UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciências Biológicas
Programa de Pós-graduação em Microbiologia

Luiz Marcelo Ribeiro Tomé

BIOLOGIA DE SISTEMAS DO FUNGO *TRAMETES VILLOSA* CCMB561: integração de análises ômicas para compreensão do metabolismo da lignocelulose

Belo Horizonte
2023

Luiz Marcelo Ribeiro Tomé

BIOLOGIA DE SISTEMAS DO FUNGO *TRAMETES VILLOSA* CCMB561: integração de análises ômicas para compreensão do metabolismo da lignocelulose

Versão final

Tese apresentada ao programa de Pós-Graduação em Microbiologia da Universidade Federal de Minas Gerais, como requisito parcial para à obtenção do título de Doutor em Microbiologia.

Orientador: Prof. Dr. Aristóteles Góes Neto
Coorientadora: Profa. Dra. Fernanda Badotti

Belo Horizonte
2023

043 Tomé, Luiz Marcelo Ribeiro.
Biologia de sistemas do fungo Trametes villosa CCMB561: integração de análises ômicas para compreensão do metabolismo da lignocelulose [manuscrito] / Luiz Marcelo Ribeiro Tomé. – 2023.
114 f. : il. ; 29,5 cm.

Orientador: Prof. Dr. Aristóteles Góes Neto. Coorientadora: Profa. Dra. Fernanda Badotti.
Tese (doutorado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa de Pós-Graduação em Microbiologia.

1. Microbiologia. 2. Genômica. 3. Transcriptoma. 4. Biomassa. 5. Biodegradação Ambiental. 6. Trametes. I. Góes Neto, Aristóteles. II. Badotti, Fernanda. III. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. IV. Título.

CDU: 579

Ficha Catalográfica elaborada pela bibliotecária Fabiane C M Reis CRB 6 2680

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PÓS-GRADUAÇÃO EM MICROBIOLOGIA

ATA DE DEFESA DE TESE

ATA DA DEFESA DE TESE DE **LUIZ MARCELO RIBEIRO TOMÉ**

Nº REGISTRO: 2019663079

Às 08:30 horas do dia **01 de fevereiro de 2023**, reuniu-se, por via remota, a Comissão Examinadora composta pelos Drs. Flavia Figueira Aburjaile (Universidade Federal de Minas Gerais-UFMG), Gonçalo Amarante Guimarães Pereira (Universidade Estadual de Campinas (UNICAMP), Izinara Rosse da Cruz (Universidade Federal de Ouro Preto (UFOP), Bruno Silva Andrade (Universidade Estadual do Sudoeste da Bahia (UESB), o Prof. Dr. Aristóteles Góes Neto (Orientador) e a Dra. Fernanda Badotti (Coorientadora), para julgar o trabalho final **"Biologia de sistemas do fungo Trametes villosa CCMB561: Integração de análises ômicas para compreensão do metabolismo da lignocelulose"** do aluno **Luiz Marcelo Ribeiro Tomé**, requisito final para a obtenção do Grau de **DOUTOR EM CIÊNCIAS BIOLÓGICAS: MICROBIOLOGIA**. Abrindo a sessão, o Presidente da Comissão, Prof. Dr. Aristóteles Góes Neto , após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra ao candidato, para a apresentação de seu trabalho. Seguiu-se a arguição pelos Examinadores, com a respectiva defesa do candidato. Logo após, a Comissão se reuniu, sem a presença do candidato e do público, para julgamento e expedição de resultado final. O candidato foi considerado **APROVADO**. O resultado final foi comunicado publicamente ao candidato pelo Presidente da Comissão. Nada mais havendo a tratar, o Presidente encerrou a reunião e lavrou a presente ata, que será assinada por todos os membros participantes da Comissão Examinadora. O candidato tem 60 (sessenta) dias, a partir desta data, para entregar a versão final da tese ao Programa de Pós-graduação em Microbiologia da UFMG e requerer seu diploma.

Belo Horizonte, 01 de fevereiro de 2023

Membros da Banca:

Profa. Dra. Flavia Figueira Aburjaile

Dr. Gonçalo Amarante Guimarães Pereira

Dra. Izinara Rosse da Cruz

Prof. Dr. Bruno Silva Andrade

De acordo:

Prof. Dr. . Aristóteles Góes Neto

(Orientador)
Dra. Fernanda Badotti
(Coorientadora)

Profa. Daniele da Glória de Souza
(Coordenadora do Programa de Pós-graduação
em Microbiologia)

Documento assinado eletronicamente por **Aristoteles Goes Neto**, **Professor do Magistério Superior**, em 01/02/2023, às 15:10, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Izinara Rosse da Cruz**, **Usuário Externo**, em 01/02/2023, às 15:32, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Fernanda Badotti**, **Usuário Externo**, em 01/02/2023, às 15:50, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Flávia Figueira Aburjaile**, **Professora do Magistério Superior**, em 01/02/2023, às 17:50, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Daniele da Gloria de Souza**, **Coordenador(a) de curso de pós-graduação**, em 02/02/2023, às 09:51, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Bruno Silva Andrade**, **Usuário Externo**, em 02/02/2023, às 10:51, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Gonçalo Amarante Guimarães Pereira**, **Usuário Externo**, em 02/02/2023, às 13:30, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **1965972** e o código CRC **2F2970F3**.

Referência: Processo nº 23072.227783/2021-61 SEI nº 1965972

A todas as pessoas que acreditam que a educação e a ciência são a base para a transformação e desenvolvimento de uma sociedade.

