignina traz sustentação para as plantas, impermeabilidade e proteção contra ataques microbianos e também é encontrada de forma abundante na natureza (PÉREZ et al., 2002) (Figura 4). A lignina é um polímero heterogêneo, altamente ramificado, aromático e constitui de 10-20% da biomassa seca (TAHERZADEH; KARIMI, 2008). Esta é composta por resíduos fenólicos como álcool *trans-p*-coumaril, *trans-p*-coniferil e *trans-p*-sinapil unidos por diferentes ligações que formam uma estrutura química difícil de ser biodegradada, considerando-se que a dificuldade de degradação é diretamente proporcional à quantidade de lignina presente na planta (TAHERZADEH; KARIMI, 2008). Apesar da molécula de lignina não estar relacionada com moléculas simples de açúcares que são utilizadas em rotas fermentativas para a produção de etanol, sua estrutura rígida

é resistente a degradações químicas e enzimáticas, trazendo um desafio, já que a remoção da lignina se faz necessária para a produção de etanol a partir do material lignocelulósico (TAHERZADEH; KARIMI, 2008).

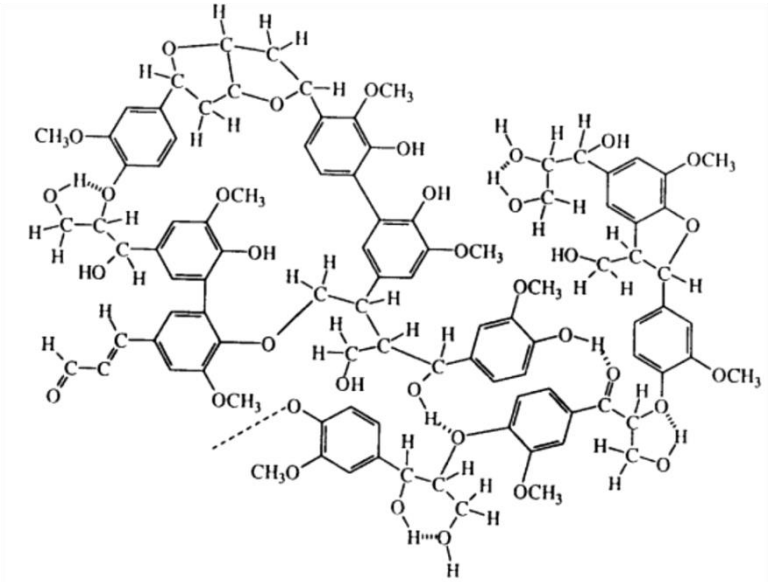

Figura 4. Estrutura química da lignina.

As enzimas lignina-peroxidase, manganês-peroxidase e lacase são capazes de degradar a estrutura heterogênea, ramificada e aromática da lignina (WANG et al., 2016). Sabe-se que essas enzimas foram reportadas em fungos de podridão branca (*Trametes hirsuta*) e parda (*Gloeophyllum trabeum*), que degradam a lignina, deixando a madeira com uma coloração clara ou mais escura, respectivamente, e em bactérias incluindo *Streptomyces viridosporus* T7A, *Nocardia autotrophica*, *Sphingobium* sp. SYK-6, *Pseudomonas putida* mt-2, *Rhodococcus* sp., *Burkholderia cepacia*, *Microbacterium* sp. e *Citrobacter* sp., porém o conhecimento sobre os mecanismos utilizados pelas bactérias ainda é limitado (WANG et al., 2016).

3.6 Consórcios Microbianos

Consórcio microbiano é um conjunto de micro-organismos diversos, que formam comunidades complexas capazes de realizarem tarefas através de suas interações, as quais não seriam possíveis se fossem feitas por um único organismo

(ZUROFF; XIQUES; CURTIS, 2013). São responsáveis pela manutenção dos ciclos biogeoquímicos do planeta através de diversas interações. Por exemplo, entre os organismos eucariontes e o ambiente em que se encontram (interações de cooperação), sendo chamadas de interações positivas das quais se enquadram mutualismo, simbiose, comensalismo ou de não cooperação, que são interações negativas como predação, parasitismo e amensalismo (GHOSH; CHOWDHURY; BHATTACHARYA, 2016; TAFFS et al., 2009). Assim, a compreensão dessas interações é de grande importância, já que traz entendimento das funções metabólicas de cada micro-organismo e de como atuam em conjunto (GHOSH; CHOWDHURY; BHATTACHARYA, 2016; TAFFS et al., 2009). Além disso, ajudam a elucidar como as comunidades microbianas respondem as alterações climáticas e ações antropogênicas, servindo de base para melhorar as aplicações de interesse biotecnológico, por exemplo, utilizando consórcios microbianos para a produção de biomateriais e biocombustíveis a partir de fontes renováveis (GHOSH; CHOWDHURY; BHATTACHARYA, 2016; TAFFS et al., 2009).

Apesar da notável eficiência de degradação da lignocelulose realizada em conjunto pelos micro-organismos na natureza e da promessa na utilização de consórcios microbianos sintéticos nas indústrias para a melhoria dos processos, cepas únicas de micro-organismos ainda são usadas com frequência para esse fim (CORTES-TOLALPA et al., 2016).

3.7 Metagenômica e Sequenciamento de Nova Geração (NGS, do inglês “*Next Generation Sequencing”*)

Estima-se que menos de 1% dos micro-organismos presentes em diversos hábitats pode ser cultivado em condições laboratoriais, apesar dos conhecimentos atuais (KUMAR et al., 2015; TURNBAUGH; GORDON, 2008).

Nesse sentido, a metagenômica revolucionou os estudos de microbiologia, pois é uma abordagem que permite analisar os genomas tanto de micro-organismos cultiváveis, como os não-cultiváveis em laboratório, possibilitando o isolamento de DNA a partir de amostras obtidas de diversos ambientes (NESME et al., 2016; SIMON; DANIEL, 2011). Assim, por meio de análise do DNA, que representa o

genoma coletivo das espécies contidas em determinado hábitat, e a identificação dos micro-organismos que estão presentes nestas amostras, torna-se possível à busca por respostas relacionadas à interação desses micro-organismos entre si (sinérgicas ou antagônicas) e com seu ambiente (NESME et al., 2016). Por exemplo, investigar se as mudanças climáticas que ocorrem no ambiente em que estão inseridos alteram a abundância dos micro-organismos ou procurar entender o papel que aquela comunidade microbiana desenvolve em seu hábitat, entre outras questões fundamentais para o completo entendimento da microbiota existente (NESME et al., 2016; SIMON; DANIEL, 2011).

A primeira proposta de estudo metagenômico ocorreu em 1985, por Pace et al. (1985), ao isolar DNA diretamente do ambiente e utilizar vetores para clonagem do DNA, seguindo com as análises das sequências 5S e 16S rRNA para a descrição da diversidade dos micro-organismos presentes no ambiente em estudo (MANDAL, et al., 2015; SIMON; DANIEL, 2011; VAKHLU; SUDAN; JOHRI, 2008). Porém o termo “metagenômica” passou a ser utilizado somente em 1998, após a publicação de Handelsman et al. (1998), ao sugeri-lo para um conjunto de genomas encontrados em determinado ambiente (HANDELSMAN et al., 1998; MANDAL, et al., 2015).

A metagenômica foi expandida graças às tecnologias de sequenciamento de alto-desempenho ou de nova geração (NGS – “*Next Generation Sequencing”*) que possibilitaram o sequenciamento de micro-organismos sem a necessidade de cultivo em laboratório, gerando grandes quantidades de dados que são analisados com o apoio de ferramentas da bioinformática (KUMAR et al., 2015; SIMON; DANIEL, 2011).

Atualmente, duas abordagens principais de estudos metagenômicos são utilizadas:

1) A primeira, relacionada ao estudo do gene 16S rRNA, que é utilizado como um marcador molecular para caracterizar uma comunidade bacteriana em seus diversos nichos ecológicos (MIZRAHI-MAN; DAVENPORT; GILAD, 2013). Neste estudo, a diversidade dos micro-organismos é inferida através da amplificação do gene 16S rRNA pela técnica de PCR (Reação em Cadeia da Polimerase), sequenciamento e comparação com um banco de dados, possibilitando traçar um

perfil dessa comunidade (OULAS et al., 2015; VARUZZA, 2013). Isso é possível devido às características do 16S rRNA, como por exemplo, presença de regiões conservadas, que permitem o desenho de *primers* para a amplificação de pelo menos 9 regiões hipervariáveis (V1-V9). Através do estudo comparativo de similaridade destas regiões é possível distinguir e classificar os micro-organismos presentes em um determinado meio (MIZRAHI-MAN; DAVENPORT; GILAD, 2013).

2) A segunda abordagem é realizada através da técnica de *shotgun* ou sequenciamento total, onde o DNA total de uma comunidade microbiana é extraído e sequenciado. O resultado desse sequenciamento é analisado por abordagens de bioinformática para identificação da diversidade microbiana e inferência funcional da comunidade microbiana presente na amostra ou no ambiente analisados (OULAS et al., 2015; VARUZZA, 2013).

3.8 *Brazilian Microbiome Project* (BMP)

Conhecer a diversidade microbiana, as interações e os aspectos funcionais em um ecossistema são de grande importância para que se possa promover a manutenção desse ecossistema e o uso sustentável da terra (PYLRO et al., 2014a). Além disso, as descobertas de micro-organismos com potencial biotecnológico pode levar a diversas aplicações, como por exemplo, processos biossintéticos (produção de novos medicamentos, síntese de biocombustíveis e etc.) ou em biodegradação (degradação de poluentes), assim como promover o desenvolvimento científico e tecnológico gerados pelo avanço do conhecimento (PYLRO et al., 2014a). Portanto, caracterizar genes desconhecidos e suas vias metabólicas que codificam biocatalisadores revela-se como uma importante abordagem para o desenvolvimento biotecnológico e para o uso racional da biodiversidade (PYLRO et al., 2014a).

Atualmente, a análise comparativa da diversidade de comunidades microbianas em um ecossistema e da abundância de importantes genes funcionais dessas comunidades, tem avançado devido à evolução das técnicas moleculares e do sequenciamento através de plataformas de nova geração (NGS) (PYLRO et al., 2014a). Assim, considerando-se que ainda há muito a se conhecer sobre a

diversidade microbiana do Brasil e a importância dos micro-organismos na manutenção dos ecossistemas, e utilizando as abordagens moleculares atuais juntamente com as análises convencionais e ferramentas da bioinformática, o *Brazilian Microbiome Project* (BMP - http://www.brmicrobiome.org/) propõe o desenvolvimento de um banco de dados contendo todas as informações relacionadas a estudos metagenômicos e do microbioma nacional (PYLRO et al., 2014a). Possui também como objetivos principais coordenar e padronizar as abordagens utilizadas na realização desses estudos metagenômicos (PYLRO et al., 2014a). Assim, com o potencial de aplicações biotecnológicas desses micro-organismos, o entendimento dos microbiomas e a disponibilidade de um banco de dados que reúna todas as informações dos consórcios microbianos estudados, podem trazer melhor compreensão de como ocorrem às interações entre os micro-organismos de determinados nichos e de como são afetados por diferentes tipos de estresse (alterações de temperatura, pH, nutrientes), promovendo o fácil acesso dessas informações à comunidade científica (PYLRO et al., 2014a).

4. MATERIAL E MÉTODOS

4.1 Coleta do Material, Isolamento e Adaptação do Consórcio Bacteriano em Meio de Cultura Contendo Bagaço de Cana-de-Açúcar

Para o desenvolvimento desse trabalho foi utilizado um consórcio bacteriano coletado do solo de área de cultivo comercial de cana-de-açúcar, com colheita mecanizada, contendo palhada em decomposição, nas áreas da Usina São Martinho, localizada no município de Pradópolis, SP, com as seguintes coordenadas: LAT: 21º19´23.51220” S / LOG: 48º09´12.24864” O / ALT: 534,100m (DE ALMEIDA, 2011). As amostras foram coletadas aleatoriamente do solo em profundidade de 0 a 20 cm, em zigue-zague, reunidas e homogeneizadas, resultando em uma amostra composta, a qual foi utilizada para o isolamento do consórcio (PAIXÃO et al., 2010). Do montante das amostras coletadas (aproximadamente 500 gr), 1 g da amostra do solo foi adicionado a 50 ml de meio

mineral Bushnell-Haas Broth (BHB), estéril, contendo 0,1 $g.ml^{-1}$ de ciclohexamida, utilizada para garantir o crescimento e estabilização de um consórcio de bactérias. O meio BHB é composto por ($g.L^{-1}$): KH_2PO_4, 1,0; $MgSO_4.7H_2O$, 0,2; $CaCl_2.2H_2O$, 0,02; $(NH_4)2HPO_4$, 0,02; KNO_3, 1; $FeCl_3$, 0,05. Essas soluções foram esterilizadas individualmente e reunidas na hora do uso (BUSHNELL; HAAS, 1941).

A amostra permaneceu nesse meio de cultivo por 12 h, a temperatura de 30°C, em constante agitação de 150 rpm. Após esse período de incubação, a suspensão ficou em repouso durante 1 h para que ocorresse a decantação da fase sólida. Para o enriquecimento, foi inoculada uma alíquota de 500 µl do sobrenadante da suspensão do solo em 50 ml de meio de cultivo BHB estéril, contendo 0,5% de bagaço de cana-de-açúcar triturado. A fonte de carbono utilizada foi esterilizada em água a 120°C, 1,5 atm, durante 20 minutos, antes de realizar a mistura com os sais.

O processo de adaptação do consórcio bacteriano ao meio contendo bagaço de cana-de-açúcar triturado foi realizado da seguinte forma: a suspensão foi incubada durante 7 dias, a 30°C em constante agitação de 150 rpm. Após esse intervalo, uma alíquota de 500 µl da suspensão era transferida para um novo frasco com meio enriquecido contendo 50 ml de meio BHB com bagaço de cana-de-açúcar.

Esse procedimento foi realizado durante 8 semanas e a cada transferência de alíquotas, 800 µl das amostras foram estocados em glicerol a temperatura de -80°C.

Para a realização desse trabalho, foi resgatada do estoque em freezer -80º C uma amostra da primeira semana de cultivo do consórcio, e posteriormente realizado o cultivo nas mesmas condições citadas acima, ao longo de 20 semanas. A cada semana, durante o repique, alíquotas de 1 ml das amostras eram utilizadas para a extração do DNA metagenômico e alíquotas de 800 µl foram estocadas em glicerol em freezer -80°C.

4.2 Extração e Quantificação do DNA Metagenômico do Consórcio

Para a extração do DNA metagenômico, utilizou-se como amostra a parte líquida da cultura referente a cada semana (das 20 semanas testadas). O DNA total foi extraído utilizando o *Kit Wizard® Genomic DNA Purification*, da Promega Corporation, seguindo as instruções do fabricante. Em seguida, o DNA

metagenômico foi submetido a uma eletroforese realizada em uma cuba (modelo Max Cell EC 360M) conduzida em tampão TBE 1X (Tris 89 mM, ácido bórico 89 mM, EDTA 2,5 mM, pH 8,3), em gel de agarose 0,8% acrescido de brometo de etídio (0,5 mg/ml), durante aproximadamente 2 h, à voltagem constante de 90V. O perfil eletroforético foi visualizado sob luz UV e documentado através de um sistema de documentação de géis, GEL DOC Universal Hood II (BIO-RAD).

As amostras foram avaliadas quanto a sua pureza em espectrofotômetro Nano Drop ND-1000 (Thermo Scientific). A concentração do DNA foi quantificada utilizando o aparelho Qubit® Fluorometer (Life Technologies) com o *kit Qubit™ DNA BR Assay* (Invitrogen®) seguindo as recomendações do fabricante.

4.3 Amplificação do Gene 16S rRNA Através da Técnica de PCR (Reação em Cadeia da Polimerase) e Purificação dos Produtos da PCR

O 16S rRNA foi amplificado utilizando os *primers* F515 e R926 descritos por Quince et al. (2011) (QUINCE et al., 2011). Esses *primers* compreendem as regiões V4-V5 do gene, contendo em torno de 400-500pb.

A Reação em Cadeia da Polimerase (PCR) foi realizada em um volume final de 20 µl contendo: enzima KAPA HIFI Hot Start (ReadyMiX PCR *Kit*), 10 µM de cada *primer*, 0,001 µg/µl de BSA e água deionizada ultra pura para completar o volume final da reação. O controle negativo foi realizado a partir dos reagentes utilizados para a PCR com adição de água deionizada ultra pura, em substituição ao DNA, completando para o volume final de 20 µl da reação.

A reação de amplificação foi realizada em termociclador GeneAmp® PCR System 9700-Applied Biosystems, sendo 3 minutos a 95ºC, seguido por 30 segundos a 95°C, 1 minuto e 30 segundos a 55°C, 45 segundos a 72°C por 35 ciclos, em seguida, 5 minutos a 72°C.

Após a amplificação, o produto da PCR foi purificado utilizando o *Kit DNA Clean & Concentrator* (Promega), seguindo as instruções do fabricante. A pureza e qualidade das amostras a serem sequenciadas foram verificadas no Agilent 2100 Bioanalyzer.

4.4 Sequenciamento do 16S rRNA em Plataforma Ion PGM™ Torrent

Após a obtenção e purificação do produto da PCR, realizou-se o sequenciamento na plataforma Ion PGM™ (Personal Genome Machine) Torrent da Thermo Fisher Scientific, no Laboratório de Bioquímica de Microrganismos e Plantas (LBMP) do Departamento de Tecnologia da Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal, SP (FCAV), na Universidade Estadual Paulista "Júlio de Mesquita Filho" (UNESP).

