UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS

BRUNO LUAN SOARES PAULA DE MELLO

Identificação e caracterização da primeira exoxilanase da família 11
de hidrolase de glicosídeo a partir do estudo do metatranscriptoma de
um consórcio derivado da compostagem

São Carlos
2017

BRUNO LUAN SOARES PAULA DE MELLO

Identificação e caracterização da primeira exoxilanase da família 11 de hidrolase de glicosídeo a partir do estudo do metatranscriptoma de um consórcio derivado da compostagem

Tese apresentada ao Programa de Pós-Graduação em Física do Instituto de Física de São Carlos da Universidade de São Paulo, para obtenção do título de Doutor em Ciências.

Área de concentração: Física Aplicada
Opção: Física Biomolecular
Orientador: Prof. Dr. Igor Polikarpov

Versão corrigida

(versão original disponível na Unidade que aloja o Programa)

São Carlos
2017

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica revisada pelo Serviço de Biblioteca e Informação do IFSC,
com os dados fornecidos pelo(a) autor(a)

Mello, Bruno Luan
Identificação e caracterização da primeira
exoxilanase da família 11 de hidrolase de glicosídeo a
partir do estudo do metatranscriptoma de um consórcio
derivado da compostagem / Bruno Luan Mello;
orientador Igor Polikarpov - versão corrigida -- São
Carlos, 2017.
103 p.

Tese (Doutorado - Programa de Pós-Graduação em
Física Biomolecular) -- Instituto de Física de São
Carlos, Universidade de São Paulo, 2017.

1. Metatranscriptômica. 2. Compostagem. 3.
Bioetanol. 4. Etanol de segunda geração. I.
Polikarpov, Igor, orient. II. Título.

À minha mãe, que sempre fez o impossível
para que eu tivesse a melhor
educação disponível.

AGRADECIMENTOS

À Dr. Francieli Colussi, por me ensinar, quando eu ainda estava cursando a graduação, as bases experimentais de tudo o que sei hoje.

À Dr. Viviane Serpa, pelos ensinamentos durante nossa passagem pelo Canadá.

A todos que convivi no Grupo de Cristalografia e, mais tarde, no Grupo de Biotecnologia Molecular. Impossível listar, mas todos me inspiraram de alguma forma para meu amadurecimento acadêmico. Em especial a Aline, Amanda, André, Danilo, Evandro, Lis, Marco, Mel, Renata e Vanessa, pelas conversas, risadas, ajuda no laboratório e discussões científicas.

À Dr. Anna Alessi, pelo empenho neste projeto, sem a qual o mesmo nível científico não teria sido alcançado.

Ao João, Josimar, Livia e Maria, pelo apoio técnico.

Ao Dr. Fábio Márcio Squina, pela infraestrutura disponibilizada, e ao Dr. Diego Mauricio Riaño-Pachón, pelos ensinamentos de bioinformática.

Ao Prof. Dr. Francisco Eduardo Contijo Guimarães e ao Prof. Dr. Eduardo Ribeiro de Azevêdo, pela disponibilidade e empenho em realizar as medidas para caracterização da biomassa.

Ao Prof. Dr. Neil Bruce e ao Prof. Dr. Simon McQueen Mason, por me receberem em seus laboratórios e estarem sempre presentes durante o desenvolvimento do projeto.

Ao Prof. Dr. Igor Polikarpov, pelas oportunidades oferecidas.

Ao Instituto de Física de São Carlos, pela oportunidade de realização do curso de doutorado.

À EMBRAPA Instrumentação (São Carlos) pelo tempo de uso no espectrômetro de ressonância magnética nuclear.

À Fundação de Amparo à Pesquisa do Estado de São Paulo, pela concessão da bolsa de doutorado e pelo apoio financeiro para a realização desta pesquisa.

Calvin and Hobbes

RESUMO

MELLO, B. L. **Identificação e caracterização da primeira exoxilanase da família 11 de hidrolase de glicosídeo a partir do estudo do metatranscriptoma de um consórcio derivado da compostagem.** 2017. 103 p. Tese (Doutorado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2017.

O uso de resíduos agrícolas como fonte de carbono para a geração de energia renovável parece uma solução promissora para reduzir nossa dependência em combustíveis fósseis. Na natureza, como na compostagem, comunidades microbianas formam redes metabólicas complexas que degradam eficientemente a biomassa disponível através de um conjunto de enzimas sinérgicas. Entretanto, a desconstrução da lignocelulose continua uma desafio para a indústria devido a natureza recalcitrante do substrato e a baixa atividade das enzimas, aumentando o preço do biocombustível produzido. Estudos de metagenômica e metatranscriptômica de comunidades microbianas complexas tornam possível acessar as funções metabólicas empregadas por consórcios lignocelulolíticos e revelar novos biocatalisadores que podem melhorar a conversão industrial da lignocelulose. Aqui, através de uma abordagem metagenômica, foi examinada a diversidade de microrganismos obtidos em condições laboratoriais quando um meio definido ou um complexo foi usado no seu crescimento. Em seguida, a comunidade microbiana derivada de compostagem foi crescida em meio mínimo com bagaço de cana-de-açúcar como única fonte de carbono. A degradação do substrato foi monitorada e o metatranscriptoma da cultura resultante foi sequenciado, seguido pela seleção e caracterização funcional de vários alvos. Durante as cinco semanas de estudo, a comunidade microbiana crescida em meio mínimo mostrou maior diversidade e enriquecimento em microrganismos capazes de degradar a lignocelulose do que a comunidade microbiana crescida no meio complexo. A partir do metatranscriptoma foi descoberta a primeira hidrolase de glicosídeo da família 11 com atividade exoxilanase (C21). A estrutura cristalográfica da C21, refinada à 1,76 Å, revelou que a atividade exoxilanase observada se deve a presença de duas alças que não estão presentes nas demais estruturas dos membros da família 11 de hidrolase de glicosídeo depositadas até então. A adição da C21 a um coquetel comercial provocou um aumento na velocidade de hidrólise do Avicel quando na presença de xilooligômeros. As análises metagenômica e metatranscriptômica da comunidade microbiana proveniente da compostagem revelaram que o uso de um meio definido pode deslocar espécies generalistas, levando a uma fonte enriquecida para explorar enzimas com aplicação

biotecnológica. Também demonstrou a diversidade de mecanismos envolvidos na degradação *in situ* da lignocelulose.

Palavras-chave: Metatranscriptômica. Compostagem. Bioetanol. Etanol de segunda geração.

ABSTRACT

MELLO, B.L. **Identification and characterization of the first exo-xylanase from glycosyl hydrolase family 11 from the study of the metatranscriptome of a compost-derived consortia.** 2017. 103 p. Tese (Doutorado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2017.

Using of the globally abundant crop residues as carbon source for energy generation seems a promising solution to reduce our dependency on fossil fuels. In nature, such as in compost habitats, microbial communities create complex metabolic networks that efficiently degrade the available plant biomass using a set of synergistic enzymes. However, deconstruction of lignocellulose remains a challenge for industry due to recalcitrant nature of the substrate and enzymes low activity, raising the price of the produced biofuel. Metagenomics and metatranscriptomics studies on complex microbial communities can assess the metabolic functions employed by the lignocellulolytic consortia and unveil novel biocatalysts that could improve industrial lignocellulose conversion. Here, using 16S rRNA amplicon metagenomic approach, we examined the diversity of microorganisms obtained in the laboratory setting when a nutrient-limited or nutrient-rich media are used. Then, a microbial community derived from compost was grown in minimal medium with sugarcane bagasse as a sole carbon source. The substrate degradation was monitored and the metatranscriptome from the resulting cultures was sequenced; several target genes were selected and functional characterized. During a five-week time course, the microbial community grown in minimal medium showed greater diversity and enrichment in lignocellulose-degrading microorganisms than the one grown in nutrient rich medium. Metatranscriptomics analysis revealed the first glycoside hydrolase from family 11 with exo-xylanase activity (C21). C21 crystal structure, refined at 1.76 Å, explained the molecular basis of exo-xylanase activity due to two extra loops previously unseen in the other reported structures from members from glycoside hydrolase family 11. A supplementation of commercial enzyme mix with C21 showed improvement in Avicel hydrolysis in the presence of inhibitory xylooligomers. The combination of metagenomic and metatranscriptomic analysis of compost-derived microbial community showed that nutrient-limited medium may displace bacterial generalist species, leading to an enriched source for mining novel

enzymes for biotechnology applications. It also unveiled a diversity of mechanisms involved in lignocellulose degradation *in situ*.

Keywords: Metatranscriptomics. Compost. Bioethanol. Second-generation ethanol.

LISTA DE FIGURAS

<i>Figura 1 - Consumo de energia no Brasil e no mundo. a) Progressão do consumo mundial de 1965 a 2015. b) Matriz energética brasileira e mundial, em 2015.....</i>	<i>24</i>
<i>Figura 2 - Representação esquemática da disposição das moléculas de celulose, hemicelulose e lignina na parede celular das plantas. Na estrutura da lignina, G, S e H representam, respectivamente, os ácidos coniferílico, sinapílico e p-cumarílico.....</i>	<i>26</i>
<i>Figura 3 - Estrutura dos principais monolignóis componentes da lignina.</i>	<i>27</i>
<i>Figura 4 - Estrutura geral do xilano, principal componente da hemicelulose de angiospermas. Na figura estão representados os seus principais constituintes.</i>	<i>27</i>
<i>Figura 5 - Representação de uma molécula de celulose, destacando a celobiose (unidade repetitiva do polímero).</i>	<i>28</i>
<i>Figura 6 - Modelo da digestão da celulose.</i>	<i>29</i>
<i>Figura 7 - Ilustração da evolução de um processo de compostagem. Na parte superior inicia-se o processo com restos de plantas e alimentos que são decompostos ao longo do tempo pelos microrganismos presentes. A parte inferior representa os estágio final, no qual o material se transforma em um fertilizante que pode ser aplicado em novas plantas.</i>	<i>33</i>
<i>Figura 8 - Número de unidades formadoras de colônia (UFC) no consórcio crescido em a) meio mínimo (MM) e b) PCS.</i>	<i>52</i>
<i>Figura 9 - Análise qualitativa da degradação do bagaço de cana-de-açúcar coletado após crescimento do consórcio microbiano em meio mínimo (MM) e PCS durante cinco semanas. É possível observar mudanças morfológicas, tais como redução no tamanho das partículas e escurecimento da biomassa, principalmente no cultivo em MM.</i>	<i>53</i>
<i>Figura 10 - Abundância relativa dos a) filos e b-d) classes bacterianos identificados no consórcio crescido em meio mínimo (MM) e PCS. Na figura, o número após o meio de cultura indica a semana na qual a amostra foi coletada. Eixo-y: abundância relativa (%).</i>	<i>55</i>
<i>Figura 11 - Análise da diversidade no consórcio microbiano. a) Diagrama de Venn apresentando o número de OTUs únicas e compartilhadas entre as comunidades crescidas em meio mínimo (MM) e PCS. b) Curvas de rarefação mostrando a diversidade filogenética calculada versus o número de fragmentos sequenciados por amostra. Outros índices de diversidade, como chao1 e observed_species, também foram testados e forneceram resultados similares. c) Box-plot do índice de diversidade Shannon calculado. Os resultados foram agrupados pelo meio de cultura usado no crescimento microbiano. O asterisco (*) denota uma diferença significativa ($p \leq 0,05$, teste t) entre os índices de diversidade obtidos para a cultura em MM e PCS. d) Análise de coordenadas principais das distâncias UniFrac ponderadas. Os consórcios crescidos no mesmo meio se agrupam. PCO-1: coordenada principal 1.</i>	<i>58</i>
<i>Figura 12 - Análise da composição do bagaço de cana-de-açúcar. a) Espectro de ressonância magnética nuclear de estado sólido do bagaço de cana-de-açúcar antes e depois do crescimento microbiano. As regiões 1 e 1' são atribuídas ao carbono C1 da celulose (103-106 ppm) e ao carbono C4 da celulose cristalina (86-92 ppm). As regiões 2 e 2' são atribuídas aos carbonos C1, C2 e C4 do anel aromático da lignina (123-142 ppm) e aos carbonos aril metoxil da lignina (50-56 ppm). As regiões 3 e 3' são atribuídas ao carbono C1 da hemicelulose e aos CH₃ em grupos acetil da hemicelulose. A abundância relativa de celulose, hemicelulose e lignina foi estimada a partir das regiões 1, 1', 2, 2', 3 e 3'. b-e) Imagens de microscopia confocal da lignina no bagaço antes e depois do crescimento microbiano. f) Espectro de emissão correspondente.....</i>	<i>65</i>
<i>Figura 13 - Imagens de microscopia eletrônica de varredura do bagaço de cana-de-açúcar. a), b) antes e c), d) depois de cinco semanas de crescimento microbiano. a), c) aumento de 300x; b), d) aumento de 1.000x.</i>	<i>67</i>

Figura 14 - Análise da profundidade de sequenciamento do metatranscriptoma por dois métodos. a) Variação da taxa de expressão calculada a partir de bibliotecas com diferentes profundidades de sequenciamento. b) Fração de novos kmers obtidos.	71
Figura 15 - Perfil metabólico dos transcritos expressos no consórcio derivado da compostagem. Classificação realizada de acordo com a base de dados KEGG orthology. Cada eixo apresenta a quantidade relativa de genes, em porcentagem, pertencentes a respectiva classe.....	73
Figura 16 - Expressão de CAZymes putativas no metatranscriptoma do consórcio derivado da compostagem. AA - atividade auxiliar; CE - esterase de carboidrato; CBM - módulo de ligação a carboidrato; GH - hidrolase de glicosil; GT - transferase de glicosil; PL - liase de polissacarídeo.	74
Figura 17 - Expressão diferencial e distribuição filogenética das famílias de hidrolase de glicosídeo (GH) detectadas no metatranscriptoma da compostagem. a) Heat map da expressão de GHs. As colunas representam o tempo de cultivo quando a amostra foi coletada. As linhas retratam as diferentes famílias de GH identificadas. As famílias de GH foram agrupadas de acordo com sua preferência por substratos.202 b) Atribuição filogenética dos fragmentos pertencentes as famílias GHs mais expressas.	76
Figura 18 - Atribuição filogenética das CAZymes expressas ao longo do tempo pelo consórcio derivado da compostagem. Expressão relativa dos a) filos bacterianos e b) reinos eucariotos. A expressão total de cada domínio é representada pela linha em cinza.	77
Figura 19 - Caracterização bioquímica das enzimas C7, C13 e C21 derivadas do metatranscriptoma da compostagem. Efeito do a) pH e b) temperatura na atividade enzimática. c) Especificidade por substrato. d) Atividade residual da proteína após incubação na temperatura de estudo.....	80
Figura 20 - Termoestabilidade, examinada por fluorimetria de escaneamento diferencial, da proteína C7 em diferentes pHs.	80
Figura 21 - Evidência experimental que a proteína C21 é uma exoxilanase. a) TLC dos produtos formados por: linhas 1 e 3: ação da C21 e C13 contra xilano, respectivamente; linha 2: branco de reação; linhas 4 a 8: ação da C21 contra xilooligômeros de duas a seis unidades de xilose (X2-X6); Ác.G: ácido glicurônico; M: padrão com X1-X6. b) HPLC Dionex dos produtos formado pela ação da C21 contra xilano.	82
Figura 22 - Evidência estrutural que a proteína C21 é uma exoxilanase. a) Alinhamento múltiplo da sequência de aminoácidos da C21 com os homólogos mais próximos presentes nos bancos de dados do NCBI-nr (segunda a sexta sequência) e PDB (cinco últimas sequências). Os resíduos responsáveis pela formação dos loops extras (EL1 e EL2) estão selecionados nas caixas vermelhas; os resíduos catalíticos estão selecionados nas caixas azuis. Esses loops extras estão presentes em outras proteínas cujas estruturas ainda não foram resolvidas. b) Estrutura secundária da C21 colorida do azul para o vermelho (do N- pro C-terminal). c) Superposição de todas as estruturas disponíveis da família GH11 (em cinza) e da C21 (em magenta). A orientação dessas estruturas foi rotacionada de 90° em relação a estrutura apresentada no item b. Domínios incomuns que não se alinham com as demais proteínas foram removidos da representação. d) Representação da superfície da C21 (em magenta) alinhada com o homólogo mais próximo do PDB (ID 1XNK, em cinza). Note que o terminal não-redutor do ligante metil 4,4"-ditio- α -xilotriose, presente na estrutura 1XNK, é impedido estericamente no subsítio -3 da C21 pelo EL2.	84
Figura 23 - Impacto de xilooligômeros na performance de um coquetel enzimático comercial. a) Hidrólise de Avicel por Accellerase® na presença de xilooligômeros e C21. b) Análise por HPLC Dionex da degradação dos xilooligômeros pela Accellerase® e c) pela Accellerase® suplementada com C21, como determinado pela resposta do detector em nanocoulomb (nC). Note que no item c, a concentração de xilobiose aumenta rapidamente após 1 hora de reação como resultado da rápida degradação pela C21 de xilooligômeros com grau de polimerização maior ou igual a quatro. Xilooligômeros com maior grau de polimerização possuem um efeito inibitório maior, portanto sua rápida degradação é benéfica. O coquetel comercial usado foi capaz de degradar os xilooligômeros adicionados dentro de 24 horas.....	86

LISTA DE TABELAS

<i>Tabela 1 - Número de fragmentos totais e após filtro de qualidade (usado para atribuição de OTUs). Na coluna comunidade, o número após o meio de cultura indica a semana na qual a amostra foi coletada.....</i>	<i>54</i>
<i>Tabela 2 - Análise de SIMPER mostrando quais filos e classes de microrganismos foram os principais responsáveis pela dissimilaridade Bray-Curtis entre as comunidades crescidas em meio mínimo (MM) e PCS.</i>	<i>56</i>
<i>Tabela 3 - 21 OTUs mais abundantes na comunidade microbiana (CM) crescida em MM. Taxas sublinhados estão presentes também no consórcio crescido em meio PCS.</i>	<i>59</i>
<i>Tabela 4 - 21 OTUs mais abundantes na comunidade microbiana (CM) crescida em meio PCS. Taxas sublinhados estão presentes também no consórcio crescido em MM.</i>	<i>61</i>
<i>Tabela 5 - métricas do sequenciamento e processamento do metatranscriptoma.</i>	<i>69</i>
<i>Tabela 6 - Abundância relativa de rDNA amplificado com oligonucleotídeos específicos para fungos e bactérias.</i>	<i>78</i>
<i>Tabela 7 - Estatísticas da coleta e refinamento dos dados de difração da proteína C21. Os valores em parênteses se referem a camada mais externa de resolução. 5% das reflexões foram aleatoriamente excluídas do refinamento e usadas para cálculo do R_{free}.</i>	<i>83</i>

LISTA DE ABREVIATURAS E SIGLAS

AA	Atividades auxiliar
CAZyme	Enzima ativa em carboidrato
CBM	Módulos de ligação a carboidrato
CE	Esterase de carboidrato
CM	Comunidade microbiana
CMC	Carboximetilcelulose
CPMAS	Rampa de rádio frequência na polarização cruzada com rotação em torno do ângulo mágico
DNS	Ácido 3,5-dinitrosalicílico
DSF	Fluorimetria de escaneamento diferencial
ELX	Loop extra X
GH	Hidrolase de glicosil
GT	Transferase de glicosil
HPLC	Cromatografia líquida de alta performance
ITS	Espaçadora transcrita interna
KEGG	Enciclopédia de Genes e Genomas de Kyoto
LIC	Clonagem independente de ligação
LNLS	Laboratório Nacional de Luz Sincrotron
MEV	Microscopia eletrônica de varredura
MM	Meio mínimo
Ni-NTA	Níquel-nitrilotriacético
NREL	Laboratório nacional de energias renováveis

ORF	Quadro de leitura aberto
OTU	Unidade taxonômica operacional
pb	Pares de base
PCoA	Análise de coordenadas principais
PCR	Reação em cadeia da polimerase
PLs	Liases de polissacarídeo
PMSF	Fluoreto de fenilmetilsulfonil
PVDF	Fluoreto de polivinilideno
SIMPER	Similaridade percentual
ssNMR	Ressonância magnética nuclear de estado sólido
TLC	Cromatografia de camada delgada
TOSS	Supressão total de bandas laterais através de rotação
UFC	Unidade formadora de colônia
V4-16S	V4 do gene do rRNA 16S
X1	Xilose
Xn	xilooligômero com n resíduos de xilose
β-ME	β-mercaptoetanol

SUMÁRIO

1	INTRODUÇÃO	21
2	REVISÃO BIBLIOGRÁFICA	23
2.1	CONSUMO E PRODUÇÃO DE ENERGIA.....	23
2.2	O BAGAÇO DE CANA-DE-AÇÚCAR COMO BIOMASSA PARA A PRODUÇÃO DE ETANOL	24
2.3	BIOMASSA	25
2.4	SACARIFICAÇÃO POR HIDRÓLISE ENZIMÁTICA.....	28
2.5	CAZY	30
2.6	METAGENÔMICA	30
2.7	METATRANSCRIPTÔMICA.....	32
2.8	COMPOSTAGEM.....	33
3	JUSTIFICATIVA.....	35
4	OBJETIVOS	37
4.1	OBJETIVO GERAL	37
4.2	OBJETIVOS ESPECÍFICOS.....	37
4.2.1	<i>Estudo do meio de cultura utilizado no crescimento</i>	<i>37</i>
4.2.2	<i>Estudo da degradação do bagaço de cana-de-açúcar por microrganismos presentes na compostagem.....</i>	<i>37</i>
4.2.3	<i>Estudo de alvos do metatranscriptoma</i>	<i>37</i>
5	MATERIAIS E MÉTODOS.....	39
5.1	COLETA DE AMOSTRAS E CULTIVO PARA ESTUDO DO MEIO DE CULTURA	39
5.2	TRIAGEM DE ORGANISMOS CELULOLÍTICOS NO CONSÓRCIO DERIVADO DA COMPOSTAGEM	39
5.3	SEQUENCIAMENTO DA REGIÃO V4 DO GENE DO rRNA 16S	39
5.4	BIOINFORMÁTICA E ANÁLISE DOS DADOS.....	40
5.5	COLETA DE AMOSTRAS E CULTIVO PARA ESTUDO DA DEGRADAÇÃO DO BAGAÇO DE CANA-DE-AÇÚCAR	41
5.6	ANÁLISE DA BIOMASSA.....	41
5.7	EXTRAÇÃO DE ÁCIDOS NUCLÉICOS	42
5.8	PCR EM TEMPO REAL DO DNA GENÔMICO ISOLADO	43
5.9	SÍNTESE DA BIBLIOTECA DE cDNA E SEQUENCIAMENTO	44
5.10	MONTAGEM E ANOTAÇÃO DO METATRANSCRIPTOMA	44
5.11	CLONAGEM, EXPRESSÃO E PURIFICAÇÃO DE GENES ALVO.....	45
5.12	ANÁLISE DAS SEQUÊNCIAS E CARACTERIZAÇÃO ENZIMÁTICA	46
5.13	ANÁLISE DA ESTABILIDADE TÉRMICA ATRAVÉS DE FLUORIMETRIA DE ESCANEAMENTO DIFERENCIAL.....	47
5.14	IDENTIFICAÇÃO DO PRODUTO DE REAÇÃO COM TLC E DIONEX HPLC	47
5.15	PRODUÇÃO DE XILOOLIGÔMEROS E QUANTIFICAÇÃO	48
5.16	INIBIÇÃO DE HIDRÓLISE POR XILOOLIGÔMEROS	48
5.17	ENSAIOS DE CRISTALIZAÇÃO E COLETA DE DADOS	49
5.18	SUBSTITUIÇÃO MOLECULAR, CONSTRUÇÃO DO MODELO E REFINAMENTO DA ESTRUTURA	49
6	RESULTADOS E DISCUSSÃO	51
6.1	TRIAGEM DE ORGANISMOS CELULOLÍTICOS NO CONSÓRCIO DERIVADO DA COMPOSTAGEM	51

6.2 ANÁLISE DA DIVERSIDADE MICROBIANA NO CONSÓRCIO DERIVADO DA COMPOSTAGEM ATRAVÉS DE SEQUENCIAMENTO EM LARGA ESCALA.....	53
6.3 MEDIDAS DE DIVERSIDADE MICROBIANA.....	63
6.4 A DISPONIBILIDADE DE NUTRIENTES AFETA A ESTRUTURA DA COMUNIDADE MICROBIANA	63
6.5 MUDANÇAS QUÍMICAS E MORFOLÓGICAS DO BAGAÇO DE CANA-DE-AÇÚCAR	64
6.6 MÉTRICAS DO METATRANSCRIPTOMA	68
6.7 ANÁLISE DA PROFUNDIDADE DE SEQUENCIAMENTO DO METATRANSCRIPTOMA	70
6.8 PERFIS FUNCIONAIS.....	72
6.9 PERFIL DE EXPRESSÃO E CLASSIFICAÇÃO FILOGENÉTICA DE CAZYMES	73
6.10 CARACTERIZAÇÃO FUNCIONAL DE CAZYMES PUTATIVAS DO METATRANSCRIPTOMA	78
6.11 C21 É A PRIMEIRA EXOXILANASE DESCRITA DA FAMÍLIA GH 11.....	81
6.12 A ATIVIDADE DA C21 MELHORA A PERFORMANCE DE COQUETÉIS ENZIMÁTICOS COMERCIAIS.....	85
7 CONCLUSÕES E PERSPECTIVAS	87
REFERÊNCIAS	89

1 INTRODUÇÃO

O aumento acelerado no consumo de combustíveis fósseis aliado a preocupações com o meio ambiente desencadearam pesquisas que viabilizem a produção de energia renovável. A lignocelulose é o componente básico da parede celular de plantas e uma das fontes de carbono mais abundantes da biosfera. Portanto, sua conversão em biocombustíveis representa uma solução promissora para a geração de energia.¹⁻³ Na natureza, a desconstrução da lignocelulose em açúcares simples, que possam ser fermentados em biocombustíveis, é realizada por comunidades microbianas através de enzimas. Como a compostagem é um ambiente onde a desconstrução da lignocelulose ocorre ativamente, diversas pesquisas passaram a explorá-lo em busca de novas enzimas.⁴⁻⁹ Isso foi possível devido ao desenvolvimento do sequenciamento em larga escala e de técnicas de extração diretas de DNA e RNA dessas comunidades.¹⁰ Dessa forma, se tornou possível acessar o reservatório de genes da chamada “maioria incultivável” (apenas cerca de 1% dos microrganismos são cultiváveis em cultura pura).¹¹⁻¹² Portanto, através de estudos de metagenômica e metatranscriptômica, aumentam-se as chances de descobrir enzimas com características industrialmente interessantes que permitam reduzir os custos de processos biotecnológicos.