AGRADECIMENTOS

A Deus, por ter me concedido força, saúde e sabedoria;

Aos meus pais, pela luta para que eu sempre tivesse uma educação de qualidade, e por acreditarem em meu potencial. Em especial a minha mãe, que mesmo distante, se faz presente todos os dias em minha vida, me apoiando e dando forças para seguir nessa jornada encantadora e difícil que é a vida;

Aos meus irmãos Murilo e Felipe, por todo o amor e torcida;

Ao meu namorado Ramon, por todo o companheirismo, amor, respeito e força durante todo o período de desenvolvimento da minha tese. Só nós dois sabemos o real significado de "ninguém solta a mão de ninguém". Até aqui nós nos apoiamos e juntos estamos compartilhando conquistas;

A minha família que está na Bahia, por toda ajuda, carinho, amor e companheirismo, apesar da distância;

A Tia Lúcia, Tio Ribeiro, Bruna e David por todo o apoio durante o mestrado e na fase inicial do meu doutorado;

Ao meu orientador, Prof. Dr. Aristóteles Góes Neto, por todo o conhecimento compartilhado durante esses mais de 10 anos de orientação. Nossa história começou com uma orientação na iniciação científica na Universidade Estadual de Feira de Santana (UEFS), que me levou a um intercâmbio nos Estados Unidos (University of Maryland), e ao mestrado e doutorado na Universidade Federal de Minas Gerais (UFMG). Muito obrigado Ari, por todas as oportunidades concedidas a mim! Você mudou a minha vida através da educação;

A minha coorientadora, Profa. Dra. Fernanda Badotti, por ter aceitado me orientar e por ter acreditado nas minhas ideias e projetos. Muito obrigado por todos os ensinamentos compartilhados;

Aos meus amigos de Belo Horizonte, em especial, Vivian, Paula e Thairine, por todo apoio e amizade que foram essenciais para concluir essa jornada;

Aos secretários do Programa de Pós-Graduação em Microbiologia, em especial, Débora e Tiago, por todo o suporte durante a realização do meu doutorado;

Aos Professores do Programa de Pós-Graduação em Microbiologia, por todo o suporte e conhecimento compartilhado;

A equipe do Laboratório de Biologia Molecular e Computacional de Fungos (LBMCF) pela parceria nesses anos.

A UFMG pela estrutura fornecida e por ser uma instituição que mudou a minha vida através da educação;

Ao professor Dr. Bertram Brenig da University of Göttingen por todo o auxilio com os sequenciamentos dos genomas e transcriptomas;

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela bolsa concedida durante todo o período do doutorado.

“A dúvida é o princípio da sabedoria.” (Aristóteles).

RESUMO

A bioconversão de resíduos lignocelulósicos é uma alternativa eficiente para a obtenção de biocombustíveis e produtos químicos. Esses resíduos constituem uma matéria-prima renovável e sustentável, e são formados majoritariamente por celulose, hemicelulose e lignina. Para o aproveitamento de todos os constituintes da lignocelulose, é necessária a deslignificação e posterior despolimerização dos carboidratos a açúcares simples. Fungos da decomposição branca são organismos que, por meio da produção de enzimas hidrolíticas e oxidativas, possuem a habilidade de degradar todos os constituintes da parede celular vegetal. Estudos realizados com o fungo *Trametes villosa* CCMB561 vêm demostrando o potencial deste isolado em produzir enzimas ligninolíticas, e assim, atuar na despolimerização de resíduos vegetais. Portanto, para revelar as bases genéticas de degradação da lignocelulose por *Trametes villosa* CCMB561, neste trabalho, o genoma do fungo foi sequenciado utilizando as plataformas de sequenciamento MinION (Oxford Nanopore) e HiSeq (Illumina). Após a montagem, o genoma teve as regiões codificantes, tRNAs, elementos transponíveis, clusters de genes de metabólitos secundários (SMGCs), e genes codificadores de *Carbohydrate-Active Enzymes* (CAZymes) anotados. Além das análises genômicas, foram conduzidas análises bioquímicas de quantificação enzimática, bem como de expressão gênica por qPCR e sequenciamento do transcriptoma. Por meio das análises, foi obtido um genoma com 46,748,415 bp, 99.1% de completude e 14,540 genes. Foram identificados 426 genes codificadores de CAZymes e 22 SMGCs. Os resultados dos ensaios bioquímicos e expressão gênica revelaram que o isolado *Trametes villosa* CCMB561 tem a capacidade de crescer em meio de cultura contendo apenas a lignina como fonte de carbono. Nesse substrato foi detectada uma super expressão de genes codificadores de Citocromo P450, que foi associada ao processo de O-desmetilação da lignina. Também foram identificados clusters de genes codificadores de CAZymes *up-regulated* em lignina e bagaço de cana-de-açúcar. Este resultado revelou que *T. villosa* CCMB561 atua na despolimerização simultânea da lignina, celulose, hemicelulose e pectina. Genes codificadores de nitroredutases e homogentisato-1,2-dioxigenase, que atuam na degradação de poluentes orgânicos, foram regulados positivamente no meio contendo apenas a lignina. Os resultados obtidos neste estudo revelaram os mecanismos genéticos de degradação da lignocelulose por *T. villosa* CCMB561 e o potencial desse isolado para a produção de enzimas de interesse biotecnológico e aplicação na conversão da biomassa lignocelulósica e degradação de poluentes orgânicos.

Palavras-chave: *Trametes villosa*¸ genômica, transcriptômica, biomassa lignocelulósica, CAZymes, bioconversão.

ABSTRACT

The bioconversion of lignocellulosic materials, a renewable and sustainable raw material, is an efficient alternative to produce biofuels and obtain chemicals. These residues consist mostly of cellulose, hemicellulose, and lignin. For the use of all the lignocellulosic constituents, both delignification and saccharification of biomass are needed for the removal/modification of lignin and subsequent depolymerization of polymeric carbohydrates to simple sugars. White-rot fungi are organisms that, by the production of hydrolytic and oxidative enzymes, have the ability to degrade all the polymers of the plant cell wall. Previous studies using *Trametes villosa* CCMB561 demonstrated the potential of this isolate to produce ligninolytic enzymes, and thus act in the depolymerization of plant residues. Hence, in order to reveal the genetic bases of lignocellulose degradation by the fungus *Trametes villosa* CCMB561, its genome was sequenced using two sequencing platforms (MinION and Illumina) and assembled using different software. After the assembly, the genome had coding regions, tRNAs, transposable elements, Secondary Metabolite Gene Clusters (SMGCs) and Carbohydrate-Active Enzymes (CAZymes) encoding genes annotated. In addition, biochemical assays for enzymatic quantification, as well as, gene expression analyses by qPCR and transcriptome sequencing were carried out. Through the analyses, a genome with 46,748,415 bp, 99.1% of completeness, and 14,540 genes were obtained. A total of 426 genes encoding CAZymes and 22 SMGCs were identified. The results of biochemical assays and gene expression analyses revealed that *T. villosa* CCMB561 has the ability to grow in culture medium with lignin as the sole carbon source. In this substrate, an overexpression of Cytochrome P450 was detected, which was associated with the initial process of O-demethylation of lignin for subsequent depolymerization by oxidases and peroxidases. Clusters of up-regulated CAZymes-encoding genes were identified in lignin and sugarcane bagasse, revealing that *T. villosa* CCMB561 acts in the simultaneous depolymerization of lignin, cellulose, hemicellulose, and pectin. Furthermore, genes encoding nitroreductases and homogentisate-1,2-dioxygenase that act in the degradation of organic pollutants were up-regulated in the lignin medium. Finally, these findings revealed the genetic mechanisms of lignocellulose breakdown by *Trametes villosa* CCMB561 and exhibits the great potential of this fungal isolate to be used for the production of enzymes with biotechnological interest and application in the conversion of lignocellulosic biomass and degradation of organic pollutants.