As amostras selecionadas para o sequenciamento foram: primeira semana (em triplicata), terceira, quinta, sétima, nona, décima primeira, décima terceira, décima quinta e décima nona (em triplicata) semanas.

Após a PCR e purificação das amostras, as bibliotecas da região V4-V5 do gene 16S rRNA foram construídas utilizando o *Prepare Amplicon Library whithout Fragmentation using the Ion Plus Fragment Kit*, seguindo as recomendações do fabricante.

O sequenciamento da região V4-V5 do gene 16S rRNA foi realizado na plataforma Ion PGM™ Torrent, utilizando o chip 318 com os *kits Ion PGM Template* OT2 400 (Life Technologies) e *Ion PGM Sequencing* 400 (Life Technologies) de acordo com as recomendações do fabricante.

4.4.1 Análises das Sequências Obtidas pela Plataforma Ion PGM™ Torrent

As sequências geradas foram exportadas para arquivos em formato fastq que são arquivos de sequências em formato texto contendo as bases sequenciadas com os respectivos valores de qualidade na escala PHRED.

A qualidade das amostras brutas, obtidas após o sequenciamento, foi avaliada utilizando a ferramenta FastQC versão 0.11.4 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/).

As análises das sequências foram realizadas de acordo com as recomendações do *Brazilian Microbiome Project* (BMP), seguindo o protocolo para análises de dados para o perfil de 16S rRNA sequenciados em plataforma Ion Torrent (PYLRO et al., 2014b), descritos a seguir.

4.4.2 Procedimento de Análise do 16S rRNA de Acordo com as Recomendações do BMP

- Filtragem e Seleção dos *Reads*

Através da ferramenta USERACH 8, as amostras *single-read* geradas após o sequenciamento obtiveram os barcodes removidos e foram filtradas, descartando os *reads* com erros total e esperados maiores que 0,5 (*-fastq_maxee 0.5*) e menores que 200 nucleotídeos (*-fastq_trunclen 200*). A identificação das sequências únicas e remoção de uma ou mais sequências idênticas também foram realizadas (*-derep_fulllength*).

- Preparação das Sequências, Formação de Matriz das Unidades Taxonômicas Operacionais (UTOs) e Classificação

A classificação da abundância e os passos descritos em seguida foram realizados com a ferramenta USEARCH 8, considerando o agrupamento de sequências mais abundantes (*-sortbysize*) e o descarte de “singletons” (*-minsize 2*).

A clusterização de UTOs (Unidades Taxonômicas Operacionais) foi realizada seguindo o método UPARSE (*-cluster_otus*), onde a identidade das sequências correspondentes a uma mesma UTO é maior ou igual a 97% (≥97%). Com esses parâmetros, foram construídas tabelas de UTOs (*usearch_global*).

Com o auxílio do programa QIIME 1.8.0 *(Quantitative Insights Into Microbial Ecology)* as UTOs criadas foram relacionadas aos seus táxons correspondentes através da ferramenta “*assign_taxonomy.py*”, associando as sequências classificadas taxonômicamente ao banco de dados *Greengenes*. Em seguida, o alinhamento das sequências foi feito através da ferramenta “*align_seqs.py*”, utilizando o banco de dados *Greengenes* (para sequências de genes 16S rRNA) como referência das sequências. As regiões que continham muitas variações foram filtradas através da ferramenta “*filter_alignment.py*” e a relação filogenética realizada através da ferramenta “*make_phylogeny.py*”.

- Estimativa de Riqueza de UTOs e Diversidade dos Micro-organismos Presentes no Consórcio Bacteriano a Partir da Análise do 16S rRNA

As UTOs foram verificadas utilizando o programa QIIME 1.8.0, através da ferramenta "*summarize-table*", para isso os arquivos UTOs gerados anteriormente foram convertidos para o formato BIOM (*The Biological Observation Matrix*). A análise da diversidade foi realizada através da ferramenta "*core_diversity_analyses.py*", gerando os resultados de *α*-diversidade, *β*-diversidade, as tabelas de frequência e os gráficos de classificação taxonômica.

4.5 Predição Metabólica Funcional do Consórcio a Partir do 16S rRNA

PICRUSt, abreviação do inglês, "*Phylogenetic Investigation of Communities by Reconstruction of Unobserved States*", é um software de bioinformática utilizado para realizar predições de potenciais funções metabólicas de metagenomas a partir de sequências de genes marcadores, como por exemplo o 16S rRNA, e usando como referência os bancos de dados IMG (*Integrated Microbial Genomes*) (MARKOWITZ et al., 2012) e Greengenes (DESANTIS et al., 2006; LANGILLE et al., 2013). Esses bancos de dados servem como base para inferir o potencial codificador de um metagenoma a partir de cada UTO (Unidade Taxonômica Operacional) identificada (LANGILLE et al., 2013). Para a classificação funcional, o PICRUSt utiliza os bancos de dados como KEGG (*Kyoto Encyclopedia of Genes and Genomes*) ou COG (*Clusters of Orthologous Groups of Proteins*) (LANGILLE et al., 2013).

As UTOs geradas através do programa QIIME versão 1.8.0, contendo as informações dos táxons presentes no consórcio bacteriano, foram utilizadas como entradas do software PICRUSt versão 1.1.0. Assim, foram geradas tabelas no formato BIOM contendo as predições funcionais do consórcio por meio de associações do gene marcador 16S rRNA com os níveis categorizados de acordo com o banco de dados KEGG (níveis 1, 2 e 3):

- Categorização nível 1: processos celulares, processamento de informação ambiental, processamento de informação genética, metabolismo, sistema do organismo e não classificados;

- Categorização nível 2: metabolismo de aminoácidos, metabolismo de carboidratos, comunicação celular, crescimento e morte celular, motilidade celular, energia de metabolismo, adaptação ambiental, metabolismo de lipídio, transporte de membrana, processamento de reparo de DNA;

- Categorização nível 3: transportadores, glicólise e gliconeogêneses, metabolismo do nucleotídeo, proteínas de motilidade bacteriana, degradação de xileno, enzimas relacionadas com aminoácidos, transportadores ABC, quimiotaxia bacteriana, energia de metabolismo, metabolismo de aminoácidos, biossíntese e biodegradação de metabólitos secundários, metabolismo de amido e sacarose, proteínas de recombinação e reparo de DNA, sistema de secreção bacteriano, metabolismo de frutose e manose, metabolismo de carboidratos e via de fixação de carbono em procariotos.

Para a classificação das sequências do gene do 16S rRNA em Unidades Taxonômicas Operacionais (UTOs), o banco de dados *Greengenes* versão gg_13_5 foi utilizado como referência, seguindo com a normalização do número de cópias de 16S com o banco de dados KEGG versão ko_13_5 (*Kyoto Encyclopedia of Genes and Genomes*), utilizado para a predição das potenciais funções metabólicas do consórcio.

Para a análise dos perfis funcionais, as seguintes comparações foram realizadas: entre a primeira (S01) e a terceira (S03) semanas, entre a primeira e nona (S09) semanas, entre a primeira e décima primeira (S11) e a primeira e décima nona (S19) semanas, que foram utilizados como entradas no programa STAMP versão 3.2.1, realizando o teste estatístico G-teste + teste de Fisher para análise estatística, o método FDR de *Benjamini-Hochberg* para correção de testes múltiplos e utilizando a filtragem por q-value < 0,05.

O software STAMP (“*Statistical Analysis of Metagenomic Profiles*”) oferece várias opções de hipóteses estatísticas, onde se pode escolher a mais apropriada para a identificação dos potenciais perfis funcionais ou outras características desejadas, assim como das diferenças biologicamente relevantes entre os conjuntos

de dados metagenômicos (PARKS et al., 2014). Os perfis podem ser comparados entre pares ou entre conjunto de dados que são organizados em dois ou mais grupos (PARKS et al., 2014).

4.6 Sequenciamento Total do Metagenoma em Plataforma Illumina® HiSeq 2500 e Análise da Qualidade dos *Reads*

Para realizar o sequenciamento total do metagenoma, as amostras das 20 semanas de repique do consórcio bacteriano foram reunidas em 3 *pools* equimolares (1º *pool* contendo amostras da primeira a sexta semanas, 2º *pool* da sétima à décima segunda semanas e 3º *pool* da décima terceira até a vigésima semana). Para isso, foram resgatadas do estoque as amostras de DNA extraído para cada semana, contendo triplicatas na primeira, décima primeira e décima nona semanas. A qualidade dessas amostras foi verificada submetendo-as a eletroforese que foi realizada em uma cuba horizontal (modelo LCH 13x15 da Loccus Biotecnologia), conduzida em tampão TBE 1X, em gel de agarose 0,8% acrescido de brometo de etídio (0,9 mg/ml), durante aproximadamente 2 h à voltagem constante de 100V. O perfil eletroforético foi visualizado sob luz UV e documentado através de um sistema de documentação de géis, GEL DOC Universal Hood II – BIO-RAD. As amostras foram avaliadas quanto a sua pureza em espectrofotômetro Nano Drop ND-1000 (Thermo Scientific).

As bibliotecas para o sequenciamento do DNA total do metagenoma foram preparadas utilizando o *kit Illumina Nextera DNA Library Prep*, no Laboratório de Bioquímica de Microrganismos e Plantas (LBMP) da UNESP-FCAV e o sequenciamento foi realizado no Laboratório de Biotecnologia Animal - LZT (ESALQ – USP), em Piracicaba, SP, em plataforma HiSeq 2500 da Illumina®, em uma HiSeq Flow Cell v4, utilizando o kit HiSeq SBS v4, em leituras *paired-end* de 2x 100 pb.

A qualidade dos *reads* sequenciados foi avaliada através da ferramenta FastQC (v. 0.11.4) e a remoção dos adaptadores foi realizada através da ferramenta Trimmomatic (v. 0.36) (BOLGER; LOHSE; USADEL, 2014). Após a retirada dos adaptadores, os três *pools* foram reunidos em um único arquivo para as etapas de identificação de um micro-organismo do consórcio, montagem e anotação.

4.7 Identificação, Recuperação, Montagem e Anotação de um Genoma Completo Proveniente do Consórcio por Abordagens de Bioinformática

A montagem dos *reads* filtrados foi realizada com os programas: MetaSPAdes (v. 3.10.1) (NURK et al., 2016) e MetaPlatanus (v. 1.0.3) (KAJITANI et al., 2014). Os resultados foram comparados a fim de se identificar qual programa apresentava os melhores índices de montagem (menor número de *contigs*, maior *contig* e N50).

A identificação taxonômica dos genomas montados foi efetuada através do programa Kraken (v. 0.10.5 – beta) (WOOD; SALZBERG, 2014) e, em seguida, as sequências de cada genoma montado foram recuperadas através do programa ZEUSS (v. 1.0.0) (ALVARENGA; FIORE; VARANI, 2017), seguindo os critérios adotados por Alvarenga et al., (2017), com a finalidade de identificação de um genoma completo presente no metagenoma. Nesse trabalho, foi considerado o genoma que apresenta maior chance de ser fechado através de técnicas de bioinformática, não representando necessariamente o micro-organismo mais prevalente presente no consórcio. Com a finalidade de verificar se esse genoma selecionado continha plasmídeos, foi utilizada a ferramenta plasmidSPAdes (ANTIPOV et al., 2016). Para ordenação e orientação da montagem dos *contigs* provenientes do genoma selecionado em *scaffolds*, foi utilizado o programa SIS (programa lançado em 03/07/12) (DIAS; DIAS; SETUBAL, 2012). Através da ferramenta SIS foi possível obter apenas um super-*scaffold*, representando o genoma fechado, e indicando a ausência de plasmídeos. O fechamento e circularização do genoma foram realizados de forma manual, através da comparação entre as extremidades de cada região de lacuna (*gaps*) presente no *super-scaffold*, frente a genomas próximos, *Dyella jiangningensis* SBZ3-12 (número de acesso do GenBank: CP007444.1) e *Dyella japonica* A8 (número de acesso do GenBank: NZ_CP008884.1), com o uso da ferramenta BLAST. O mapeamento e a finalização da sequência final do genoma foram realizados através do programa CLC Genomics Workbench (v. 9.0) (CLC Bio - http://www.clcbio.com). A anotação do genoma foi efetuada através das ferramentas RAST (*Rapid Annotation using Subsystem Technology*), versão 2.0 (AZIZ et al., 2008). Demais elementos genômicos como, por exemplo, regiões de Profagos, Sequências de Inserção (ISs) e

Ilhas Genômicas foram identificados com as ferramentas phiSPY (AKHTER; AZIZ; EDWARDS, 2012), ISsaga (VARANI et al., 2011) e OligoWords (GANESAN et al., 2008), respectivamente. A identificação de Glicosil Hidrolases foi realizada a partir da anotação manual e comparação com o banco de dados UNIPROT (THE UNIPROT CONSORTIUM, 2017), com base nos resultados gerados na anotação pela plataforma RAST. Para visualização genômica circular e análises comparativas foram utilizadas as ferramentas DNA Plotter (CARVER et al., 2009) e BLAST, respectivamente.

5. RESULTADOS

5.1 Diversidade da Comunidade Bacteriana Presente no Consórcio Através da Análise do Sequenciamento do 16S rRNA Realizado pela Plataforma Ion PGM™

O sequenciamento na plataforma Ion PGM™ Torrent do 16S rRNA do consórcio bacteriano foi realizado a partir de cada semana analisada, gerando um total de 3.433.255 de *reads*, com fragmentos variando entre 456 a 508 pb (pares de bases) (Tabela 2), contendo 1.191.937 de sequências após a filtragem realizada seguindo o protocolo do BMP.

Tabela 2. Total de *reads* gerados após o sequenciamento do 16S rRNA do consórcio bacteriano na plataforma Ion PGM™, antes de realizar a filtragem seguindo o protocolo do BMP.

Semanas	Total de *reads*	Tamanho médio (pb)
S 01.1	184.407	508
S 01.2	270.516	493
S 01.3	296.945	468
S 03	235.042	468
S 05	355.458	479
S 07	380.696	487
S 09	16.642	468
S 11	114.111	478
S 13	89.760	456
S 15	379.077	460
S 17	414.543	472
S 19.1	269.391	480
S 19.2	157.567	486
S 19.3	269.100	468

Nota: as semanas estão representadas pelo prefixo “S” seguido do número que indica qual semana foi selecionada para o sequenciamento. A primeira e a décima nona semanas foram analisadas em triplicatas.

5.2 Estimativa de Riqueza de UTOs do Consórcio Bacteriano, Realizada Através do Programa QIIME de Acordo com o Índice de Chao-1

A estimativa de riqueza de UTOs do consórcio bacteriano foi realizada para as semanas S01, S03, S05, S07, S09, S11, S13, S15, S17 e S19, através do programa QIIME 1.8.0 de acordo com o índice de Chao-1.

Os valores de riqueza das amostras demonstram que a maior riqueza ou o maior número de contagem de UTOs foi apresentado nas amostras da nona, décima primeira, décima terceira e décima nona semanas (Figura 5).

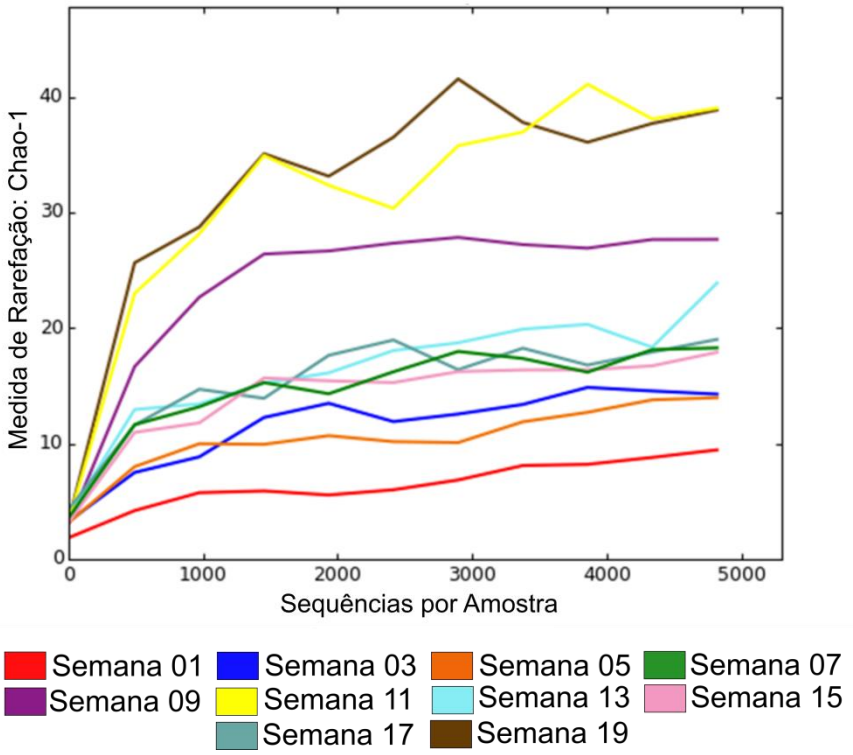

Figura 5. Valores de riqueza de UTOs do consórcio bacteriano estimados pelo índice Chao-1 através do programa QIIME, demonstrando maior riqueza de UTOs presentes no consórcio na nona, décima primeira, décima terceira e décima nona semanas avaliadas.