Apesar de muitos microrganismos de vida livre serem capazes de digerir a biomassa de plantas, esse processo continua um desafio para a indústria. Um dos maiores obstáculos para a produção em escala industrial de biocombustível de segunda geração é a desconstrução da lignocelulose. Isso porque o substrato é naturalmente recalcitrante e as enzimas atualmente disponíveis possuem baixa atividade relativa, tornando o processo ineficiente.¹³ A maioria dos estudos de metagenômica que visam identificar novas enzimas para solucionar esse gargalo se baseiam na coleta de amostras diretamente do meio ambiente. Dessa forma, os dados resultante frequentemente apresentam um baixo número de genes alvo.

Em estudos, já foi realizado o sequenciamento e a caracterização funcional de comunidades microbianas provenientes de diferentes ambientes ativos na digestão da lignocelulose. Exemplos incluem o rúmen bovino,¹³⁻¹⁴ intestino de animais, cupim e abelha,¹⁵⁻²¹ solo,²² rio²³ e compostagem.²⁴ Visando melhorar a taxa de detecção de propriedades desejadas, técnicas de enriquecimento podem ser aplicadas. Nessas técnicas, uma amostra ambiental é inoculada em um meio líquido suplementado com algo que force o enriquecimento (substrato específico, xenobióticos etc). Assim, estudos metagenômicos de consórcios microbianos do solo, compostagem ou sedimentos marinhos suplementados com biomassa de planta mostraram um enriquecimento em enzimas envolvidas na degradação da

biomassa.²⁵⁻²⁷ Esses estudos revelaram a capacidade lignocelulolítica de comunidades microbianas presentes nesses ecossistemas e expandiram nosso entendimento sobre a decomposição da biomassa. Eles fornecem oportunidades de melhorar processos biotecnológicos tais como o tratamento de esgoto, remediação de solo, degradação de restos agroindustriais, produção de biocombustíveis de segunda geração etc.²⁸⁻³⁰

No presente trabalho foi analisado o efeito de crescer uma comunidade microbiana derivada de compostagem em um meio mínimo (MM), sem fonte de carbono, ou em um meio complexo, com fonte de carbono. Ambos meios foram suplementados com bagaço de cana-de-açúcar. Através da técnica de metagenômica e seleção em placa, foi determinado como a composição da comunidade microbiana e a atividade lignocelulósica são afetadas pela disponibilidade de recursos no meio de cultura. Através da análise dos resultados, ficou demonstrado que a comunidade microbiana crescida no MM possui, na verdade, maior diversidade e maior proliferação de microrganismos capazes de degradar a lignocelulose em comparação com a comunidade microbiana crescida no meio rico em nutrientes. Em seguida, esse MM, suplementado com bagaço de cana-de-açúcar como única fonte de carbono, foi usado para cultivar novamente uma comunidade microbiana proveniente de compostagem. A desconstrução da biomassa lignocelulósica pelo consórcio microbiano foi monitorada através de microscopia eletrônica de varredura, ressonância magnética nuclear de estado sólido e microscopia confocal. Para obter informação funcional da comunidade microbiana nesse ambiente, o sequenciamento do metatranscriptoma foi realizado. Os fragmentos sequenciados foram filtrados e usados para montagem *de novo* do metatranscriptoma, no qual foram preditos 283.356 fragmentos codificantes. Esse estudo revelou que o consórcio microbiano emprega um grande número de enzimas ativas em carboidratos (CAZymes, do inglês, *carbohydrate-active enzymes*) para digerir a lignocelulose e permitir o seu crescimento. 27 genes com alta expressão e que mostraram homologia a CAZymes foram selecionados para estudos de expressão, os quais permitiram a identificação da primeira exoxilanase da família 11 de hidrolase de glicosídeo (GH, do inglês, *glycoside hydrolase*). Logo, nesse trabalho foi identificada um modo de ação enzimática inédito que pode facilitar a conversão da lignocelulose.

2 REVISÃO BIBLIOGRÁFICA

2.1 Consumo e produção de energia

Com o aumento da população mundial, a população moderna cada vez mais dependente de eletricidade e o desenvolvimento econômico dos países subdesenvolvidos, o consumo de energia no mundo tem aumentado ano a ano, como mostrado na figura 1a. No Brasil, por exemplo, houve um aumento de 50% entre 2003 e 2015.³¹ As fontes de energia de hidrocarbonetos como o carvão, petróleo e gás natural fornecem uma limitada e não renovável fonte de abastecimento. Mesmo assim, 86 e 65% da energia produzida no mundo e no Brasil, respectivamente, foi derivada dessas fontes em 2015 (Figura 1b).³¹ Portanto, para suprir a necessidade crescente por energia, torna-se imprescindível investir em fontes energéticas renováveis.³²

a

b

Figura 1 - Consumo de energia no Brasil e no mundo. a) Progressão do consumo mundial de 1965 a 2015. b) Matriz energética brasileira e mundial, em 2015.

Fonte: Adaptada de BP STATISTICAL REVIEW OF WORLD ENERGY.³¹

2.2 O bagaço de cana-de-açúcar como biomassa para a produção de etanol

Devido ao baixo custo de colheita, transporte e armazenagem, além da alta concentração de carboidratos e baixo conteúdo relativo de lignina, o bagaço de cana-de-açúcar é uma das principais matérias-primas utilizadas para a produção de etanol em países tropicais.³³ No Brasil, sua produção foi de 666 milhões de toneladas na safra de 2016 - mais que o dobro do produzido em 2003.³⁴ Ele é composto por cerca de 43% de celulose, 25% de hemicelulose e 23% de lignina.³⁵ A riqueza em celulose desse resíduo, faz dele um importante

recurso para a produção de glicose através do processo de sacarificação.³⁶ A partir do bagaço de cana-de-açúcar seria possível aumentar a produção de etanol, por hectare plantado, em pelo menos 43%.³⁷ Adicionalmente, a energia gerada a partir dele não acrescenta dióxido de carbono na atmosfera, o principal gás causador do aquecimento global.³⁸ Portanto, a biomassa de plantas é uma importante fonte energética renovável. Além disso, o Brasil e outros países tropicais possuem abundantes recursos de biomassa e estudos que visem o aproveitamento dessas fontes na produção de energia mostram-se de grande importância.

2.3 Biomassa

O termo biomassa é definido como todo o recurso renovável de origem animal ou vegetal que pode ser transformado em energia. Conforme esquematizado na figura 2, a biomassa vegetal é composta primariamente pelos polímeros celulose, hemicelulose e lignina, que se ligam covalentemente e interagem através de interações fracas. Essa estrutura, chamada de lignocelulose, é o recurso mais abundante na natureza.³⁹⁻⁴⁰ Entretanto, a biomassa de plantas é uma fonte de difícil exploração devido a dificuldade de degradação de seus componentes.⁴¹ O peso seco de uma parede celular consiste tipicamente de 35 a 50% de celulose, 20 a 35% de hemicelulose e 10 a 25% de lignina. Tais valores variam com a espécie da planta e o tipo celular.⁴²

Figura 2 - Representação esquemática da disposição das moléculas de celulose, hemicelulose e lignina na parede celular das plantas. Na estrutura da lignina, G, S e H representam, respectivamente, os ácidos coniferílico, sinapílico e *p*-cumarílico.

Fonte: Adaptada de RUBIN.⁴³

A lignina é um heteropolímero ramificado composto por diferentes unidades aromáticas. As principais unidades são os monolignóis álcool coniferílico, sinapílico e *p*-cumarílico, cujas estruturas estão desenhadas na figura 3.⁴⁴ Sua função primária é manter a matriz lignocelulósica coesa, conferindo suporte estrutural à planta, impermeabilidade, alta resistência à degradação e proteção contra a invasão de patógenos e predadores.^{43,45-46} Essa característica se deve ao fato da lignina formar uma estrutura tridimensional hidrofóbica que se colapsa entre as moléculas de celulose, pectina e hemicelulose, formando ligações cruzadas entre elas.⁴⁵ Consequentemente, esse polímero cria dificuldades no uso da lignocelulose como matéria-prima para a produção de biocombustíveis.⁴⁶

Figura 3 - Estrutura dos principais monolignóis componentes da lignina.

Fonte: Adaptada de BOERJAN.⁴⁷

A hemicelulose é a segunda fonte de carbono renovável mais abundante na natureza.⁴⁸ Trata-se de um polímero heterogêneo e ramificado de pentoses (xilose, arabinose), hexoses (manose, glicose, galactose) e ácidos de açúcar. A hemicelulose de angiospermas (*hardwood*, como cana-de-açúcar, milho, trigo, *switchgrass*) é composta principalmente por xilano, enquanto a de gimnospermas (*softwoods*) possui principalmente glucomananos.⁴⁹ O xilano possui um esqueleto homopolimérico formado por resíduos de 1,4-β-D-xilose. A esse esqueleto podem estar ligados arabinose e os ácidos acético/ferúlico/p-cumárico/(4-O-metil)glicurônico (figura 4). A frequência e composição dessas ramificações dependem da fonte do xilano, sendo que o grau de polimerização de *hardwoods* (150 a 200) é maior que de *softwoods* (70 a 130).⁴⁹

Figura 4 - Estrutura geral do xilano, principal componente da hemicelulose de angiospermas. Na figura estão representados os seus principais constituintes.

Fonte: Adaptada de DÉJEAN.⁵⁰

A celulose constitui o polímero mais abundante da Terra⁵¹ e tem sido utilizada pelo homem há séculos. De um terço a metade do tecido das plantas é constituído por celulose, a qual é constantemente ressintetizada pela fotossíntese. Sua biossíntese anual é estimada em cerca de 10^{11} toneladas.⁵¹ Ela é constituída por um homopolímero linear de moléculas de glicose, unidas por ligações β -1,4 e com alto peso molecular (figura 5). Nos vegetais, a molécula de celulose é arranjada em fibrilas (várias moléculas de celulose paralelas e ligadas entre si por ligações de hidrogênio e forças de van der Waals). Tais fibrilas ocorrem ligadas à lignina e à hemicelulose.⁵² Elas possuem áreas ordenadas, nas quais as ligações de hidrogênio ocorrem de forma regular, e outras desordenadas. Essas são denominadas, respectivamente, por áreas cristalinas e amorfas. A porção cristalina das fibrilas de celulose são mais resistentes à hidrólise enzimática.⁵³

Figura 5 - Representação de uma molécula de celulose, destacando a celobiose (unidade repetitiva do polímero).

Fonte: Adaptada de STONE.⁵⁴

2.4 Sacarificação por hidrólise enzimática

Na parede celular de plantas, a celulose, hemicelulose e lignina estão fortemente associadas formando uma estrutura rígida e resistente ao ataque enzimático.⁵⁵⁻⁵⁶ Para converter esse material em etanol são necessárias as etapas de pré-tratamento, hidrólise, fermentação e destilação.⁵⁷

O pré-tratamento ideal de materiais lignocelulósicos teria a finalidade de separar e recuperar as frações de lignina, hemicelulose e celulose, permitindo que o componente celulósico seja hidrolisado com baixas cargas enzimáticas.⁵⁸ Embora isso pareça irreal, o pré-tratamento, de fato, é capaz de reduzir a estabilidade das fibras, tornando o substrato mais acessível às enzimas que convertem esse substrato em açúcares fermentescíveis.⁵⁹ Diferentes pré-tratamentos afetam de forma diferente os componentes da lignocelulose. Assim, estudos já mostraram que dependendo do método utilizado, é possível obter aumentos significativos no rendimento de hidrólise do bagaço de cana-de-açúcar.⁶⁰ Entretanto, as técnicas de pré-

tratamento termoquímicas, em geral, produzem químicos que são inibidores para as enzimas da etapa de hidrólise subsequente.⁶¹

O processo de conversão enzimática da celulose em açúcares solúveis é chamado de sacarificação. Apesar da estrutura simples da celulose, a sacarificação é um processo complexo que envolve vários tipos de celulases e enzimas acessórias, conforme esquematizado na figura 6.⁶² Tais enzimas podem ser agrupadas em: endoglucanases, exoglucanases, β -glicosidases, monooxigenases líticas de polissacarídeos e proteínas que promovem o afrouxamento da fibra (expansinas, swollenina e domínios de ligação a carboidratos).⁶³⁻⁶⁵ Já que a celulose é um polímero muito grande e não atravessa a membrana plasmática, todos os organismos que degradam a celulose secretam tais enzimas. O complexo enzimático secretado age de forma sinérgica.⁶⁶

Figura 6 - Modelo da digestão da celulose.

Fonte: Adaptada de SILVEIRA.⁶⁷

Mesmo com o pré-tratamento da biomassa, grandes quantidades de enzimas são necessárias para a conversão enzimática da celulose e hemicelulose em etanol. Desta forma, os principais obstáculos enfrentados na produção em larga escala de etanol a partir de biomassa lignocelulósica referem-se ao grande consumo de energia durante o pré-tratamento e ao alto custo das enzimas utilizadas.⁶⁸⁻⁶⁹ Essa barreira poderá ser ultrapassada com a combinação de abordagens que aumentem a digestibilidade inerente da biomassa utilizada e com a descoberta de novas enzimas que auxiliem na digestão da matéria-prima de forma mais efetiva.

2.5 CAZy

As enzimas que agem na degradação, modificação ou criação de ligações glicosídicas são organizadas na base de dados CAZy em famílias de acordo com a sua sequência de aminoácidos (<http://www.cazy.org>). Como a sequência de aminoácidos está relacionada com a estrutura tridimensional da proteína, membros de uma mesma família (com raras exceções) possuem a maquinaria catalítica, mecanismo molecular e estereoquímica do produto liberado conservadas. Por outro lado, uma família geralmente possui membros com especificidade por diferentes substratos.⁷⁰⁻⁷¹

As enzimas do CAZy que agem na clivagem ou construção de carboidratos complexos estão organizadas nas classes chamadas de GHs, liases de polissacarídeo (PLs, do inglês, *polysaccharide lyases*), esterases de carboidrato (CEs, do inglês, *carbohydrate esterases*), transferases de glicosil (GTs, do inglês, *glycosyltransferases*), além dos módulos de ligação a carboidrato (CBMs, do inglês, *carbohydrate-binding modules*). Mais recentemente a base de dados incluiu uma nova classe, chamada de atividades auxiliares (AAs), para incluir enzimas que agem na conversão da lignina e oxidação de carboidratos. A razão para incluir as enzimas que atuam na lignina é que, apesar delas não agirem diretamente em carboidratos, elas facilitam o acesso das demais enzimas aos carboidratos da parede celular de plantas.⁷²

Celulose é o polissacarídeo mais abundante encontrado na natureza.⁵¹ As celulasas podem ser divididas em celobiohidrolases, endoglucanases e β -glicosidasas. Celobiohidrolases típicas são encontradas nas famílias GH 6, 7 e 48. Já as endoglucanases são encontradas em mais de 10 famílias, sendo que as famílias GH 5, 7, 9, 12, 45 e 48 são representativas. Xilano é o segundo polissacarídeo mais abundante da natureza e é o principal tipo de hemicelulose. Xilanases típicas são encontradas nas famílias GH10 e 11.⁷³⁻⁷⁴ A escolha de alvos no metatranscriptoma para estudo posterior foi feita escolhendo-se representantes das principais GHs discutidas acima.

2.6 Metagenômica

Apesar de não enxergarmos os microrganismos, eles são essenciais para a manutenção da vida na terra. De fato, o próprio organismo humano depende de microrganismos que vivem em sua pele, olho, mucosa oral e trato gastrointestinal para sobreviver. Essa associação pode ser chamada de superorganismo.⁷⁵ As funções desempenhadas pelos microrganismos são conduzidas através de comunidades complexas – compostas por entidades que se adaptam de

maneira rápida e flexível a mudanças ambientais. Entretanto, historicamente, o estudo dos microrganismos se focou no estudo de espécies isoladas em cultura pura. Assim, o entendimento dessas comunidades complexas dependia do entendimento dos membros individuais, o qual é obtido sob circunstâncias não usuais e não naturais de crescimento em meios de cultura puros, ou seja, sem contexto ecológico.⁷⁶ Além disso, foi estimado que cerca de apenas 1% dos microrganismos são cultiváveis em cultura pura.^{11-12,77}

Dessa forma, surgiu a necessidade do desenvolvimento de novas técnicas de estudo de comunidades microbianas complexas, tal como a metagenômica. O termo metagenômica inclui a caracterização de comunidades ou de parte de seus membros, em nível de genoma, através de técnicas independentes de cultivo. Também inclui o estudo de genes específicos dessas comunidades.⁷⁶ Portanto, a metagenômica abrange o estudo de marcadores filogenéticos de comunidades microbianas, tais como o gene que codifica o rRNA 16S ou a região espaçadora transcrita interna (ITS, do inglês, *internal transcribed spacer*). Esses genes fornecem um panorama de quais taxas de procariotos e fungos, respectivamente, estão presentes em uma dada amostra.^{11,78} A investigação dos taxas presentes em uma amostra, baseada no sequenciamento da região V4 do gene do rRNA 16S (V4-16S) e da região ITS tem uma grande importância, já que elas podem ser usadas na documentação da biodiversidade inexplorada, presente em comunidades microbianas complexas.⁷⁹

Tais estudos de metagenômica, entretanto, só se tornaram possíveis devido a melhorias na tecnologia de sequenciamento do DNA nos últimos anos. O impacto desses estudos no número de proteínas conhecidas foi demonstrado através do primeiro projeto de metagenômica em larga escala de um ambiente microbiologicamente rico: o mar de Sargasso. Nesse projeto foram identificadas 1,2 milhão de quadros de leitura abertos (ORFs, do inglês, *open reading frames*) que possivelmente codificam para proteínas.⁸⁰ Até 95% das proteínas identificadas nesse projeto se mostraram únicas, ou seja, não possuíam homólogas conhecidas. Além disso, a maioria delas era proveniente de espécies que nunca foram estudadas antes.⁸¹ Portanto, muitas enzimas que poderiam tornar processos biotecnológicos mais eficientes atualmente são desconhecidas pela falta de estudos de metagenômica. Por exemplo, até há pouco tempo a degradação do polímero mais abundante na natureza, a celulose, havia sido investigada apenas em alguns organismos modelos, tais como o *Clostridium thermocellum* e o *Trichoderma reesei*.⁸²

2.7 Metatranscriptômica

Enquanto metagenômica estuda o genoma de uma comunidade microbiana, metatranscriptômica estuda o subconjunto de genes que são transcritos sob certas condições ambientais por essa comunidade.⁸² Até há pouco tempo, os estudos de perfis de expressão eram desenvolvidos com organismos isolados, cujo genoma já havia sido estudado, pois eram desenvolvidos utilizando *microarrays* de DNA.⁸³ Eles forneceram compreensão de como certos fenótipos (resistência a radiação, patogênese ou resistência a choque térmico) estão correlacionados com a expressão gênica.⁸⁴⁻⁸⁶ Entretanto, a necessidade de conhecer previamente o genoma do organismo estudado e o uso de *microarrays* limitava sua aplicação.