Keywords: *Trametes villosa*, genomics, transcriptomics, lignocellulosic biomass, CAZymes, bioconversion.

LISTA DE ABREVIATURAS E SIGLAS

±	Mais ou menos
%	Por cento
°C	Graus Celsius
Ɛ	Coeficiente de Extinção
µL	Microlitro
AA	Atividade Auxiliar
AAO	Aril-álcool oxidase
ABS	Absorbância
ABTS	2,2′-Azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) diammonium salt
AL	Água Lignina
AlcOx	Álcool oxidase
ALG	Água Lignina Glicose
ALSA	Água Lignina Sulfato de Amônio
ALSM	Água Lignina Sulfato de Mangânes
BG	Bagaço de cana
BLAST	*Basic Local Alignment Search Tool*
CAZymes	*Carbohydrate-Active Enzymes*
CBM	Módulos de ligação a carboidrato
CDHs	Celobiose desidrogenases
cDNA	Ácido Dexorribonucleico complementar
CE	Carboidrato Esterases
CYP450	Citrocromo P450
DNA	Ácido dexorribonucleico
DyPs	*Dye-decolorizing peroxidases*
FAD	Flavina-Adenina Dinucleotídeo
GFF	*General Feature Format*
GH	Glicosídeo Hidrolases
GOX	Glicose 1-oxidase
GT	Glicosil Transferases
H_2O_2	Peróxido de hidrogênio
JGI	*Joint Genome Institute*
Lac	Lacase

LiP	Lignina Peroxidase
LPMOs	Monoxigenases Polissacarídicas Líticas Dependentes de Cobre
LTR	Repetições Terminais Longas
Mb	Megabases
MEA	Extrato de Malte
ML	Máxima Verossimilhança
Mn	Manganês
MnP	Manganês Peroxidase
$MnSO_4$	Sulfato de Manganês
mRNA	Ácido Ribonucleico mensageiro
ncRNA	Ácido Ribonucleico não codificante
NCBI	National Center for Biotechnology Information
NGS	Sequenciamento de Nova-geração
nm	Nanômetros
PAH's	Hidrocarbonetos Aromáticos Policíclicos
pb	Pares de base
PL	Polissacarídeo Liases
POD	Peroxidases de classe II
PTMs	Modificações pós-traducionais
qPCR	*Real time quantitative* PCR
RBBR	*Remazol Brilliant Blue R*
RNA	Ácido ribonucleico
rRNA	Ácido ribonucleico ribossomal
TE	Elementos Transponíveis
tRNA	Ácido Ribonucleico Transportador
U/L	Unidade por Litro
VP	Versátil Peroxidase

SUMÁRIO

1. FUNDAMENTAÇÃO TEÓRICA

1.1 Fungos

Os fungos são seres eucarióticos microscópicos e/ou macroscópicos que estão presentes em uma grande variedade de nichos naturais e artificiais sob a forma micelial (fungos filamentosos e multicelulares) ou leveduriforme (fungos unicelulares). Estes organismos constituem o Reino Fungi, um grupo taxonômico altamente heterogêneo do ponto de vista morfológico, fisiológico e ecológico (ESPÓSITO; AZEVEDO, 2010; RAJA *et al.*, 2017; WALKER; WHITE, 2005).

Em relação a fisiologia, os fungos são classificados como heterotróficos, predominantemente aeróbios e, em algumas condições, anaeróbios facultativos, obtêm nutrientes por absorção e podem utilizar diferentes fontes de carbono para o seu crescimento, como hexoses (glicose e manose), pentoses (xilose e arabinose), dissacarídeos (maltose e celobiose), polissacarídeos (celulose e hemicelulose), hidrocarbonetos aromáticos (lignina), entre outros compostos (ESPÓSITO; AZEVEDO, 2010; WALKER; WHITE, 2005). Algumas espécies de fungos são classificadas como oligotróficas possuindo a capacidade de viver em ambientes com quantidades limitadas de nutrientes, como em regiões polares, solos de ambientes extremos e cavernas (JIANG; CAI; LIU, 2017; WALKER; WHITE, 2005).

Em relação aos papéis ecológicos, os fungos podem ser classificados como sapróbios, simbióticos ou parasitas. Os sapróbios atuam na decomposição da matéria orgânica, como madeira e serapilheira, possuindo papel central no ciclo do carbono e grande importância nos ecossistemas florestais. Os fungos simbióticos estabelecem relações mutualísticas com outros organismos, como as plantas, podendo esta associação ser mutuamente benéfica, como ocorre nas micorrizas. Já os fungos patogênicos/parasitas infectam plantas, animais e, até mesmo, outros fungos (ESPÓSITO; AZEVEDO, 2010; SCHMIDT-DANNERT, 2016).

Estimativas sugerem que o Reino Fungi possui de 2 a 5 milhões de espécies que são classificadas em 12 filos e seis grandes grupos: os sub-reinos Dikarya (que inclui os filos Ascomycota, Basidiomycota e Entorrhizomycota) e Chytridiomyceta (que inclui os filos Chytridiomycota, Monoblepharidomycota e Neocallimastigomycota); os filos Mucoromycota, Zoopagomycota e Blastocladiomycota; e o grupo Opisthosporidia (que inclui os filos Aphelidiomycota, Cryptomycota/Rozellomycota e Microsporidia) (JAMES *et al.*, 2020; LI *et al.*, 2021).