5.2.1 Agrupamento de UTOs das Amostras das Semanas Avaliadas Baseado em Análise de Coordenadas Principais (PCoA) Utilizando as Distâncias Métricas UniFrac Ponderada e Não-Ponderada

A composição bacteriana do consórcio ao longo das semanas analisadas foi comparada utilizando as distâncias métricas UniFrac Ponderada e UniFrac Não-Ponderada, baseada em Análise de Coordenadas Principais (PCoA), gerados pelo programa QIIME.

A métrica UniFrac avalia a distância entre comunidades biológicas utilizando de informações filogenéticas entre os membros para calcular as medidas de dissimilaridade (LOZUPONE; KNIGHT, 2005). A distância métrica Não-Ponderada é uma medida qualitativa da *β*-diversidade que avalia ausência e presença de

sequências, enquanto a distância métrica Ponderada é quantitativa, baseada em informações de abundância relativa entre os organismos (LOZUPONE et al., 2007).

Assim, os resultados da distribuição de UTOs obtidas a partir de sequências do 16S rRNA foram apresentados através de análise de PCoA, em uma matriz de distância contendo os índices UniFrac Não-Ponderado e Ponderado.

O gráfico de PCoA gerado utilizando a distância métrica UniFrac Não-Ponderada apresenta PC 1 (51,6%), PC 2 (19,08%) e PC 3 (11,53%). As UTOs atribuídas à primeira semana apresentam-se distantes das outras semanas, principalmente com relação a nona, décima primeira, décima terceira e décima nona semanas. Esses resultados sugerem que o consórcio bacteriano apresenta variações nas classificações de UTOs ao longo das semanas avaliadas (Figura 6).

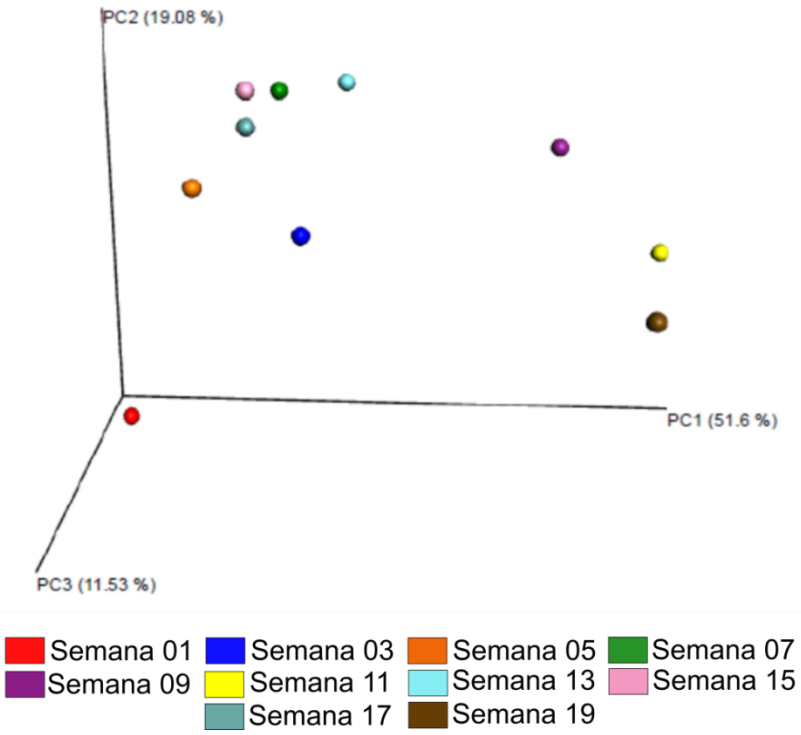

Figura 6. Análise de Coordenadas Principais (PCoA) utilizando a distância métrica UniFrac Não-Ponderada gerada pelo programa QIIME.

O gráfico de PCoA gerado utilizando a distância métrica UniFrac Ponderada apresenta PC 1 (88,07%), PC 2 (5,81%) e PC 3 (2,91%). Esses resultados sugerem

o agrupamento das UTOs atribuídas a cada semana conforme as abundâncias relativas presentes no consórcio (Figura 7).

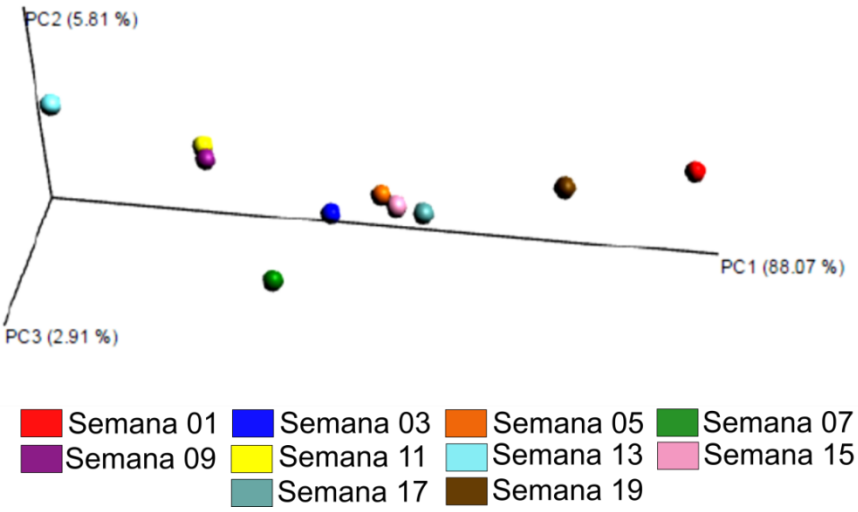

Figura 7. Análise de Coordenadas Principais (PCoA) utilizando a distância métrica UniFrac Ponderada gerada pelo programa QIIME.

Esses gráficos de PCoA sugerem que há uma variação na estrutura da comunidade bacteriana, já que as amostras se apresentam dispersas. A abundância relativa de UTOs apresenta-se similar entre a nona e décima primeira semanas, porém o resultado da análise de PCoA utilizando o índice UniFrac Ponderado sugere que há variações de abundância relativa entre as semanas, principalmente entre a primeira e a décima terceira semanas. Interessantemente, a abundância relativa de UTOs entre a primeira e a décima nona semanas mostra-se próxima. A hipótese é de que, apesar de ocorrer variação de abundância relativa de UTOs ao longo das semanas de repique, há uma estabilização entre a nona e a décima primeira semanas e, apesar das diferenças entre a primeira e a décima terceira semanas, ao final do repique (décima nona semana), as UTOs mais abundantes apresentam-se próximas às UTOs atribuídas no ínico do consórcio, sugerindo que os organismos mais abundantes permanecem presentes no consórcio ao longo das semanas avaliadas, apesar das variações de abundância relativa que ocorrem.

5.3 Diversidade Bacteriana Presente no Consórcio

Com relação à diversidade bacteriana presente no consórcio, os principais filos encontrados foram: *Proteobacteria,* que foi representado em todas as semanas analisadas, com a abundância mínima de 96,67% detectada na sétima semana, chegando a 99,56% na décima sétima semana; o filo *Planctomycetes*, identificado a partir da terceira semana, alcançando a abundância máxima de 3,18% na sétima semana e *Actinobacteria,* o terceiro filo predominante do consórcio bacteriano, que foi detectado em todas as semanas analisadas, apresentando sua maior abundância na primeira semana (1,02%).

Os filos que apresentaram abundância pouco significativa (< 1% em cada semana) foram classificados em Outros Filos, sendo eles: *Bacteroidetes*, *Chlamydiae*, *Cyanobacteria, Firmicutes* e *Fusobacteria*. Além disso, houve uma quantidade de organismos não classificados que não ultrapassaram a abundância de 0,03% ao longo das semanas analisadas. Assim, filos Não Classificados e Outros Filos foram agrupados em um único conjunto.

A abundância relativa de cada organismo variou ao longo das semanas, como mostra a Tabela 3.

Tabela 3. Abundância relativa dos principais filos presentes no consórcio bacteriano em cada semana analisada. O filo *Proteobacteria* foi representado em todas as semanas. A classificação Outros Filos contêm filos que foram detectados ao longo das semanas, porém em abundância pouco significativa, sendo: *Bacteroidetes*, *Chlamydiae*, *Cyanobacteria*, *Firmicutes* e *Fusobacteria*. Filos Não Classificados não ultrapassaram a abundância de 0,03% ao longo das semanas.

	Semanas									
Filos	**01**	**03**	**05**	**07**	**09**	**11**	**13**	**15**	**17**	**19**
Proteobacteria	98,97%	98,89%	99,32%	96,67%	99,13%	99,03%	99,26%	99,50%	99,56%	99,11%
Planctomycetes	-	1,00%	0,61%	3,18%	0,58%	0,40%	0,57%	0,42%	0,33%	0,29%
Actinobacteria	1,02%	0,10%	0,06%	0,12%	0,10%	0,24%	0,12%	0,07%	0,07%	0,39%
Não Classificados e Outros Filos	0,01%	0,01%	-	0,03%	0,19%	0,33%	0,05%	0,01%	0,03%	0,21%

Com relação às principais classes presentes no consórcio bacteriano, *Betaproteobacteria* e *Gammaproteobacteria*, pertencentes ao filo *Proteobacteria*, foram reportadas em todas as semanas analisadas.

A classe *Betaproteobacteria* apresentou abundância de 98,94% na primeira semana, diminuindo até a décima terceira semana (49,53%) e voltando a aumentar a abundância a partir da décima quinta até a décima nona semana de repique (88,13%).

Gammaproteobacteria apresentou abundância mínima de 0,02% apenas na primeira semana. A abundância máxima foi alcançada na décima terceira semana, onde apresentou 41,42%, diminuindo em seguida para 8,17% na décima nona semana.

Outra classe pertencente ao filo *Proteobacteria* também foi reportada, *Alphaproteobacteria*, porém em menor abundância quando comparada às classes *Beta* e *Gammaproteobacteria*, chegando à abundância máxima de 8,31% na décima terceira semana.

A classe *Planctomycetia* (filo *Planctomycetes*) foi detectada a partir da terceira semana, alcançando a abundância máxima na sétima semana (3,18%) e a classe *Actinobacteria* (filo *Actinobacteria*) também foi reportada entre as principais classes presentes no consórcio, com sua maior representação na primeira semana (1,02%).

Houve detecção de classes em abundâncias pouco significativas (< 1% do organismo em cada semana) que foram classificadas em Outras Classes, incluindo: *Flavobacteriia* (filo *Bacteroidetes*), *Chlamydiia* (filo *Chlamydiae*), *Synechococcales* (filo *Cyanobacteria*), *Bacilli* e *Clostridia* (filo *Firmicutes*), *Fusobacteriia* (filo *Fusobacteria*) e *Deltaproteobacteria* (filo *Proteobacteria*).

Bactérias não classificadas não ultrapassaram a abundância de 0,03% ao longo das semanas analisadas, assim, organismos Não Classificados e Outras Classes foram representados em uma única classificação.

As variações das abundâncias das UTOs detectadas e classificadas ao nível de classe em cada semana analisada são apresentadas na Tabela 4.

Tabela 4. Abundância relativa das principais classes presentes no consórcio bacteriano em cada semana analisada. A classe representada em todas as semanas foi *Betaproteobacteria*, pertencentes ao filo *Proteobacteria*. Não Classificados não ultrapassaram a abundância de 0,03% ao longo das semanas e a classificação Outras Classes incluem as que foram detectadas presentes, porém em abundâncias pouco significativas, portanto foram agrupadas.

	Semanas									
Classes	**01**	**03**	**05**	**07**	**09**	**11**	**13**	**15**	**17**	**19**
Betaproteobacteria	98,94%	71,82%	75,26%	66,62%	61,16%	60,57%	49,53%	76,87%	77,73%	88,13%
Gammaproteobacteria	0,02%	27,03%	23,77%	28,72%	36,04%	37,01%	41,42%	21,53%	19,81%	8,17%
Alphaproteobacteria	0,01%	0,05%	0,29%	1,34%	1,93%	1,43%	8,31%	1,10%	2,02%	2,81%
Planctomycetia	-	1,00%	0,61%	3,18%	0,58%	0,40%	0,57%	0,42%	0,33%	0,29%
Actinobacteria	1,02%	0,10%	0,06%	0,12%	0,10%	0,24%	0,12%	0,07%	0,07%	0,39%
Não Classificados e Outras Classes	0,01%	0,01%	-	0,03%	0,19%	0,35%	0,05%	0,01%	0,03%	0,21%

Assim, através das análises de diversidade foram detectadas bactérias até o nível de gênero presentes no consórcio bacteriano degradador da biomassa. Os principais gêneros de bactérias identificados no consórcio foram: *Burkholderia*, pertencente à família *Burkholderiaceae* (Ordem *Burkholderiales*), presente na primeira semana analisada com abundância de quase 99%, diminuindo até a décima terceira semana (49,50%) e voltando a aumentar significativamente da décima quinta (76,79%) até a última semana analisada (87,66%).

Também foi detectado um gênero assinalado como desconhecido, pertencente à família *Rhodanobacteraceae* (Ordem *Xanthomonadales*), que demonstra abundância inversamente proporcional ao gênero *Burkholderia*, ou seja, quando *Burkholderia* está abundante em determinada semana, o gênero desconhecido pertencente à família *Rhodanobacteraceae* se apresenta em menor abundância e quando *Burkholderia* apresenta menor abundância, a do gênero pertencente à *Rhodanobacteraceae* aumenta. A análise do sequenciamento do DNA total do metagenoma permitiu a identificação do gênero *Dyella* como pertencente a essa família (maiores detalhes nos próximos tópicos). Apenas na décima terceira semana ambos os gêneros permanecem presentes com abundâncias próximas, sendo *Burkholderia* detectado a 49,50% e o gênero pertencente à família *Rhodanobacteraceae* a 41,37%.

Assim, o gênero pertencente à família *Rhodanobacteraceae* apresentou abundância máxima de 41,37% na décima terceira semana e mínima de 0,004% na primeira semana.

Outros três gêneros foram identificados no consórcio, sendo *Rhizobium* (Ordem *Rhizobiales*, família *Rhizobiaceae*), um gênero assinalado como desconhecido, pertencente à família *Isosphaeraceae* (Ordem *Planctomycetales*) e o gênero *Cryocola* (Ordem *Micrococcales*, família *Microbacteriaceae*). *Rhizobium* apresentou abundância máxima de 7% na décima terceira semana, tempo em que o gênero *Burkholderia* esteve menos abundante e em que o gênero pertencente à família *Rhodanobacteraceae* esteve mais abundante.

O gênero desconhecido pertencente à família *Isosphaeraceae* foi identificado em maior abundância na sétima semana (3,18%) e *Cryocola* foi detectado a

aproximadamente 1% na primeira semana, diminuindo drasticamente a partir da terceira semana.

Não classificados não ultrapassaram a abundância de 0,03% ao longo das semanas avaliadas e foram agrupados com gêneros detectados em abundâncias pouco significativas (< 1% em cada semana). Assim, foram nomeados como Não Classificados e Outros Gêneros.

As abundâncias relativas de cada gênero podem ser visualizadas na Figura 8 e em maiores detalhes na Tabela 5.

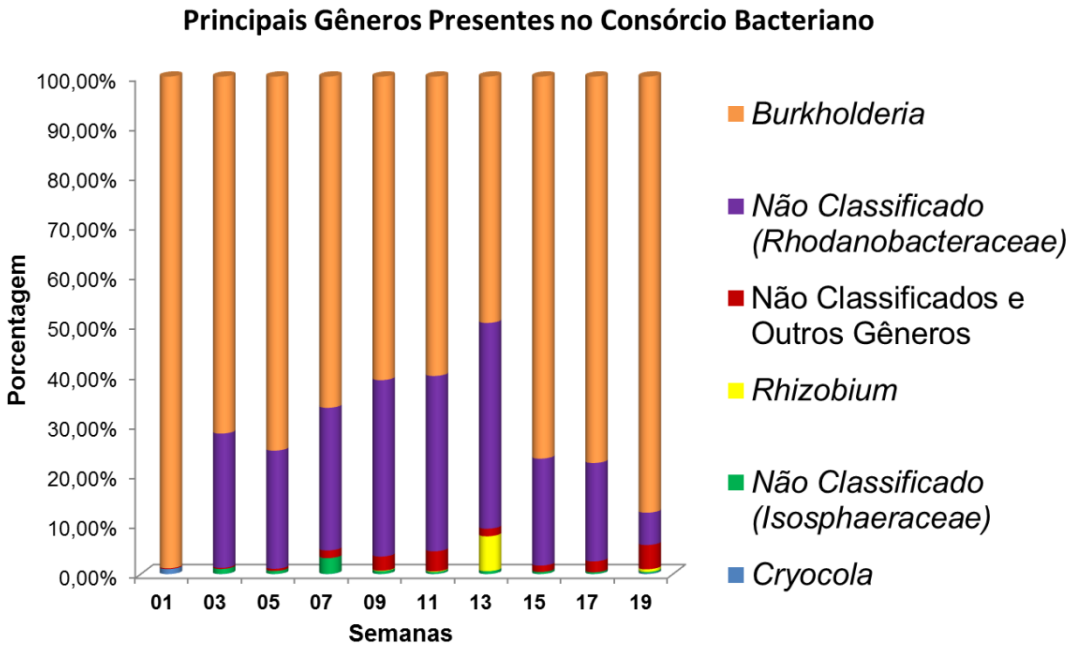

Figura 8. Diversidade dos principais gêneros presentes no consórcio bacteriano, obtida a partir de dados da matriz de UTOs, em cada semana analisada. Os gêneros *Burkholderia* e o Não Classificado, pertencente à família *Rhodanobacteraceae*, foram abundantemente representados. Além desses, foram identificados: gênero *Rhizobium*, *Cryocola* e um gênero assinalado como desconhecido (família *Isosphaeraceae)*. Não classificados ou detectados em abundâncias pouco significativas foram agrupados e nomeados como Não Classificados e Outros Gêneros.