Por outro lado, técnicas de extração de mRNA de bactérias, arqueobactérias e eucariotos presentes em amostras ambientais foram descritas recentemente.⁸⁷⁻⁸⁸ Além disso, o desenvolvimento de equipamentos com tecnologia de sequenciamento de nova geração (454 da Roche, MiSeq da Illumina e SOLiD da Applied Biosystems) aumentou a capacidade de sequenciamento e diminuiu o seu custo. Assim, esses avanços tecnológicos permitiram o desenvolvimento da técnica de metatranscriptômica, como demonstrado para comunidades microbianas que habitam o mar⁸⁹ e o solo.⁹⁰

Estudos de metagenômica e metatranscriptômica também foram aplicados para a descoberta de enzimas associadas à digestão da lignocelulose.^{18, 91-96} Por exemplo, Takasaki et al.⁹¹ identificou 56.084 ORFs. Desses, 9.449 foram classificados quanto a função, sendo 129 CAZymes. Xia et al.⁹² descobriu 31.499 ORFs a partir de 1.2 gigabases sequenciados de um consórcio de microrganismos termofílicos. Nesse conjunto de dados foram encontradas 253 CAZymes termoestáveis, onde metade possuía menos de 50% de similaridade a proteínas não redundantes da base de dados do NCBI (<http://www.ncbi.nlm.nih.gov>). Outro estudo sequenciou 268 gigabases do metagenoma de bactérias do rúmen de vaca.⁹³ A partir dessa enorme quantidade de dados adquiridos, foram encontrados 27.755 CAZymes. 90 foram escolhidas para clonagem e caracterização funcional. 57% eram ativas contra algum dos substratos celulósicos testados e 20% mantinham a atividade contra biomassa que havia sido pré-tratada com líquidos iônicos. Essa forma de pré-tratamento é bastante efetiva, entretanto conhecida por inibir a atividade enzimática.⁹⁷ Logo, esse trabalho demonstrou a possibilidade do uso de metagenômica na descoberta de enzimas resistentes a biomassa oriunda desse tipo de pré-tratamento, permitindo o seu uso.

2.8 Compostagem

O solo cobre quase toda a área terrestre do planeta e possui uma função ecológica indispensável no ciclo global do carbono, nitrogênio e enxofre. Devido a sua complexidade físico-química, o solo possui uma vasta biodiversidade filogenética e funcional.⁹⁸ Entretanto, essa diversidade é pouco explorada, pois estima-se que apenas 1% dos microrganismos presentes nele podem ser cultivados em cultura pura.⁹⁹ Compostagem pode ser definida como um processo no qual diferentes materiais orgânicos (por exemplo, sobras de alimento, estrume, galhos, folhas de plantas etc) são empilhados e deixados apodrecer até que possam ser usados como adubo ou descartados.¹⁰⁰ Esse processo, esquematizado na figura 7, é promovido por microrganismos capazes de quebrar matéria orgânica rica em lignocelulose, muitos dos quais provenientes do solo. Portanto, trata-se de um ambiente rico para a prospecção de enzimas que ajam na conversão da lignocelulose.

Figura 7 - Ilustração da evolução de um processo de compostagem. Na parte superior inicia-se o processo com restos de plantas e alimentos que são decompostos ao longo do tempo pelos microrganismos presentes. A parte inferior representa os estágio final, no qual o material se transforma em um fertilizante que pode ser aplicado em novas plantas.

Fonte: BARR.¹⁰¹

3 JUSTIFICATIVA

A preocupação com o aquecimento global, o aumento no custo do petróleo bruto, sua reserva finita e questões de segurança nacional mostram claramente que não podemos depender unicamente do petróleo como matéria-prima para combustíveis, materiais e químicos. Como alternativa, a biomassa lignocelulósica de plantas pode ser usada como matéria-prima sustentável para produção das *comodities* derivadas do petróleo através do processo de biorefinaria.¹⁰²⁻¹⁰³

No Brasil, há um grande potencial no aumento da produtividade de etanol a partir do bagaço de cana-de-açúcar, já que o caldo extraído e utilizado na produção de etanol de primeira geração armazena apenas um terço da energia presente na planta.¹⁰⁴ O bagaço que sobra após a extração do caldo é queimado para a produção de eletricidade, mas poderia ser usado de forma muito mais rentável como matéria-prima para produção de etanol de segunda geração.^{37,104}

As abordagens recentes na hidrólise da lignocelulose empregam o uso de coquetéis com enzimas de fungos que, apesar de efetivos na hidrólise de polissacarídeos, requerem que a biomassa passe por um pré-tratamento severo para remover a lignina, diminuir a cristalinidade da celulose e aumentar a acessibilidade às enzimas.¹⁰⁵ Além disso, esses coquetéis devem ser usados em altas concentrações.¹⁰⁶ A combinação desses dois fatores torna o etanol e os químicos produzidos a partir da biomassa lignocelulósica muito caros para competir com aqueles oriundos do petróleo.

Por outro lado, comunidades microbianas promovem a sacarificação da biomassa eficientemente na natureza. Como cerca de apenas 1% desses microrganismos são hábeis de serem cultivados em laboratório em cultura pura,¹¹ a descoberta de novas enzimas envolvidas na digestão da lignocelulose, que sejam mais eficientes e que atuem através de novos mecanismos, pode ser potencializada com o uso de abordagens de metagenômica ou metatranscriptômica. Também fica claro que a descoberta de tais enzimas é essencial para tornar o etanol de segunda geração economicamente viável.

4 OBJETIVOS

4.1 Objetivo geral

O objetivo geral deste trabalho é a descoberta de novas enzimas oriundas de uma comunidade microbiana da compostagem que atuem na digestão da biomassa lignocelulósica e que sejam mais eficientes, estáveis e/ou atuem através de novos mecanismos. Espera-se que as enzimas estudadas, e as presentes no metatranscriptoma sequenciado, possam contribuir para que no futuro seja produzido um etanol de segunda geração economicamente viável.

4.2 Objetivos específicos

4.2.1 Análise do meio de cultura utilizado no crescimento

- Crescer microrganismos proveniente de uma compostagem em um meio mínimo (definido) e em meio complexo;
- Identificar qualitativamente as diferenças na degradação do bagaço de cana-de-açúcar;
- Identificar a capacidade celulolítica das comunidades microbianas;
- Identificar a taxonomia dos microrganismos presentes.

4.2.2 Análise da degradação do bagaço de cana-de-açúcar por microrganismos presentes na compostagem

- Crescer os microrganismos proveniente de uma compostagem;
- Caracterizar química e morfológicamente a biomassa após o crescimento microbiano;
- Determinar a quantidade relativa de procariotos e fungos na comunidade microbiana;
- Obter o metatranscriptoma dessas comunidades.

4.2.3 Estudo de alvos do metatranscriptoma

- Escolher alvos interessantes;
- Obter os alvos através de expressão em *Escherichia coli* Rosetta™ (DE3);
- Caracterizar bioquimicamente esses alvos;
- Entender as bases moleculares da atividade observada através da resolução da estrutura tridimensional.

5 MATERIAIS E MÉTODOS

5.1 Coleta de amostras e cultivo para estudo do meio de cultura

Amostras de compostagem foram coletadas entre janeiro e abril de 2013 da composteira do projeto USP Recicla (campus de São Carlos) durante a fase mesofílica final em locais 30 cm abaixo da superfície. A amostra foi homogeneizada e 250 mg foi usado para inocular 25 mL de meio de cultura. Os meios testados foram o meio PCS (0,1% extrato de levedura, 0,5% peptona, 0,5% CaCO_3 , 0,5% NaCl pH 8,0)¹⁰⁷ e o meio mínimo (KCl 0,52 g/L, KH_2PO_4 0,815 g/L, K_2HPO_4 1,045 g/L, MgSO_4 1,35 g/L, NaNO_3 1,75 g/L, elementos traço).¹⁰⁸ A cultura foi suplementada com 2% (m/v) de bagaço de cana-de-açúcar e incubada a 30 °C com agitação de 150 rpm por até 5 semanas. O bagaço de cana-de-açúcar foi gentilmente cedido pelo Grupo Cosan (Ibaté, São Paulo). Antes do uso, o bagaço foi lavado com água quente (50 °C) e secado a 50 °C.

5.2 Triagem de organismos celulolíticos no consórcio derivado da compostagem

Uma diluição seriada (com passos de 10) da cultura foi espalhada em placas de Petri contendo o respectivo meio de cultura suplementado com ágar e 1% (m/v) de carboximetilcelulose (CMC). Placas com o meio mínimo foram também suplementadas com 1% (m/v) de dextrose para permitir o crescimento de microrganismos incapazes de degradar CMC. As placas foram incubadas a 30 °C por 1 semana e o número de unidades formadoras de colônia (UFCs) foi registrado. Em seguida, as colônias foram removidas das placas através de lavagem com água. As placas lavadas foram coradas com 1% (m/v) de vermelho de congo por 15 minutos e descoradas com solução descorante (1 M de NaCl , 5 mM de NaOH) por 20 minutos. Os organismos que secretam endocelulases promovem a quebra do CMC a sua volta, resultando na formação de halos brancos, já que o vermelho de congo interage com a celulose.

5.3 Sequenciamento da região V4 do gene do rRNA 16S

A cultura microbiana foi centrifugada a $2.000\times g$ por 10 minutos e 50 mg da fração insolúvel foi usado para extração de DNA genômico. A extração foi feita com o kit comercial FastDNA Spin Kit for Soil (MP Biomedicals, Solon, OH, USA), de acordo com

recomendações do fabricante. A região hipervariável V4-16S de bactéria foi amplificada usando 50 ng de DNA metagenômico como molde e os oligonucleotídeos universais bacterianos

515F (5'-

TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGTGCCAGCMGCCGCGGTAA-3')

e 806R (5'-

GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGGACTACHVGGGTWTCTAAT-

3'), modificados pela adição de adaptadores para permitir o sequenciamento na plataforma Illumina (sublinhado). A amplificação foi realizada pela reação em cadeia da polimerase (PCR, do inglês, *polymerase chain reaction*) com as seguintes condições: desnaturação inicial a 94 °C por 2 minutos; 25 ciclos de 94 °C por 30 segundos, 57 °C por 60 segundos, 72 °C por 40 segundos; elongação final a 72 °C por 10 minutos. Os fragmentos foram purificados com o kit comercial Wizard SV gel and PCR Clean-UP System (Promega, Madison, WI, EUA). Os adaptadores P5 e P7, necessários para sequenciamento na plataforma Illumina, e índices para permitir a distinção entre as amostras foram adicionados com o kit comercial Nextera DNA Library Prep Kit (Illumina, Carlsbad, CA, EUA), conforme instruções do fabricante. A concentração de DNA foi medida com o kit comercial KAPA Library Quantification Kit (Kapa Byosystems, Wilmington, MA, EUA). As amostras foram misturadas em quantidades equimolares e sequenciadas no equipamento MiSeq (Illumina), produzindo fragmentos *paired-end* com 2x150 pares de base (pb).

5.4 Bioinformática e análise dos dados

Os fragmentos *paired-end* foram fundidos com o programa FLASH v1.2.7.¹⁰⁹ Os próximos passos foram feitos usando o programa QIIME v1.8.0,¹¹⁰ exceto onde indicado de outra forma. Os fragmentos foram filtrados a partir da base com fator de qualidade menor do que 20, menores do que 75 pb ou que possuem mais de uma base indeterminada. Unidades taxonômicas operacionais (OTUs, do inglês, *operational taxonomic units*) foram atribuídas a grupos com 97% ou mais de identidade com o algoritmo UCLUST.¹¹¹ O algoritmo UCHIME¹¹² foi utilizado para detectar e filtrar quimeras. Grupos taxonômicos foram atribuídos às OTUs utilizando o banco de dados do GreenGenes¹¹³ lançado em maio de 2013 com o método RDP Classifier. OTUs vistos apenas uma vez foram descartados para reduzir ruído e variabilidade. As diversidades alfa e beta foram calculadas após o conjunto de dados ser rarefeito para 40.000 e 10.000 fragmentos por amostra, respectivamente. Análise de coordenadas principais (PcoA, do inglês, *principal coordinate analysis*) foi usada para

visualizar o agrupamento das amostras baseado na matriz de distâncias UniFrac ponderada. PAST v2.11¹¹⁴ foi usado para calcular o índice de diversidade Shannon (com os dados rarefeitos para 10.000 fragmentos por amostra) e a análise de similaridade percentual (SIMPER). Através da análise de SIMPER é possível determinar quais OTUs contribuem mais para a dissimilaridade entre as comunidades.

5.5 Coleta de amostras e cultivo para estudo da degradação do bagaço de cana-de-açúcar

Amostras de compostagem foram coletadas em fevereiro de 2014 no mesmo local e estágio que citado no item 5.1. Amostra homogeneizada foi usada para inocular (1%, m/v) meio mínimo preparado como descrito anteriormente. A cultura foi suplementada com 3% (m/v) de bagaço de cana-de-açúcar e incubada a 30 °C com agitação de 150 rpm por até 5 semanas. O bagaço usado também foi lavado e secado a 50 °C. Semanalmente, 3 amostras eram retiradas para análise. A biomassa e os microrganismos presentes foram separados do sobrenadante por centrifugação a 3,000×g por 5 minutos a temperatura ambiente.

5.6 Análise da biomassa

Antes das análises, o bagaço de cana-de-açúcar obtido após o crescimento da comunidade microbiana foi lavado, secado e moído a um pó fino, usando um moinho de bolas (TissueLyser II, Qiagen, Hilden, Alemanha) por 60 segundos a 30 Hz. O bagaço bruto, não utilizado no cultivo, foi usado como controle.

Experimentos de ressonância magnética nuclear de estado sólido (ssNMR, do inglês, *solid state nuclear magnetic resonance*) foram realizados em um espectrômetro Varian INOVA (Varian, Palo Alto, California, EUA) na frequência de 100,5 e 400 MHz para ¹³C e ¹H, respectivamente. Uma sonda Jackobsen de 7 mm de dupla ressonância com rotação da amostra em torno do ângulo mágico foi usada. A frequência de rotação de 4,5 kHz foi controlada por um sistema pneumático que garante precisão em torno de 2 Hz. Rampa de rádio frequência na polarização cruzada com rotação em torno do ângulo mágico (CPMAS, do inglês, *cross polarization under magic angle spinning*) combinado com supressão total de bandas laterais através de rotação (TOSS, do inglês, *total suppression of spinning sidebands*) foi usado para excitação do ¹³C. Foram usados pulsos $\pi/2$ com duração de 5 μ s para o ¹³C e ¹H. Cada medida foi repetida aproximadamente 4.000 vezes. Todos os espectros foram

normalizados com respeito a sua área. O deslocamento químico foi atribuído a grupamentos de acordo com estudos publicados.^{60,115-122} A proporção relativa de celulose, hemicelulose e lignina foi estimada pela área de picos específicos. As proporções iniciais desses componentes foi extraída de Lima et al.³⁶

Microscopia confocal foi realizada em um microscópio Zeiss LSM 780 com laser Chamaleon (Coherent) de Ti:safira como fonte de excitação por dois fótons. O bagaço moído foi hidratado por 24 horas e observado com lentes objetivas C-apocromáticas (Zeiss) com aumento de 20X, abertura numérica de 1.2 e imersão em água. Os espectros foram obtidos pela média de três imagens e nenhuma variação significativa foi observada. A resolução espacial foi de aproximadamente 200 nm (considerando a abertura numérica e o comprimento de onda de excitação). A potencia nominal do laser de 405 nm foi de 20% (aproximadamente 200 mW).

Microscopia eletrônica de varredura (MEV) foi realizada em um microscópio modelo JSM-6390 LV (Jeol, Tóquio, Japão) operando com uma voltagem de aceleração de 5 kV. O bagaço moído foi hidratado por 24 horas; uma gota foi aplicada diretamente ao pedestal de amostra e deixada secar a temperatura ambiente por 12 h. Após secas, as amostras foram recobertas com ouro em um metalizador modelo MED 020 (Bal-tec, Liechtenstein). As imagens foram obtidas no vácuo. Pelo menos 10 imagens por amostra foram obtidas de diferentes locais para garantir a reprodutibilidade dos resultados.

5.7 Extração de ácidos nucléicos

A fração insolúvel da cultura após centrifugação foi usada para extração de ácidos nucléicos seguindo um protocolo modificado de Griffiths et al.¹²³ 0,5 g de amostra foi adicionado a tubos de 2 mL contendo 0,5 g de esferas de vidro de 0,1 e 0,5 mm de diâmetro (cada) previamente tratadas com ácido clorídrico. Em seguida, foram adicionados 0,5 mL de tampão de extração (brometo de hexadeciltrimetilamônio 10% m/v, NaCl 700 mM misturado com um volume igual de tampão fosfato de potássio 240 mM, pH 8.0) e 0,5 mL de fenol:clorofórmio:álcool isoamílico (25:24:1, pH 8,0). Os tubos foram submetidos a agitação de 5,5 m/s por 30 segundos no equipamento Bead Ruptor 24 (Omni, Kennesaw, GA, EUA) para promover a lise celular. A fase aquosa foi transferida para um novo tubo e extraída com um volume igual de clorofórmio:álcool isoamílico (24:1) seguido de centrifugação a 17.000×g por 5 minutos a 4 °C. Os ácidos nucléicos totais foram precipitados com dois volumes de solução de precipitação (polietileno glicol 6.000 30% m/v, NaCl 1,6 M) por duas

horas a temperatura ambiente seguido de centrifugação a $17.000\times g$ por 20 minutos a $4\text{ }^{\circ}\text{C}$. Os ácidos nucleicos foram lavados duas vezes com etanol 70% (v/v), diluídos em $50\text{ }\mu\text{L}$ de água e armazenados a $-80\text{ }^{\circ}\text{C}$ até o uso. A qualidade do material obtido foi determinada no equipamento LabChip GXII (PerkinElmer, Waltham, MA, EUA).

Todas as soluções e frascos de vidros foram tratados com 0,1% de DEPC por 16 horas e, então, autoclavados para criar um ambiente livre de RNases. Todo material plástico usado possuía certificação de serem livres de RNase e DNase.

5.8 PCR em tempo real do DNA genômico isolado

Os ácidos nucleicos extraídos da cultura foram diluídos para $200\text{ ng}/\mu\text{L}$ e tratados com 1:100 (v/v) de RNase A:ácido nucleico (Thermo Fisher Scientific, Waltham, MA, EUA) por 15 minutos a $37\text{ }^{\circ}\text{C}$. Em seguida, o DNA metagenômico foi extraído com fenol:clorofórmio:álcool isoamílico (25:24:1), conforme descrito anteriormente, sendo que a precipitação foi realizada com 1/10 do volume de acetato de sódio 3 M, pH 5,2 e três volumes de etanol. As amostras foram incubadas em gelo por 30 minutos e centrifugadas a $17.000\times g$ por 30 minutos a $4\text{ }^{\circ}\text{C}$. O sobrenadante foi descartado. O DNA genômico foi lavado com etanol 70% e diluído em água.

PCR em tempo real foi realizado utilizando o DNA metagenômico purificado como molde e os oligonucleotídeos universais para bactéria (515F e 806R) e fungos (ITS1 e 5.8S). A sequência dos oligonucleotídeos usados foram 515F ($5'\text{-CCAGCMGCCGCGGTAA-}3'$), 806R ($5'\text{-GGACTACHVGGGTWTCTAAT-}3'$),⁷⁹ ITS1 ($5'\text{-TCCGTAGGTGAACCTGCG-}3'$) e 5.8S ($5'\text{-CGCTGCGTTCTTCATCG-}3'$).¹²⁴ As reações de PCR foram realizadas com soluções de $20\text{ }\mu\text{L}$ contendo: 2.5 ng de DNA, 300 nM de cada par de oligonucleotídeo e $10\text{ }\mu\text{L}$ do KAPA SYBR® FAST qPCR Master Mix (KAPA Biosystems, Wilmington, MA, EUA), o qual contém os nucleotídeos, polimerase, tampão e corante verde SYBR necessários para a reação. As condições utilizadas no termociclador foram as seguintes: desnaturação inicial a $95\text{ }^{\circ}\text{C}$ por 5 minutos; 35 ciclos de $95\text{ }^{\circ}\text{C}$ por 30 segundos e $60\text{ }^{\circ}\text{C}$ por 45 segundos, de acordo com as recomendações do fabricante. As medidas foram realizadas no equipamento CFX96 Real-Time System (Bio-Rad, Hercules, CA, EUA). Todas as reações foram realizadas em triplicata em dois experimentos independentes.

5.9 Síntese da biblioteca de cDNA e sequenciamento

Os ácidos nucleicos extraídos anteriormente foram diluídos para 200 ng/μL e tratados com DNase I (Invitrogen, Waltham, MA, EUA), conforme instruções do fabricante. As réplicas biológicas foram combinadas e o rRNA de procariotos foi removido utilizando o kit comercial RiboZero Magnetic Kit Bacteria (Epicentre). O RNA remanescente foi purificado com o kit comercial RNA Clean & Concentrator-5 (ZymoResearch, Irvine, CA, EUA). Em seguida, foi usado o kit comercial TruSeq Stranded Total RNA Sample Preparation kit (Illumina) para remover o rRNA de eucariotos e gerar uma biblioteca de cDNA de ~450 pb. O sequenciamento da biblioteca de cDNA de cada tempo de cultura foi realizado no equipamento MiSeq com um cartucho de 500 ciclos v2 (Illumina).

5.10 Montagem e anotação do metatranscriptoma

Os fragmentos sequenciados foram cortados e filtrados para remover sequências dos adaptadores e com baixa qualidade (com uma pontuação de qualidade média das últimas quatro bases menor do que 15) usando o software Trimmomatic.¹²⁵ Sequências menores do que 100 pb foram descartadas. O software SortMeRNA¹²⁶ foi usado para fusionar os fragmentos *paired-end* e remover as sequências contaminantes de rRNA, as quais foram identificadas usando as bases de dados de referência Silva¹²⁷ e Rfam¹²⁸ com *e-value* de corte de 1×10^{-5} . Os fragmentos remanescentes foram montados pelo pacote de montagem *de novo* Trinity.¹²⁹ Genes foram preditos no metatranscriptoma com o plugin TransGeneScan.¹³⁰ Níveis de expressão foram calculados com o eXpress¹³¹ usando Bowtie2¹³² para mapear os fragmentos. A origem filogenética dos ORFs preditos foi analisada usando o MEGAN v6¹³³ e o algoritmo de ancestral comum mais próximo com parâmetros padrões. O software HTSeq¹³⁴ foi usado para subamostrar os fragmentos sequenciados para permitir gerar curvas de rarefação mostrando como os níveis de expressão calculados variam com a profundidade de sequenciamento. Curvas de rarefação também foram calculadas usando o software BBMap (sourceforge.net/projects/bbmap/) para determinar a adequação da profundidade de sequenciamento.

CAZymes foram anotadas utilizando a ferramenta de busca HMMER¹³⁵ contra a base de dados do dbCAN.¹³⁶ Os níveis de expressão de cada família GH foi apresentada em um heat map gerado usando a função heatmap.2 do pacote gplots no R.¹³⁷ Distâncias euclidianas e o método de agrupamento Ward foram usados. Os fragmentos montados foram classificados

funcionalmente na Enciclopédia de Genes e Genomas de Kyoto (KEGG, do inglês, *Kyoto encyclopedia of genes and genomes*)¹³⁸ através da ferramenta GhostKOALA¹³⁹ usando valores padrões.

5.11 Clonagem, expressão e purificação de genes alvo

Baseado em altos valores de expressão, 27 genes preditos que codificam homólogos a CAZymes foram selecionados para estudos de expressão. Os códons foram otimizados para expressão em *Escherichia coli* (<https://www.idtdna.com/CodonOpt>) e sintetizados (GenScript, Jiangsu, China) após a predição e remoção de peptídeos sinais e hélices transmembrana. Sequências adaptadoras foram adicionadas as extremidades 5' (CAGGGCGCCATG) e 3' (TAACCGCGTCGGGTC) para permitir a clonagem pelo método da clonagem independente de ligação (LIC).¹⁴⁰

Técnicas padrão de biologia molecular foram aplicadas.¹⁴¹ Os fragmentos de DNA foram clonados no plasmídeo pETTRXA-1a/LIC¹⁴² e transformados em células de *E. coli* Rosetta (DE3) pLys (Merck, Darmstadt, Germany). Testes de expressão e solubilidade em pequena escala foram realizados como descrito anteriormente.¹⁴² Células recombinantes na presença de glicerol 20% (v/v) foram armazenadas a -80 °C.