Em torno de 97% das espécies de fungos já descritas pertencem ao sub-reino Dikarya, que é composto pelos filos Ascomycota, Basidiomycota e, mais recentemente, pelo filo Entorrhizomycota (JAMES *et al.*, 2020). Os filos Ascomycota e Basidiomycota representam os maiores filos do reino Fungi, sendo que a grande maioria (aproximadamente 70%) das espécies de fungos descritas até o momento pertencem ao filo Ascomycota (SCHMIDT-DANNERT, 2016). Espécies pertencentes a esses dois filos são objeto de diferentes pesquisas científicas e vem ganhando notoriedade devido às aplicações biotecnológicas e industriais relacionadas a importante função que desempenham na degradação da biomassa lignocelulósica (DÍAZ *et al.*, 2020; FLOUDAS *et al.*, 2012; MÄKINEN *et al.*, 2019; MIYAUCHI *et al.*, 2018; PAËS *et al.*, 2019; QIAN *et al.*, 2019).

1.2. Composição e estrutura da biomassa lignocelulósica

A biomassa lignocelulósica é uma matéria-prima renovável, sustentável, sendo considerada um dos maiores estoques de carbono e energia na superfície terrestre. Esta é composta majoritariamente por celulose (40-50%), hemiceluloses (25-30%) e lignina (15-25%), e, em menor quantidade, por pectina, proteínas, extrativos e cinzas (HOUFANI *et al.*, 2020; KRAH *et al.*, 2018; SRIVASTAVA *et al.*, 2015). A composição da lignocelulose pode variar entre espécies de plantas, tecidos e até mesmo entre plantas da mesma espécie, dependendo do genótipo e condições ambientais.

Dentre os polissacarídeos presentes na biomassa, a celulose é o menos complexo e o mais abundante (VAN DEN BRINK; DE VRIES, 2011). Este é um polímero natural e linear formado por resíduos de D-glicose unidos por ligações β-1,4-glicosídicas. Nas plantas, as moléculas de celulose se ligam paralelamente através de pontes de hidrogênio e arranjam-se em fibrilas que estão conectadas a hemicelulose e a lignina (ESPÓSITO; AZEVEDO, 2010; HOUFANI *et al.*, 2020; VAN DEN BRINK; DE VRIES, 2011).

A hemicelulose representa o segundo polissacarídeo mais abundante da biomassa vegetal, sendo caracterizada quimicamente como heteropolímeros ramificados compostos por pentoses (D-xilose e D-arabinose) e/ou hexoses (D-manose, D-glicose e D-galactose). A hemicelulose é classificada de acordo com o monossacarídeo que constitui a cadeia principal, logo, hemicelulose composta principalmente por resíduos de D-xilose (unidos por ligações β-1,4-glicosídicas) é classificada como xilana, enquanto a composta por monômeros de D-manose (unidos por ligações β-1,4-glicosídicas) é classificada como galactomananas, e a formada por resíduos de D-glicose (unidos por ligações β-1,4-glicosídicas) classificada como xiloglucano.

As hemiceluloses são frequentemente acetiladas e suas ramificações compostas por monômeros de L-arabinose, D-xilose, D-galactose, ácido urônico, ácido ferúlico e fucose. A xilose é o açúcar mais abundante da hemicelulose, contudo, a composição desta última pode variar de acordo com a espécie vegetal e o tecido da planta (ESPÓSITO; AZEVEDO, 2010; HOUFANI *et al.*, 2020; VAN DEN BRINK; DE VRIES, 2011).

A lignina é o terceiro componente da biomassa vegetal, e a segunda maior fonte de carbono na terra, caracterizada como um polímero macromolecular tridimensional, fenólico, amorfo e heterogêneo que confere rigidez, impermeabilidade à água e resistência mecânica e microbiológica aos tecidos vegetais. Este polímero é composto por três álcoois aromáticos do tipo fenilpropanóides: o p-cumarílico, o coniferílico e o sinapílico. Estes álcoois são os precursores para a formação de três tipos de ligninas que são classificadas como, lignina p-hidroxifenil (H), guaiacil (G) e siringil (S) (ESPÓSITO; AZEVEDO, 2010; LUO *et al.*, 2020). As unidades aromáticas da lignina são unidas por ligações carbono-carbono, ligações éter, e ligações éster que fortalecem a estrutura da macromolécula (LUO *et al.*, 2020). Por fim, este é o polímero mais recalcitrante e externo da parede celular vegetal, estando intimamente associado a celulose e hemicelulose por ligações covalentes e pontes de hidrogênio (ESPÓSITO; AZEVEDO, 2010; HOUFANI *et al.*, 2020; LUO *et al.*, 2020; NISHIMURA *et al.*, 2018; ZHANG *et al.*, 2015).

A biomassa lignocelulósica possui composição altamente variável e heterogênea, o que faz desta uma fonte potencial para obtenção de diferentes compostos químicos. Através do fracionamento polimérico e conversão da lignina, hemicelulose e celulose, é possível obter uma variedade de produtos comercializáveis, como bioetanol, vanilina, xileno, ácido vanílico, entre outros (BILAL *et al.*, 2018; ISIKGOR; BECER, 2015; WANG, F. *et al.*, 2021).

Além da obtenção de produtos com alto valor agregado, a utilização da biomassa lignocelulósica contribui para o fomento da bioeconomia, economia circular e sustentabilidade, uma vez que a biomassa inclui resíduos agrícolas, agroindustriais e resíduos sólidos municipais que podem ser utilizados como fonte para obtenção de energias renováveis e produtos de base biológica (CARUS; DAMMER, 2018; SHERWOOD, 2020).

Apesar de todo o caráter sustentável e aplicações econômicas, o aproveitamento eficiente da biomassa lignocelulósica ainda é um desafio, devido as suas características estruturais e recalcitrância. O aproveitamento completo da lignocelulose implica na remoção da lignina (deslignificação), para consecutiva despolimerização da hemicelulose e celulose à açúcares fermentescíveis (HOUFANI *et al.*, 2020). Para esse fim, a biomassa precisa ser submetida a um pré-tratamento físico (trituração e moagem), químico (tratamento alcalino ou com ácidos), físico-químico (processamento hidrotérmico) e/ou biológico (enzimas produzidas por microrganismos) (BRETHAUER; STUDER, 2015; JÖNSSON; MARTÍN, 2016). Com relação ao tratamento biológico, diversos estudos têm investigado a utilização de enzimas fúngicas para a bioconversão de materiais lignocelulósicos em produtos químicos (DING; WANG; LI, 2019; ISIKGOR; BECER, 2015; SUKUMARAN *et al.*, 2009; WANG, F. *et al.*, 2021).