Tabela 5. Abundâncias relativas dos principais gêneros presentes no consórcio bacteriano em cada semana analisada. Os gêneros abundantemente representados foram *Burkholderia* e um gênero Não Classificado, pertencente à família *Rhodanobacteraceae*. Além desses, os gêneros identificados foram: *Rhizobium*, um gênero assinalado como desconhecido, pertencente à família *Isosphaeraceae* e o gênero *Cryocola*. Não classificados e os detectados em abundâncias pouco significativas (< 1%), anotados como Outros Gêneros foram agrupados (Não Classificados e Outros Gêneros).

	Semanas									
Gêneros	**01**	**03**	**05**	**07**	**09**	**11**	**13**	**15**	**17**	**19**
Burkholderia	98,82%	71,75%	75,18%	66,57%	61,01%	60,16%	49,50%	76,79%	77,66%	87,66%
Não Classificado (*Rhodanobacteraceae*)	0,004%	27,00%	23,75%	28,67%	35,46%	35,25%	41,37%	21,50%	19,79%	6,50%
Não Classificados e Outros Gêneros	0,16%	0,21%	0,44%	1,55%	2,84%	4,04%	1,53%	1,28%	2,19%	4,87%
Rhizobium	0,001%	-	0,001%	0,003%	0,08%	0,11%	7,01%	0,001%	0,001%	0,53%
Não Classificado (*Isosphaeraceae*)	-	1,00%	0,61%	3,18%	0,58%	0,40%	0,57%	0,42%	0,33%	0,29%
Cryocola	1,02%	0,04%	0,02%	0,03%	0,02%	0,03%	0,02%	0,01%	0,03%	0,15%

5.4 Predição do Potencial Metabólico Funcional do Consórcio Bacteriano a Partir do 16S rRNA

Através dos testes estatísticos realizados utilizando o software STAMP (“*Statistical Analysis of Metagenomic Profiles”*) foi possível predizer, comparar e apresentar as principais diferenças entre as proporções relativas (%) do potencial perfil metabólico funcional do consórcio bacteriano nas diferentes semanas analisadas.

Os resultados de predição para genes potencialmente funcionais do consórcio bacteriano contendo bagaço de cana-de-açúcar, mostraram os subsistemas expressos ao longo das semanas analisadas, baseados em atividades do metabolismo, processamento de informação ambiental, processamento de informação genética e processos celulares, através da classificação dos subsistemas no nível 1, realizada utilizando o banco de dados KEGG como referência.

A predição das potenciais funções metabólicas foi realizada através das comparações entre as seguintes semanas:

Entre a primeira (S01) e terceira (S03) semanas (Figura 9).

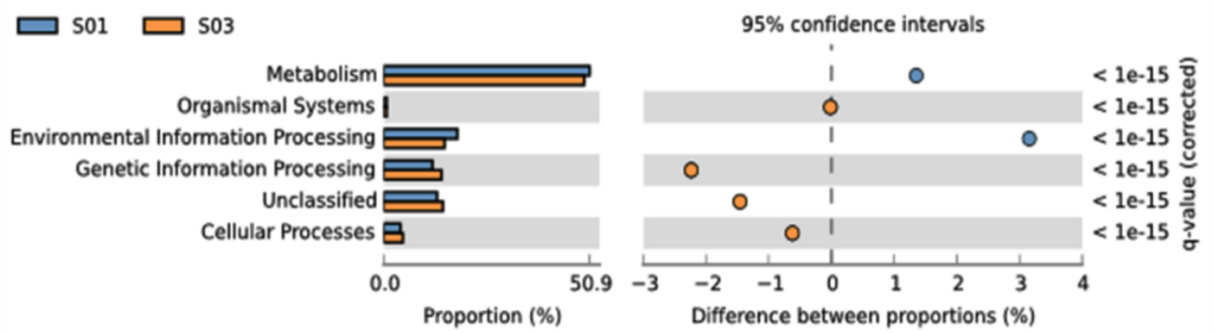

Figura 9. Análise comparativa entre a primeira (representada em azul) e terceira (representada em laranja) semanas de cultivo dos potenciais perfis metabólicos funcionais do consórcio bacteriano. Os dados foram analisados estatisticamente utilizando o software STAMP para diferenças entre proporções (%), utilizando G-teste + teste de Fisher para análise estatística, filtragem por q-value < 0,05 e correção de testes múltiplos usando o método FDR de *Benjamini-Hochber*. O banco de dados KEGG (nível 1) foi utilizado como referência.

Os resultados indicam um enriquecimento para “Metabolismo” e “Processamento de Informação Ambiental” durante a primeira semana em relação à terceira semana. Além disso, as comparações entre essas semanas indicam também o enriquecimento para “Processamento de Informação Genética” e de “Processos Celulares” para a terceira semana em relação à primeira.

Esses resultados também foram observados quando se comparou a primeira semana (S01) com a nona (S09) e décima nona (S19) semanas e são apresentados nas Figuras 10 e 11.

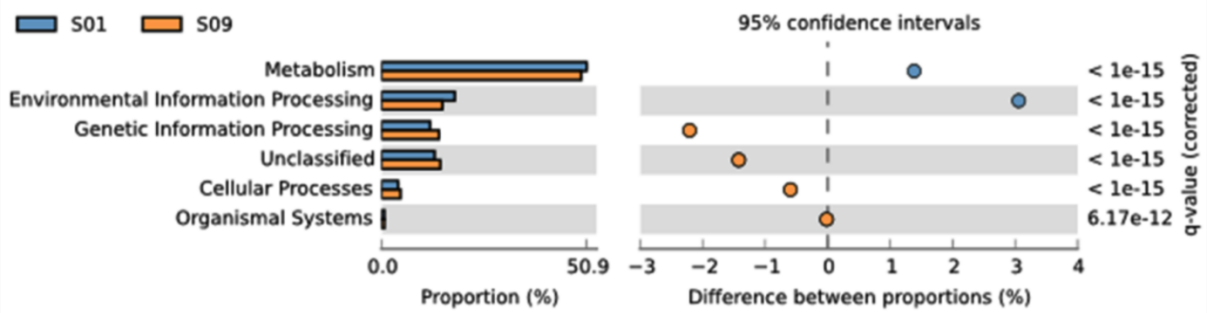

Figura 10. Análise comparativa entre a primeira (representada em azul) e nona (representada em laranja) semanas de cultivo dos potenciais perfis metabólicos funcionais do consórcio bacteriano. Os dados foram analisados estatisticamente utilizando o software STAMP para diferenças entre proporções (%), utilizando G-teste + teste de Fisher para análise estatística, filtragem por q-value < 0,05 e correção de testes múltiplos usando o método FDR de *Benjamini-Hochber*. O banco de dados KEGG (nível 1) foi utilizado como referência.

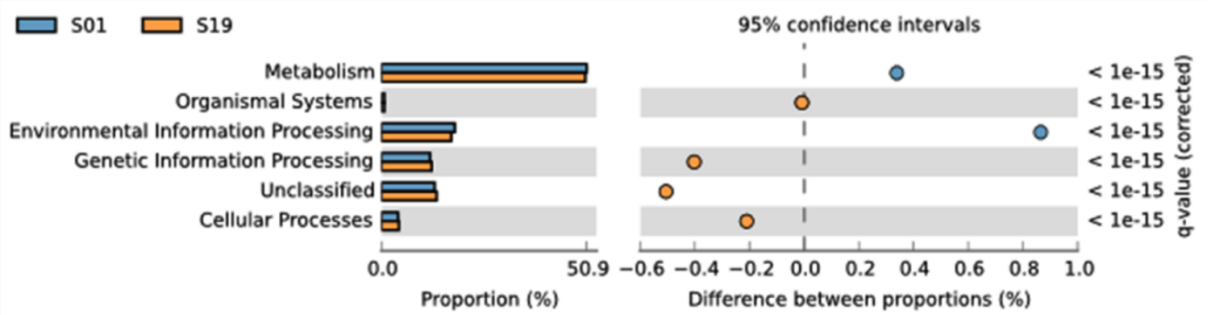

Figura 11. Análise comparativa entre a primeira (representada em azul) e décima nona (representada em laranja) semanas de cultivo dos potenciais perfis metabólicos funcionais do consórcio bacteriano. Os dados foram analisados estatisticamente utilizando o software STAMP para diferenças entre proporções (%), utilizando G-teste + teste de Fisher para análise estatística, filtragem por q-value < 0,05 e correção de testes múltiplos usando o método FDR de *Benjamini-Hochber*. O banco de dados KEGG (nível 1) foi utilizado como referência.

Também foram realizadas predições das potenciais funções metabólicas através da classificação dos subsistemas no nível 3, utilizando o banco de dados KEGG como referência.

Foram encontrados genes que apresentam perfis funcionais para diversos subsistemas específicos como Transportadores, Transportadores do tipo ABC, Glicólise/Gliconeogênese, Proteínas de Motilidade Bacteriana, Vias de Fixação de Carbono em Procarioto, Proteínas de Recombinação e Reparo de DNA, Metabolismo Energético, Sistema de Secreção Bacteriana, entre outros.

A Figura 12, ilustra os resultados de comparação entre a primeira (S01) e terceira (S03) semanas analisadas do consórcio bacteriano, onde é possível observar enriquecimento para os subsistemas “Transportadores” e “Transportadores ABC” para a primeira semana em relação à terceira. Os resultados também apresentam enriquecimento de “Proteínas de Motilidade Bacteriana” e “Proteínas de Recombinação e Reparação do DNA” na terceira semana (S03).

Em geral, esses resultados podem indicar que durante o início do cultivo do consórcio com bagaço de cana-de-açúcar, os micro-organismos presentes passam por um processo de adaptação, envolvendo a sinalização e possível metabolização do material lignocelulósico presente no meio de cultivo, através dos transportadores de açúcares.

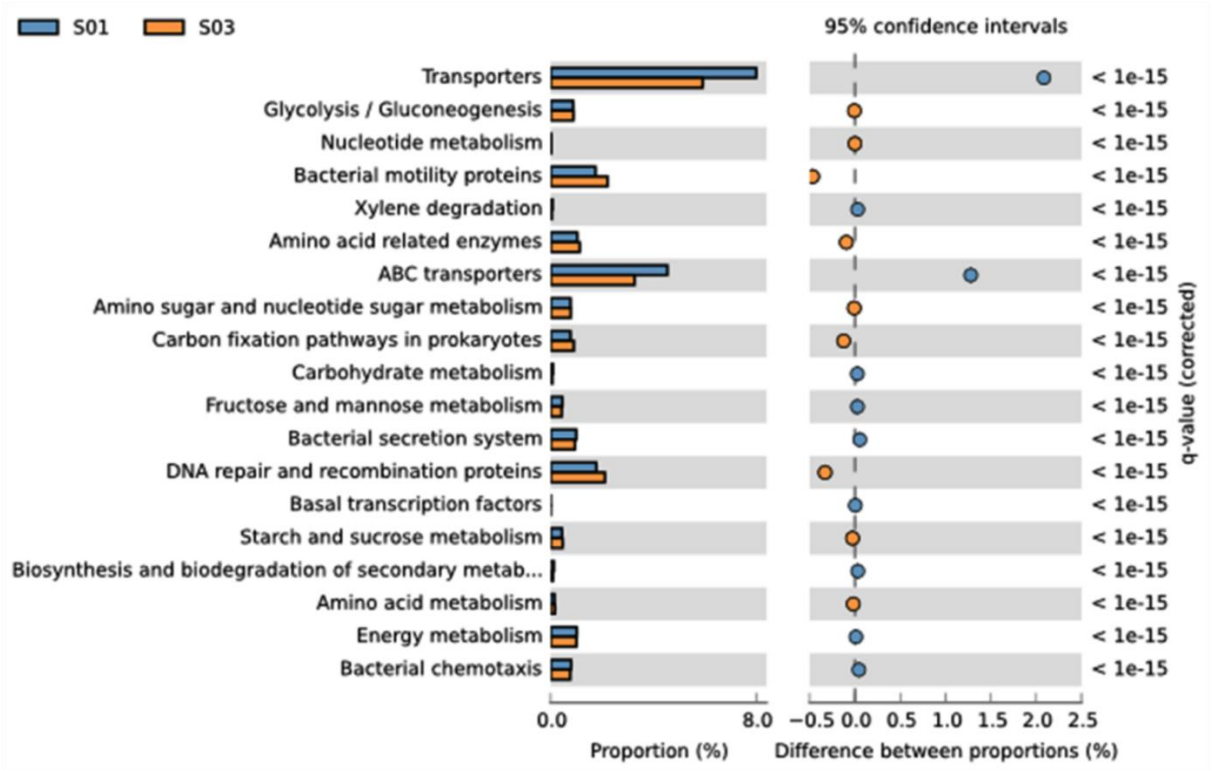

Figura 12. Análise comparativa entre a primeira (representada em azul) e terceira (representada em laranja) semanas de cultivo dos potenciais perfis metabólicos funcionais do consórcio bacteriano. Os dados foram analisados estatisticamente utilizando o software STAMP para diferenças entre proporções (%), utilizando G-teste + teste de Fisher para análise estatística, filtragem por q-value < 0,05 e correção de testes múltiplos usando o método FDR de *Benjamini-Hochber*. O banco de dados KEGG (nível 3) foi utilizado como referência.

Os resultados das análises comparativas entre a primeira (S01) e a décima primeira semana (S11), assim como entre a primeira (S01) e décima nona semanas (S19) foram similares ao anterior e são apresentados, na Figura 13 e Figura 14, respectivamente.

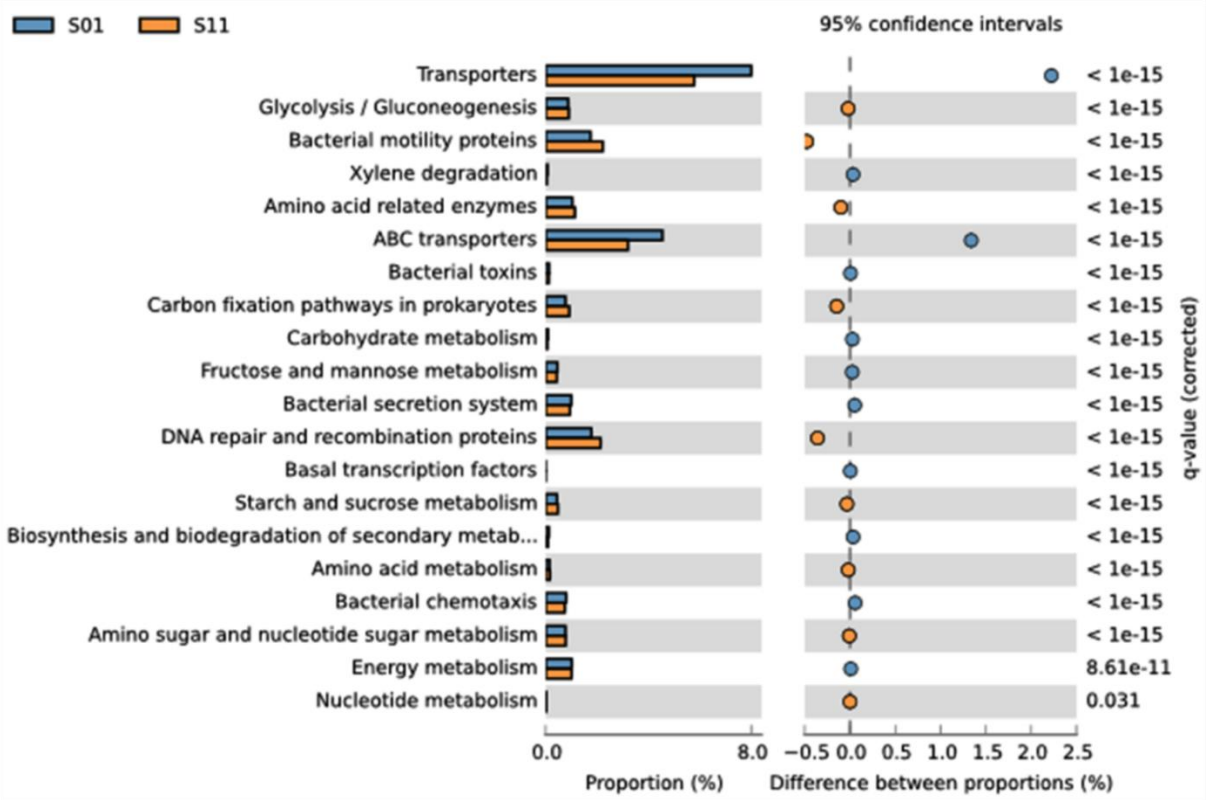

Figura 13. Análise comparativa entre a primeira (representada em azul) e décima primeira (representada em laranja) semanas de cultivo dos potenciais perfis metabólicos funcionais do consórcio bacteriano. Os dados foram analisados estatisticamente utilizando o software STAMP para diferenças entre proporções (%), utilizando G-teste + teste de Fisher para análise estatística, filtragem por q-value < 0,05 e correção de testes múltiplos usando o método FDR de *Benjamini-Hochber*. O banco de dados KEGG (nível 3) foi utilizado como referência.