Células transformantes foram crescidas por 16 horas em meio LB na presença de canamicina (50 µg/mL) e clorofenicol (34 µg/mL) a 37 °C com agitação de 150 rpm e usadas para inocular (1%, v/v) 1 L de meio LB com os mesmos antibióticos. A cultura foi incubada a 37 °C com agitação de 150 rpm até que a densidade óptica a 600 nm atingisse 0,6. A temperatura de incubação foi abaixada para 17 °C e a expressão conduzida por 16 horas após indução com 0,5 mM de IPTG. As células foram precipitadas por centrifugação a 9.000×g por 20 minutos e ressuspensas em 20 mL de tampão de lise (Tris-HCl 20 mM, NaCl 300 mM, imidazol 5 mM, glicerol 5% (v/v), β-mercaptoetanol (β-ME) 10 mM, fluoreto de fenilmetilsulfonil (PMSF, do inglês, *phenylmethylsulfonyl fluoride*) 1 mM, lisozima 100 µg/mL, pH 8.0) com ou sem sarcosil 0,6% (m/v) dependendo da proteína. A suspensão celular foi incubada em gelo por 1 hora, submetida a sonicação por 6 minutos e centrifugada a 23.000×g por 30 minutos. O sobrenadante foi aplicado a 2 mL de resina de níquel-nitrilotriacético (Ni-NTA) (Qiagen, Hilden, Germany) previamente equilibrada com 10 volumes de tampão de lise. A resina foi lavada com 4 volumes de tampão de lavagem (Tris-HCl 20 mM, NaCl 300 mM, imidazol 5 mM, glicerol 5% (v/v), β-ME 10 mM, PMSF 1 mM, pH 8.0) e as proteínas ligadas foram digeridas pela adição de tampão de clivagem (Tris-HCl

20 mM, NaCl 300 mM, glicerol 5% (v/v), β -ME 10 mM, PMSF 1 mM, pH 8.0) e 1:50 TEV:proteína (medido por absorbância em 280 nm). Após incubar por 16 horas a 10 °C com homogeneização, a proteína recombinante foi eluída no tampão de clivagem. As proteínas ainda ligadas foram eluídas com 4 volumes de tampão de eluição (Tris-HCl 20 mM, NaCl 300 mM, imidazol 300 mM, glicerol 5% (v/v), β -ME 10 mM, PMSF 1 mM, pH 8.0). A proteína foi adicionalmente purificada usando uma coluna Superdex™ 75 16/60 (GE Healthcare Biosciences Corporation, Picataway, EUA) previamente equilibrada com tampão de purificação (Tris-HCl 20 mM, NaCl 200 mM, pH 8,0). A pureza da proteína foi determinada por eletroforese em gel de poliacrilamida-dodecilsulfato de sódio corado com azul de Comassie.¹⁴³

5.12 Análise das sequências e caracterização enzimática

Alinhamento múltiplo de sequências de aminoácidos foi realizado com o Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo>)¹⁴⁴ e visualizados no servidor Boxshade (http://www.ch.embnet.org/software/BOX_form.html). A atividade enzimática foi determinada pela quantidade de açúcares redutores liberados do substrato (Megazyme, Ireland; Sigma-Aldrich, St. Louis, MO, EUA) usando o método do ácido 3,5-dinitrosalicílico (DNS, do inglês, *3,5-dinitrosalicylic acid*).¹⁴⁵ Xilano foi usado na concentração final de 1% (m/v) e os demais substratos (arabinano, arabinogalactano, arabinoxilano, beta-glucano, arabinano desramificado, laminarina, galactomanano, manano, liquenano, xiloglucano), de 0,5%. Glicose foi usada como padrão. Todos os ensaios foram realizados em quadruplicata. Os ensaios de atividade foram inicialmente realizados em pH 5,0 por 30 minutos a 30 °C contra uma variedade de substratos. Em seguida, os testes foram realizados contra o substrato ótimo em tampão ABF 20 mM (acetato de sódio, borato de sódio, fosfato de sódio dibásico, 20 mM cada; pH ajustado com HCl/NaOH) com o pH variando de 2 a 10. A temperatura de reação foi variada de 30 a 85 °C usando o substrato e pH ótimos. Por fim, a atividade foi testada contra diferentes polissacarídeos nas condições de pH e temperatura ótimos da enzima. A atividade residual foi testada incubando a enzima em tampão no pH ótimo de atividade e na temperatura estudada por até 48 horas. Alíquotas foram removidas e ensaios de atividade nas condições ótimas foram realizados.

5.13 Análise da estabilidade térmica através de fluorimetria de escaneamento diferencial

Experimentos de estabilidade térmica foram realizados no equipamento CFX96 Real-Time System (Bio-Rad, Hercules, CA, EUA) com comprimento de onda de excitação e emissão de 490 e 530 nm, respectivamente. Reações de 20 μ L foram preparadas com 0,2 mg/mL de proteína em diferentes tampões e sonda Sypro orange (Invitrogen, Waltham, MA, EUA) 2.000 vezes diluída. Essa mistura foi adicionada a placas de PCR de parede fina com 96 poços (Bio-Rad, Hercules, CA, USA) seladas com fita de vedação de qualidade óptica (Bio-Rad). Todos os tampões utilizados foram preparados na concentração de 50 mM. A temperatura foi variada de 25 a 90 °C, com incremento de 1 °C por minuto. A análise e determinação da temperatura de *melting* foi realizada usando o software GraphPad Prism v5.0 (GraphPad Software, La Jolla, CA, EUA).

5.14 Identificação do produto de reação com TLC e Dionex HPLC

Os produtos da reação enzimática foram analisados por cromatografia de camada delgada (TLC, do inglês, *thin-layer chromatography*) em sílica gel 60 F254 (Merck, Darmstadt, Alemanha) com butanol:ácido acético:água (2:1:1) como eluente. As placas foram reveladas por carbonização através da exposição a ácido sulfúrico 10% (v/v) diluído em etanol e incubação a 105 °C. Uma mistura de xilose (X1) e xilooligossacarídeos com 2 a 6 resíduos de xilose (X2-X6) foram usados como padrão.

Os produtos de reação também foram analisados com o sistema de cromatografia líquida de alta performance (HPLC, do inglês, *high performance liquid chromatography*) Dionex ICS-5000 equipado com um detector eletroquímico, uma coluna de troca de ânions CarboPac PA1 (4x250 mm) e um cartucho de guarda (Thermo Fisher Scientific, Waltham, MA, EUA). O programa seguinte foi usado: fluxo de 1 mL/min; 30 °C; 0 a 30 min: 100 mM NaOH isocrático; [segmento 1] 5 a 20 min: de 0 a 20 mM acetato de sódio, [segmento 2] 20 a 24 min: de 20 a 100 mM acetato de sódio, [segmento 3] 24 a 30 min: 100 mM acetato de sódio isocrático.

5.15 Produção de xilooligômeros e quantificação

Xilooligômeros foram produzidos seguindo um protocolo modificado de Qing et al.¹⁴⁶ Em resumo, xilano de bétula 5% (m/v) foi selado em um reator de aço inoxidável de 150 mL. O reator foi transferido para um banho de areia a 330 °C. Após a temperatura alcançar o equilíbrio a 200 °C, ele foi incubado por 10 minutos e rapidamente transferido para um banho de água com gelo. Os sólidos foram removidos por centrifugação a 4.000×g por 10 minutos e o sobrenadante foi filtrado através de uma membrana hidrofóbica de fluoreto de polivinilideno (PVDF) com poros de 0,45 µm (Merck, Kenilworth, NJ, EUA). HPLC Dionex foi empregado para verificar a distribuição de xilooligossacarídeos. Os xilooligômeros foram hidrolisados em ácido sulfúrico 4% (v/v) por uma hora a 121 °C, baseado em protocolos padrões do laboratório nacional de energias renováveis (NREL, do inglês, *national renewable energy laboratory*).¹⁴⁷ A concentração de xilooligômeros totais foi determinada através de HPLC (Shimadzu LC-20AT, Quioto, Japão) equipado com detectores de índice de refração e ultra violeta-visível e uma coluna aminex HPX-87H (Bio-Rad Laboratories, Hercules, EUA). Padrões de xilose foram tratados em paralelo e usados para calcular a concentração da amostra.

5.16 Inibição de hidrólise por xilooligômeros

A hidrólise enzimática de Avicel 2% (m/v) (Sigma-Aldrich, St. Louis, MO, EUA) foi realizada em tampão citrato de sódio 50 mM pH 5,0 com ou sem xilooligômeros 8 g/L. As reações também foram feitas com ou sem 0,1 g/L da enzima C21. Azida de sódio 0,02% (m/v) foi usada pra impedir crescimento microbiano. As reações foram iniciadas pela adição de Accellerase 1500 (DuPont, Wilmington, DE, EUA) e Novozyme 188 (Sigma-Aldrich, St. Louis, MO, EUA) diluídas 25 (para 5 FPU/g de substrato) e 250 vezes (10 CBU/g de substrato) respectivamente. Os frascos foram incubado a 35 °C com agitação de 150 rpm. Brancos de substrato (sem enzima) e brancos de enzima (sem substrato) também foram feitos. Alíquotas de 0,5 mL foram removidas após 1, 4, 7, 24, 48, 72 e 96 horas de hidrólise. As alíquotas foram imediatamente fervidas por 10 minutos para inativar as enzimas, centrifugadas a 10.000×g por 1 minuto, filtradas através de membrana PVDF com poros de 0,45 µm e congeladas a -20 °C. Dois experimentos independentes foram realizados. O rendimento da conversão de celulose foi analisado por HPLC, como discutido anteriormente,

utilizando padrões de glicose. A configuração do experimento foi baseada em protocolos padrões do NREL.¹⁴⁸

5.17 Ensaios de cristalização e coleta de dados

A triagem das condições de cristalização foi realizada utilizando o robô de cristalização Honeybee (Genomic Solutions, Ann Arbor, MI, EUA) e kits comerciais. Cristais foram obtidos para a proteína C21 em múltiplas condições. Cristais crescidos a 18 °C em 0.1 M BIS-TRIS propano pH 7.5, PEG 3350 20% (m/v), 0.2 M NaI foram brevemente transferidos para uma solução crioprotetora (solução de cristalização adicionada de etilenoglicol 15% v/v) e rapidamente vitrificados em nitrogênio gasoso a 100 K. Os dados de difração foram coletados na linha MX2¹⁴⁹ do Laboratório Nacional de Luz Sincrotron (LNLS, Campinas) usando radiação sincrotron com comprimento de onda configurado para 1,459 Å, detector PILATUS2M (Dectris, Taefernweg, Suíça) e oscilação de 0.5° por imagem. A redução e integração dos dados foi realizada com o XDS.¹⁵⁰

5.18 Substituição molecular, construção do modelo e refinamento da estrutura

O modelo estrutural foi construído por substituição molecular com o programa PHASER¹⁵¹ usando as coordenadas do PDB ID 1XNK²⁰³ como modelo. Os programas PHENIX¹⁵² e Coot¹⁵³ foram usados alternadamente para rodar ciclos de construção de modelo e refinamento. A qualidade estereoquímica do modelo foi validada com o programa MolProbity.¹⁵⁴ PyMOL (the PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC) foi usado para visualizar e gerar as representações estruturais.

6 RESULTADOS E DISCUSSÃO

6.1 Triagem de organismos celulolíticos no consórcio derivado da compostagem

Para estimar a densidade de células microbianas cultiváveis que cresceram na cultura derivada da compostagem, placas de Petri com o respectivo meio foram inoculadas com uma diluição seriada do sobrenadante do cultivo. Como mostrado na figura 8, após uma semana, a cultura crescida no meio PCS possuía densidade celular 18 vezes maior que aquela crescida em meio mínimo. Na segunda semana, as culturas mostraram uma queda brusca na densidade celular em relação a primeira semana, sendo que o número de UFC/mL da cultura em MM e PCS caiu 10 e 5 vezes, respectivamente. Por outro lado, a proporção de microrganismos com atividade celulase, conforme medido pelo número de colônias que exibiram halos nas placas com CMC, foi bastante menor na cultura em PCS do que em MM.

Portanto, o meio PCS, complexo, fornece um rico conjunto de nutrientes provenientes do extrato de levedura e peptona, tais como vitaminas e aminoácidos, que sustenta o crescimento microbiano a altas densidades celulares. Entretanto, a proporção de microrganismos que degradam a lignocelulose foi baixa. Já o MM, suplementado com bagaço de cana-de-açúcar como única fonte de carbono, contém elementos básicos para o crescimento microbiano, como íons, eletrólitos (fornecidos através da mistura de elementos traços) e nitrogênio. O consórcio microbiano crescido em MM apresentou uma maior proporção de microrganismos celulolíticos do que o crescido em meio PCS.

Logo, embora a cultura em MM apresente uma densidade celular menor que a cultura em meio PCS, a fração de organismos com atividade celulolítica é maior. Consequentemente, para atender os objetivos deste estudo, o MM se mostrou mais interessante a princípio. Além disso, vale notar que uma densidade celular menor não é indicativo de menor diversidade de espécies.

Figura 8 - Número de unidades formadoras de colônia (UFC) no consórcio crescido em a) meio mínimo (MM) e b) PCS.

Fonte: Elaborada pelo autor.

Essa observação foi confirmada pela análise qualitativa da degradação do bagaço de cana-de-açúcar (figura 9). O bagaço de cana-de-açúcar remanescente após o cultivo foi filtrado, secado e examinado visualmente. Análises quantitativas da biomassa, através de sua hidrólise e separação dos açúcares por HPLC, foram previamente usadas para avaliar mudanças na composição de materiais lignocelulósicos.¹⁶⁸⁻¹⁶⁹ Entretanto, neste estudo não foi possível separar o bagaço das células microbianas devido a forte associação entre eles. Desta forma, as células sofreram hidrólise durante o método de hidrólise para quantificação por HPLC, liberando açúcares da parede celular e da matriz bacteriana. Esses açúcares microbianos afetaram a acurácia do método e valores inconsistentes foram obtidos.

A biomassa restante após cultura em MM mostrou redução no tamanho das partículas e escurecimento da biomassa. As culturas crescidas em PCS, que oferecia outras fontes de

carbono mais facilmente metabolizáveis que o bagaço de cana-de-açúcar, mostrou baixa degradação do bagaço mesmo após cinco semanas de incubação. O tamanho e morfologia da biomassa se mantiveram inalterados nessa cultura. Esses resultados indicam a baixa abundância de microrganismos especializados na digestão da lignocelulose no cultivo em meio PCS.

Ou seja, a degradação da lignocelulose foi suprimida no meio PCS por causa da disponibilidade de outras fontes de energia mais facilmente acessíveis. Já a comunidade crescida em MM apresentou e maior eficiência de degradação do bagaço de cana-de-açúcar durante as cinco semanas de estudo.

Figura 9 - Análise qualitativa da degradação do bagaço de cana-de-açúcar coletado após crescimento do consórcio microbiano em meio mínimo (MM) e PCS durante cinco semanas. É possível observar mudanças morfológicas, tais como redução no tamanho das partículas e escurecimento da biomassa, principalmente no cultivo em MM.

Fonte: Elaborada pelo autor.

6.2 Análise da diversidade microbiana no consórcio derivado da compostagem através de sequenciamento em larga escala

Para avaliar a complexidade da comunidade microbiana, foi aplicado o sequenciamento do gene V4-16S usando métodos e padrões recomendados pelo Earth Microbiome Project.¹⁷⁰ Essa técnica é amplamente usada para estimar a abundância relativa de microrganismos em comunidades microbianas. Um total de 1.224.357 fragmentos foram obtidos (~170.000 fragmentos por amostra). O número mínimo de fragmentos para uma dada amostra foi de 46.999, com exceção da amostra extraída após cinco semanas de crescimento

em meio PCS, para a qual foram obtidos 11.063 fragmentos. O número de fragmentos por amostra e o número de OTUs atribuídas se encontram na tabela 1.

Tabela 1 - Número de fragmentos totais e após filtro de qualidade (usado para atribuição de OTUs). Na coluna comunidade, o número após o meio de cultura indica a semana na qual a amostra foi coletada

Comunidade	Número de fragmentos	Número de fragmentos usados na atribuição de OTUs	Número de OTUs atribuídas	Número de OTUs após rarefação
MM-1	444.182	405.578	411	195
MM-2	211.159	191.892	373	194
MM-3	87.256	78.782	347	238
MM-5	55.113	50.223	322	237
PCS-1	46.999	42.216	235	169
PCS-2	290.779	272.638	303	158
PCS-3	77.266	72.558	228	155
PCS-5	11.603	10.480	183	162

Fonte: Elaborada pelo autor.

Os filos e classes de microrganismos presentes na cultura derivada da compostagem foram examinados, como apresentado na figura 10. Observa-se que a abundância relativa dos diferentes filos depende do meio de cultura usado. O consórcio crescido em MM foi dominado pelos filos Proteobacteria, Bacteroidetes e Actinobacteria, representando cerca de 58, 18 e 13% da diversidade total na primeira semana de cultivo. As OTUs atribuídas a Bacteroidetes e Actinobacteria mostraram um crescimento da abundância relativa para 25 e 18%, respectivamente, durante as semanas intermediárias, seguido por uma diminuição na quinta semana. Além disso, OTUs atribuídas a Acidobacteria, Chloroflexi, Firmicutes e Planctomycetes representavam cerca de 6% das OTUs na primeira semana e apresentaram um aumento relativo até a quinta semana, evidenciando que a diversidade de filos bacterianos nessa cultura aumentou com o tempo. O filo Actinobacteria foi fortemente dominado pela classe Actinobacteria, com abundância relativa maior que 87% durante as cinco semanas. Em contraste, o filo Bacteroidetes mostrou maior variação ao longo do tempo, sendo que os representantes das classes Cytophagia, Flavobacteria e Saprospirae foram os mais dominantes. Além disso, as classes Alphaproteobacteria e Gamaproteobacteria representaram mais de 88% dos membros do filo Proteobacteria durante todo o experimento.

Figura 10 - Abundância relativa dos a) filos e b-d) classes bacterianos identificados no consórcio crescido em meio mínimo (MM) e PCS. Na figura, o número após o meio de cultura indica a semana na qual a amostra foi coletada. Eixo-y: abundância relativa (%).

Fonte: Elaborada pelo autor.

A comunidade microbiana crescida em meio PCS foi dominada por OTUs pertencentes aos filos Bacteroidetes e Proteobacteria, que representaram 49% e 41% de toda a diversidade na primeira semana de cultivo. O filo Bacteroidetes apresentou mais de 44% de abundância durante todo o experimento, demonstrando forte dominância na cultura. Filos que apresentaram aumento na abundância relativa foram Chloroflexi (de 0,3% pra 14%) e Firmicutes (de 1% pra 5%). Dentro do filo Bacteroidetes, membros da classe Saprospirae apresentaram dominância, embora os membros das classes Cytophagia e Flavobacteria tenham apresentado aumento na abundância ao longo do tempo. O filo Proteobacteria foi fortemente dominado por Alphaproteobacteria e Gamaproteobacteria, similarmente a cultura em MM.

Portanto, membros do filo Proteobacteria, os quais são conhecidos por sua diversidade metabólica,¹⁷¹ dominaram ambos consórcios. Membros dos filos Bacteroidetes e Actinobacteria, também presentes em ambas comunidades, estão vinculados com a degradação de polímeros complexos na compostagem.^{4,9}

Tabela 2 - Análise de SIMPER mostrando quais filos e classes de microrganismos foram os principais responsáveis pela dissimilaridade Bray-Curtis entre as comunidades crescidas em meio mínimo (MM) e PCS.

Filo	Classe	Abundância média no MM (%)	Abundância média no PCS (%)	Contribuição para a dissimilaridade (%)
Bacteroidetes	Saprospirae	4,4	26,9	22,4
	Sphingobacteria	0,4	8,5	8,1
	Cytophagia	3,3	8,7	6,4
	Flavobacteria	4,6	5,5	4,2
	Outros	2,9	2,8	3,0
Proteobacteria	γ-proteobacteria	15,1	6,0	9,4
	α-proteobacteria	19,9	19,3	7,8
	Outros	3,5	2,3	2,2
Actinobacteria	Actinobacteria	13,2	1,5	11,6
	Outros	0,8	0,7	0,6
Acidobacteria	-	7,4	1,6	6,5
Planctomycetes	-	7,3	1,5	5,8
Chloroflexi	-	5,2	4,8	5,2
Firmicutes	-	3,7	2,0	2,3

Fonte: Elaborada pelo autor.

A SIMPER foi usada para analisar quais filos foram os principais responsáveis pelas diferenças observadas entre as comunidades crescidas em MM e PCS, conforme mostrado na tabela 2. As OTUs classificadas em gênero foram agrupadas nos filos e classes correspondentes apenas após a análise. Os filos Bacteroidetes (classes Saprospirae e

Sphingobacteria), Proteobacteria (classes Gamaproteobacteria e Alphaproteobacteria) e Actinobacteria (classe Actinobacteria) foram responsáveis por aproximadamente 60% de toda a dissimilaridade observada entre os consórcios crescidos em MM e PCS, conforme medido pelo índice de distância Bray-Curtis. Essas linhagens estavam entre as mais abundantes na cultura e correspondiam de 37% a 83% da abundância relativa, dependendo da amostra. O gênero dos principais responsáveis pela diferença entre as comunidades também foram analisados, conforme mostrado na tabela 3. Dentre as 21 OTUs mais abundantes em cada meio, cinco filotipos estavam presentes em ambos consórcios. Em ambas culturas, 70% das OTUs não puderam ser classificadas até o nível de gênero, indicando a presença de microrganismos pertencentes a gêneros ainda não descritos. Dentre os gêneros descritos, a comunidade crescida em MM possuía abundância de representantes dos gêneros *Cellulomonas*, *Paenibacillus* e *Cellvibrio*. Isolados desses gêneros foram capazes de degradar materiais lignocelulósicos.¹⁷²⁻¹⁷⁴ Já no meio PCS, houve abundância de espécies que necessitam de fontes de carbono simples para o crescimento, como *Sphingobacterium* e *Asticcacaulis*. Em um estudo do metagenoma de um consórcio derivado do solo, foi relatado que *Sphingobacterium* sp. produz principalmente CAZymes das famílias 2 e 3 e foi o principal responsável pela desconstrução da fração hemicelulósica da parede celular de plantas.²⁵ Portanto, a alta abundância relativa de *Sphingobacterium* sp. indica a presença de oligossacarídeos no meio. Como nas culturas em meio PCS, de acordo com inspeção visual, a degradação do bagaço de cana-de-açúcar foi limitada, esses oligossacarídeos provavelmente são derivados do extrato de levedura.

As 21 OTUs mais abundantes foram responsáveis por 57% e 75% de todas as OTUs presentes nas culturas em MM e PCS, respectivamente. Esse resultado mostra que relativamente poucas espécies dominaram as culturas e indica que a diversidade microbiana na comunidade crescida em meio PCS é menor que a crescida em MM. A hipótese de maior diversidade da comunidade crescida em MM foi confirmada pela presença de um maior número de OTUs únicas obtidas a partir da comunidade crescida nesse meio, representando 41% de todos os filotipos identificados (figura 11a). Além disso, o número de OTUs necessários para representar 50% de todos os fragmentos sequenciados foi de 17 na cultura em MM. Para a cultura em meio PCS, esse número foi de apenas 8 OTUs, como pode ser visto nas tabelas 3 e 4.