1.3. Fungos decompositores da madeira

Os fungos decompositores da madeira têm se destacado como uma excelente alternativa para aplicação no pré-tratamento e bioconversão de materiais lignocelulósicos devido à capacidade de produzirem enzimas lignocelulolíticas que atuam na despolimerização de todos os constituintes da parede celular vegetal. Estes organismos são denominados em relação ao tipo de decomposição que exercem, a saber: branca, marrom e macia, de acordo com a preferência pelos substratos e padrão de decomposição (ESPÓSITO; AZEVEDO, 2010; SISTA KAMESHWAR; QIN, 2018a).

Os fungos da decomposição branca (*white-rot fungi*) possuem a capacidade de degradar todos os constituintes da madeira (lignina, hemicelulose e celulose) de forma simultânea (ex.: *Phanerochaete chrysosporium* e *Irpex lacteus*) ou seletiva (ex.: *Ceriporiopsis subvermispora*, *Physisporinus rivulosus* e *Phlebia* spp.), com algumas espécies apresentando ambos os padrões de degradação (ex.: *Heterobasidium annosum*) (ANDRIANI *et al.*, 2020; ESPÓSITO; AZEVEDO, 2010; SISTA KAMESHWAR; QIN, 2018a). Os fungos desse grupo são considerados os mais efetivos na degradação de materiais lignocelulósicos, principalmente devido a capacidade de despolimerizar eficientemente a lignina, facilitando o acesso das enzimas responsáveis pela degradação da hemicelulose e celulose (ESPÓSITO; AZEVEDO, 2010).

Os fungos da decomposição marrom (*brow-rot fungi*) degradam preferencialmente, e em maior velocidade, a celulose e a hemicelulose, porém, acessam os polissacarídeos modificando a lignina em alguma extensão (ESPÓSITO; AZEVEDO, 2010; SISTA KAMESHWAR; QIN, 2018a).

Os Ascomicetos e os fungos mitospóricos são classificados majoritariamente como causadores da decomposição macia/branda (*soft-rot fungi*) e estes podem degradar/modificar os polímeros da parede celular vegetal, porém em uma velocidade baixa (ESPÓSITO; AZEVEDO, 2010; SISTA KAMESHWAR; QIN, 2018a).

1.4. Fungos do gênero *Trametes*

Entre os Basidiomicetos, os fungos pertencentes a classe Agaricomycetes e ao gênero *Trametes* (ordem Polyporales) têm demonstrado alto potencial para utilização na indústria biotecnológica devido a sua habilidade de produzir enzimas hidrolíticas e oxidativas que atuam eficientemente na degradação da lignocelulose (ANDRIANI *et al.*, 2020; ATILANO-CAMINO *et al.*, 2020; KNEŽEVIĆ *et al.*, 2016; LIU, Y. *et al.*, 2019; MÄKINEN *et al.*, 2019; PAËS *et al.*, 2019; VASINA *et al.*, 2017).

Em relação a fisiologia e morfologia, os fungos do gênero *Trametes* são classificados como fungos da decomposição branca, e são caracterizados morfologicamente pela formação de basidiomas pileados com himenófero poróide, sistema hifal trimítico e produção de basidiósporos lisos e de parede fina que não reagem na presença do reagente de Melzer (JUSTO; HIBBETT, 2011).

Trametes hirsuta, Trametes trogii e Trametes versicolor estão entre as espécies mais estudadas do gênero, e diversas pesquisas científicas vêm demonstrando a habilidade destes fungos em produzir de forma simultânea e sequencial enzimas ligninolíticas, hemicelulolíticas e celulolíticas (ANDRIANI *et al.*, 2020; GAI *et al.*, 2014; LIU, Y. *et al.*, 2019; MACHADO *et al.*, 2020; TIŠMA *et al.*, 2021). Também já foi reportado que espécies desse gênero possuem a habilidade de crescer em diferentes substratos lignocelulósicos e, ao mesmo tempo, atuar na deslignificação e sacarificação (AKYOL *et al.*, 2019; ANDRIANI *et al.*, 2020; JEYA *et al.*, 2009; TIŠMA *et al.*, 2021; VASINA *et al.*, 2017).

A espécie *Trametes villosa* é pouco estuda quando comparada a outras espécies do gênero, contudo, o potencial desse fungo de produzir lacases, peroxidases e celulases em substratos lignocelulósicos já foi demonstrado (CARNEIRO; LOPES; SILVA; SANTOS; SOUZA; *et al.*, 2017; CONIGLIO *et al.*, 2019, 2020; SILVA *et al.*, 2014; YAMANAKA *et al.*, 2008). Em relação ao isolado CCMB561, pesquisas anteriores já testaram a habilidade deste de produzir Manganês Peroxidase (MnP) e Lignina Peroxidase (LiP) em bagaço de cana. CARNEIRO e colaboradores (2017) utilizando meios de cultura contendo bagaço, detectaram a atividade de LiP e MnP nos dias 7, 14 e 21. Em um outro estudo, a enzima MnP foi produzida em meio de cultura contendo apenas o bagaço de cana, em um sistema de fermentação em estado sólido, e esta foi aplicada na deslignificação de fibras de sisal, casca de coco e bagaço de cana (SILVA et al., 2014).

1.5. Enzimas oxidativas e hidrolíticas produzidas por fungos que atuam na despolimerização da lignocelulose

Diversos estudos vêm sendo realizados com o intuito de contribuir para o entendimento dos complexos mecanismos de despolimerização da biomassa lignocelulósica por fungos, sendo que dois sistemas principais são descritos. O primeiro está relacionado com a produção de enzimas oxidativas/ligninolíticas que atuam na despolimerização da lignina, e o segundo com a produção de enzimas hidrolíticas que agem de forma específica na degradação da celulose e hemicelulose (ESPÓSITO; AZEVEDO, 2010; LEVASSEUR *et al.*, 2013).

As enzimas que atuam na desconstrução da parede celular vegetal são divididas em diferentes famílias, classificadas como enzimas ativas de carboidrato (*Carbohydrate-Active enzymes*). Estas compõem o banco de dados CAZy (http://www.cazy.org/), no qual estão todas as enzimas que possuem atividade na clivagem, biossíntese ou modificação dos carboidratos e que são classificadas em Glicosídeo Hidrolases (GH), Carboidrato Esterases (CE), Glicosil Transferases (GT), Polissacarídeo Liases (PL) e Módulos de ligação a carboidrato (CBM) (CANTAREL *et al.*, 2009; LOMBARD *et al.*, 2014). Neste banco de dados também estão incluídas as enzimas que atuam diretamente na clivagem da lignina, classificadas como enzimas de "Atividade Auxiliar" (AA), uma vez que auxiliam as enzimas GH, CE e PL a obterem acesso aos carboidratos (LEVASSEUR *et al.*, 2013; WONG, 2009).