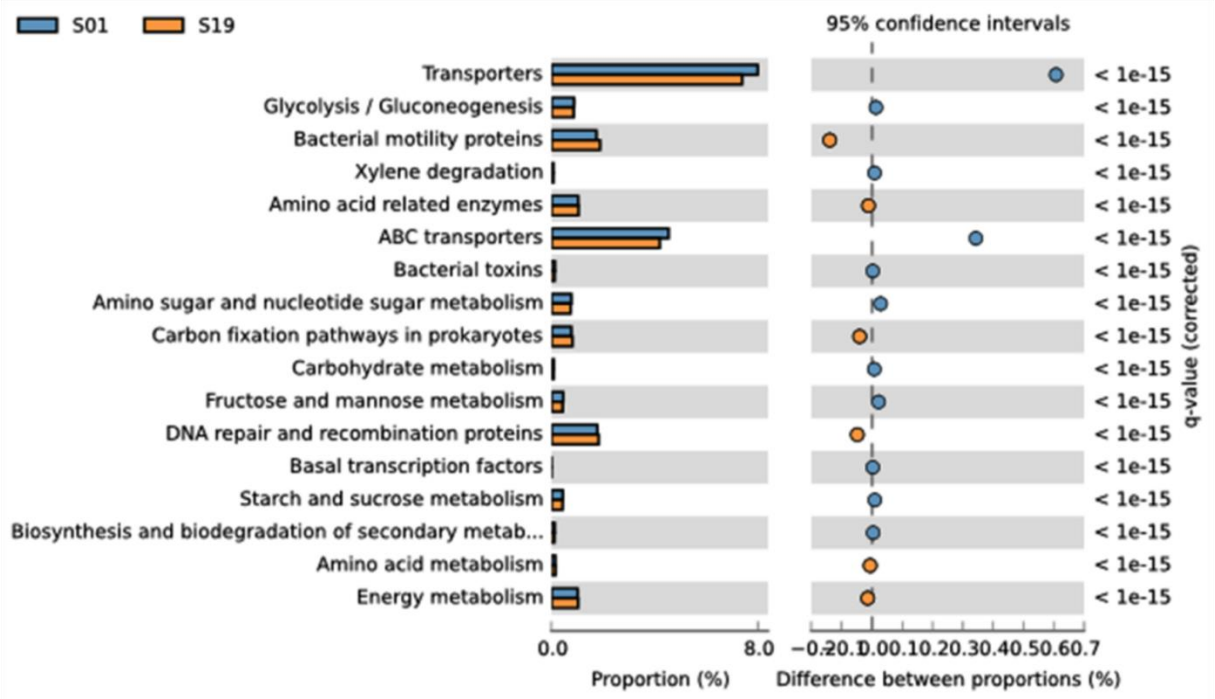

Figura 14. Análise comparativa entre a primeira (representada em azul) e décima nona (representada em laranja) semanas de cultivo dos potenciais perfis metabólicos funcionais do consórcio bacteriano. Os dados foram analisados estatisticamente utilizando o software STAMP para diferenças entre proporções (%), utilizando G-teste + teste de Fisher para análise estatística, filtragem por q-value < 0,05 e correção de testes múltiplos usando o método FDR de *Benjamini-Hochber*. O banco de dados KEGG (nível 3) foi utilizado como referência.

5.5 Sequenciamento do DNA Total do Metagenoma em Plataforma Illumina® HiSeq 2500

O sequenciamento dos três *pools* do DNA metagenômico utilizando a plataforma Ilumina® Hiseq 2500, gerou um arquivo *paired-end*, no formato fastq, para cada *pool*. Foram gerados aproximadamente 8 Gb de dados. Essas sequências foram filtradas, totalizando 69.482.643 sequências *paired-end* e 8.535.551 sequências *single-end*.

5.6 Identificação, Montagem e Anotação do Genoma da *Dyella jiangningensis* FCAV SCS01

Apesar dos resultados da análise do 16S rRNA indicarem que *Burkholderia* é o gênero mais abundante presente no consórcio, a análise do sequenciamento do metagenoma total indicou a presença de diferentes espécies ou linhagens pertencentes à *Burkholderia*, dificultando o processo de montagem e recuperação desses genomas. Os genomas pertencentes a esse gênero serão explorados em trabalhos futuros.

Porém, foi possível recuperar o genoma de uma *Rhodanobacteraceae* que, de acordo com a análise do 16S rRNA sequenciado em plataforma Ion PGM™, representa a segunda família mais abundante presente no consórcio. Análises comparativas entre o 16S rRNA obtido por meio do sequenciamento na plataforma Ion PGM™ e do 16S rRNA recuperado por meio do sequenciamento do metagenoma total, realizado na plataforma HiSeq 2500, foram realizadas utilizando a ferramenta BLAST, revelando que as sequências pertencem ao mesmo gênero.

Esse genoma foi montado em 7 *scaffolds* e finalizado manualmente através de análises comparativas utilizando as ferramentas BLAST e SIS, frente a dois micro-organismos que foram utilizados como referência para ordenação da montagem: *Dyella jiangningensis* SBZ 3-12 (número de acesso do GenBank: CP007444.1) e *Dyella japonica* A8 (número de acesso do GenBank: NZ_CP008884.1). Através do uso dessas ferramentas, foi possível finalizar a sequência genômica em uma estrutura cromossômica circular, permitindo, então, a realização subsequente da anotação genômica e caracterização de enzimas relacionadas à degradação da biomassa.

Durante o processo de montagem e anotação, foi possível verificar que o genoma do consórcio pertencente ao gênero *Dyella* sp. apresenta 99% de similaridade da sequência do 16S rRNA e 90,77% de identidade de suas sequências codificadoras com a bactéria *Dyella jiangningensis* SBZ 3-12, sendo, portanto nomeada como *Dyella jiangningensis* FCAV SCS01.

5.7 Características do Genoma de *Dyella jiangningensis* FCAV SCS01

As características genômicas da bactéria *Dyella jiangningensis* FCAV SCS01 são apresentadas na tabela 6.

Tabela 6. Características Genômicas da *Dyella jiangningensis* FCAV SCS01.

Características Genômicas	
Tamanho do Genoma (pb)	4,758,639
DNA Codificante (pb)	4,164,575
DNA G+C (pb)	65.25
DNA *scaffolds*	7
- maior *contig*	2,446,838
- menor *contig*	146,103
- bases não identificadas (Ns)	586 (0.01%)
N50 *contig* (pb)	2,446,838
Total de Genes	4,250
Genes codificadores de proteínas	4,194
Genes de RNA	56
ncRNA	20
Repetição CRISPR	0
Elementos Genéticos Móveis	19 (395,605pb)
- sequências de inserção (transposases)	5 (5,712bp)
- regiões de profago	2 (63,205bp)
- ilhas genômicas	12 (326,688bp)
Genes com Predição Funcional	3,193
- função hipotética ou desconhecida	1,001
Genes com Subsistemas atribuídos	2,825
Genes atribuídos pelo COG	3,779
Genes com domínios Pfam	3,449
Genes com UniProtKB	2,393
Genes com peptídeos sinal	675
Genes com hélices transmembrana	1,034

5.8 Número de Genes Associados aos Subsistemas Gerados pela Plataforma RAST

A anotação realizada pela plataforma RAST também permitiu a identificação de genes relacionados aos subsistemas classificados (Tabela 7).

Tabela 7. Descrição dos subsistemas classificados através da plataforma RAST contendo o número de atribuições e porcentagens para cada subsistema.

Número de Genes	%	Descrição
1,369	32.64	Não estão em subsistemas
371	8.84	Aminoácidos e Derivados
363	8.65	Carboidratos
278	6.62	Metabolismo de Proteína
258	6.15	Cofatores, Vitaminas, Grupos Prostéticos, Pigmentos
162	3.86	Metabolismo de RNA
148	3.52	Transporte de Membrana
143	3.40	Respiração
141	3.36	Ácidos Graxos, Lipídeos e Isoprenóides
132	3.14	Resposta ao estresse
127	3.02	Parede Celular e Cápsula
126	3.01	Virulência, Doença e Defesa
103	2.45	Metabolismo de DNA
102	2.43	Motilidade e Quimiotaxia
66	1.57	Nucleosídeos e Nucleotídeos
54	1.28	Diversos
50	1.19	Metabolismo de Fósforo
48	1.14	Sinalização e Regulação celular
32	0.76	Divisão celular e Ciclo celular
32	0.76	Metabolismo de Enxofre
25	0.59	Metabolismo de Nitrogênio
24	0.57	Metabolismo de Componentes Aromáticos
17	0.40	Aquisição de ferro e Metabolismo
16	0.38	Metabolismo de Potássio
5	0.11	Metabolismo Secundário
1	0.02	Dormência e Esporulação
1	0.02	Fagos, Profagos, Elementos transposons, Plasmídeos

5.9 Visualização Genômica Circular de *Dyella jiangningensis* FCAV SCS01

O genoma da bactéria *Dyella jiangningensis* FCAV SCS01 é apresentado na Figura 15 através da ferramenta DNA Plotter.

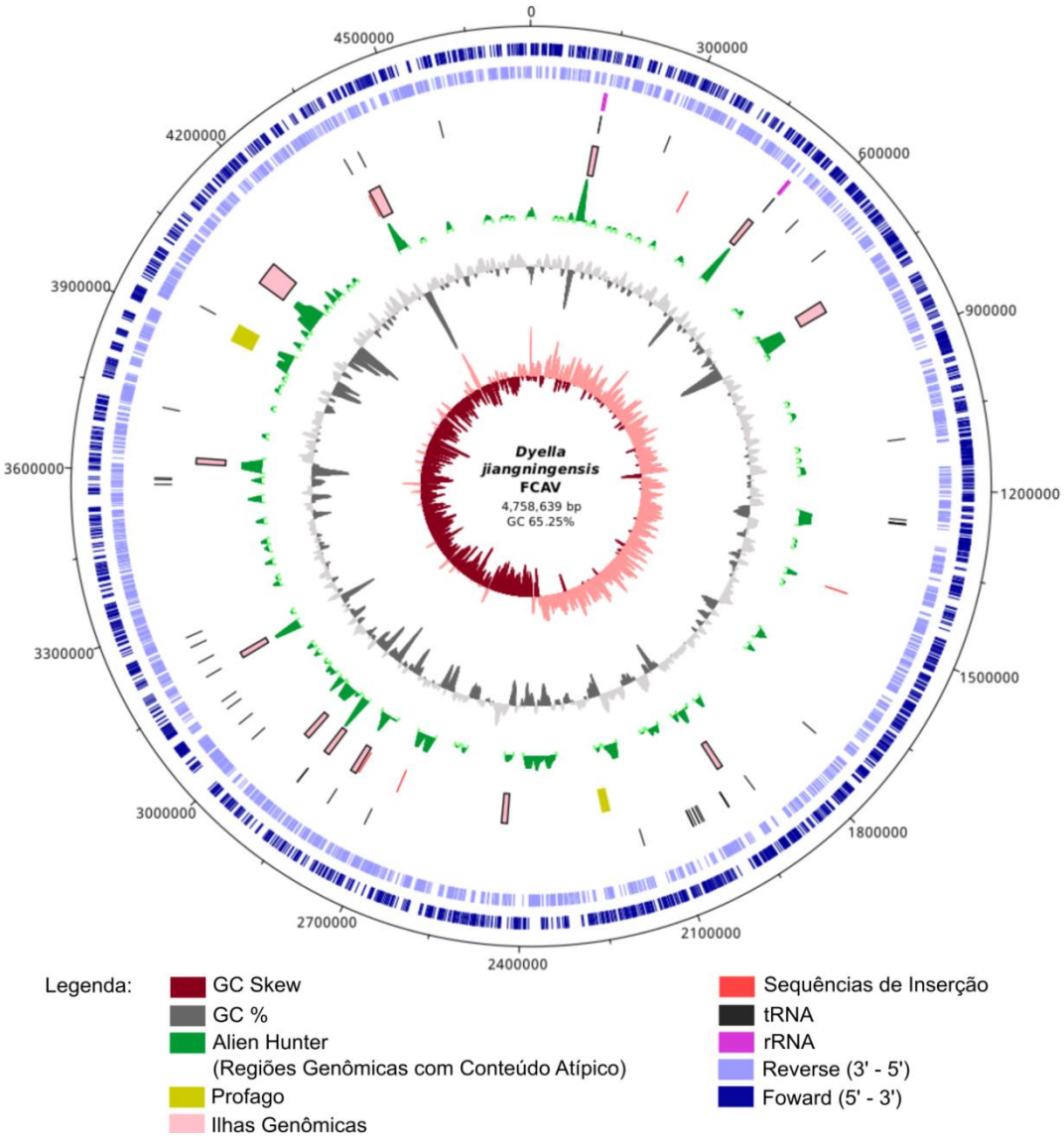

Figura 15. Genoma da bactéria *Dyella jiangningensis* FCAV SCS01 gerado através da ferramenta DNA Plotter. A partir dos círculos presentes de dentro para fora da figura: GC Skew ((G-C) / (G+C)) (vermelho escuro), conteúdo GC % (cinza), Alien Hunter, Regiões Genômicas com Conteúdo Atípico (verde), regiões de Profago (amarelo), Ilhas Genômicas (rosa claro), Sequências de Inserção (vermelho), tRNA (cinza escuro), rRNA (rosa escuro), regiões codificadoras no sentido 3’ → 5’ (azul claro) e 5’ → 3’ (azul escuro).

5.10 Enzimas com Potencial Degradador de Biomassa Encontradas no Genoma de *Dyella jiangningensis* FCAV SCS01

Através da anotação do genoma realizada pela plataforma RAST, foi possível fazer buscas de enzimas com potencial degradador de biomassa relacionadas aos subsistemas classificados.

Foram identificadas 36 enzimas pertencentes à família de Glicosil Hidrolases e 6 enzimas pertencentes à classificação de Atividades Auxiliares (AA), corroborando o potencial de *Dyella jiangningensis* FCAV SCS01 em degradar a biomassa lignocelulósica. A Tabela 8 apresenta a lista de enzimas encontradas no genoma de *Dyella jiangningensis* FCAV SCS01. As enzimas foram separadas conforme os substratos em que catalisam a hidrólise.

Tabela 8. Enzimas com potencial degradador de biomassa encontradas no genoma de *Dyella jiangningensis* FCAV SCS01 através da classificação de subsistemas realizadas pela plataforma RAST. A tabela apresenta o nome, número EC da enzima, as famílias GHn e AA onde são encontradas e a quantidade de enzimas identificadas no genoma da bactéria.

Enzimas	**Número EC**	**Família GH**	**Quantidade de Enzimas**
(Hemi)celulose			
α-amilase	3.2.1.1	13, 14, 57, 119	1
α-galactosidase precursor	3.2.1.22	4, 27, 31, 36, 57, 97, 110	1
α-glicosidase	3.2.1.20	4, 13, 31, 63, 97, 122	2
α-glucuronidase	3.2.1.139	67	1
α-L-fucosidase	3.2.1.51	29, 95	3
α-N-acetilgalactosaminidase	3.2.1.49	36	1
α-N-acetilglicosaminidase	3.2.1.50	89	1
α-1,2-manosidase	3.2.1.24	31, 38, 92	7
β-galactosidase	3.2.1.23	1, 2, 3, 35, 42, 50	3
β-manosidase	3.2.1.25	1, 2, 5	2
β-xilosidase	3.2.1.37	3, 30, 39, 43, 52, 54	2
Glicoamilase	3.2.1.3	15, 97	1
Maltodextrina glicosidase	3.2.1.20	4, 13	2
Malto-oligosiltrealose trealohidrolase	3.2.1.141	13	1
Trealase	3.2.1.28	13, 15, 37, 65	1
Xilanase	3.2.1.8	5, 8, 10, 11, 43	1
Quitina			
β-hexosaminidase	3.2.1.52	3, 5, 18, 20, 84, 116	4
Quitinase	3.2.1.14	18, 19	2
Total de GHs	-	-	**36**
Lignina	**Número EC**	**Família AA**	**Quantidade de Enzimas**
Catalase	1.11.1.6	AA 2	2
Peroxidase	1.11.1.7	AA 2	1
Tiol Peroxidase	1.11.1.15	AA 2	3
Total de AAs	-	-	**6**

6. DISCUSSÃO

Os resultados do presente estudo indicam que os gêneros *Burkholderia* (*Burkholderiaceae*) e *Dyella* (*Rhodanobacteraceae*) são os mais abundantes no consórcio estudado.

As análises do DNA total do metagenoma sequenciado permitiram recuperar o genoma completo da bactéria *Dyella jiangningensis* FCAV SCS01. Apesar de *Burkholderia* ser o gênero mais abundante, as análises do sequenciamento do metagenoma total sugerem que nesse consórcio há presença de diferentes cepas desse gênero, dificultando o processo de montagem e recuperação desse genoma. Diversos outros micro-organismos foram identificados em menores abundâncias no consórcio, e o perfil metabólico e de predição funcional sugerem um enriquecimento para os subsistemas relacionados ao transporte, em particular transportadores ABC. Em linhas gerais, os resultados obtidos também corroboram o potencial que ferramentas de sequenciamento de alto rendimento, metagenômica e bioinformática, apresentam para o estudo de comunidades microbianas com potencial biotecnológico, mostrando-se como uma valiosa alternativa para a investigação de novos alvos para a pesquisa em bioenergia.