Figura 11 - Análise da diversidade no consórcio microbiano. a) Diagrama de Venn apresentando o número de OTUs únicas e compartilhadas entre as comunidades crescidas em meio mínimo (MM) e PCS. b) Curvas de rarefação mostrando a diversidade filogenética calculada *versus* o número de fragmentos sequenciados por amostra. Outros índices de diversidade, como *chao1* e *observed_species*, também foram testados e forneceram resultados similares. c) Box-plot do índice de diversidade Shannon calculado. Os resultados foram agrupados pelo meio de cultura usado no crescimento microbiano. O asterisco (*) denota uma diferença significativa ($p \leq 0,05$, teste t) entre os índices de diversidade obtidos para a cultura em MM e PCS. d) Análise de coordenadas principais das distâncias UniFrac ponderadas. Os consórcios crescidos no mesmo meio se agrupam. PCO-1: coordenada principal 1.

Fonte: Elaborada pelo autor.

Tabela 3 - 21 OTUs mais abundantes na comunidade microbiana (CM) crescida em MM. Taxas sublinhadas estão presentes também no consórcio crescido em meio PCS.

Filo	Classe	Ordem	Família	Gênero	Abundância média na CM crescida em MM (%)	Abundância média na CM crescida em PCS (%)	% cumulativa no MM
Actinobacteria	Actinobacteria	Actinomycetales	Cellulomonadaceae	Cellulomonas	6,7	0,0	6,7
<u>Acidobacteria</u>	<u>Chloracidobacteria</u>	<u>RB41</u>	<u>Ellin6075</u>	<u>desconhecido</u>	5,8	1,3	12,5
Actinobacteria	Actinobacteria	Actinomycetales	desconhecido	desconhecido	4,8	0,5	17,3
Bacteroidetes	Flavobacteriia	Flavobacteriales	Cryomorphaceae	desconhecido	3,6	0,0	21,0
Firmicutes	Bacilli	Bacillales	Paenibacillaceae	Paenibacillus	3,2	0,1	24,1
Proteobacteria	γ -proteobacteria	Alteromonadales	Alteromonadaceae	desconhecido	3,0	0,3	27,1
<u>Bacteroidetes</u>	<u>Saprospirae</u>	<u>Saprospirales</u>	<u>Chitinophagaceae</u>	<u>desconhecido</u>	2,6	5,5	29,7
Proteobacteria	γ -proteobacteria	Xanthomonadales	Xanthomonadaceae	desconhecido	2,3	0,1	32,0
<u>Proteobacteria</u>	<u>α-proteobacteria</u>	<u>Caulobacterales</u>	<u>Caulobacteraceae</u>	<u>Asticcacaulis</u>	2,3	7,6	34,3
Proteobacteria	α -proteobacteria	Caulobacterales	Caulobacteraceae	desconhecido	2,3	0,8	36,6
Proteobacteria	α -proteobacteria	Rhizobiales	desconhecido	desconhecido	2,2	0,9	38,8
Proteobacteria	γ -proteobacteria	Xanthomonadales	Sinobacteraceae	desconhecido	2,1	0,3	40,9
outros reinos	-	-	-	-	2,0	1,8	43,0

(continua)

Tabela 3 - 21 OTUs mais abundantes na comunidade microbiana (CM) crescida em MM. Taxas sublinhados estão presentes também no consórcio crescido em meio PCS.

(continuação)

Filo	Classe	Ordem	Família	Gênero	Abundância média na CM crescida em MM (%)	Abundância média na CM crescida em PCS (%)	% cumulativa no MM
Planctomycetes	Planctomycetia	Pirellulales	Pirellulaceae	desconhecido	2,0	0,6	45,0
Planctomycete	Planctomycetia	Pirellulales	Pirellulaceae	A17	1,9	0,0	46,8
Chloroflexi	Anaerolineae	Caldilineales	Caldilineaceae	desconhecido	1,8	0,9	48,7
<u>Proteobacteria</u>	<u>α-proteobacteria</u>	<u>Rhizobiales</u>	<u>Hyphomicrobiaceae</u>	<u>Devosia</u>	1,7	3,0	50,4
<u>Bacteroidetes</u>	<u>desconhecido</u>	<u>desconhecido</u>	<u>desconhecido</u>	<u>desconhecido</u>	1,7	2,0	52,1
Planctomycetes	Planctomycetia	Planctomycetales	Planctomycetaceae	Planctomyces	1,7	0,6	53, 8
Proteobacteria	γ -proteobacteria	Xanthomonadales	Xanthomonadaceae	desconhecido	1,6	0,5	55,4
Proteobacteria	α -proteobacteria	Rhodobacterales	Hyphomonadaceae	desconhecido	1,6	0,0	57,0

Fonte: Elaborada pelo autor.

Tabela 4 - 21 OTUs mais abundantes na comunidade microbiana (CM) crescida em meio PCS. Taxas sublinhados estão presentes também no consórcio crescido em MM.

Filo	Classe	Ordem	Família	Gênero	Abundância média na CM crescida em MM (%)	Abundância média na CM crescida em PCS (%)	% cumulativa no PCS
Bacteroidetes	Saprospirae	Saprospirales	Chitinophagaceae	desconhecido	0,9	10,2	10,2
Bacteroidetes	Saprospirae	Saprospirales	Saprospiraceae	desconhecido	0,6	10,0	20,2
<u>Proteobacteria</u>	<u>α-proteobacteria</u>	<u>Caulobacterales</u>	<u>Caulobacteraceae</u>	<u>Asticcacaulis</u>	2,3	7,6	27,9
Bacteroidetes	Cytophagia	Cytophagales	desconhecido	desconhecido	1,5	6,2	34,1
<u>Bacteroidetes</u>	<u>Saprospirae</u>	<u>Saprospirales</u>	<u>Chitinophagaceae</u>	<u>desconhecido</u>	2,6	5,5	39,6
Bacteroidetes	Sphingobacteria	Sphingobacterales	Sphingobacteriaceae	desconhecido	0,1	4,5	44,1
Bacteroidetes	Sphingobacteria	Sphingobacterales	Sphingobacteriaceae	Sphingobacterium	0,0	3,6	47,7
<u>Proteobacteria</u>	<u>α-proteobacteria</u>	<u>Rhizobiales</u>	<u>Hyphomicrobiaceae</u>	<u>Devosia</u>	1,7	3,0	50,7
Bacteroidetes	Flavobacteria	Flavobacteriales	Weeksellaceae	desconhecido	0,0	3,0	53,6
Chloroflexi	Anaerolineae	SBR1031	A4b	desconhecido	0,7	2,7	56,3
Proteobacteria	γ -proteobacteria	Alteromonadales	Alteromonadaceae	Cellvibrio	1,5	2,6	58,9
Bacteroidetes	Cytophagia	Cytophagales	Cytophagaceae	Leadbetterella	0,9	2,3	61,2
<u>Bacteroidetes</u>	<u>desconhecido</u>	<u>desconhecido</u>	<u>desconhecido</u>	<u>desconhecido</u>	1,7	2,0	63,2

(continua)

Tabela 4 - 21 OTUs mais abundantes na comunidade microbiana (CM) crescida em meio PCS. Taxas sublinhadas estão presentes também no consórcio crescido em MM.

(continuação)

Filo	Classe	Ordem	Família	Gênero	Abundância média na CM crescida em MM (%)	Abundância média na CM crescida em PCS (%)	% cumulativa no PCS
Firmicutes	Bacilli	Bacillales	Planococcaceae	desconhecido	0,0	1,7	68,4
Verrucomicrobia	Spartobacteria	Chthoniobacterales	Chthoniobacteraceae	heteroC45_4W	0,1	1,5	70,0
Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	desconhecido	0,4	1,5	71,5
<u>Acidobacteria</u>	<u>Chloracidobacteria</u>	<u>RB41</u>	<u>Ellin6075</u>	<u>desconhecido</u>	5,8	1,3	72,8
Proteobacteria	α -proteobacteria	Rhodobacterales	Rhodobacteraceae	desconhecido	1,2	1,2	74,0
Proteobacteria	α -proteobacteria	Rhizobiales	Hyphomicrobiaceae	desconhecido	0,9	1,1	75,1

Fonte: Elaborada pelo autor.

6.3 Medidas de diversidade microbiana

Curvas de diversidade alfa (figura 11b) mostram uma leve diminuição na riqueza de espécies com o tempo de cultivo. Mesmo com 40.000 fragmentos por amostra, as curvas não alcançaram uma assíntota horizontal, indicando que a diversidade microbiana nas amostras era alta e que a profundidade de sequenciamento empregada deveria ser maior para que todos os microrganismos presentes fossem identificados. O índice Shannon (figura 11c) confirmou as observações de que a diversidade microbiana era maior nos consórcios crescidos em MM em relação aos crescidos em meio PCS.

Foi usada a PCoA com distâncias UniFrac ponderadas (figura 11d) para estimar a diferença filogenética entre as amostras. A PCoA mostrou que os consórcios crescidos em MM e PCS foram classificados em dois grupos distintos. A coordenada principal 1 (PCO-1) foi suficiente para separar totalmente os grupos crescidos nos diferentes meios. Dentro de cada grupo, a PCO-2 permitiu que as amostras fossem separadas pelo tempo de cultivo.

6.4 A disponibilidade de nutrientes afeta a estrutura da comunidade microbiana

As bactérias que são capazes de crescer em condições laboratoriais representam uma pequena fração da diversidade microbiana total encontrada na natureza. Estudos estimam que apenas cerca de 1% dos microrganismos provenientes do solo podem se proliferar em meios de cultura padrões.¹⁵⁵⁻¹⁵⁷ Muitas bactérias necessitam de moléculas específicas, naturalmente presentes em seu ambiente,¹⁵⁸⁻¹⁵⁹ mas difíceis de replicar em laboratório. Além disso, relações sintróficas em comunidades microbianas complexas são um fator limitante no cultivo separado de muitos microrganismos. Entretanto, essa parte incultivável das comunidades microbianas são capazes de sintetizar produtos desconhecidos que desempenham um papel importante no ciclo do carbono, nitrogênio e nutrientes,¹⁶⁰ como na quebra da lignocelulose.

Já foi demonstrado que consórcios microbianos são capazes de degradar a lignocelulose mais rapidamente do que monoculturas.^{39,161-163} Entretanto, estudos diferentes usaram diferentes tipos de meio de cultura no crescimento desses consórcios. Por exemplo, estudos da degradação da lignocelulose foram realizados empregando meios de cultura complexos^{107,164-165} ou definidos.^{2,166-167}

Nos resultados descritos até a presente seção, foram analisadas as diferenças entre as comunidades microbianas cultivadas em um meio complexo e um MM. Foi mostrado que a fração de organismos com atividade celulase foi maior na comunidade crescida em MM,

assim como verificado por uma maior degradação da biomassa. Também foi mostrado que as comunidades microbianas puderam ser agrupadas, através de medidas de distâncias filogenéticas, de acordo com o meio utilizado no seu crescimento. Além disso, as comunidades crescidas em ambos meios apresentaram abundância em gêneros não descritos. Entretanto, baseado na medida de diversidade alfa, a diversidade bacteriana foi significativamente maior na comunidade crescida em MM ($p \leq 0,05$). De acordo com Mitri et al.¹⁷⁵ a disponibilidade de nutrientes está intimamente relacionada com a perda de diversidade por grupos celulares. A densidade de células no meio nutricionalmente mais rico (PCS) era maior que na comunidade crescida em MM. Portanto, células de crescimento rápido se expandiram rapidamente, provocando perda de diversidade da cultura, conforme discutido por Mitri et al. Baseado nesses resultados, conclui-se que o uso de um MM é mais adequado para estudo de consórcios microbianos em laboratório, pois fornece um ambiente onde a taxa de crescimento é lenta e a densidade celular é baixa, semelhante a ambientes naturais. Além disso, não fornece fontes de carbono, possibilitando o enriquecimento de microrganismos capazes de degradar o substrato estudado.

Portanto, através dos experimentos acima, foi determinado que o MM é mais interessante para o estudo da degradação do bagaço de cana-de-açúcar pela comunidade microbiana derivada da compostagem. Um novo consórcio microbiano foi estabelecido em MM e a degradação do bagaço foi estudada através dos experimentos discutidos abaixo.

6.5 Mudanças químicas e morfológicas do bagaço de cana-de-açúcar

A figura 12 mostra o espectro de ssNMR do bagaço de cana-de-açúcar destacando as regiões atribuídas especificamente a celulose, hemicelulose e lignina. Após os espectros serem normalizados pela sua respectiva área, as regiões específicas foram integradas para estimar as frações de celulose, hemicelulose e lignina na amostra, como descrito por Lima et al.³⁶ O gráfico da porcentagem relativa é apresentado como função da semana de crescimento (figura 12a). Devido ao crescimento microbiano e atividade enzimática da cultura, foi observada uma clara redução na quantidade relativa de celulose. A quantidade relativa de hemicelulose permaneceu aproximadamente constante, com alguma variação atribuída a incertezas experimentais. Como a amostra perde massa com o tempo, conclui-se que a hemicelulose foi removida na mesma taxa. Por outro lado, a quantidade relativa de lignina aumentou na mesma proporção que a celulose diminuiu. Portanto, apesar das outras frações terem sido removidas, a lignina não foi efetivamente degradada.

A concentração e o arranjo da lignina na biomassa foi investigada por microscopia confocal com excitação por dois fótons. A figura 12b-f, mostra que nenhuma mudança no espectro de emissão foi observado mesmo após 5 semanas de crescimento microbiano. Esse resultado concorda com o que foi observado por ssNMR, pois indica que a lignina não está sendo removida significativamente. Além disso, também sugere que nem ao menos a estrutura ou a organização da lignina está sendo modificada.

Figura 12 - Análise da composição do bagaço de cana-de-açúcar. a) Espectro de ressonância magnética nuclear de estado sólido do bagaço de cana-de-açúcar antes e depois do crescimento microbiano. As regiões 1 e 1' são atribuídas ao carbono C1 da celulose (103-106 ppm) e ao carbono C4 da celulose cristalina (86-92 ppm). As regiões 2 e 2' são atribuídas aos carbonos C1, C2 e C4 do anel aromático da lignina (123-142 ppm) e aos carbonos aril metoxil da lignina (50-56 ppm). As regiões 3 e 3' são atribuídas ao carbono C1 da hemicelulose e aos CH_3 em grupos acetil da hemicelulose. A abundância relativa de celulose, hemicelulose e lignina foi estimada a partir das regiões 1, 1', 2, 2', 3 e 3'. b-e) Imagens de microscopia confocal da lignina no bagaço antes e depois do crescimento microbiano. f) Espectro de emissão correspondente.

Fonte: Elaborada pelo autor.

Imagens de MEV do bagaço antes do crescimento microbiano mostram uma superfície continua com fibras coesivas e bem definidas (figura 13). Após 5 semanas de crescimento microbiano, é possível observar uma perda acentuada na estrutura da biomassa, com separação das fibras, exposição de buracos e diminuição do tamanho das partículas.

a **b**

(continua)

(continuação)

Figura 13 - Imagens de microscopia eletrônica de varredura do bagaço de cana-de-açúcar. a), b) antes e c), d) depois de cinco semanas de crescimento microbiano. a), c) aumento de 300x; b), d) aumento de 1.000x.

Fonte: Elaborada pelo autor.

Baseado nas técnicas apresentadas acima, foi mostrado que a lignina não foi significativamente removida ou modificada pela comunidade microbiana. A lignina é um heteropolímero de unidades fenólicas altamente recalcitrante aos ataques químico e microbiológico. Apenas um pequeno e especializado grupo de fungos e bactérias expressam enzimas capazes de degradar esse polímero.¹⁷⁶ Esse resultado concorda com um estudo anterior que mostra que a perda de massa da biomassa utilizada como fonte de carbono para o crescimento microbiano foi principalmente devido a degradação das frações celulósica e hemicelulósica.¹ Apesar do conteúdo de lignina não ter sido alterado, mudanças morfológicas drásticas foram observadas na biomassa. Portanto, os microrganismos que se estabeleceram na cultura tiveram que encontrar um modo de extrair a celulose e hemicelulose do bagaço, para sustentar o seu crescimento, sem remover o escudo de lignina. Isso provavelmente limitou o seu acesso ao substrato, e deve ser resultado da falta de uma maquinaria eficiente para a degradação da lignina.

6.6 Métricas do metatranscriptoma

O sequenciamento do metatranscriptoma do consórcio derivado da compostagem gerou um total de 66.307.103 fragmentos, conforme apresentado na tabela 5. Desses, 63% passaram pelo filtro de qualidade e remoção de DNA ribossomal e foram usados para montagem *de novo* do metatranscriptoma. Foram montados 302.961 transcritos e 283.356 ORFs foram detectados. 2,5% dos ORFs identificados foram preditos como CAZymes.

Tabela 5 - métricas do sequenciamento e processamento do metatranscriptoma.

	Semana 1	Semana 2	Semana 3	Semana 4	Semana 5	Total
Total de fragmentos gerados	20.119.184	14.715.430	7.767.801	5.879.612	17.825.076	66.307.103
Fragmentos após filtro de qualidade	15.916.667	12.072.389	6.757.570	5.111.343	14.535.407	54.393.376
mRNA	12.645.050	10.652.024	3.608.818	3.971.178	11.004.487	41.881.555
Transcritos montados	104.408	64.923	2.7821	34.023	71.786	302.961
tamanho médio	947	888	789	817	858	912
N50	963	921	820	847	884	887
ORFs preditos	104.425	59.885	23.156	31.198	64.692	283.356
CAZymes	3.012	1.523	534	654	1.473	7.196

Fonte: Elaborada pelo autor.

6.7 Análise da profundidade de sequenciamento do metatranscriptoma

Análises de rarefação do metatranscriptoma foram realizadas por dois métodos diferentes. O primeiro método usado se baseia na hipótese de que as estatísticas calculadas são afetadas pela profundidade de sequenciamento. Ou seja, quando o sequenciamento se torna redundante, as estatísticas ficam estáveis.¹⁷⁷ Para isso, a taxa de expressão (FPKM) calculada quando se utiliza todos os fragmentos sequenciados foi comparada com a calculada quando apenas uma fração deles é usada. Mesmo selecionando 90% dos fragmentos sequenciados, a taxa de expressão de alguns genes apresentou erro relativo superior a 90%, indicando que a saturação não foi atingida. O segundo método extrai um *kmer* (pedaço de sequência) de cada fragmento e confere se essa sequência já foi obtida anteriormente. Para cada 25.000 fragmentos, o percentual de novos *kmers* é calculado e um gráfico é desenhado. Portanto, o sequenciamento está saturado quando atinge-se o zero. Como podemos ver no gráfico 14, a saturação não foi atingida. Esse método também mostra que a biblioteca da semana 1 é mais diversa que as outras, já que uma proporção maior de novos *kmers* é obtida para uma mesma profundidade de sequenciamento. Apesar do metatranscriptoma ter sido montado com boas estatísticas, ambas análises indicam que a profundidade de sequenciamento não foi suficiente para alcançar a saturação.

(continua)

(continuação)

b

Figura 14 - Análise da profundidade de sequenciamento do metatranscriptoma por dois métodos. a) Variação da taxa de expressão calculada a partir de bibliotecas com diferentes profundidades de sequenciamento. b) Fração de novos *kmers* obtidos.

Fonte: Elaborada pelo autor.

O primeiro desafio quando se deseja realizar o sequenciamento de um metagenoma ou metatranscriptoma é determinar a profundidade de sequenciamento que deve ser empregada para responder as perguntas propostas. Versluis et al. mostrou que sequenciar 4 M de fragmentos por amostra foi suficiente pra identificar a maioria dos genes de resistência no metatranscriptoma do intestino de diversos animais.¹⁷⁸ Simmons et al. mostrou que 2 M de fragmentos foi suficiente para alcançar uma assíntota quando se considera o número de genes expressos pelo número de fragmentos amostrados no metatranscriptoma de uma comunidade lignocelulolítica.² Por outro lado, Jian et al. sugeriu que sequenciar até 20 M fragmentos por microbioma ainda fornece uma boa razão entre custo de sequenciamento e descoberta de enzimas.¹⁷⁹ Logo, a profundidade de sequenciamento necessária também é dependente do ambiente e dos objetivos do estudo. Neste projeto, a maioria das amostras foram sequenciadas a uma profundidade de 10 M fragmentos, o qual é um valor intermediário dentre os discutidos acima.

6.8 Perfis funcionais

O metatranscriptoma obtido foi comparado com a base de dados do KEGG *orthology* para avaliar o potencial metabólico da comunidade microbiana cultivada. Analisando a distribuição dos ORFs nas diferentes classes, observamos um perfil similar durante as cinco semanas de estudo (figura 15). A classe mais abundante foi Tradução [XIX], com uma abundância relativa em torno de 8,5% durante todo o experimento. As classes Metabolismo de energia [VII], Transdução de sinais [XVIII], Metabolismo de carboidratos [III] e Metabolismo de aminoácidos [I] também foram abundantes, apresentando uma média de 6 a 8% de abundância relativa cada uma. A classe [III] apresentou uma queda de 10% pra 6% de abundância da primeira para a terceira semana de cultivo e se estabilizou até a quinta semana. O mesmo ocorreu com a classe [VII], sugerindo que a cultura está expressando CAZymes e se dividindo mais ativamente durante as duas primeiras semanas de cultivo.

O perfil funcional obtido é similar ao de outras comunidades microbianas capazes de degradar a lignocelulose, tal como compostagem enriquecida com *switchgrass*²⁷ e palha de arroz,¹ bagaço de cana em contato com o solo³ e intestino posterior de um inseto térmita.¹⁵ Além de genes gerais necessários para a manutenção da vida, o perfil funcional se mostrou abundante em proteínas associadas ao metabolismo de carboidratos. Portanto, essa reconstrução metabólica é outra evidência do alto potencial dessa comunidade microbiana em desconstruir a biomassa, especialmente nas duas primeiras semana de crescimento.

I - Metabolismo de aminoácido; **II** - Câncer; **III** - Metabolismo de carboidrato; **IV** - Mobilidade celular; **V** - Comunicação celular (procariotos); **VI** - Sistema endócrino; **VII** - Metabolismo de energia; **VIII** - Enovelamento, etiquetamento e degradação; **IX** - Crescimento e morte celular; **X** - Sistema imune; **XI** - Doenças infecciosas; **XII** - Metabolismo de lipídeo; **XIII** - Transporte de membrana; **XIV** - Metabolismo de cofatores e vitaminas; **XV** - Metabolismo de outros aminoácidos; **XVI** - Doenças neurodegenerativas; **XVII** - Metabolismo de nucleotídeos; **XVIII** - Transdução de sinal; **XIX** - Tradução; **XX** - Transporte e catabolismo.

Figura 15 - Perfil metabólico dos transcritos expressos no consórcio derivado da compostagem. Classificação realizada de acordo com a base de dados KEGG *orthology*. Cada eixo apresenta a quantidade relativa de genes, em porcentagem, pertencentes a respectiva classe.

Fonte: Elaborada pelo autor.