1.5.1. Enzimas oxidativas ligninolíticas

As principais enzimas envolvidas na despolimerização da lignina fazem parte das famílias AA1 a AA5. Na família AA1 estão as Lacases (Lac, 1.10.3.2), oxirredutases multicobre (fenoloxidases) que atuam diretamente sobre estruturas fenólicas e compostos aromáticos, como a lignina. As lacases catalisam reações de oxidação através da extração de um elétron e concomitante redução do oxigênio em água. Esta classe de enzimas atua em uma ampla variedade de substratos sem a necessidade de cofatores, por isso, há um grande interesse nestas para utilização industrial (DASHTBAN *et al.*, 2010; ESPÓSITO; AZEVEDO, 2010; LEVASSEUR *et al.*, 2013; WONG, 2009).

Enzimas da família AA2 são classificadas como heme peroxidases (PODs) de classe II, e são dependentes de peróxido de hidrogênio (H_2O_2) como cofator para ativação e degradação da lignina. As enzimas Lignina Peroxidase (LiP, EC 1.11.1.14), Manganês Peroxidase (MnP, EC 1.11.1.13) e Versátil Peroxidase (VP, EC 1.11.1.16) compõe a família AA2 (LEVASSEUR *et al.*, 2013).

As LiPs são relativamente inespecíficas, oxidam substratos aromáticos fenólicos e atuam clivando ligações C-C e ligações éter de substratos de modelo de lignina não-fenólica com alto potencial redox. MnPs são enzimas dependentes de Mn (II) e atuam oxidando o Mn^{+2} para Mn^{+3}, que por sua vez, oxida uma variedade de compostos fenólicos. VPs são enzimas híbridas, que combinam propriedades catalíticas de LiP e MnP, sendo, desta forma, capazes de oxidar substratos fenólicos, não-fenólicos e Mn^{+2} (DASHTBAN *et al.*, 2010; LEVASSEUR *et al.*, 2013).

As enzimas pertencentes a classe AA2 apresentam atividade em uma ampla variedade de substratos e podem ser aplicadas nas indústrias de alimentos (produção de vanilina), papel e celulose (branqueamento de polpa de kraft), têxtil (degradação de tintas têxteis e branqueamento), biocombustíveis (pré-tratamento de biomassa lignocelulósica) e também na biorremediação de hidrocarbonetos aromáticos policíclicos (PAH's) e outros xenobióticos (DASHTBAN *et al.*, 2010; ESPÓSITO; AZEVEDO, 2010; LEVASSEUR *et al.*, 2013; WONG, 2009).

Além das enzimas ligninolíticas, outras enzimas extracelulares produzidas por fungos também são necessárias no processo de despolimerização da lignina, como aquelas que atuam produzindo o peróxido de hidrogênio (H_2O_2), requerido para a atividade das peroxidases de classe II (PODs). Na família AA3 estão as Glicose-Metanol-Colina (GMC) oxidorredutases, que são flavoproteínas contendo um domínio de ligação a Flavina-Adenina Dinucleotídeo (FAD). Entre as principais enzimas que produzem H_2O_2 nessa família, estão: Piranose 2-oxidase (P2O, EC 1.1.3.10), Álcool oxidase (AlcOx, EC 1.1.3.13), Glicose 1-oxidase (GOX, EC 1.1.3.4) e Aril-álcool oxidase (AAO, EC 1.1.3.7) (DASHTBAN *et al.*, 2010; LEVASSEUR *et al.*, 2013).

De forma complementar ao processo enzimático, fungos da podridão branca e marrom, por meio da produção de H_2O_2 e da reação de Fenton, podem gerar radicais hidroxila livres que atacam a lignina e os polissacarídeos da biomassa por meio de uma série de clivagens não-específicas que viabilizam a penetração das enzimas lignocelulolíticas na parede celular vegetal (DASHTBAN *et al.*, 2010). Neste processo de geração de radicais hidroxila, também estão incluídas as Celobiose desidrogenases (CDHs), que fazem parte da família AA3. CDHs agem gerando radicais hidroxila por meio de reações de Fenton, que, por sua vez, irão oxidar a lignina de forma não-enzimática (DASHTBAN *et al.*, 2010; LEVASSEUR *et al.*, 2013).

Na família AA5 estão incluídas as enzimas Glioxal oxidase (GLOX, EC 1.2.3.15) e Galactose oxidase (GAO, EC 1.1.3.9) que possuem papel central na geração de H_2O_2

(DASHTBAN *et al.*, 2010; ESPÓSITO; AZEVEDO, 2010; LEVASSEUR *et al.*, 2013; WONG, 2009).

1.5.2. Enzimas hidrolíticas

As enzimas hidrolíticas constituem uma grande classe de enzimas que atuam catalisando reações de hidrólise em diferentes polímeros. Dentre estas, estão as enzimas que atuam na despolimerização da hemicelulose, celulose e pectina, que estão distribuídas nas famílias Glicosídeo Hidrolase (GH), Carboidrato Esterase (CE) e Polissacarídeo Liase (PL). Como a composição da lignocelulose é altamente variável, diferentes enzimas são produzidas pelos fungos para despolimerizar os carboidratos da parede celular vegetal (HOUFANI *et al.*, 2020; VAN DEN BRINK; DE VRIES, 2011) .