6.1 Provável Papel do Gênero *Burkholderia* Presente no Consórcio

O gênero *Burkholderia* apresenta abundância de 98,82% na primeira semana, diminuindo para 49,50% na décima terceira semana, sua menor abundância detectada. O gênero *Dyella* demonsta comportamento distinto, aumentando de 0,004% na primeira semana, para 41,37% na décima terceira semana. A abundância do gênero *Burkholderia* volta a aumentar, após a décima quinta semana, alcançando 87,66% na décima nona semana, enquanto que *Dyella* volta a diminuir, chegando a 6,50%. Isso pode demonstrar que existe uma relação inversamente proporcional entre esses organismos dentro desse consórcio bacteriano, e que ambos podem apresentar interação entre si (sinérgica ou antagônica) e, portanto fundamental na degradação de material lignocelulósico.

O gênero *Burkholderia* é composto por bactérias Gram-negativas bacilares com a capacidade de crescer em substratos que contenham uma única fonte de carbono, como glicerol, galactose, glicose, entre outros (ROSENBERG et al., 2014). As espécies atribuídas a esse gênero ocupam diversos ambientes, desde solo, raízes, até animais, e podem também desempenhar diferentes papéis em seu hábitat (COENYE; VANDAMME, 2003). Bem como as espécies *B. kururiensis*, *B. phenoliruptrix*, *B. sartisoli*, *B. unamae* e *B. xenovorans* são relatadas como capazes de degradar diversos monômeros aromáticos associados à lignina (AKITA et al., 2016).

Segundo Akita et al. (2016), espécies do gênero *Burkholderia*, isoladas de solo de floresta tropical, possuem capacidade de crescer em meio líquido contendo lignina alcalina como única fonte de carbono e de degradar monômeros aromáticos associados à lignina, principalmente *p*-hidroxibenzenos, assim como a linhagem *B. cepacia* KK01, que apresenta capacidade de degradar eficientemente o-cresol, guaiacol, fenol e ácido vanílico (AKITA et al., 2016).

Estudos anteriores sugerem que *Burkholderia* spp., isoladas do solo de floresta tropical possam apresentar potencial de degradação da lignina devido à presença de dois genes relacionados às enzimas multicobre oxidases e 22 genes supostos para peroxidases ou catalases (WOO et al., 2014b).

O *draft* do genoma da cepa *Burkholderia* sp. LIG30 (JGVW00000000.1), apresenta 5,5 Mb e conteúdo GC de 66,4%, codificando 4.996 genes. Para 73% destes genes foi possível atribuir uma função através do banco de dados COGs (*Cluster of Orthologous Groups*) (WOO et al., 2014b). Dentre os genes que apresentaram função, 305 foram associados à biossíntese de metabólitos secundários, transporte e catabolismo (WOO et al., 2014b). Além disso, espécies pertencentes ao gênero *Burkholderia*, também podem secretar enzimas extracelulares que são capazes de degradar os componentes da biomassa lignocelulósica, pois apresentam atividades enzimáticas de *β*-D-glicosidase (EC 3.2.1.21), celobio-hidrolase (EC 3.2.1.91), *β*-xilopiranosidase (EC 3.2.1.37) e peroxidase (EC 1.11.1.7) (WOO et al., 2014a). Portanto, estudos de Akita et al. (2016) e Woo et al. (2014a), corroboram a hipótese do gênero *Burkholderia* se

apresentar como um agente degradador da biomassa e com capacidade de degradar material lignocelulósico proveniente do bagaço da cana-de-açúcar.

6.2 O Genoma da *Dyella jiangningensis* FCAV SCS01 e seu Possível Papel na Degradação da Biomassa

O gênero *Dyella* foi proposto em 2015 por Xie & Yokota, após isolarem três linhagens de bactérias de solo de jardim e analisá-las através do método de taxonomia polifásica, ou seja, fazendo testes considerando não somente as sequências de 16S rRNA, mas também as características fisiológicas, a composição de ácidos graxos, o conteúdo GC e a hibridização DNA-DNA entre os isolados (XIE; YOKOTA, 2005). Apesar dos isolados serem relatados como próximos a membros dos gêneros *Fulvimonas*, *Frateuria* e *Rhodanobacter* (organismos pertencentes à família *Rhodanobacteraceae*), com base nesse estudo, foi possível concluir que os isolados não pertenciam a esses gêneros (NAUSHAD et al., 2015; XIE; YOKOTA, 2005). Portanto, foi proposta a classificação de um novo gênero e uma nova espécie, a bactéria denominada *Dyella japonica* (XIE; YOKOTA, 2005).

Já a linhagem de SBZ3-12, classificada em 2013 por Zhao e colaboradores, como uma nova espécie, pertencente ao gênero *Dyella*, foi inicialmente isolada da superfície de rocha contendo potássio e, após análises taxonômicas seguindo a abordagem polifásica, foi classificada como *Dyella jiangningensis*, em homenagem ao local de onde foi isolada, distrito de Jiangning, na China (ZHAO et al., 2013). Os testes realizados revelaram que a linhagem de SBZ3-12, da espécie de *Dyella jiangningensis*, apresentou as seguintes características: faz parte do grupo de bactérias Gram-negativas, aeróbias, que possui motilidade, não são formadoras de esporos e que possui forma de haste (ZHAO et al., 2013). Além disso, foram realizados testes de atividades enzimáticas que revelaram que a *Dyella jiangningensis* SBZ3-12 possui atividades positivas para *α*-galactosidase, *β*-galactosidase, *α*-glicosidase, *β*-glicosidase e N-acetil-*β*-glicosaminidase (ZHAO et al., 2013).

Os resultados da montagem do organismo selecionado a partir do sequenciamento do metagenoma total do consórcio bacteriano e as análises

seguintes classificaram o gênero pertencente à família *Rhodanobacteraceae* como *Dyella*, sendo que a sequência do 16S rRNA apresenta 99% de similaridade com a bactéria *Dyella jiangningensis* SBZ3-12. Assim, foi atribuído à espécie *Dyella jiangningensis* FCAV SCS01.

Segundo a anotação funcional gerada pela plataforma RAST, a bactéria em estudo, *Dyella jiangningensis* FCAV SCS01, isolada de um consórcio bacteriano proveniente de solo contendo bagaço de cana-de-açúcar em decomposição, possui um total de 36 genes atribuídos à família de GHs que podem codificar para enzimas relacionadas à degradação da biomassa, incluindo as enzimas relacionadas às mesmas atividades apresentadas por *Dyella jiangningensis* SBZ3-12, com exceção de *β*-glicosidase e N-acetil-*β*-glicosaminidase.

A enzima *α*-galactosidase (EC 3.2.1.22) catalisa hidrólises de ligações *α*-D-galactosídicas encontradas em oligossacarídeos, liberando resíduos de *α*-D-galactose (CAZy). Segundo Bayer et al. (2013), essa enzima faz parte de um conjunto enzimático capaz de degradar heteropolissacarídeos ramificados, glicomananas e galactoglicomananas, presentes em madeiras duras e moles (“*hardwood”* e “*softwood”*, respectivamente) (BAYER; SHOVAM; LAMED, 2013). Além disso, os autores citam que a enzima *β*-galactosidase realiza atividades de degradação de polissacarídeos da parede celular vegetal e faz parte das famílias GH 1, GH 2 e GH 42 (BAYER; SHOVAM; LAMED, 2013). A enzima *β*-galactosidase catalisa hidrólises em ligações *β*-D-galactosídicas, liberando resíduos de *β*-D-galactose (CAZy).

Dyella jiangningensis FCAV SCS01 apresenta um potencial precursor de *α*-glicosidase (EC 3.2.1.20), enzima que catalisa hidrólise de ligações *α*-D-glicosídicas, liberando resíduos de D-glicoses (CARVALHO et al., 2010). Essa enzima tem capacidade de degradar amido, especialmente na fase final de degradação (REISER; GASPERÍK, 1995), assim como de degradar celulose (HALGREN; BANOWETZ, 2012).

Além disso, as outras potenciais enzimas reportadas na bactéria *Dyella jiangningensis* FCAV SCS01 também possuem atividades interessantes para a degradação de material lignocelulolítico, como:

Xilanases, que atuam hidrolisando as cadeias principais da hemicelulose e *β*-xilosidase, que hidrolisam xilooligossacarídeos em xilose. Essas enzimas também podem se associar às hemicelulases e *α*-L-arabinofuranosidase, que clivam ligações *α*-L-arabinofuranosídicas em polissacarídeos que contenham arabinoxilana ou outro componente de L-arabinose (BAYER; SHOVAM; LAMED, 2013); e *β*-manosidase, que possuem a capacidade de clivar as ligações de D-mananas e D-galacto-mananas presentes em heteropolissacarídeos ramificados (glicomananas e galactoglicomananas), encontrados em madeiras duras e moles (BAYER, E. A.; SHOVAM, Y.; LAMED, R., 2013).

Portanto, esses resultados sugerem que a bactéria *Dyella jiangningensis* FCAV SCS01 apresenta potencial em degradar a biomassa lignocelulósica.

6.3 Considerações Sobre os Outros Gêneros Bacterianos Identificados no Consórcio

Espécies atribuídas à família *Isosphaeraceae*, pertencentes aos gêneros *Isosphaera* (*Isosphaera pallida* ATCC 43644) e *Singulisphaera* (*Singulisphaera acidiphila* DSM 18658) codificam genes pertencentes às famílias Glicosil Hidrolases: *Isosphaera pallida* ATCC 43644 possui uma celulase (GH5) e uma glicoamilase (GH15), e *Singulisphaera acidiphila* DSM 18658, glicoamilase, *β*-galactosidase/*β*-glucuronidase (GH2), quitinase (GH18), *β*-galactosidase (GH1), uma possível *β*-xilosidase (GH43) e possível glicosilase (GH43, GH62, GH32, GH68) (CAZy - http://www.cazy.org/).

O gênero *Rhizobium*, é capaz de produzir enzimas com potencial degradador dos componentes da parede celular vegetal, pois apresenta enzimas celulolíticas e pectinolíticas que agem rompendo as ligações glicosídicas (HUANG; SHENG; ZHANG, 2012). Essas enzimas são importantes no processo de infecção simbiótica primária sobre as raízes da leguminosa hospedeira para que a bactéria simbionte possa penetrar na célula e formar o nódulo na raiz (ROBLEDO et al., 2008), *Rhizobium* de vida livre geralmente são isoladas do solo, da rizosfera e de raízes de plantas (BERGE et al., 2009).

O gênero *Cryocola*, é descrito como desnitrificante natural, sendo comumente isolado do solo com cultivo de cana-de-açúcar contendo tratamento com fertilizante sintético, nitrogênio, utilizado para aumentar o rendimento da plantação (YEOH et al., 2016). Micro-organismos presentes em solos contendo fertilizante sintético de nitrogênio podem utilizá-lo como fonte de energia através da metabolização do nitrogênio, podendo causar a diminuição desse componente no solo (YEOH et al., 2016).

Linhagens de *Cryocola antiquus*, isoladas de solo de tundra subártica, na Finlândia, foram inoculadas em placas de ágar contendo uma mistura de fontes de carbono (D-frutose, D-glicose, sucrose, D-manitol, ácido succínico e ácido-DL-málico em razões molar de 2∶2: 2∶2: 1∶1) em ausência de fontes de nitrogênio, e foi possível observar o desenvolvimento dessas bactérias, apresentando crescimento lento (HARA et al., 2012).

Assim, esses gêneros identificados no consórcio bacteriano podem apresentar capacidade de degradar a biomassa lignocelulósica ou de utilizar os açúcares liberados pelo processo de degradação como fonte de energia.

6.4 Predição do Perfil Metabólico do Consórcio Através da Análise do 16S rRNA Sugere que as Bactérias são Capazes de Utilizar os Açúcares Disponíveis no Meio de Cultivo Contendo Bagaço de Cana-de-açúcar Triturado

Através da predição de potenciais funções da comunidade bacteriana a partir do 16S rRNA e das análises estatísticas dos perfis funcionais, foi possível observar que os perfis funcionais do consórcio bacteriano foram supostamente enriquecidos para os subsistemas “Transportadores” e “Transportadores ABC”.

Esses resultados sugerem que, durante o início do repique do consórcio com bagaço de cana-de-açúcar, os micro-organismos presentes passam por um processo de adaptação, envolvendo a sinalização e possível metabolização do material lignocelulósico presente no meio de cultivo. Demais funções de *house-keeping* permanecem equivalentes ao longo das semanas estudadas.

Além disso, o presente trabalho sugere que as bactérias podem utilizar os açúcares disponíveis no meio de cultivo, através dos transportadores de açúcares,

que estão classificados em três categorias principais, sendo o maior grupo formado por transportadores ABC, ou cassete de ligação a ATP, que requer uma molécula de ATP, transportando açúcares como a xilose e arabinose (ACTON, 2013). O segundo grupo de transportadores, classificado como proteínas pertencentes à superfamília dos facilitadores maioritários (MFS – "*Major Facilitator Super Family"*), facilita o transporte de açúcares (como galactose, glicose, arabinose e xilose) para dentro da célula sem a necessidade de utilizar uma molécula de ATP para cada molécula de açúcar transportada (ACTON, 2013). Já o terceiro grupo de transportadores, classificado como sistema de fosfotransferase (PTS – "*Phosphotransferase System"*), um transporte ativo utilizado por bactérias, é responsável por captar açúcares como a glicose para o interior da célula (ACTON, 2013). Assim, os resultados sugerem que as bactérias são capazes de utilizar os açúcares disponíveis no meio de cultivo contendo bagaço de cana-de-açúcar triturado.

7. CONCLUSÃO

Através das análises foi possível concluir que os principais gêneros com potencial degradador da biomassa, encontrados no consórcio bacteriano proveniente do solo, contendo bagaço de cana-de-açúcar em decomposição, foram *Burkholderia* e *Dyella*, representados abundantemente ao longo das semanas de cultivo estudadas, apresentando uma relação inversamente proporcional entre ambos dentro desse consórcio bacteriano, sugerindo uma interação entre si (sinérgica ou antagônica) no processo de degradação da biomassa. As análises de diversidade e análises de predição funcional sugerem que as bactérias identificadas nesse consórcio possuem capacidade de desconstruir a biomassa lignocelulósica e consequentemente liberar açúcares. Revelando, assim, que o consórcio bacteriano em estudo tem potencial para utilização da biomassa lignocelulósica.

O organismo selecionado foi identificado como uma nova linhagem de *Dyella jiangningensis* FCAV SCS01 e apresentou enzimas com potencial de degradação do material lignocelulítico e potencial para diversas aplicações biotecnológicas.

8. REFERÊNCIAS

ACTON, Q. A. **Advances in Proteobacteria Research and Application**. Ed. ScholarlyEditions, 2013. p. 175 - 181.

AKHTER, S.; AZIZ, R. K.; EDWARDS, R. A. PhiSpy: a novel algorithm for finding prophages in bacterial genomes that combines similarity- and composition-based strategies. **Nucleic Acids Research**, v. 40, n. 16, p. 126, set. 2012. Disponível em: <10.1093/nar/gks406>

AKITA, H.; KIMURA, Z.; YOSSOF, M. Z. M.; NAKASHIMA, N; HOSHINO, T. Isolation and characterization of *Burkholderia* sp. strain CCA53 exhibiting ligninolytic potential. **SpringerPlus**, v. 5, n. 1, p. 596, 11 maio 2016. Disponível em: <10.1186/s40064-016-2237-y>

ALVARENGA, D. O.; FIORE, M. F.; VARANI, A. M. A Metagenomic Approach to Cyanobacterial Genomics. **Frontiers in Microbiology**, v. 8, p. 809, 2017. Disponível em: <10.3389/fmicb.2017.00809>

ALVES-PRADO, H. F.; PAVEZZI, F. C.; LEITE, R. S.; DE OLIVEIRA, V. M.; SETTE, L. D.; DASILVA, R. Screening and production study of microbial xylanase producers from Brazilian Cerrado. **Applied Biochemistry and Biotechnology**, v. 161, n. 1–8, p. 333–346, maio 2010. Disponível em: <10.1007/s12010-009-8823-5>

ANTIPOV, D.; HARTWICK, N.; SHEN, M.; RAIKO, M.; LAPIDUS, A.; PEVZNER, P. A. plasmidSPAdes: assembling plasmids from whole genome sequencing data. **Bioinformatics (Oxford, England)**, v. 32, n. 22, p. 3380–3387, 15 nov. 2016. Disponível em: <10.1093/bioinformatics/btw493>

AQUINO, A. F.; BIDÔ, E. S; GALVÃO, M. L. DE M.; OLIVEIRA, V. N. O Etanol da Cana-de-açúcar: Possibilidades Energéticas da Região de Ceará-Mirim-RN. **Holos**, v. 1, n. 30, p. 105–125, 2014. Disponível em: <http://dx.doi.org/10.15628/holos.2014.713>

AZIZ, R. K.; BARTELS, D.; BEST, A. A.; DEJONGH, M.; DISZ, T.; EDWARDS, R. A.; FORMSMA, K.; GERDES, S.; GLASS, E. M.; KUBAL, M.; MEYER, F.; OLSEN, G. J.; OLSON, R.; OSTERMAN, A. L.; OVERBEEK, R. A.; MCNEIL, L. K.; PAARMANN, D.; PACZIAN, T.; PARRELLO, B.; PUSCH, G. H.; REICH, C.; STEVENS, R.; VASSIEVA, O.; VONSTEIN, V.; WILKE, A.; ZAGNITKO, O. The RAST Server: rapid annotations using subsystems technology. **BMC genomics**, v. 9, p. 75, 8 fev. 2008. Disponível em: <10.1186/1471-2164-9-75>

BAYER, E. A.; SHOVAM, Y.; LAMED, R. Lignocellulose-Decomposing Bacteria and Their Enzyme Systems. In: ROSENBERG, E. et al. (Eds.). **The Prokaryotes: Prokaryotic Physiology and Biochemistry**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013. p. 215–266.