6.9 Perfil de expressão e classificação filogenética de CAZymes

O repertório de enzimas capazes de degradar a lignocelulose foi investigado através da busca de genes do metatranscriptoma por homólogos no banco de dados do CAZy. As classes de CAZymes putativas mais expressas foram respectivamente GHs, CBMs, GTs, CEs, AAs e PLs (figura 16). A expressão de CAZymes, de forma geral, foi maior na primeira semana, seguida de uma queda na segunda semana e estabilização até a quinta semana. Apenas os transcritos atribuídos a GTs apresentaram um perfil contrastante, no qual a maior expressão se deu na quinta semana. Esse resultado está de acordo com o perfil de classes do KEGG, que

revelou um enriquecimento em genes envolvidos no metabolismo de carboidratos e energia na primeira semana da cultura.

O perfil de CAZymes encontrados é similar ao identificado em uma comunidade microbiana proveniente de solo florestal e selecionada em palha de trigo.²⁵ Um relativo baixo número de GHs foi encontrado no metatranscriptoma (0,87%), mas valores similares (de 0,71% a 0,97%) foram descritos para outras comunidades ativas na degradação da lignocelulose.^{2-3,14-15,18} Enzimas pertencentes ao grupo de AA, que contêm enzimas responsáveis pela modificação oxidativa da lignina, representaram apenas 3,3% das CAZymes previstas, explicando a ausência observada de modificação no conteúdo de lignina do bagaço de cana-de-açúcar.

Figura 16 - Expressão de CAZymes putativas no metatranscriptoma do consórcio derivado da compostagem. AA - atividade auxiliar; CE - esterase de carboidrato; CBM - módulo de ligação a carboidrato; GH - hidrolase de glicosil; GT - transferase de glicosil; PL - liase de polissacarídeo.

Fonte: Elaborada pelo autor.

Como resumido na figura 17, as famílias de GH mais expressas foram aquelas envolvidas na degradação da celulose e hemicelulose (GH 3, 5, 11, 13, 43, 74 e 109), as quais apresentaram diminuição gradual de expressão. Genes anotados como GH 19 e 25 apresentaram um comportamento contrário, com expressão maior nas últimas semanas do estudo. Essas famílias consistem de membros com atividade lisozima e quitinase. Portanto, a maior expressão dessas famílias pode estar associada a um aumento de interações competitivas entre as espécies devido a escassez de substrato nas últimas semanas de cultivo. Os fragmentos que codificam as famílias GHs mais expressas foram classificados

filogeneticamente pelo MEGAN com o algoritmo LCA. Membros do filo Proteobacteria foram responsáveis por expressar a maioria das GH 5, 11 e 19; membros do filo Bacteroidetes expressaram a maioria das GH 13, 23, 43 e 109. Os transcritos classificados como GH 3 foram expressos principalmente por membros do filo Verrumicrobia. Foi predito que algumas famílias de GH foram expressas principalmente por eucariotos, tais como a GH 13 (altamente expressa por membros do reino Animalia) e GH 25 (expressa principalmente por membros sem reino definido).

A família GH 3 foi a mais expressa no metatranscriptoma. Ela é composta por enzimas que hidrolisam oligossacarídeos, atividade essencial para alcançar alta eficiência na conversão da lignocelulose¹⁸⁰ e para o uso dos açúcares no metabolismo.

Figura 17 - Expressão diferencial e distribuição filogenética das famílias de hidrolase de glicosídeo (GH) detectadas no metatranscriptoma da compostagem. a) Heat map da expressão de GHs. As colunas representam o tempo de cultivo quando a amostra foi coletada. As linhas retratam as diferentes famílias de GH identificadas. As famílias de GH foram agrupadas de acordo com sua preferência por substratos.²⁰² b) Atribuição filogenética dos fragmentos pertencentes as famílias GHs mais expressas.

Fonte: Elaborada pelo autor.

Dentre as 7.196 CAZymes previstas, 75% puderam ser classificadas filogeneticamente em nível de filo pelo MEGAN. A distribuição de taxas foi calculada pela expressão (FPKM) relativa de CAZymes atribuídas aos diferentes taxas (figura 18). Dentre os procariotos, os microrganismos mais ativamente envolvidos na modificação de carboidratos pertenciam aos filos Bacteroidetes e Proteobacteria, similar ao observado por Simmons et al. no estudo de uma compostagem enriquecida com palha de arroz.² Esses filos somaram mais de 65% das CAZymes expressas durante todo o período estudado. Enquanto as CAZymes de Bacteroidetes mostraram um aumento de 29 para 44%, as de Proteobacteria mostraram uma

diminuição de 42 para 24%. O filo Firmicutes mostrou um incremento gradual de 1 para 5%. O filo Proteobacteria é metabolicamente diverso e comumente domina habitats oxigenados.³ Além disso, Bacteroidetes foram responsáveis por produzirem o maior conjunto de CAZymes em diferentes ambientes estudados.^{3,25}

A expressão de CAZymes atribuídas a procariotos diminuiu de 20.171 para 6.465 FPKM durante as cinco semanas de cultura. Por outro lado, os genes atribuídos a eucariotos mostraram um aumento drástico de expressão de 287 para 8.405 FPKM (cerca de 29 vezes de aumento). Surpreendentemente, na quinta semana, 82% das CAZymes expressas por eucariotos foram atribuídas ao reino Animalia e apenas 3% ao reino Fungi. Interessante observar o aumento de expressão de CAZymes por eucariotos e, em especial, por animais. Nemátodos, protistas e outros grupos de eucariotos estão naturalmente presentes em uma compostagem,¹⁸¹ mas é improvável que suas células sobrevivessem por semanas em cultura líquida com os métodos utilizados. Uma explicação é que o algoritmo LCA não atribuiu a filogenia corretamente devido a falta de genomas de referência para as espécies presentes. Entretanto, esse resultado é um indicativo de que eucariotos estão presentes e possuem um papel fundamental na degradação da lignocelulose. De fato, evidências de que eucariotos invertebrados são capazes de degradar a parede celular de plantas vem aumentando recentemente.¹⁸²⁻¹⁸⁵

Figura 18 - Atribuição filogenética das CAZymes expressas ao longo do tempo pelo consórcio derivado da compostagem. Expressão relativa dos a) filós bacterianos e b) reinos eucariotos. A expressão total de cada domínio é representada pela linha em cinza.

Fonte: Elaborada pelo autor.

Para quantificar a fração de bactérias e fungos no consórcio microbiano, PCR em tempo real foi realizado. Membros do reino Bacteria dominaram a comunidade, mas a fração de fungos/bactérias aumentou gradualmente de 5 para 20% nas primeiras três semanas e se estabilizou até o final do experimento (tabela 6). Portanto, quando considerados todos os membros da comunidade (não apenas a expressão de CAZymes), bactérias dominam a cultura em relação aos fungos.

Tabela 6 - Abundância relativa de rDNA amplificado com oligonucleotídeos específicos para fungos e bactérias.

Semanas de crescimento	Porcentagem relativa de rDNA de fungo/bactéria
0	11 ± 2
1	4.8 ± 0.4
2	9 ± 1
3	21 ± 3
4	20 ± 3
5	22 ± 1

Fonte: Elaborada pelo autor.

6.10 Caracterização funcional de CAZymes putativas do metatranscriptoma

Como o principal objetivo deste trabalho era caracterizar novas enzimas com atividade lignocelulolítica, genes preditos que codificam homólogos a CAZymes foram selecionados para expressão heteróloga em *Escherichia coli* e caracterização funcional. 27 genes altamente expressos no metatranscriptoma e preditos possuir atividade celobiohidrolase, endoglucanase ou xilanase foram selecionados. Após teste de expressão, sete proteínas (26% de eficiência) foram obtidas na fração solúvel de culturas de *E. coli* e três proteínas apresentaram atividade enzimática após teste contra uma variedade de substratos. A baixa eficiência de expressão de proteínas solúveis demonstra os desafios em caracterizar proteínas obtidas a partir de metatranscriptomas.

Análises de bioinformática mostraram que a proteína C7 possui 49% de identidade com uma β -1,4-glucano celobiohidrolase da família GH6 do organismo *Sorangium cellulosum*. Testes contra vários substratos mostraram que essa proteína possui atividade contra β -glucano e liquenano. Nenhuma atividade foi detectada contra papel de filtro e CMC. Essa enzima apresentou atividade máxima em pH 10,0 e foi capaz de reter 70% ou mais de atividade até que o pH caísse para 4,0, conforme mostrado na figura 19a. Testes buscando a

temperatura ótima, em pH 6,0 e 10,0, revelaram que a enzima possui atividade máxima a 50 °C e 45 °C (figura 19b), respectivamente. O experimento de fluorimetria de escaneamento diferencial (DSF, do inglês, *differential scanning fluorimetry*) (figura 20) mostrou que essa enzima possui termoestabilidade máxima entre os pHs 6,0 e 9,0, confirmando seu comportamento alcalifílico. As condições ótimas de pH e temperatura foram usadas para testar sua especificidade. Os resultados mostraram que ela possui atividade específica contra β -glucano e liquenano de 2,0 e 1,5 U/mg respectivamente (figura 19c). Também foi determinado que após 24 horas em pH 6,0 a 45 °C, a proteína C7 mantém 50% de sua atividade inicial (figura 19d). Em pH 10,0 a 40 °C, ela retém 100% de sua atividade mesmo após 96 horas. Devido a atividade específica maior contra β -glucano, essa enzima é uma β -1,3-1,4-endoglucanase com preferencia por substratos com uma maior proporção de ligações β -1,3 do que β -1,4. Outras enzimas com essa especificidade já foram reportadas.^{186; 187} Além disso, ela é altamente tolerante a ambientes alcalinos, uma característica essencial para aplicação na indústria de detergentes.¹⁸⁸

Figura 19 - Caracterização bioquímica das enzimas C7, C13 e C21 derivadas do metatranscriptoma da compostagem. Efeito do a) pH e b) temperatura na atividade enzimática. c) Especificidade por substrato. d) Atividade residual da proteína após incubação na temperatura de estudo.

Fonte: Elaborada pelo autor.

Figura 20 - Termoestabilidade, examinada por fluorimetria de escaneamento diferencial, da proteína C7 em diferentes pHs.

Fonte: Elaborada pelo autor.

A proteína C13 foi predita como membro da família GH 10 e possui 91% e 89% de identidade a uma endoglucanase e uma endo- β -1,4-xilanase de *S. cellulorum*, respectivamente. Ela apresentou atividade contra xilano, o qual foi usado como substrato para determinar as condições ótimas de pH e temperatura. A C13 apresentou atividade máxima em pH 6,0 e reteve mais de 50% de sua atividade em uma ampla extensão de pHs (pH de 3,0 a 10,0, figura 19a). A temperatura ótima foi de 65 °C em pH 6,0 (figura 19b). Essas condições foram usadas para testar a atividade específica dessa enzima. As maiores atividade foram encontradas contra xilano (25 U/mg) e arabinoxilano (11 U/mg), indicando que ela é uma endoxilanase (figura 19c). Nenhuma atividade foi encontrada contra CMC. O estudo da atividade residual, realizado a 50 °C em pH 6,0, revelou que essa enzima retém mais de 60% de sua atividade após 96 horas de incubação, confirmando sua considerável termoestabilidade (figura 19d). A especificidade por substratos dessa enzima e seu perfil de hidrólise são de uma endoxilanase típica.¹⁸⁹⁻¹⁹¹

A proteína C21 foi predita como membro da família GH 11 e possui 77% de identidade com uma GH não caracterizada de *Marinimicrobium agarilyticum*. O homólogo caracterizado mais próximo é uma β -1,4-xilanase de *S. cellulorum*, com 40% de identidade. A C21 apresentou atividade máxima contra xilano em pH 6,0 e 35 °C (figuras 19a-b). Ela reteve mais de 60% de sua atividade em todos os pH testados, mas a atividade observada caiu rapidamente em temperaturas acima de 40 °C. Quando incubada a 35 °C em pH 6,0, a C21 reteve mais de 90% de sua atividade mesmo após 96 horas (figura 19d). A triagem contra diferentes polissacarídeos revelou que ela é específica por xilano (figura 19c). Usando esse substrato, a C21 apresentou uma alta atividade específica de 320 U/mg.

6.11 C21 é a primeira exoxilanase descrita da família GH 11

A enzima C21 não apresentou atividade contra substratos que possuem um esqueleto de xilano, como arabinoxilano e AZCL-xilano, um substrato específico para endoxilanasas devido a sua estrutura com ligações cruzadas que aprisionam moléculas de corante no seu interior.¹⁹² Para investigar o modo de ação da C21 sobre o xilano, os seus produtos de hidrólise foram analisados por TLC e HPLC Dionex (figura 21). Os resultados de TLC mostraram que a C21 age no xilano e xilooligômeros liberando apenas xilobiose como produto de reação. Os resultados de HPLC Dionex confirmaram esse padrão de hidrólise. Além disso, a enzima não apresentou atividade contra 4-nitrofenil- β -D-xilopiranosídeo, confirmando que ela não é capaz de hidrolisar a xilobiose. Logo, essa enzima não apresentou

atividade contra substratos típicos de uma endoxilanase. Ainda mais, o perfil de hidrólise apresentado é típico de uma exoenzima, liberando apenas um produto em vez de vários oligômeros de tamanhos aleatórios. Por isso, foi especulado que essa enzima não é uma endoxilanase.

Figura 21 - Evidência experimental que a proteína C21 é uma exoxilanase. a) TLC dos produtos formados por: colunas 1 e 3: ação da C21 e C13 contra xilano, respectivamente; coluna 2: branco de reação; colunas 4 a 8: ação da C21 contra xilooligômeros de duas a seis unidades de xilose (X2-X6); Ác.G: ácido glicurônico; M: padrão com X1-X6. b) HPLC Dionex dos produtos formado pela ação da C21 contra xilano.

Fonte: Elaborada pelo autor.

Para compreender o mecanismo de ação dessa enzima, sua estrutura cristalográfica foi resolvida a 1,76 Å de resolução. Cristais foram obtidos em múltiplas condições, sendo que cristais crescidos em BIS-TRIS propano 100 mM pH 7,5, PEG 3350 20% (m/V), NaI 200 mM a 18 °C produziram padrões de difração de melhor qualidade. As estatísticas de coleta e refinamento se encontram na tabela 7. O cristal possuía simetria tetragonal e pertencia ao grupo espacial $P4_32_12$. Havia apenas 1 molécula por unidade assimétrica e os dados foram coletados e resolvidos à 1,76 Å de resolução. Os dados foram coletados com alta multiplicidade devido a alta simetria do cristal, sendo obtida uma completeza de 97% na alta resolução. O modelo final foi construído do primeiro ao último resíduo da proteína cristalizada, a qual foi clonada sem os primeiros 13 resíduos: preditos como uma região desordenada e descritos anteriormente como um peptídeo sinal.¹⁹³ O R_{work} e R_{free} ficaram na faixa dos 20%. O modelo final foi validado, não apresentando desvios grandes nos comprimentos e ângulos de ligações e sem nenhum aminoácido em regiões não permitidas do gráfico de Ramachandran. O modelo foi depositado no PDB (código 5VQJ).

Tabela 7 - Estatísticas da coleta e refinamento dos dados de difração da proteína C21. Os valores em parênteses se referem a camada mais externa de resolução. 5% das reflexões foram aleatoriamente excluídas do refinamento e usadas para cálculo do R_{free} .

Coleta de dados	
Comprimento de onda/linha de luz	1,459/MX2, LNLS
Grupo espacial	P4 ₃ 2 ₁ 2
Dimensões da célula unitária (Å)	64.33; 64.33; 105.87
Moléculas por unidade assimétrica	1
Coeficiente de Matthews (Å ³ /Da)	2,24
Conteúdo de solvente (%)	45,0
Resolução (Å)	41,79 (1,76)
Número de reflexões únicas	22.593 (2.163)
Mosaicidade (°)	0.205
Multiplicidade	24,0 (19,2)
Completeza	99,4 (97,0)
I/sigma (I)	24,72 (2,49)
R_{meas} (%)	12,4
CC _{1/2}	0,999 (0,824)
Refinamento	
Número de resíduos de aminoácidos	216
Número de moléculas de águas	281
R_{work} (%)	19,4 (28,5)
R_{free} (%)	21,8 (30,8)
RMSD (ligações) (Å)	0.069
RMSD (ângulos) (°)	1.144
B-fatores médios (Å ²)	19.5
Ramachandran favoráveis (%)	96.73
Ramachandran não permitidos (%)	0
Clashscore	11,97
PDB ID	5VQJ

Fonte: Elaborada pelo autor.

Como mostrado na figura 22, a proteína C21 possui enovelamento β -jelly-roll, típico da família GH 11.¹⁹⁴⁻¹⁹⁸ Ela é constituída por 15 fitas- β , nomeadas de A1 a A6 e B1 a B9, e uma α -hélice. As estruturas secundárias possuem aproximadamente os mesmos comprimentos que nas demais proteínas dessa família.¹⁹⁴ A curvatura das fitas- β B1 a B9 formam uma fenda na qual os resíduos catalíticos estão localizados (nucleófilo Glu98 na fita- β B6 e doador de prótons Glu200 na fita- β B4). Apesar de já terem sido depositadas 28 estruturas de proteínas da família GH 11 no PDB, a proteína C21 possui dois loops extras (EL1 e EL2) não vistos nas demais estruturas. Eles são criados por resíduos adicionais entre as fitas- β B5 e B6 e no interior da fita- β B4, respectivamente. O alinhamento múltiplo com os homólogos mais

próximos do NCBI-nr revelou que muitas outras proteínas, cujas estruturas ainda não foram resolvidas, também devem possuir esses loops extras. Isso é uma evidência da necessidade de desenvolver e utilizar ferramentas para ajudar a escolher alvos com características potencialmente diferentes, em vez de selecioná-los aleatoriamente, o que, muitas vezes, leva ao estudo de “mais do mesmo”.

Figura 22 - Evidência estrutural que a proteína C21 é uma exoxilanase. a) Alinhamento múltiplo da sequência de aminoácidos da C21 com os homólogos mais próximos presentes nos bancos de dados do NCBI-nr (segunda a sexta sequência) e PDB (cinco últimas sequências). Os resíduos responsáveis pela formação dos loops extras (EL1 e EL2) estão selecionados nas caixas vermelhas; os resíduos catalíticos estão selecionados nas caixas azuis. Esses loops extras estão presentes em outras proteínas cujas estruturas ainda não foram resolvidas. b) Estrutura secundária da C21 colorida do azul para o vermelho (do N- pro C-terminal). c) Superposição de todas as estruturas disponíveis da família GH11 (em cinza) e da C21 (em magenta). A orientação dessas estruturas foi rotacionada de 90° em relação a estrutura apresentada no item b. Domínios incomuns que não se alinharam com as demais proteínas foram removidos da representação. d) Representação da superfície da C21 (em magenta) alinhada com o homólogo mais próximo do PDB (ID 1XNK, em cinza). Note que o terminal não-redutor do ligante metil 4,4''-ditio- α -xilotriose, presente na estrutura 1XNK, é impedido estericamente no subsítio -3 da C21 pelo EL2.

Fonte: Elaborada pelo autor.

O EL2 bloqueia a interação do substrato no subsítio -3, que acomoda a extremidade não-redutora do xilano. Esse loop é estabilizado por um número de ligações de hidrogênio entre a cadeia principal/lateral de resíduos do loop e interações de van der Waals, especialmente entre as cadeias laterais da Pro192 e Phe186. Ligações de hidrogênio também

são estabelecidas com o EL1 e com a volta que conecta as fitas- β B7 e B8. A presença do EL1 parece essencial para manter o EL2 em posição por impedimento estérico. Portanto, a presença desses loops extras bloqueiam uma das extremidades do sítio ativo da C21, tornando essa enzima uma exoxilanase que age a partir da extremidade não-redutora. Esse estudo é o primeiro a relatar uma enzima da família GH11 com atividade exoxilanase. Ao contrário das exoxilanases da família GH8, que apresentam preferência por xilooligômeros solúveis,^{199; 200} a C21 é altamente ativa em substratos poliméricos insolúveis.

6.12 A atividade da C21 melhora a performance de coquetéis enzimáticos comerciais

A hidrólise enzimática de Avicel 2% por Accellerase® (5 FPU/g de substrato, suplementada ou não com a C21) foi analisada na presença e ausência de xilooligômeros. A figura 23 mostra que os xilooligômeros inibiram fortemente a atividade Accellerase®, especialmente nos pontos iniciais. Comparando com a reação sem adição de xilooligômeros, a atividade diminuiu 96,5% na primeira hora de reação. A inibição diminuiu ao longo do tempo, passando a 45% após 96 horas de reação. Portanto, as xilanases presentes no coquetel enzimático provavelmente foram capazes de degradar, em certa medida, os xilooligômeros adicionados e diminuíram a inibição observada. Quando a C21 foi adicionada ao coquetel, a queda de atividade variou entre 83% e 53% no intervalo de uma a 96 horas de reação. Portanto, a inibição nos estágios iniciais foi muito menor. Na primeira hora de reação, por exemplo, a diferença entre a reação suplementada e a não suplementada com a exoxilanase C21 foi de 380 vezes. HPLC Dionex mostrou que os xilooligômeros com grau de polimerização maior do que quatro foram exauridos da reação não suplementada com C21 apenas após 24 horas. Enquanto isso, 4 horas foram suficientes para que eles não fossem mais detectados na reação suplementada com C21. Como a concentração de xilooligômeros reduziu em ambas reações nos estágios finais de hidrólise, ambas alcançaram aproximadamente a mesma conversão após 96 horas.

Outros estudos já relataram que xilooligômeros são fortes inibidores de celulasas, enquanto xilose e xilobiose apresentam um efeito inibitório menor.¹⁴⁶ Entretanto, coquetéis enzimáticos comerciais possuem baixa atividade xilanase, o que provoca um acúmulo de xilooligômeros em processos industriais.²⁰¹ Suplementar coquetéis enzimáticos com a C21 provou ser eficiente para aumentar a performance enzimática enquanto há xilooligômeros na reação. Portanto, em processos industriais onde os xilooligômeros se acumulam,²⁰¹ coquetéis

suplementados com a C21 tem o potencial de aumentar a taxa de conversão ou diminuir o tempo de reação para alcançar a mesma conversão.

Figura 23 - Impacto de xilooligômeros na performance de um coquetel enzimático comercial. a) Hidrólise de Avicel por Accellerase® na presença de xilooligômeros e C21. b) Análise por HPLC Dionex da degradação dos xilooligômeros pela Accellerase® e c) pela Accellerase® suplementada com C21, como determinado pela resposta do detector em nanocoulomb (nC). Note que no item c, a concentração de xilobiose aumenta rapidamente após 1 hora de reação como resultado da rápida degradação pela C21 de xilooligômeros com grau de polimerização maior ou igual a quatro. Xilooligômeros com maior grau de polimerização possuem um efeito inibitório maior, portanto sua rápida degradação é benéfica. O coquetel comercial usado foi capaz de degradar os xilooligômeros adicionados dentro de 24 horas. Em consequência, após 96 horas, as reações com e sem adição de C21 alcançaram aproximadamente o mesmo rendimento de hidrólise. Entretanto, nos estágios iniciais, a adição da C21 aumentou grandemente a taxa de produção de glicose.

Fonte: Elaborada pelo autor.