A degradação da celulose ocorre através da ação sinérgica de três tipos de enzimas, as β-1,4-endoglicanases (famílias GH5_5, GH5_22, GH9, GH12, GH45, GH131), as exoglicanases/celobiohidrolases (famílias GH6 e GH7) e as β-glicosidases (famílias GH1 e GH3) (ROTH; HOELTZ; BENITEZ, 2020). As endoglicanases agem de forma randômica na região interna amorfa da fibra da celulose, gerando oligossacarídeos de diferentes tamanhos, com extremidades redutoras e não-redutoras. As exoglicanases agem nas extremidades redutoras e não-redutoras dos oligossacarídeos e liberam unidades de glicose e/ou celobiose. Por fim, as β-glicosidases agem hidrolisando a ligação entre as duas unidades de glicose da celobiose, gerando monômeros de glicose (ESPÓSITO; AZEVEDO, 2010; HOUFANI *et al.*, 2020; MANAVALAN; MANAVALAN; HEESE, 2015; VAN DEN BRINK; DE VRIES, 2011). Além da ação das enzimas hidrolíticas, a celulose também é degradada de forma oxidativa através de clivagens catalisadas por enzimas da família AA9 (anteriormente GH61), que são classificadas como mono-oxigenases polissacarídicas líticas dependentes de cobre (LPMOs) (LEVASSEUR *et al.*, 2013).

A hidrolise da hemicelulose é realizada por um grupo de enzimas que age na estrutura principal e nas cadeias laterais deste polissacarídeo. β-1,4-endoxilanases e β-1,4-xilosidases agem na hidrólise das xilanas, β-1,4-endoglicanases e β-1,4-glicosidases na despolimerização do xiloglucano, enquanto β-1,4-endomananase e β-1,4-manosidase na hidrólise das mananas. Além das enzimas que agem na estrutura principal das hemiceluloses, também há enzimas que atuam despolimerizando as cadeias laterais, tais como: α-arabinofuranosidase, α-fucosidase, α-1,4-galactosidase, α-glicuronidase, acetilesterases, α-xylosidase e β-1,4-galactosidase (HOUFANI *et al.*, 2020; VAN DEN BRINK; DE VRIES, 2011). Dentre as enzimas acessórias, as acetilesterases possuem papel crucial no processo de despolimerização da hemicelulose, uma vez que estas irão atuar na desacetilação das unidades de xilana, permitindo a ação das hidrolases (BIELY *et al.*, 2014; SISTA KAMESHWAR; QIN, 2018b).

Em relação as aplicações biotecnológicas, as celulases e hemicelulases podem ser utilizadas em diferentes setores da economia, como na indústria de alimentos, papel e celulose, têxtil e farmacêuticas (HOUFANI *et al.*, 2020). Além dessas aplicações, nos últimos anos, tem aumentado o interesse na aplicação dessas enzimas para hidrólise (pré-tratamento) da biomassa lignocelulósica para obtenção de bioetanol (fermentação dos açúcares simples) e biogás (digestão anaeróbica) (BUNDHOO; MUDHOO; MOHEE, 2013).

1.6. Utilização de ferramentas ômicas na compreensão do metabolismo da lignocelulose

A bioconversão de resíduos florestais, agrícolas e industriais para a obtenção de compostos químicos e energias renováveis é altamente desejável. Contudo, inovações tecnológicas viáveis para a conversão da biomassa ainda precisam ser alcançadas e implementadas para que o desenvolvimento de biorrefinarias, "química verde" e economia circular sejam impulsionados.

Nos últimos anos, com o desenvolvimento de novas tecnologias de sequenciamento de DNA e RNA, e ferramentas de bioinformática, a biotecnologia e biologia molecular tem sofrido evoluções substanciais, e novas descobertas sobre a biologia de diferentes organismos têm sido alcançadas. Por meio da utilização do sequenciamento de nova-geração (NGS) e da aplicação integrada das ferramentas "ômicas", incluindo genômica, transcriptômica e proteômica é possível compreender de forma global e, em larga escala, os processos celulares e vias metabólicas dos organismos (WILKEN *et al.*, 2019).

A genômica tem como principal objetivo investigar a estrutura, função e evolução dos genomas (DE CARVALHO *et al.*, 2019). Na interface micologia e biotecnologia, a genômica tem sido aplicada para identificar genes de resposta ao estresse, genes codificadores de enzimas com aplicação biotecnológica, realizar predição de proteínas, identificar regiões codificantes e não-codificantes e identificar vias metabólicas (SHARMA, 2016; WILKEN *et al.*, 2019). O primeiro fungo a ter seu genoma sequenciado foi o *Saccharomyces cerevisiae* S288C, o que representou um marco histórico na Biologia e Genômica de fungos (GOFFEAU *et al.*, 1996). Desde então, iniciativas como o projeto "*1000 fungal genomes*" tem contribuído para o sequenciamento de mais genomas de fungos, de forma a fornecer informações precisas sobre a biologia desse grupo de organismos.

No que se refere a degradação da lignocelulose, muitos estudos têm focado no sequenciamento de genomas de fungos da decomposição branca de forma a explorar toda a maquinaria metabólica responsável pela degradação da parede celular vegetal, principalmente a lignina. O primeiro genoma de um basidiomiceto a ser sequenciado foi o de *Phanerochaete chrysosporium* RP78 (MARTINEZ *et al.*, 2004). As análises genômicas desse fungo permitiram identificar genes e famílias multigênicas que estão diretamente relacionadas com a degradação da lignina e dos carboidratos da parede celular vegetal. Além disso, foi possível identificar retrotransposons que afetam a expressão de genes que atuam na degradação da lignina (MARTINEZ *et al.*, 2004).

Apesar da maioria dos estudos de genômica serem focados em apenas um organismo, estes também podem ser realizados utilizando a informação de vários genomas de forma integrada, com o objetivo de entender as relações evolutivas entre diferentes espécies e/ou entre organismos da mesma espécie. A genômica comparativa é uma ferramenta que vem sendo amplamente difundida nos últimos anos, e que pode auxiliar na identificação de genes conservados, expansão de famílias gênicas, variações no ordenamento dos genes, identificação de duplicações e deleções gênicas, dentre outros (DE CARVALHO *et al.*, 2019; FLOUDAS *et al.*, 2012; MIYAUCHI *et al.*, 2018). FLOUDAS et al., 2012 utilizaram a genômica comparativa para compreender os mecanismos evolutivos de degradação da lignina. Utilizando 31 genomas de fungos de diferentes filos, os autores identificaram que fungos da decomposição branca apresentam expansão no número de cópias de genes codificadores de peroxidases, enquanto fungos da decomposição marrom tiveram uma contração dessa família gênica.

O sequenciamento de genomas também vem sendo utilizado para inferir relações filogenéticas em escala genômica (filogenômica). Em geral, são utilizadas matrizes contendo milhares de genes ortólogos de cópia única (DELSUC; BRINKMANN; PHILIPPE, 2005) ou proteínas que possibilitam avaliar as relações evolutivas dos organismos a partir de seus genomas, e, assim, testar sinais filogenéticos com maior acurácia (STAJICH, 2017; WANG, Z. *et al.*, 2018).