BERGE, O.; LODHI, A.; BRANDELET, G.; SANTAELLA, C.; RONCATO, M. A.; CHRISTEN, R.; HEULIN, T.; ACHOUAK, W. *Rhizobium alamii* sp. nov., an exopolysaccharide-producing species isolated from legume and non-legume rhizospheres. **International Journal of Systematic and Evolutionary Microbiology**, v. 59, n. Pt 2, p. 367–372, fev. 2009. Disponível em: <10.1099/ijs.0.000521-0>

BOLGER, A. M.; LOHSE, M.; USADEL, B. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics (Oxford, England)**, v. 30, n. 15, p. 2114–2120, 1 ago. 2014. Disponível em: <https://doi.org/10.1093/bioinformatics/btu170>

BRETHAUER, S.; STUDER, M. H. Biochemical Conversion Processes of Lignocellulosic Biomass to Fuels and Chemicals - A Review. **Chimia**, v. 69, n. 10, p. 572–581, 2015. Disponível em: <10.2533/chimia.2015.572>

BUSHNELL, L. D.; HAAS, H. F. The Utilization of Certain Hydrocarbons by Microorganisms. **Journal of Bacteriology**, v. 41, n. 5, p. 653–673, maio 1941.

CARVALHO, A. F. A.; COUTINHO, P. M.; RANCUREL, C.; BERNARD, T.; LOMBARD, V.; HENRISSAT, B. Purification and characterization of the *α*-glucosidase produced by thermophilic fungus *Thermoascus aurantiacus* CBMAI 756. **The Journal of Microbiology**, v. 48, n. 4, p. 452–459, 1 ago. 2010. Disponível em: <10.1007/s12275-010-9319-2>

CARVER, T.; THOMSON, N.; BLEASBY, A.; BERRIMAN, M.; PARKHILL, J. DNAPlotter: circular and linear interactive genome visualization. **Bioinformatics**, v. 25, n. 1, p. 119–120, 1 jan. 2009. Disponível em: <10.1093/bioinformatics/btn578>

CASTRO, A. M. DE; PEREIRA JR, N. Produção, propriedades e aplicação de celulases na hidrólise de resíduos agroindustriais. **Química Nova**, v. 33, n. 1, p. 181–188, 2010. Disponível em: <http://dx.doi.org/10.1590/S0100-40422010000100031>

CGEE. **Bioetanol Combustível: Uma Oportunidade Para O Brasil**. Brasília: CGEE - Centro de Gestão e Estudos Estratégicos, 2009.

CHAKRABORTY, S.; AGGARWAL, V.; MUKHERJEE, D.; ANDRAS, K. Biomass to biofuel: a review on production technology. **Asia-Pacific Journal of Chemical Engineering**, v. 7, p. S254–S262, 1 ago. 2012. Disponível em: <10.1002/apj.1642>

COENYE, T.; VANDAMME, P. Diversity and significance of *Burkholderia* species occupying diverse ecological niches. **Environmental Microbiology**, v. 5, n. 9, p. 719–729, set. 2003. Disponível em: <10.1046/j.1462-2920.2003.00471.x>

CONAB. **Acompanhamento da Safra Brasileira: Cana-de-Açúcar**; Safra 2015/16. v. 2, n. 4, p. 71, 2016.

COOPER, G. M.; HAUSMAN, R. E. **A Célula - 3ed: Uma Abordagem Molecular**. Artmed Editora, 2016.

CORTES-TOLALPA, L.; JIMÉNEZ, D. J.; DE LIMA BROSSI, M. J.; SALLES, J. F.; VAN ELSAS, J. D. Different inocula produce distinctive microbial consortia with similar lignocellulose degradation capacity. **Applied Microbiology and Biotechnology**, v. 100, n. 17, p. 7713–7725, set. 2016. Disponível em: <10.1007/s00253-016-7516-6>

CORTEZ, L. A. B. **Bioetanol de Cana-de-Açúcar: P&D Para Produtividade e Sustentabilidade**. São Paulo: Blucher, 2010.

COSTA, J. A. V.; DE MORAIS, M. G. The role of biochemical engineering in the production of biofuels from microalgae. **Bioresource Technology**, v. 102, n. 1, p. 2–9, jan. 2011. Disponível em: <https://doi.org/10.1016/j.biortech.2010.06.014>

D'ALMEIDA, M. L. O.; KOGA, M. E. T.; FERREIRA, D. C.; PIGOZZO, R. J. B. Composição Química da Madeira e Matérias-Primas Fibrosas. In: SENAI (SP). **Celulose.** São Paulo: SENAI, 2013. cap. 1, p. 11–58. Série Informações Tecnológicas: Área de Celulose e Papel.

DE ALMEIDA, M. L. M. **Diversidade Bacteriana de Um Consórcio Degradador de Resíduo Celulósico da Biomassa.** 2011. Relatório Científico Final - Câmpus de Jaboticabal: Universidade Estadual Paulista "Júlio de Mesquita Filho" - Faculdade de Ciências Agrárias e Veterinárias (FCAV), Jaboticabal, 2011.

DESANTIS, T. Z.; HUGENHOLTZ, P.; LARSEN, N.; ROJAS, M.; BRODIE, E. L.; KELLER, K.; HUBER, T.; DALEVI, D.; HU, P.; ANDERSEN, G. L. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. **Applied and Environmental Microbiology**, v. 72, n. 7, p. 5069–5072, jul. 2006. Disponível em: <10.1128/AEM.03006-05>

DIAS, Z.; DIAS, U.; SETUBAL, J. C. SIS: a program to generate draft genome sequence scaffolds for prokaryotes. **BMC Bioinformatics**, v. 13, p. 96, 2012. Disponível em: <10.1186/1471-2105-13-96>

DING, S.-Y.; HIMMEL, M. E. The maize primary cell wall microfibril: a new model derived from direct visualization. **Journal of Agricultural and Food Chemistry**, v. 54, n. 3, p. 597–606, 8 fev. 2006. Disponível em: <10.1021/jf051851z>

EIA. **International Energy Outlook 2016, With Projections to 2040 - Cap. 8 - 9, Energy-related CO2 emissions.** U. S. Energy Information Administration, 2016. Disponível em: <https://www.eia.gov/outlooks/ieo/pdf/0484(2016).pdf>. Acesso em: 4 maio. 2017.

EMBRAPA. **Biomassa Para Química Verde**. 1. ed. Brasília: Embrapa Agroenergia; Sílvio Vaz Júnior, 2013.

GANESAN, H. et al. The SeqWord Genome Browser: an online tool for the identification and visualization of atypical regions of bacterial genomes through oligonucleotide usage. **BMC Bioinformatics**, v. 9, p. 333, 7 ago. 2008. Disponível em: <10.1186/1471-2105-9-333>

GHOSH, S.; CHOWDHURY, R.; BHATTACHARYA, P. Mixed consortia in bioprocesses: role of microbial interactions. **Applied Microbiology and Biotechnology**, v. 100, n. 10, p. 4283–4295, maio 2016. Disponível em: <10.1007/s00253-016-7448-1>

GÍRIO, F. M.; CARVALHEIRO, F.; DUARTE, L. C.; MARQUES, S.; BOGEL-LUKASIK, R. Hemicelluloses for fuel ethanol: A review. **Bioresource Technology**, v. 101, n. 13, p. 4775–4800, jul. 2010. Disponível em: <10.1016/j.biortech.2010.01.088>

GU, T.; HELD, M. A.; FAIK, A. Supercritical CO2 and ionic liquids for the pretreatment of lignocellulosic biomass in bioethanol production. **Environmental Technology**, v. 34, n. 13–14, p. 1735–1749, 1 jul. 2013. Disponível em: <10.1080/09593330.2013.809777>

HAGHIGHI MOOD, S. et al. Lignocellulosic biomass to bioethanol, a comprehensive review with a focus on pretreatment. **Renewable and Sustainable Energy Reviews**, v. 27, p. 77–93, 1 nov. 2013. Disponível em: <https://doi.org/10.1016/j.rser.2013.06.033>

HALGREN, A.; BANOWETZ, G. M. Life cycle expression analysis of three cell wall degradation-related genes in ethylene-treated grass. **Plant Growth Regulation**, v. 66, n. 2, p. 167–177, 1 mar. 2012. Disponível em: <10.1007/s10725-011-9641-7>

HANDELSMAN, J.; RONDON, M. R.; BRADY, S. F.; CLARDY, J.; GOODMAN, R. M. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. **Chemistry & Biology**, v. 5, n. 10, p. R245-249, out. 1998. Disponível em: <https://doi.org/10.1016/S1074-5521(98)90108-9>

HARA, S.; ISODA, R.; TAHVANAINEN, T.; HASHIDOKO, Y. Trace Amounts of Furan-2-Carboxylic Acids Determine the Quality of Solid Agar Plates for Bacterial Culture. **PLOS ONE**, v. 7, n. 7, p. e41142, 27 jul. 2012. Disponível em: <https://doi.org/10.1371/journal.pone.0041142>

HO, D. P.; NGO, H. H.; GUO, W. A mini review on renewable sources for biofuel. **Bioresource Technology**, v. 169, p. 742–749, out. 2014. Disponível em: <https://doi.org/10.1016/j.biortech.2014.07.022>

HUANG, S.; SHENG, P.; ZHANG, H. Isolation and Identification of Cellulolytic Bacteria from the Gut of *Holotrichia parallela* Larvae (Coleoptera: Scarabaeidae). **International Journal of Molecular Sciences**, v. 13, n. 3, p. 2563–2577, 23 fev. 2012. Disponível em: <10.3390/ijms13032563>

IEA. **Energy and Climate Change - World Energy Outlook Special Report; Cap. 4.** IEA - International Energy Agency, 2015. Disponível em: <https://www.iea.org/publications/freepublications/publication/WEO2015SpecialReportonEnergyandClimateChange.pdf>. Acesso em: 4 jun. 2017

IPEA. Biocombustível no Brasil: Etanol e Biodiesel. **Eixos do Desenvolvimento Brasileiro.** n. 53, p. 57, 2010.

KAJITANI, R.; TOSHIMOTO, K.; NOGUCHI, H.; TOYODA, A.; OGURA, Y.; OKUNO, M.; YBANA, M.; HARADA, M.; NAGAYASU, E.; MARUYAMA, H.; KOHARA, Y.; FUJIYAMA, A.; HAYASHI, T.; ITOH, T. Efficient de novo assembly of highly heterozygous genomes from whole-genome shotgun short reads. **Genome Research**, v. 24, n. 8, p. 1384–1395, 1 ago. 2014. Disponível em: <10.1101/gr.170720.113>

KANG, Q.; APPELS, L.; TAN, T.; DEWIL, R. Bioethanol from Lignocellulosic Biomass: Current Findings Determine Research Priorities. **The Scientific World Journal**, v. 2014, 2014. Disponível em: <http://dx.doi.org/10.1155/2014/298153>

KOHLHEPP, G. Análise da situação da produção de etanol e biodiesel no Brasil. **Scielo.** v. 24, n. 68, p. 223–253, 2010. Disponível em: <http://dx.doi.org/10.1590/S0103-40142010000100017>

KUMAR, S.; KRISHNANI, K. K.; BRUSHAN, B.; BRAHMANE, M. P. Metagenomics: Retrospect and Prospects in High Throughput Age. **Biotechnology Research International**, v. 2015, p. 1-13, 17 nov. 2015. Disponível em: <http://dx.doi.org/10.1155/2015/121735>

LANGILLE, M. G. I.; ZANEVELD, J.; CAPORASO, G.; MCDONALD, D.; KNIGHTS, D.; REYES, J. A.; CLEMENTE, J. C.; BURKEPILE, D. E.; THURBER, R. L. V.; KNIGHT, R.; BEIKO, R. G.; HUTTENHOWER, C. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. **Nature Biotechnology**, v. 31, n. 9, p. 814–821, set. 2013. Disponível em: <10.1038/nbt.2676>

LE FLOCH, A.; JOURDES, M.; TEISSEDRE, P.-L. Polysaccharides and lignin from oak wood used in cooperage: Composition, interest, assays: A review. **Carbohydrate Research**, v. 417, p. 94–102, 19 nov. 2015. Disponível em: <https://doi.org/10.1016/j.carres.2015.07.003>

LEAL, M. R. L. V.; LARSON, E. D.; WILLIAMS, R. H. A review of biomass integrated-gasifier/gas turbine combined cycle technology and its application in sugarcane industries, with an analysis for Cuba. **Energy for Sustainable Development.** v. 5, n. 1, p. 54–76, 2001. Disponível em: <https://doi.org/10.1016/S0973-0826(09)60021-1>

LIMA, M. T. **Análise Funcional de um Consórcio Microbiano de Solo e Prospecção de Genes Envolvidos na Desconstrução da Biomassa**. 2014. Dissertação de Mestrado - Câmpus de Jaboticabal: Faculdade de Ciências Agrárias e Veterinárias - Universidade Estadual Paulista "Júlio de Mesquita Filho", 2014.

LOMBARD, V. et al. The carbohydrate-active enzymes database (CAZy) in 2013. **Nucleic Acids Research**, v. 42, n. Database issue, p. D490-495, jan. 2014. Disponível em: <10.1093/nar/gkt1178>

LOZUPONE, C. A.; HAMADY, M.; KELLEY, S. T.; KNIGHT, R. Quantitative and Qualitative β Diversity Measures Lead to Different Insights into Factors That Structure Microbial Communities. **Applied and Environmental Microbiology**, v. 73, n. 5, p. 1576–1585, 1 mar. 2007. Disponível em: <10.1128/AEM.01996-06>

LOZUPONE, C.; KNIGHT, R. UniFrac: a New Phylogenetic Method for Comparing Microbial Communities. **Applied and Environmental Microbiology**, v. 71, n. 12, p. 8228–8235, dez. 2005. Disponível em: <10.1128/AEM.71.12.8228-8235.2005>

LYND, L. R.; WEIMER, P. J.; VAN ZYL, W. H.; PRETORIUS, I. S. Microbial Cellulose Utilization: Fundamentals and Biotechnology. **Microbiology and Molecular Biology Reviews**, v. 66, n. 3, p. 506–577, 1 set. 2002. Disponível em: <10.1128/MMBR.66.3.506–577.2002>

MANDAL, S. D.; PANDA, A. K; BISHT, S. S.; KUMAR, N. S. Microbial Ecology in the Era of Next Generation Sequencing. **Journal of Next Generation Sequencing & Applications**, 3 dez. 2015. Disponível em: <10.4172/ 2469-9853.S1-001>

MARKOWITZ, V. M.; CHEN, I. M.; PALANIAPPAN, K.; CHU, K.; SZETO, E.; GRECHKIN, Y.; RATNER, A.; JACOB, B.; HUANG, J.; WILLIAMS, P.; HUNTEMANN, M.; ANDERSON, I.; MAVROMATIS, K.; IVANOVA, N. N.; KYRPIDES, N. C. IMG: the Integrated Microbial Genomes database and comparative analysis system. **Nucleic Acids Research**, v. 40, n. Database issue, p. D115-122, jan. 2012. Disponível em: <10.1093/nar/gkr1044>

MATATIELE, P.; GULUMIAN, M. A cautionary approach in transitioning to “green” energy technologies and practices is required. **Reviews on Environmental Health**, v. 31, n. 2, p. 211–223, 1 jun. 2016. Disponível em: <10.1515/reveh-2016-0004>

MENDES, K. C. S. **Estudo do processo de pré-tratamento da palha de cana-de-açúcar para otimização da produção de Xilose com reduzida formação de produtos de degradação.** 2013. Dissertação de Mestrado - Centro de Tecnologia e Geociências da Universidade Federal de Pernambuco (UFPE), Recife, 2013.

MHUANTONG, W.; CHAROENSAWAN, V.; KANOKRATANA, P.; TANGPHATSORNRUANG, S.; CHAMPREDA, V. Comparative analysis of sugarcane bagasse metagenome reveals unique and conserved biomass-degrading enzymes among lignocellulolytic microbial communities. **Biotechnology for Biofuels**, v. 8, p. 16, 2015. Disponível em: <10.1186/s13068-015-0200-8>

MIZRAHI-MAN, O.; DAVENPORT, E. R.; GILAD, Y. Taxonomic Classification of Bacterial 16S rRNA Genes Using Short Sequencing Reads: Evaluation of Effective Study Designs. **PLOS ONE**, v. 8, n. 1, 1 jul. 2013. Disponível em: <https://doi.org/10.1371/journal.pone.0053608>

MOSIER, N.; WYMAN, C.; DALE, B.; ELANDER, R.; LEE, Y. Y.; HOLTZAPPLE, M.; LADISCH, M. Features of promising technologies for pretreatment of lignocellulosic biomass. **Bioresource Technology**, v. 96, n. 6, p. 673–686, abr. 2005. Disponível em: <https://doi.org/10.1016/j.biortech.2004.06.025>

MUNIR, R.; LEVIN, D. B. Enzyme Systems of Anaerobes for Biomass Conversion. In: HATTI-KAUL, R.; MAMO, G.; MATTIASSON, B. (Eds.). **Anaerobes in Biotechnology**. Advances in Biochemical Engineering/Biotechnology. Springer International Publishing, 2016. p. 113–138.