7 CONCLUSÕES E PERSPECTIVAS

Em resumo, neste trabalho foi estudado como os recursos nutricionais presentes no meio de cultivo afetam a estrutura de comunidades microbianas. Através do sequenciamento da região V4-16S, foram encontradas diferenças significativas nas taxas microbianas dominantes do consórcio, dependendo do meio de cultivo. Foi sugerido que um número de microrganismos são requeridos para degradar efetivamente substratos recalcitrantes, tal como o bagaço de cana-de-açúcar. Assim, utilizar um meio sem outras fontes de carbono possibilitou o crescimento de uma comunidade tão ou mais diversa que aquela crescida em um meio com fontes de carbono facilmente metabolizáveis. Além disso, provocou o enriquecimento de grupos especializados na digestão da biomassa. Consequentemente, meios com limitações de nutrientes devem ser usados em aplicações biotecnológicas quando se deseja o enriquecimento em espécies com atividades específicas.

Uma vez determinado qual meio é mais adequado para esta aplicação, a degradação do bagaço de cana-de-açúcar por um consórcio microbiano derivado da compostagem e crescido em meio mínimo foi estudada. A ampla degradação da biomassa observada foi principalmente resultado da remoção das frações hemicelulósica e celulósica, sendo que não foram detectadas modificações na estrutura da lignina. Os dados de RNA-seq mostraram um grande conjunto de CAZymes expressas, com participação crescente de microrganismos eucariotos. A expressão de genes selecionados do metatranscriptoma revelou enzimas com características industrialmente interessantes. Em especial, foi demonstrado que a proteína C21 degrada o xilano por um mecanismo *exo* nunca antes relatado para proteínas da família GH 11. Essa proteína foi capaz de melhorar a performance de coquetéis enzimáticos e, portanto, é uma enzima promissora para a complementação de coquetéis utilizados na sacarificação da biomassa. Desta forma, os resultados obtidos evidenciaram a habilidade da comunidade microbiana obtida em degradar a biomassa e mostraram que o seu estudo é promissor para a descoberta de enzimas biotecnologicamente interessantes.

Como perspectivas fica a continuidade do estudo da biblioteca metatranscriptômica gerada e dos microrganismos isolados da compostagem. Milhares de alvos estão disponíveis para estudo, sendo que o trabalho apresentado aqui representa apenas uma pequena fração do que pode ser encontrado. Além disso, a experiência adquirida trouxe conhecimento de como melhor escolher os alvos para estudos futuros de forma a obter uma taxa de sucesso maior. Muitas outras classes de enzimas presentes ainda devem ser estudadas. Ainda mais, durante o cultivo, foram isolados microrganismos bacterianos e fúngicos que podem ser submetidos a

triagem em placa para detectar bons degradadores de celulose, lignina ou outros substratos. O estudo desses microrganismos poderá levar a descoberta de novas espécies e enzimas com alta atividade e/ou adaptadas a condições desejadas.

REFERÊNCIAS

- 1 WANG, C. et al. Metagenomic analysis of microbial consortia enriched from compost: new insights into the role of Actinobacteria in lignocellulose decomposition. **Biotechnology for Biofuels**, v. 9, p. 22, 2016. doi: 10.1186/s13068-016-0440-2.
- 2 SIMMONS, C. W. et al. Metatranscriptomic analysis of lignocellulolytic microbial communities involved in high-solids decomposition of rice straw. **Biotechnology for Biofuels**, v. 7, n. 1, p. 495, 2014.
- 3 MHUANTONG, W. et al. 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. doi.org/10.1186/s13068-015-0200-8.
- 4 RYCKEBOER, J. et al. A survey of bacteria and fungi occurring during composting and self-heating processes. **Annals of Microbiology**, v. 53, n. 4, p. 349-410, 2003
- 5 ISHII, K.; TAKII, S. Comparison of microbial communities in four different composting processes as evaluated by denaturing gradient gel electrophoresis analysis. **Journal of Applied Microbiology**, v. 95, n. 1, p. 109-119, 2003.
- 6 STEGER, K. et al. Microbial community growth and utilization of carbon constituents during thermophilic composting at different oxygen levels. **Microbial Ecology**, v. 50, n. 2, p. 163-171, 2005.
- 7 TAKEBAYASHI, S. et al. Water availability is a critical determinant of a population shift from Proteobacteria to Actinobacteria during start-up operation of mesophilic fed-batch composting. **Microbes and Environments**, v. 22, n. 3, p. 279-289, 2007.
- 8 VARGAS-GARCÍA, M. C. et al. Microbial population dynamics and enzyme activities in composting processes with different starting materials. **Waste Management**, v. 30, n. 5, p. 771-778, 2010.
- 9 PARTANEN, P. et al. Bacterial diversity at different stages of the composting process. **BMC Microbiology**, v. 10, p. 94, Mar. 2010. doi.org/10.1186/1471-2180-10-94.
- 10 MARTIN, F. **The ecological genomics of fungi**. New Jersey: Wiley Blackwell, 2014. 384 p.
- 11 PACE, N. R. A molecular view of microbial diversity and the biosphere. **Science**, v. 276, n. 5313, p. 734-740, May 1997.
- 12 SOGIN, M. L. et al. Microbial diversity in the deep sea and the underexplored "rare biosphere". **Proceedings of the National Academy of Sciences of the United States of America**, v. 103, n. 32, p. 12115-12120, Aug. 2006.
- 13 HESS, M. et al. Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. **Science**, v. 331, n. 6016, p. 463-467, Jan. 2011.

- 14 BRULC, J. M. et al. Gene-centric metagenomics of the fiber-adherent bovine rumen microbiome reveals forage specific glycoside hydrolases. **Proceedings of the National Academy of Sciences of the United States of America**, v. 106, n. 6, p. 1948-1953, Feb. 2009.
- 15 WARNECKE, F. et al. Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. **Nature**, v. 450, n. 7169, p. 560-565, Nov. 2007.
- 16 ZHU, L. et al. Evidence of cellulose metabolism by the giant panda gut microbiome. **Proceedings of the National Academy of Sciences of the United States of America**, v. 108, n. 43, p. 17714-17719, Oct. 2011.
- 17 ENGEL, P.; MARTINSON, V. G.; MORAN, N. A. Functional diversity within the simple gut microbiota of the honey bee. **Proceedings of the National Academy of Sciences of the United States of America**, v. 109, n. 27, p. 11002-11007, July 2012.
- 18 POPE, P. B. et al. Adaptation to herbivory by the Tammar wallaby includes bacterial and glycoside hydrolase profiles different from other herbivores. **Proceedings of the National Academy of Sciences of the United States of America**, v. 107, n. 33, p. 14793-14798, Aug. 2010.
- 19 LAMENDELLA, R. et al. Comparative fecal metagenomics unveils unique functional capacity of the swine gut. **BMC Microbiology**, v. 11, p. 103, May 2011. doi.org/10.1186/1471-2180-11-103.
- 20 QI, M. et al. Snapshot of the eukaryotic gene expression in muskoxen rumen--a metatranscriptomic approach. **PLoS One**, v. 6, n. 5, p. e20521, 2011.
- 21 TARTAR, A. et al. Parallel metatranscriptome analyses of host and symbiont gene expression in the gut of the termite *Reticulitermes flavipes*. **Biotechnology for Biofuels**, v. 2, p. 25, Oct. 2009. doi: 10.1186/1754-6834-2-25.
- 22 URICH, T. et al. Simultaneous assessment of soil microbial community structure and function through analysis of the meta-transcriptome. **PLoS One**, v. 3, n. 6, p. e2527, 2008.
- 23 GHAI, R. et al. Metagenomics of the water column in the pristine upper course of the Amazon river. **PLoS One**, v. 6, n. 8, p. e23785, 2011.
- 24 MARTINS, L. F. et al. Metagenomic analysis of a tropical composting operation at the são paulo zoo park reveals diversity of biomass degradation functions and organisms. **PLoS One**, v. 8, n. 4, p. e61928, 2013.
- 25 JIMÉNEZ, D. J.; CHAVES-MORENO, D.; VAN ELSAS, J. D. Unveiling the metabolic potential of two soil-derived microbial consortia selected on wheat straw. **Scientific Reports**, v. 5, p. 13845, Sept. 2015. doi:10.1038/srep13845
- 26 HOLLISTER, E. B. et al. Mesophilic and thermophilic conditions select for unique but highly parallel microbial communities to perform carboxylate platform biomass conversion. **PLoS One**, v. 7, n. 6, p. e39689, June 2012. doi: 10.1371/journal.pone.0039689.

- 27 ALLGAIER, M. et al. Targeted discovery of glycoside hydrolases from a switchgrass-adapted compost community. **PLoS One**, v. 5, n. 1, p. e8812, Jan. 2010.
- 28 CURTIS, T. P.; HEAD, I. M.; GRAHAM, D. W. Theoretical ecology for engineering biology. **Environmental Science & Technology**, v. 37, n. 3, p. 64A-70A, Fev. 2003.
- 29 RITTMANN, B. E. et al. A vista for microbial ecology and environmental biotechnology. **Environmental Science & Technology**, v. 40, n. 4, p. 1096-103, Fev. 2006.
- 30 CASTILLO, J. M.; ROMERO, E.; NOGALES, R. Dynamics of microbial communities related to biochemical parameters during vermicomposting and maturation of agroindustrial lignocellulose wastes. **Bioresource Technology**, v. 146, p. 345-54, Oct. 2013.
- 31 BP Statistical Review of World Energy. 2016. Disponível em: <<http://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html>>. Acesso em: 23 jan. 2017.
- 32 WEBER, C. et al. Trends and challenges in the microbial production of lignocellulosic bioalcohol fuels. **Applied Microbiology and Biotechnology**, v. 87, n. 4, p. 1303-1315, July 2010.
- 33 PANDEY, A. et al. Biotechnological potential of agro-industrial residues. I: sugarcane bagasse. **Bioresource Technology**, v. 74, n. 1, p. 69-80, Aug. 2000.
- 34 BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. Secretaria de Produção e Agroenergia. Declaração de Conformidade de Atividade Agropecuária. **Estatísticas e dados básicos de economia agrícola**. 2016. Disponível em: <http://www.agricultura.gov.br/arq_editor/PASTA_DE_NOVEMBRO.pdf>. Acesso em: 23 jan. 2017.
- 35 RODRIGUES, F. D. A.; GUIARDELLO, R. Evaluation of a sugarcane bagasse acid hydrolysis technology. **Chemical Engineering & Technology**, v. 31, n. 6, p. 883-892, Jun. 2008.
- 36 LIMA, M. A. et al. Evaluating the composition and processing potential of novel sources of Brazilian biomass for sustainable biorenewables production. **Biotechnology for Biofuels**, v. 7, n. 1, p. 10, Jan. 2014.
- 37 SOMERVILLE, C. et al. Feedstocks for lignocellulosic biofuels. **Science**, v. 329, n. 5993, p. 790-792, Aug. 2010.
- 38 MIRZA, U. K.; AHMAD, N.; MAJEED, T. An overview of biomass energy utilization in Pakistan. **Renewable & Sustainable Energy Reviews**, v. 12, n. 7, p. 1988-1996, Sept. 2008.
- 39 LYND, L. R. et al. Microbial cellulose utilization: fundamentals and biotechnology. **Microbiology and Molecular Biology Reviews**, v. 66, n. 3, p. 506-577, Sept. 2002.
- 40 PERCIVAL ZHANG, Y. H.; HIMMEL, M. E.; MIELENZ, J. R. Outlook for cellulase improvement: screening and selection strategies. **Biotechnology Advances**, v. 24, n. 5, p. 452-81, 2006 Sept. 2006.

- 41 GLOSTER, T. M. et al. Characterization and three-dimensional structures of two distinct bacterial xyloglucanases from families GH5 and GH12. **Journal of Biological Chemistry**, v. 282, n. 26, p. 19177-19189, June 2007.
- 42 SAHA, B. C. Lignocellulose biodegradation and applications in biotechnology. In: SAHA, B. C.; HAYASHI, K. (Ed.). **Lignocellulose biodegradation**. Washington: American Chemical Society, 2004. cap. 1, p. 2-34. (ACS Symposium Series, v. 889).
- 43 RUBIN, E. M. Genomics of cellulosic biofuels. **Nature**, v. 454, n. 7206, p. 841-845, Aug. 2008.
- 44 VANHOLME, R. et al. Lignin biosynthesis and structure. **Plant Physiology**, v. 153, n. 3, p. 895-905, July 2010.
- 45 GRABBER, J. H. How do lignin composition, structure, and cross-linking affect degradability? a review of cell wall model studies this paper was originally presented at the lignin and forage digestibility symposium, 2003 CSSA annual meeting, Denver, CO. **Crop Science**, v. 45, n. 3, p. 820-831, 2005.
- 46 BADGER, P. C. Ethanol from cellulose: a general review. In: JANICK, J.; WHIPKEY, A. (Ed.). **Trends in new crops and new uses**. Alexandria, VA: SHS Press, 2002.
- 47 BOERJAN, W.; RALPH, J.; BAUCHER, M. Lignin biosynthesis. **Annual Review of Plant Biology**, v. 54, p. 519-546, 2003.
- 48 SAHA, B. C. Hemicellulose bioconversion. **Journal of Industrial Microbiology and Biotechnology**, v. 30, n. 5, p. 279-291, May 2003.
- 49 MCMILLAN, J. D. Pretreatment of lignocellulosic biomass. In: Himmel, M. E.; Baker, J. O.; Overend, R. P. (Ed.). **Enzymatic conversion of biomass for fuels production**. Michigan: American Chemical Society, 1994. cap. 15, p. 292-324. (ACS Symposium Series, v. 566).
- 50 DÉJEAN, G. et al. The xylan utilization system of the plant pathogen *Xanthomonas campestris* pv *campestris* controls epiphytic life and reveals common features with oligotrophic bacteria and animal gut symbionts. **New Phytologist**, v. 198, n. 3, p. 899-915, Maio 2013.
- 51 SUN, J. X. et al. Isolation and characterization of cellulose from sugarcane bagasse. **Polymer Degradation and Stability**, v. 84, n. 2, p. 331-339, May 2004.
- 52 MATULOVA, A. et al. Degradation of wheat straw by fibrobacter succinogenes S85: a liquid-and solid-state nuclear magnetic resonance study. **Applied and Environmental Microbiology**, v. 71, n. 3, p. 1247-1253, Mar. 2005.
- 53 WADA, M. et al. The structure of celluloses. **Powder Diffraction**, v. 23, n. 2, p. 92-95, 2008.
- 54 STONE, B. Cellulose: structure and distribution. In: **Encyclopedia of life sciences**. Chichester: John Wiley & Sons Ltd, 2005. doi: 10.1038/npg.els.0003892.

55 HALL, G. S. WOOD: chemistry, ultrastructure, reactions. by FENGEL, D, WEGENER, G. **Nature**, v. 310, n. 5977, p. 521-521, 1984. doi:10.1038/310521b0.

56 DEPARTMENT OF ENERGY OF THE UNITED STATES. U. S. Department of Energy's Bioenergy Research Centers: an overview of the Science. 2010. Disponível em: < http://www.doe.genomestolife.org/centers/BRCbrochure2010webFinalURLs_HR.pdf>. Acesso em: 30 nov. de 2014.

57 MOSIER, N. et al. Features of promising technologies for pretreatment of lignocellulosic biomass. **Bioresource Technology**, v. 96, n. 6, p. 673-686, Apr. 2005.

58 CHANDRA, R. P. et al. The influence of lignin on steam pretreatment and mechanical pulping of poplar to achieve high sugar recovery and ease of enzymatic hydrolysis. **Bioresource Technology**, v. 199, p. 135-41, Jan. 2016.

59 CHANG, V. S.; NAGWANI, M.; HOLTZAPPLE, M. T. Lime pretreatment of crop residues bagasse and wheat straw. **Applied Biochemistry and Biotechnology**, v. 74, n. 3, p. 135-159, Sept. 1998.

60 REZENDE, C. A. et al. Chemical and morphological characterization of sugarcane bagasse submitted to a delignification process for enhanced enzymatic digestibility. **Biotechnology for Biofuels**, v. 4, n. 1, p. 54, 2011.

61 WARD, O. P.; SINGH, A. Bioethanol technology: developments and perspectives. **Advances in Applied Microbiology**, v. 51, p. 53-80, Feb. 2002. doi: 10.1016/S0065-2164(02)51001-7

62 KLEIN-MARCUSCHAMER, D. et al. The challenge of enzyme cost in the production of lignocellulosic biofuels. **Biotechnology and Bioengineering**, v. 109, n. 4, p. 1083-1087, Apr. 2012.

63 BHAT, M. K.; BHAT, S. Cellulose degrading enzymes and their potential industrial applications. **Biotechnology Advances**, v. 15, n. 3-4, p. 583-620, 1997.

64 JOHANSEN, K. S. Lytic Polysaccharide Monooxygenases: The Microbial Power Tool for Lignocellulose Degradation. **Trends in Plant Science**, v. 21, n. 11, p. 926-936, Nov. 2016.

65 EKWE, E. et al. Non-hydrolytic cellulose active proteins: research progress and potential application in biorefineries. **Industrial Biotechnology**, v. 9, n. 3, p. 123-131, 2013. doi.org/10.1089/ind.2013.0010

66 BEGUIN, P.; AUBERT, J. P. The biological degradation of cellulose. **Fems Microbiology Reviews**, v. 13, n. 1, p. 25-58, Jan. 1994.

67 SILVEIRA, M. H. L. et al. The essential role of plant cell wall degrading enzymes in the success of biorefineries: current status and future challenges. In: DA SILVA, S. S.; CHANDEL, A. K. (Ed.). **Biofuels in Brazil: fundamental aspects, recent developments, and future perspectives**. Cham: Springer International Publishing, 2014. p.151-172.

68 GREGG, D. J.; BOUSSAID, A.; SADDLER, J. N. Techno-economic evaluations of a generic wood-to-ethanol process: Effect of increased cellulose yields and enzyme recycle. **Bioresource Technology**, v. 63, n. 1, p. 7-12, Jan. 1998.

69 LYND, L. R. et al. Fuel ethanol from cellulosic biomass. **Science**, v. 251, n. 4999, p. 1318-1323, Mar. 1991.

70 HENRISSAT, B.; DAVIES, G. Structural and sequence-based classification of glycoside hydrolases. **Current Opinion in Structural Biology**, v. 7, n. 5, p. 637-644, Oct. 1997.

71 HENRISSAT, B. A classification of glycosyl hydrolases based on amino acid sequence similarities. **Biochemical Journal**, v. 280, Pt 2, p. 309-316, Dec. 1991.

72 LEVASSEUR, A. et al. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. **Biotechnology for Biofuels**, v. 6, n. 1, p. 41, 2013.

73 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, June 2012.

74 SERPA, V. I.; POLIKARPOV, I. Enzymes in bioenergy. In: BUCKERIDGE, M. S.; GOLDMAN, G. H. (Ed.). **Routes to cellulosic ethanol**. New York: Springer, 2011. p.97-113.

75 ZHU, B.; WANG, X.; LI, L. Human gut microbiome: the second genome of human body. **Protein & Cell**, v. 1, n. 8, p. 718-725, Aug. 2010.

76 NATIONAL RESEARCH COUNCIL COMMITTEE ON METAGENOMICS: challenges and functional, A. the National Academies Collection: reports funded by National Institutes of Health. In: **The new science of metagenomics: revealing the secrets of our microbial planet**. Washington (DC): National Academies Press; National Academy of Sciences, 2007.

77 TORSVIK, V.; OVREAS, L.; THINGSTAD, T. F. Prokaryotic diversity - magnitude, dynamics, and controlling factors. **Science**, v. 296, n. 5570, p. 1064-1066, May 2002.

78 WHITE, T. J. et al. **Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics**. San Diego, California: Academic Press, 1990.

79 CAPORASO, J. G. et al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. **Proceedings of the National Academy of Sciences of the United States of America**, v. 108, p. 4516-4522, Mar. 2011.

80 VENTER, J. C. et al. Environmental genome shotgun sequencing of the Sargasso Sea. **Science**, v. 304, n. 5667, p. 66-74, Apr. 2004.

81 GODZIK, A. Metagenomics and the protein universe. **Current Opinion in Structural Biology**, v. 21, n. 3, p. 398-403, June 2011.

82 WARNECKE, F.; HESS, M. A perspective: metatranscriptomics as a tool for the discovery of novel biocatalysts. **Journal of Biotechnology**, v. 142, n. 1, p. 91-95, June 2009.

- 83 DARBY, A. C.; HALL, N. Fast forward genetics. **Nature Biotechnology**, v. 26, n. 11, p. 1248-1249, Nov. 2008.
- 84 AUDIA, J. P.; PATTON, M. C.; WINKLER, H. H. DNA microarray analysis of the heat shock transcriptome of the obligate intracytoplasmic pathogen *Rickettsia prowazekii*. **Applied and Environmental Microbiology**, v. 74, n. 24, p. 7809-7812, Dec. 2008.
- 85 LIU, Y. Q. et al. Transcriptome dynamics of *Deinococcus radiodurans* recovering from ionizing radiation. **Proceedings of the National Academy of Sciences of the United States of America**, v. 100, n. 7, p. 4191-4196, Apr. 1 2003.
- 86 QIU, J. et al. DNA microarray-based global transcriptional profiling of *Yersinia pestis* in multicellularity. **Journal of Microbiology**, v. 46, n. 5, p. 557-563, Oct. 2008.
- 87 BAILLY, J. et al. Soil eukaryotic functional diversity, a metatranscriptomic approach. **Isme Journal**, v. 1, n. 7, p. 632-642, Nov. 2007.
- 88 PORETSKY, R. S. et al. Analysis of microbial gene transcripts in environmental samples. **Applied and Environmental Microbiology**, v. 71, n. 7, p. 4121-4126, July 2005.
- 89 GILBERT, J. A. et al. Detection of large numbers of novel sequences in the metatranscriptomes of complex marine microbial communities. **Plos One**, v. 3, n. 8, Aug. 2008.
- 90 LEININGER, S. et al. Archaea predominate among ammonia-oxidizing prokaryotes in soils. **Nature**, v. 442, n. 7104, p. 806-809, Aug. 2006.
- 91 TAKASAKI, K. et al. Discovery of glycoside hydrolase enzymes in an avicel-adapted forest soil fungal community by a metatranscriptomic approach. **Plos One**, v. 8, n. 2, Feb. 2013.
- 92 XIA, Y. et al. Mining of novel thermo-stable cellulolytic genes from a thermophilic cellulose-degrading consortium by metagenomics. **Plos One**, v. 8, n. 1, Jan. 2013.
- 93 HESS, M. et al. Metagenomic discovery of biomass-degrading genes and genomes from Cow Rumen. **Science**, v. 331, n. 6016, p. 463-467, Jan. 2011.
- 94 WARNECKE, F. et al. Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. **Nature**, v. 450, n. 7169, p. 560-U17, Nov. 2007.
- 95 POPE, P. B. et al. Metagenomics of the Svalbard Reindeer Rumen microbiome Reveals abundance of polysaccharide utilization loci. **Plos One**, v. 7, n. 6, June 2012.
- 96 BRULC, J. M. et al. Gene-centric metagenomics of the fiber-adherent bovine rumen microbiome reveals forage specific glycoside hydrolases. **Proceedings of the National Academy of Sciences of the United States of America**, v. 106, n. 6, p. 1948-1953, Feb. 2009.
- 97 ZHAO, H. et al. Regenerating cellulose from ionic liquids for an accelerated enzymatic hydrolysis. **Journal of Biotechnology**, v. 139, n. 1, p. 47-54, Jan. 2009.