NAIK, S. N.; GOUD, V.; ROUT, P. K.; DALAI, A. K. Production of first and second generation biofuels: A comprehensive review. **Renewable and Sustainable Energy Reviews**, v. 14, n. 2, p. 578–597, fev. 2010. Disponível em: <https://doi.org/10.1016/j.rser.2009.10.003>

NAUSHAD, S.; ADEOLU, M.; WONG, S.; SOHAIL, M.; SCHELLHORN, H. E.; GUPTA, R; S. A phylogenomic and molecular marker based taxonomic framework for the order *Xanthomonadales*: proposal to transfer the families *Algiphilaceae* and *Solimonadaceae* to the order *Nevskiales* ord. nov. and to create a new family within the order *Xanthomonadales*, the family *Rhodanobacteraceae* fam. nov., containing the genus *Rhodanobacter* and its closest relatives. **Antonie Van Leeuwenhoek**, v. 107, n. 2, p. 467–485, fev. 2015. Disponível em: <10.1007/s10482-014-0344-8>

NESME, J.; ACHOUAK, W.; AGATHOS, S. N.; BAILEY, M.; BALDRIAN, P.; BRUNEL, D.; FROSTEGARD, A.; HEULIN, T.; JANSSON, J. K.; JURKEVITCH, E.; KRUUS, K. L.; KOWALCHUK, G. A.; LAGARES, A.; LAPPIN-SCOTT, H. M.; LEMANCEAU, P.; LE PASLIER, D.; MANDIC-MULEC, I.; MURRELL, J. C.; MYROLD, D. D.; NALIN, R.; NANNIPIERI, P.; NEUFELD, J. D.; O'GARA, F.; PARNELL, J. J.; PÜHLER, A.; PYLRO, V.; RAMOS, J. L.; ROESCH, L. F. W.; SCHLOTER, M.; SCHLEPER, C.; SCZYRBA, A.; SESSITSCH, A.; SJÖLING, S.; SORENSEN, J.; SORENSEN, S. J.; TEBBE, C. C.; TOPP, E.; TSIAMIS, G.; VAN ELSAS, J. D.; VAN KEULEN, G.; WIDMER, F.; WAGNER, M.; ZHANG, T.; ZHANG, X.; ZHAO, L.; ZHU, Y-G.; VOGEL, T. M.; SIMONET, P. Back to the Future of Soil Metagenomics. **Frontiers in Microbiology**, v. 7, 10 fev. 2016. Disponível em: <10.3389/fmicb.2016.00073>

NURK, S.; DMITRY, M.; ANTON, K.; PAVEL, P. metaSPAdes: a new versatile de novo metagenomics assembler. **Cornel University Library**, eprint arXiv:1604.03071, 11 abr. 2016.

ONU. **ONU: Com mais de 1 bilhão de veículos, planeta precisa de modelo mais sustentável de transporteONU Brasil**, 22 jun. 2015. Disponível em: <https://nacoesunidas.org/com-mais-de-1-bilhao-de-veiculos-planeta-precisa-de-modelo-mais-sustentavel-de-transporte-afirma-onu/>. Acesso em: 6 abr. 2017

OULAS, A.; PAVLOUDI, C.; POLYMENAKOU, P.; PAVLOPOULOS, G. A.; PAPANIKOLAOU, N.; KOTOULAS, G.; ARVANITIDIS, C.; ILIOPOULOS, I. Metagenomics: Tools and Insights for Analyzing Next-Generation Sequencing Data Derived from Biodiversity Studies. **Bioinformatics and Biology Insights**, v. 9, p. 75–88, 5 maio 2015. Disponível em: <10.4137/BBI.S12462>

PAIXÃO, D. A. A.; DIMITROV, M. R.; PEREIRA, R. M.; ACCORSINI, F. R.; VIDOTTI, M. B.; LEMOS, E. G. DE M. Molecular analysis of the bacterial diversity in a specialized consortium for diesel oil degradation. **Revista Brasileira de Ciência do Solo**, v. 34, n. 3, p. 773–781, jun. 2010. Disponível em: <http://dx.doi.org/10.1590/S0100-06832010000300019>

PARKS, D. H.; TYSON, G. W.; HUGENHOLTZ, P.; BEIKO, R. G. STAMP: statistical analysis of taxonomic and functional profiles. **Bioinformatics**, v. 30, n. 21, p. 3123–3124, 1 nov. 2014. Disponível em: <10.1093/bioinformatics/btu494>

PÉREZ, J.; MUÑOZ-DORADO, J.; DE LA RUBIA, T.; MARTÍNEZ, J. Biodegradation and biological treatments of cellulose, hemicellulose and lignin: an overview. **International Microbiology**, v. 5, n. 2, p. 53–63, 27 abr. 2002. Disponível em: <10.1007/s10123-002-0062-3>

PERRONE, O. M.; COLOMBARI, F. M.; ROSSI, J. S.; MORETTI, M. M. S.; BORDIGNON, S. E.; NUNES, C. C. C.; GOMES, E.; BOSCOLO, M.; DA-SILVA, R. Ozonolysis combined with ultrasound as a pretreatment of sugarcane bagasse: Effect on the enzymatic saccharification and the physical and chemical characteristics of the substrate. **Bioresource Technology**, v. 218, p. 69–76, out. 2016. Disponível em: <https://doi.org/10.1016/j.biortech.2016.06.072>

POLIZELI, M. L. T. M.; RIZZATTI, A. C.; MONTI, R.; TERENZI, H. F. JORGE, J. A.; AMORIM, D. S. Xylanases from fungi: properties and industrial applications. **Applied Microbiology and Biotechnology**, v. 67, n. 5, p. 577–591, jun. 2005. Disponível em: <10.1007/s00253-005-1904-7>

PYLRO, V. S; ROESCH, L. F.; ORTEGA, J. M.; DO AMARAL, A. M.; TÓTOLA, M. R.; HIRSCH, P. R.; ROSADO, A. S.; GÓES-NETO, A.; DA COSTA DA SILVA, A. L.; ROSA, C. A.; MORAIS, D. K.; ANDREOTE, F. D.; DUARTE, G. F.; DE MELO, I. S.; SELDIN, L.; LAMBAIS, M. R.; HUNGRIA, M.; PEIXOTO, R. S.; KRUGER, R. H.; TSAI, S. M.; AZEVEDO, V.; COMMITTEE ORGANIZATION B. M. P. Brazilian Microbiome Project: revealing the unexplored microbial diversity--challenges and prospects. **Microbial Ecology**, v. 67, n. 2, p. 237–241, fev. 2014a. Disponível em: <10.1007/s00248-013-0302-4>

PYLRO, V. S.; ROESCH, L. F.; MORAIS, D. K.; CLARK, I. M.; HIRSCH, P. R.; TÓTOLA, M. R. Data analysis for 16S microbial profiling from different benchtop sequencing platforms. **Journal of Microbiological Methods**, v. 107, p. 30–37, dez. 2014b. Disponível em: <10.1016/j.mimet.2014.08.018>

QUINCE, C.; LANZEN, A.; DAVENPORT, R. J.; TURNBAUGH, P. J. Removing Noise From Pyrosequenced Amplicons. **BMC Bioinformatics**, v. 12, p. 38, 2011. Disponível em: <10.1186/1471-2105-12-38>

RAVINDRAN, R.; JAISWAL, A. K. A comprehensive review on pre-treatment strategy for lignocellulosic food industry waste: Challenges and opportunities. **Bioresource Technology**, Pretreatment of Biomass. v. 199, p. 92–102, jan. 2016. Disponível em: <https://doi.org/10.1016/j.biortech.2015.07.106>

REISER, V.; GASPERÍK, J. Purification and characterization of the cell-wall-associated and extracellular alpha-glucosidases from *Saccharomycopsis fibuligera*. **Biochemical Journal**, v. 308, n. Pt 3, p. 753–760, 15 jun. 1995.

ROBLEDO, M.; JIMÉNEZ-ZURDO, J. I.; VELÁSQUEZ, E.; TRUJILLO, M. E.; ZURDO-PIÑEIRO, J. L.; RAMÍREZ-BAHENA, M. H.; RAMOS, B.; DÍAZ-MÍNGUES, J. M.; DAZOO, F.; MARTÍNEZ-MOLINA, E.; MATEOS, P. F. *Rhizobium* cellulase CelC2 is essential for primary symbiotic infection of legume host roots. **Proceedings of the National Academy of Sciences**, v. 105, n. 19, p. 7064–7069, 13 maio 2008. Disponível em: <10.1073/pnas.0802547105>

ROSENBERG, E.; DELONG, E. F.; THOMPSON, F.; LORY, S.; STACKEBRANDT, E. **The Prokaryotes - Alphaproteobacteria and Betaproteobacteria**. 4. ed. New Delhi, India: Springer-Verlag Berlin Heidelberg, 2014.

RUBIN, E. M. Genomics of cellulosic biofuels. **Nature**, v. 454, n. 7206, p. 841–845, 14 ago. 2008. Disponível em: <doi:10.1038/nature07190>

SANT'ANNA, C.; COSTA, L. T.; ABUD, Y.; BIANCATTO, L.; MIGUENS, F. C.; DE SOUZA, W. Sugarcane cell wall structure and lignin distribution investigated by confocal and electron microscopy. **Microscopy Research and Technique**, v. 76, n. 8, p. 829–834, ago. 2013. Disponível em: <10.1002/jemt.22235>

SANTOS, DA S. F. R. **Produção e Caracterização de Celulases e Hemicelulases por Linhagens Fúngicas Mesófilas Isoladas do Cerrado Sul-Matogrossense**. 2014. Dissertação de Mestrado - Universidade Federal da Grande Dourados, Faculdade de Ciências Exatas e Tecnológicas, 2014.

SARITHA, M.; ARORA, A.; LATA. Biological Pretreatment of Lignocellulosic Substrates for Enhanced Delignification and Enzymatic Digestibility. **Indian Journal of Microbiology**, v. 52, n. 2, p. 122–130, jun. 2012. Disponível em: <10.1007/s12088-011-0199-x>

SIMON, C.; DANIEL, R. Metagenomic Analyses: Past and Future Trends. **Applied and Environmental Microbiology**, v. 77, n. 4, p. 1153–1161, 15 fev. 2011. Disponível em: <10.1128/AEM.02345-10>

SWEENEY, M. D.; XU, F. Biomass Converting Enzymes as Industrial Biocatalysts for Fuels and Chemicals: Recent Developments. **Catalysts**, v. 2, n. 2, p. 244–263, 12 abr. 2012. Disponível em: <10.3390/catal2020244>

TAFFS, R.; ASTON, J. E.; BRILEYA, K.; JAY, Z.; KLATT, C. G.; MCGLYNN, S.; MALLETTE, N.; MONTROSS, S.; GERLACH, R.; INSKEEP, W. P.; WARD, D. M.; CARLSON, R. P. In silico approaches to study mass and energy flows in microbial consortia: a syntrophic case study. **BMC Systems Biology**, v. 3, p. 114, 2009. Disponível em: <10.1186/1752-0509-3-114>

TAHERZADEH, M. J.; KARIMI, K. Pretreatment of Lignocellulosic Wastes to Improve Ethanol and Biogas Production: A Review. **International Journal of Molecular Sciences**, v. 9, n. 9, p. 1621–1651, 1 set. 2008. Disponível em: <10.3390/ijms9091621>

TANAKA, S.; LIN, Y. Ethanol Fermentation from Biomass Resources: Current State and Prospects. **Applied Microbiology and Biotechnology**, v. 69, p. 627–642, 2006. Disponível em: <10.1007/s00253-005-0229-x>

THE UNIPROT CONSORTIUM. UniProt: the universal protein knowledgebase. **Nucleic Acids Research**, v. 45, n. D1, p. D158–D169, 4 jan. 2017. Disponível em: <https://doi.org/10.1093/nar/gkw1099>

TURNBAUGH, P. J.; GORDON, J. I. An Invitation to the Marriage of Metagenomics and Metabolomics. **Cell**, v. 134, n. 5, p. 708–713, 5 set. 2008. Disponível em: <10.1016/j.cell.2008.08.025>

TYE, Y. Y.; LEE, K. T.; NADIAH, W. W.; LEH, C. P. The world availability of non-wood lignocellulosic biomass for the production of cellulosic ethanol and potential pretreatments for the enhancement of enzymatic saccharification. **Renewable and Sustainable Energy Reviews**, v. 60, p. 155–172, jul. 2016. Disponível em: <https://doi.org/10.1016/j.rser.2016.01.072>

UNCTAD. **Second Generation Biofuel Markets: State of Play, Trade and Developing Country Perspectives**. 2016

VAKHLU, J.; SUDAN, A. K.; JOHRI, B. N. Metagenomics: Future of microbial gene mining. **Indian Journal of Microbiology**, v. 48, n. 2, p. 202–215, jun. 2008. Disponível em: <10.1007/s12088-008-0033-2>

VARANI, A. M.; SIGUIER, P.; GOURBEYRE, E.; CHARNEAU, V.; CHANDLER, M. ISsaga is an ensemble of web-based methods for high throughput identification and semi-automatic annotation of insertion sequences in prokaryotic genomes. **Genome Biology**, v. 12, n. 3, p. R30, 2011. Disponível em: <10.1186/gb-2011-12-3-r30>

VARUZZA, L. **Introdução à análise de dados de sequenciadores de nova geração. Versão 2.1.0**, 2013.

VIEIRA, M. C. A.; LIMA, J. F.; BRAGA, N. M. **Setor Sucroalcooleiro Brasileiro: Evolução e Perspectivas.** BNDES, 2007.

WANG, L.; NIE, Y.; TANG, Y-Q.; SONG, X-M.; CAO, K.; SUN, L-Z.; WANG, Z-J.; WU, X-L. Diverse Bacteria with Lignin Degrading Potentials Isolated from Two Ranks of Coal. **Frontiers in Microbiology**, v. 7, 9 set. 2016. Disponível em: <10.3389/fmicb.2016.01428>

WONGWILAIWALIN, S.; LAOTHANACHAREON, T.; MHUANTONG, W.; TANGPHATSORNRUANG, S.; EURWILAICHITR, L.; IGARASHI, Y.; CHAMPREDA, V. Comparative metagenomic analysis of microcosm structures and lignocellulolytic enzyme systems of symbiotic biomass-degrading consortia. **Applied Microbiology and Biotechnology**, v. 97, n. 20, p. 8941–8954, out. 2013. Disponível em: <10.1007/s00253-013-4699-y>

WOO, H. L.; HAZEN, T. C.; SIMMONS, B. A.; DEANGELIS, K. M. Enzyme activities of aerobic lignocellulolytic bacteria isolated from wet tropical forest soils. **Systematic and Applied Microbiology**, v. 37, n. 1, p. 60–67, fev. 2014a. Disponível em: <https://doi.org/10.1016/j.syapm.2013.10.001>

WOO, H. L.; UTTURKAR, S.; KLINGEMAN, D.; SIMMONS, B. A.; DEANGELIS, K. M.; BROWN, S. D.; HAZEN, T. C. Draft Genome Sequence of the Lignin-Degrading *Burkholderia* sp. Strain LIG30, Isolated from Wet Tropical Forest Soil. **Genome Announcements**, v. 2, n. 3, 19 jun. 2014b. Disponível em: <10.1128/genomeA.00637-14>

WOOD, D. E.; SALZBERG, S. L. Kraken: ultrafast metagenomic sequence classification using exact alignments. **Genome Biology**, v. 15, n. 3, p. R46, 2014. Disponível em: <10.1186/gb-2014-15-3-r46>

XIE, C.-H.; YOKOTA, A. *Dyella japonica* gen. nov., sp. nov., a *γ-proteobacterium* isolated from soil. **International Journal of Systematic and Evolutionary Microbiology**, v. 55, n. 2, p. 753–756, 2005. Disponível em: <10.1099/ijs.0.63377-0>

YANG, B.; DAI, Z.; DING, S-Y.; WYMAN, C. E. Enzymatic hydrolysis of cellulosic biomass. **Biofuels**, v. 2, n. 4, p. 421–449, 1 jul. 2011. Disponível em: <10.4155/BFS.11.116>

YEOH, Y. K.; PAUNGFOO-LONHIENNE, C.; DENNIS, P. G.; ROBINSON, N.; RAGAN, M. A.; SCHIMIDT, S.; HUGENHOLTZ, P. The core root microbiome of sugarcanes cultivated under varying nitrogen fertilizer application. **Environmental Microbiology**, v. 18, n. 5, p. 1338–1351, maio 2016. Disponível em: <10.1111/1462-2920.12925>

ZHAO, F.; GUO, X. Q.; WANG, P.; HE, L. Y.; HUANG, Z.; SHENG, X. F. *Dyella jiangningensis* sp. nov., a *γ-proteobacterium* isolated from the surface of potassium-bearing rock. **International Journal of Systematic and Evolutionary Microbiology**, v. 63, n. Pt 9, p. 3154–3157, set. 2013. Disponível em: <10.1099/ijs.0.048470-0>

ZUROFF, T. R.; XIQUES, S. B.; CURTIS, W. R. Consortia-mediated bioprocessing of cellulose to ethanol with a symbiotic *Clostridium* phytofermentans/yeast co-culture. **Biotechnology for Biofuels**, v. 6, p. 59, 2013. Disponível em: <10.1186/1754-6834-6-59>