- 98 URICH, T. et al. Simultaneous assessment of soil microbial community structure and function through analysis of the meta-transcriptome. **Plos One**, v. 3, n. 6, June 2008.
- 99 DAVIS, K. E. R.; JOSEPH, S. J.; JANSSEN, P. H. Effects of growth medium, inoculum size, and incubation time on culturability and isolation of soil bacteria. **Applied and Environmental Microbiology**, v. 71, n. 2, p. 826-834, Feb. 2005.
- 100 WEI, H. et al. Tracking dynamics of plant biomass composting by changes in substrate structure, microbial community, and enzyme activity. **Biotechnology for Biofuels**, v. 5, p. 14, Apr. 2012.
- 101 BARR, B. Composter reviews: 13 best bins and tumblers of 2016. 2016. Disponível em: <<https://www.ecokarma.net/composting/composter-reviews/>>. Acesso em: 23 jan. 2017.
- 102 KEIJERS, E. R.; YILMAZ, G.; VAN DAM, J. E. The cellulose resource matrix. **Carbohydrate Polymers**, v. 93, n. 1, p. 9-21, Mar. 2013.
- 103 XI, Y. L. et al. Ultrasonic pretreatment and acid hydrolysis of sugarcane bagasse for succinic acid production using *Actinobacillus succinogenes*. **Bioprocess and Biosystems Engineering**, v. 36, n. 11, p. 1779-1785, Nov. 2013.
- 104 CANTARELLA, H. et al. Sugarcane. In: CHITTARANJAN, K.; JOSHI, C. P.; SHONNARD, D. R. (Ed.). **Handbook of bioenergy crop plants**. New York: CRC Press, 2012. cap. 21. p. 523-561.
- 105 BOND, A. et al. Shock treatment of corn stover. **Biotechnology progress**. No prelo. doi: 10.1002/btpr.2437.
- 106 KUBICEK, C. P.; KUBICEK, E. M. Enzymatic deconstruction of plant biomass by fungal enzymes. **Current Opinion in Chemical Biology**, v. 35, p. 51-57, Dec. 2016.
- 107 WONGWILAIWALIN, S. et al. Analysis of a thermophilic lignocellulose degrading microbial consortium and multi-species lignocellulolytic enzyme system. **Enzyme and Microbial Technology**, v. 47, n. 6, p. 283-290, Nov. 2010.
- 108 HUTNER, S. H. et al. Some Approaches to the Study of the Role of Metals in the Metabolism of Microorganisms. **Proceedings of the American Philosophical Society**, v. 94, n. 2, p. 152-170, 1950.
- 109 MAGOC, T.; SALZBERG, S. L. FLASH: fast length adjustment of short reads to improve genome assemblies. **Bioinformatics**, v. 27, n. 21, p. 2957-2963, Nov. 2011.
- 110 CAPORASO, J. G. et al. QIIME allows analysis of high-throughput community sequencing data. **Nature Methods**, v. 7, n. 5, p. 335-336, May 2010.
- 111 EDGAR, R. C. Search and clustering orders of magnitude faster than BLAST. **Bioinformatics**, v. 26, n. 19, p. 2460-2461, Oct. 2010.
- 112 EDGAR, R. C. et al. UCHIME improves sensitivity and speed of chimera detection. **Bioinformatics**, v. 27, n. 16, p. 2194-2200, Aug. 2011.

- 113 DESANTIS, T. Z. et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. **Applied and Environmental Microbiology**, v. 72, n. 7, p. 5069-5072, July 2006.
- 114 HAMMER, O.; HARPER, D. A. T.; RYAN, P. D. PAST: paleontological statistics software package for education and data analysis. **Palaeontologia Electronica**, v. 4, n. 1, 2001.
- 115 WICKHOLM, K.; LARSSON, P. T.; IVERSEN, T. Assignment of non-crystalline forms in cellulose I by CP/MAS ¹³C NMR spectroscopy. **Carbohydrate Research**, v. 312, n. 3, p. 123-129, 1998.
- 116 TEMPLETON, D. W. et al. Compositional analysis of lignocellulosic feedstocks. 2. Method uncertainties. **Journal of Agricultural and Food Chemistry**, v. 58, n. 16, p. 9054-9062, Aug. 2010.
- 117 FOCHER, B. et al. Effects of structural features of cotton cellulose on enzymatic hydrolysis. **Journal of Applied Polymer Science**, v. 26, n. 6, p. 1989-1999, 1981.
- 118 HALLAC, B. B. et al. Biomass characterization of *Buddleja davidii*: a potential feedstock for biofuel production. **Journal of Agricultural and Food Chemistry**, v. 57, n. 4, p. 1275-1281, Feb. 2009.
- 119 EL HAGE, R. et al. Effects of process severity on the chemical structure of *Miscanthus* ethanol organosolv lignin. **Polymer Degradation and Stability**, v. 95, n. 6, p. 997-1003, 2010.
- 120 SANNIGRAHI, P.; MILLER, S. J.; RAGAUSKAS, A. J. Effects of organosolv pretreatment and enzymatic hydrolysis on cellulose structure and crystallinity in Loblolly pine. **Carbohydrate Research**, v. 345, n. 7, p. 965-970, May 2010.
- 121 FOSTON, M. B.; HUBBELL, C. A.; RAGAUSKAS, A. J. Cellulose isolation methodology for NMR analysis of cellulose ultrastructure. **Materials**, v. 4, n. 11, p. 1985-2002, Nov. 2011.
- 122 MARTÍNEZ, A. T. et al. Solid-state NMR Studies of Lignin and Plant Polysaccharide Degradation by Fungi. **Holzforschung**, v. 45, Supplement 1, p. 49-54, Jan. 1991
- 123 GRIFFITHS, R. I. et al. Rapid method for coextraction of DNA and RNA from natural environments for analysis of ribosomal DNA- and rRNA-based microbial community composition. **Applied and Environmental Microbiology**, v. 66, n. 12, p. 5488-5491, Dec. 2000.
- 124 FIERER, N. et al. Assessment of soil microbial community structure by use of taxon-specific quantitative PCR assays. **Applied and Environmental Microbiology**, v. 71, n. 7, p. 4117-4120, July 2005.
- 125 BOLGER, A. M.; LOHSE, M.; USADEL, B. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, n. 15, p. 2114-2120, Aug. 2014.

126 KOPYLOVA, E.; NOE, L.; TOUZET, H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. **Bioinformatics**, v. 28, n. 24, p. 3211-3217, Dec. 2012.

127 PRUESSE, E. et al. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. **Nucleic Acids Research**, v. 35, n. 21, p. 7188-7196, 2007.

128 GARDNER, P. P. et al. Rfam: wikipedia, clans and the "decimal" release. **Nucleic Acids Research**, v. 39, p. D141-D145, Jan. 2011.

129 GRABHERR, M. G. et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. **Nature Biotechnology**, v. 29, n. 7, p. 644, July 2011.

130 ISMAIL, W. M.; YE, Y.; TANG, H. Gene finding in metatranscriptomic sequences. **BMC Bioinformatics**, v. 15, Suppl 9, p. S8, 2014.

131 ROBERTS, A.; PACHTER, L. Streaming fragment assignment for real-time analysis of sequencing experiments. **Nature Methods**, v. 10, n. 1, p. 71-3, Jan. 2013.

132 LANGMEAD, B.; SALZBERG, S. L. Fast gapped-read alignment with Bowtie 2. **Nature Methods**, v. 9, n. 4, p. 357-9, Apr. 2012.

133 HUSON, D. H. et al. MEGAN analysis of metagenomic data. **Genome Research**, v. 17, n. 3, p. 377-386, Mar. 2007.

134 ANDERS, S.; PYL, P. T.; HUBER, W. HTSeq--a Python framework to work with high-throughput sequencing data. **Bioinformatics**, v. 31, n. 2, p. 166-169, Jan. 2015.

135 FINN, R. D.; CLEMENTS, J.; EDDY, S. R. HMMER web server: interactive sequence similarity searching. **Nucleic Acids Research**, v. 39, p. W29-W37, July 2011.

136 YIN, Y. et al. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. **Nucleic Acids Research**, v. 40, n. W1, p. W445-W451, July 2012.

137 TEAM, R. D. C. R: A Language and environment for statistical computing 2014. Disponível em: <<https://www.R-project.org>>. Acesso em: 23 jan. 2017.

138 KANEHISA, M.; GOTO, S. KEGG: kyoto encyclopedia of genes and genomes. **Nucleic Acids Research**, v. 28, n. 1, p. 27-30, Jan. 2000.

139 KANEHISA, M.; SATO, Y.; MORISHIMA, K. BlastKOALA and GhostKOALA: KEGG Tools for functional characterization of genome and metagenome sequences. **Journal of Molecular Biology**, v. 428, n. 4, p. 726-731, Feb. 2016.

140 ASLANIDIS, C.; DEJONG, P. J. Ligation-independent cloning of PCR products (LIC-PCR). **Nucleic Acids Research**, v. 18, n. 20, p. 6069-6074, Out. 1990.

141 MICHAEL, R. G.; JOSEPH, S. **Molecular cloning: a laboratory manual**. Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 2012.

- 142 CAMILO, C. M.; POLIKARPOV, I. High-throughput cloning, expression and purification of glycoside hydrolases using Ligation-Independent Cloning (LIC). **Protein Expression and Purification**, v. 99, p. 35-42, July 2014.
- 143 LAEMMLI, U. K. Cleavage of structural proteins during assembly of head of bacteriophage-T4. **Nature**, v. 227, n. 5259, Aug. 1970. doi:10.1038/227680a0.
- 144 SIEVERS, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. **Molecular Systems Biology**, v. 7, p. 539, Oct. 2011.
- 145 GHOSE, T. K. Measurement of cellulase activities. **Pure and Applied Chemistry**, v. 59, n. 2, p. 257-268, Feb. 1987.
- 146 QING, Q.; YANG, B.; WYMAN, C. E. Xylooligomers are strong inhibitors of cellulose hydrolysis by enzymes. **Bioresource Technology**, v. 101, n. 24, p. 9624-9630, Dec. 2010.
- 147 SLUITER, A. et al. **Determination of sugars, byproducts, and degradation products in liquid fraction process samples.** NREL/TP-510-42623. Disponível em: <<http://www.nrel.gov/docs/gen/fy08/42623.pdf>>. Acesso em: 23 jan. 2017.
- 148 SELIG, M.; WEISS, N.; JI, Y. **Enzymatic saccharification of lignocellulosic biomass. NREL/TP-510-42629.** Disponível em: <<http://www.nrel.gov/docs/gen/fy08/42629.pdf>>. Acesso em: 23 jan. 2017.
- 149 GUIMARÃES, B. G. et al. The MX2 macromolecular crystallography beamline: a wiggler X-ray source at the LNLS. **Journal of Synchrotron Radiation**, v. 16, n. Pt 1, p. 69-75, Jan. 2009.
- 150 KABSCH, W. XDS. **Acta Crystallographica D**, v. 66, n. Pt 2, p. 125-132, Feb. 2010.
- 151 MCCOY, A. J. et al. Phaser crystallographic software. **Journal of Applied Crystallography**, v. 40, n. Pt 4, p. 658-674, Aug. 2007.
- 152 ADAMS, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. **Acta Crystallographica D**, v. 66, n. Pt 2, p. 213-221, Feb. 2010.
- 153 EMSLEY, P.; COWTAN, K. Coot: model-building tools for molecular graphics. **Acta Crystallographica D**, v. 60, p. 2126-2132, Dec. 2004.
- 154 CHEN, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. **Acta Crystallographica D**, v. 66, n. Pt 1, p. 12-21, Jan. 2010.
- 155 AMANN, R. I.; LUDWIG, W.; SCHLEIFER, K. H. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. **Microbiological Reviews**, v. 59, n. 1, p. 143-169, Mar. 1995.
- 156 MCCAIG, A. E. et al. Impact of cultivation on characterisation of species composition of soil bacterial communities. **FEMS Microbiology Ecology**, v. 35, n. 1, p. 37-48, Mar. 2001.

- 157 PHAM, V. H.; KIM, J. Cultivation of unculturable soil bacteria. **Trends in Biotechnology**, v. 30, n. 9, p. 475-484, Sept. 2012.
- 158 GRABER, J. R.; BREZNAK, J. A. Folate cross-feeding supports symbiotic homoacetogenic spirochetes. **Applied and Environmental Microbiology**, v. 71, n. 4, p. 1883-1889, Apr. 2005.
- 159 TRIPP, H. J. et al. SAR11 marine bacteria require exogenous reduced sulphur for growth. **Nature**, v. 452, n. 7188, p. 741-744, Apr. 2008.
- 160 STEWART, E. J. Growing unculturable bacteria. **Journal of Bacteriology**, v. 194, n. 16, p. 4151-4160, Aug. 2012.
- 161 EITEMAN, M. A.; LEE, S. A.; ALTMAN, E. A co-fermentation strategy to consume sugar mixtures effectively. **Journal of Biological Engineering**, v. 2, p. 3, Feb. 2008. doi: 10.1186/1754-1611-2-3.
- 162 SZAMBELAN, K.; NOWAK, J.; CZARNECKI, Z. Use of *Zymomonas mobilis* and *Saccharomyces cerevisiae* mixed with *Kluyveromyces fragilis* for improved ethanol production from Jerusalem artichoke tubers. **Biotechnology Letters**, v. 26, n. 10, p. 845-848, May 2004.
- 163 SCHWARZ, W. H. The cellulosome and cellulose degradation by anaerobic bacteria. **Applied Microbiology and Biotechnology**, v. 56, n. 5-6, p. 634-649, Sept. 2001.
- 164 HUI, W. et al. Bioconversion of un-pretreated lignocellulosic materials by a microbial consortium XDC-2. **Bioresource Technology**, v. 136, p. 481-487, May 2013.
- 165 LV, Z. et al. Characterization of extracellular and substrate-bound cellulases from a mesophilic sugarcane bagasse-degrading microbial community. **Process Biochemistry**, v. 43, n. 12, p. 1467-1472, Dec. 2008.
- 166 DE LIMA BROSSI, M. J. et al. Soil-derived microbial consortia enriched with different plant biomass reveal distinct players acting in lignocellulose degradation. **Microbial Ecology**, v. 71, n. 3, Oct. 2015. doi: 10.1007/s00248-015-0683-7.
- 167 TAKASAKI, K. et al. Discovery of glycoside hydrolase enzymes in an avicel-adapted forest soil fungal community by a metatranscriptomic approach. **PLoS One**, v. 8, n. 2, p. e55485, 2013.
- 168 JIN, M. et al. Consolidated bioprocessing (CBP) performance of *Clostridium* phytofermentans on AFEX-treated corn stover for ethanol production. **Biotechnology and Bioengineering**, v. 108, n. 6, p. 1290-1297, June 2011.
- 169 MOXLEY, G.; ZHANG, Y. H. P. More accurate determination of acid-labile carbohydrates in lignocellulose by modified quantitative saccharification. **Energy & Fuels**, v. 21, n. 6, p. 3684-3688, 2007.
- 170 GILBERT, J. A.; JANSSON, J. K.; KNIGHT, R. The earth microbiome project: successes and aspirations. **BMC Biology**, v. 12, n. 1, p. 69, 2014.

- 171 HUGENHOLTZ, P.; GOEBEL, B. M.; PACE, N. R. Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. **Journal of Bacteriology**, v. 180, n. 18, p. 4765-4774, Sept. 1998.
- 172 VENTORINO, V. et al. Exploring the microbiota dynamics related to vegetable biomasses degradation and study of lignocellulose-degrading bacteria for industrial biotechnological application. **Science Reports**, v. 5, p. 8161, Feb. 2015. doi: 10.1038/srep08161.
- 173 NIELSEN, P.; SØRENSEN, J. Multi-target and medium-independent fungal antagonism by hydrolytic enzymes in *Paenibacillus polymyxa* and *Bacillus pumilus* strains from barley rhizosphere. **FEMS Microbiology Ecology**, v. 22, n. 3, p. 183-192, 1997.
- 174 MERGAERT, J. et al. Taxonomic study of *Cellvibrio* strains and description of *Cellvibrio ostraviensis* sp. nov., *Cellvibrio fibrivorans* sp. nov. and *Cellvibrio gandavensis* sp. nov. **International Journal of Systematic and Evolutionary Microbiology**, v. 53, n. Pt 2, p. 465-471, Mar. 2003.
- 175 MITRI, S.; CLARKE, E.; FOSTER, K. R. Resource limitation drives spatial organization in microbial groups. **ISME Journal**, v. 10, n. 6, p. 1471-1482, June 2016.
- 176 DASHTBAN, M. et al. Fungal biodegradation and enzymatic modification of lignin. **International Journal of Biochemistry and Molecular Biolog**, v. 1, n. 1, p. 36-50, 2010.
- 177 WANG, Y. et al. Evaluation of the coverage and depth of transcriptome by RNA-Seq in chickens. **BMC Bioinformatics**, v. 12, Supplement 10, p. S5, 2011.
- 178 VERSLUIS, D. et al. Mining microbial metatranscriptomes for expression of antibiotic resistance genes under natural conditions. **Scientific Reports**, v. 5, p. 11981, July 2015. doi:10.1038/srep11981.
- 179 JIANG, Y. et al. Metatranscriptomic analysis of diverse microbial communities reveals core metabolic pathways and microbiome-specific functionality. **Microbiome**, v. 4, p. 2, Jan. 2016. doi: 10.1186/s40168-015-0146-x
- 180 BERLEMONT, R.; MARTINY, A. C. Phylogenetic distribution of potential cellulases in bacteria. **Applied and Environmental Microbiology**, v. 79, n. 5, p. 1545-1554, Mar. 2013.
- 181 DAMON, C. et al. Metatranscriptomics reveals the diversity of genes expressed by eukaryotes in forest soils. **PLoS One**, v. 7, n. 1, p. e28967, 2012.
- 182 PAUCHET, Y. et al. Diversity of beetle genes encoding novel plant cell wall degrading enzymes. **PLoS One**, v. 5, n. 12, p. e15635, 2010.
- 183 COLBOURNE, J. K. et al. The ecoresponsive genome of *Daphnia pulex*. **Science**, v. 331, n. 6017, p. 555-561, Feb. 2011.
- 184 DANCHIN, E. G. et al. Multiple lateral gene transfers and duplications have promoted plant parasitism ability in nematodes. **Proceedings of the National Academy of Sciences of the United States of America**, v. 107, n. 41, p. 17651-17656, Oct. 2010.

- 185 KERN, M. et al. Structural characterization of a unique marine animal family 7 cellobiohydrolase suggests a mechanism of cellulase salt tolerance. **Proceedings of the National Academy of Sciences of the United States of America**, v. 110, n. 25, p. 10189-10194, June 2013.
- 186 BACIC, A.; FINCHER, G. B.; STONE, B. A. **Chemistry, biochemistry, and biology of (1-3)-[beta]-glucans and related polysaccharides**. Amsterdam: Academic, 2009.
- 187 WOOD, P. J.; WEISZ, J.; BLACKWELL, B. A. Structural studies of (1-3),(1-4)-beta-d-glucans by C(13)-nuclear magnetic-resonance spectroscopy and by rapid analysis of cellulose-like regions using high-performance anion-exchange chromatography of oligosaccharides released by lichenase. **Cereal Chemistry**, v. 71, n. 3, p. 301-307, May 1994.
- 188 SARETHY, I. P. et al. Alkaliphilic bacteria: applications in industrial biotechnology. **Journal of Industrial Microbiology and Biotechnology**, v. 38, n. 7, p. 769-790, July 2011.
- 189 KNOB, A.; CARMONA, E. C. Purification and characterization of two extracellular xylanases from *Penicillium sclerotiorum*: a novel acidophilic xylanase. **Applied Biochemistry and Biotechnology**, v. 162, n. 2, p. 429-443, Sept. 2010.
- 190 CHANG, L. et al. Characterization of a bifunctional xylanase/endoglucanase from yak rumen microorganisms. **Applied Microbiology and Biotechnology**, v. 90, n. 6, p. 1933-1942, June 2011.
- 191 AMEL, B. D. et al. Characterization of a purified thermostable xylanase from *Caldicoprobacter algeriensis* sp. nov. strain TH7C1(T). **Carbohydrate Research**, v. 419, p. 60-68, Jan. 2016.
- 192 MCCLEARY, B. V. Measurement of polysaccharide degrading enzymes using chromogenic and colorimetric substrates. **Chemistry in Australia**, v. 58, p. 398-401, 1991.
- 193 WATANABE, M. et al. Xylanase (GH11) from *Acremonium cellulolyticus*: homologous expression and characterization. **AMB Express**, v. 4, p. 27, Apr. 2014. doi: 10.1186/s13568-014-0027-x.
- 194 KATAOKA, M. et al. Crystal structure of *Talaromyces cellulolyticus* (formerly known as *Acremonium cellulolyticus*) GH family 11 xylanase. **Applied Biochemistry and Biotechnology**, v. 174, n. 4, p. 1599-1612, Oct. 2014.
- 195 JÄNIS, J. et al. Determination of thioxyl-oligosaccharide binding to family 11 xylanases using electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry and X-ray crystallography. **FEBS Journal**, v. 272, n. 9, p. 2317-2333, May 2005.
- 196 VARDAKOU, M. et al. Understanding the structural basis for substrate and inhibitor recognition in eukaryotic GH11 xylanases. **Journal of Molecular Biology**, v. 375, n. 5, p. 1293-305, Feb. 2008.
- 197 WAKARCHUK, W. W. et al. Mutational and crystallographic analyses of the active site residues of the *Bacillus circulans* xylanase. **Protein Science**, v. 3, n. 3, p. 467-75, Mar. 1994.

- 198 PAËS, G.; BERRIN, J. G.; BEAUGRAND, J. GH11 xylanases: Structure/function/properties relationships and applications. **Biotechnology Advances**, v. 30, n. 3, p. 564-592, May 2012.
- 199 VALENZUELA, S. V. et al. The glycoside hydrolase family 8 reducing-end xylose-releasing exo-oligoxylanase rex8A from *paenibacillus barcinonensis* BP-23 is active on branched xylooligosaccharides. **Applied and Environmental Microbiology**, v. 82, n. 17, p. 5116-5124, Sept. 2016.
- 200 JUTURU, V.; WU, J. C. Microbial xylanases: engineering, production and industrial applications. **Biotechnology Advances**, v. 30, n. 6, p. 1219-1227, Nov-Dec. 2012.
- 201 KUMAR, R.; WYMAN, C. E. Effect of xylanase supplementation of cellulase on digestion of corn stover solids prepared by leading pretreatment technologies. **Bioresource Technology**, v. 100, n. 18, p. 4203-4213, Sept. 2009.
- 202 BERLEMONT, R. et al. Cellulolytic potential under environmental changes in microbial communities from grassland litter. **Frontiers in Microbiology**, v. 5, p. 639, Nov. 2014. doi: 10.3389/fmicb.2014.00639.
- 203 JANIS, J. et al. Determination of thioxylo-oligosaccharide binding to family 11 xylanases using electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry and X-ray crystallography. **FEBS Journal**, v. 272, n. 9, p. 2317-2333, May 2015.