Nas ciências ômicas, outra ferramenta que também tem sido aplicada para o estudo de genes, funções gênicas e vias metabólicas, é a transcriptômica. Através do sequenciamento em larga escala do RNA é possível explorar o perfil transcricional dos genes, a expressão diferencial e criar redes de expressão e co-regulação gênica (WILKEN *et al.*, 2019). O sequenciamento do transcriptoma (RNA-Seq) pode incluir tanto os RNAs mensageiros (mRNA), quanto os RNAs não-codificantes (ncRNAs) (DE CARVALHO *et al.*, 2019). No que concerne ao metabolismo da lignocelulose, pesquisadores já sequenciaram o transcriptoma de diferentes fungos degradadores da madeira para identificar os genes que são diferencialmente expressos em substratos lignocelulósicos. Por exemplo, HAGE et al., 2021 integraram as informações de 50 genomas de fungos pertencentes a ordem Polyporales e buscaram por adaptações genômicas e funcionais relacionadas a decomposição da madeira através da análise de composição dos genomas e do perfil de transcrição em diferentes fontes de carbono.

Nas pesquisas de bioconversão de resíduos lignocelulósicos e produção de biocombustíveis, o sequenciamento do transcriptoma pode ajudar na descoberta de novos genes que agem na despolimerização da biomassa, identificar RNAs (microRNAs e ncRNA) que atuam na regulação da expressão gênica, genes de resposta ao estresse (biótico e abiótico) e genes que estão sendo diferencialmente expressos (DE CARVALHO *et al.*, 2019; NIKOLAIVITS *et al.*, 2021).

Por fim, a proteômica é uma ferramenta que tem como objetivo identificar e quantificar em larga escala todas as proteínas que são traduzidas em um organismo e suas interações (TYERS; MANN, 2003). Com os avanços nas técnicas de separação de proteínas e espectrometria de massas, atualmente é possível analisar milhares de proteínas ao mesmo tempo, identificar isoformas, obter as características estruturais e funcionais, identificar modificações pós-traducionais (PTMs) e estudar interações entre proteínas (LIU, X. *et al.*, 2012; TYERS; MANN, 2003). Na biotecnologia, a proteômica tem sido utilizada para identificar proteínas e enzimas que são secretadas para o meio extracelular, compreender como um organismo responde a diferentes fontes de carbono, identificar novas proteínas, entender a resposta a fatores bióticos e abióticos, entre várias outras aplicações (DE CARVALHO *et al.*, 2019; PAËS *et al.*, 2019). De forma geral, o que se pretende com os estudos de proteômica na biotecnologia é compreender processos celulares e metabólicos que possam contribuir para a otimização de processos industriais.

2. JUSTIFICATIVA E RELEVÂNCIA

A matriz energética global consiste majoritariamente na utilização de combustíveis fósseis, como petróleo, carvão e gás natural. Essas fontes de carbono, além de serem não-renováveis, contribuem para a emissão de gases do efeito estufa e causam grande impacto ambiental, portanto, existe a necessidade crescente em expandir o uso de fontes de energia renovável. A biomassa lignocelulósica, que inclui resíduos agrícolas, industriais e sólidos urbanos, ganha destaque como uma promissora fonte energética sustentável, uma vez que é abundante, amplamente disponível e pode ser utilizada como matéria-prima para a obtenção de biocombustíveis e uma ampla diversidade de produtos químicos.

Embora os materiais lignocelulósicos sejam baratos e abundantes, um grande desafio para seu uso é a degradação da lignina, devido a sua característica recalcitrante, e a necessidade consecutiva de despolimerização da celulose e hemicelulose à açúcares fermentescíveis. Neste cenário, os fungos da decomposição branca surgem como uma alternativa eficiente para o pré-tratamento biológico destes materiais, uma vez que estes organismos possuem a maquinaria bioquímica necessária para a quebra da lignina e despolimerização dos carboidratos da parede celular vegetal. Dentre os fungos decompositores da madeira, o isolado *Trametes vilosa* CCMB561 vem se destacando, uma vez que estudos anteriores aplicando testes bioquímicos e análises de expressão gênica por qPCR, demonstraram o potencial desse isolado em produzir manganês peroxidase (MnP) e lignina peroxidase (LiP) em meio de cultura contendo substratos lignocelulósicos como fonte de carbono.

Estudos recentes já demonstraram que a biologia de sistemas é uma estratégia eficiente para a compreender de forma detalhada os genes, vias metabólicas e as interações dos componentes biológicos de um organismo. Contudo, até o momento, nenhum trabalho utilizou esta estratégia para avaliar os mecanismos de degradação da lignina, celulose, hemicelulose e pectina pelo fungo *T. vilosa*. Portanto, por meio da realização deste trabalho, analisamos de forma integrada a informação genética contida no DNA e aquela transcrita em RNA utilizando o isolado *T. villosa* CCMB561. Este estudo visou contribuir com conhecimentos relacionados à identificação, regulação e expressão de genes que atuam na degradação da biomassa lignocelulósica. Para desta forma, viabilizar o avanço do uso de tecnologias limpas baseado na utilização de resíduos agroindustriais.

3. OBJETIVOS

3.1 Objetivo Geral

Compreender, de forma integrada, por meio da abordagem genômica e transcriptômica os mecanismos de degradação da lignocelulose pelo fungo *Trametes villosa* CCMB561 de forma a ampliar a utilização de resíduos lignocelulósicos.

3.2 Objetivos Específicos

1. Sequenciar o genoma do isolado *T. villosa* CCMB561 utilizando as plataformas de sequenciamento de nova geração (NGS) Illumina HiSeq (*short-reads*) e Oxford Nanopore MinION (*long-reads*);
2. Testar diferentes softwares e pipelines de montagem de genomas;
3. Anotar os genes e realizar a predição de transcritos e proteínas;
4. Realizar a anotação de elementos transponíveis (TEs);
5. Realizar análise de enriquecimento e ontologia gênica (GO);
6. Realizar a predição de cluster de genes responsáveis pela produção de metabólitos secundários (SMGCs);
7. Anotar os genes codificadores de *Carbohydrate-Active Enzymes* (CAZymes);
8. Realizar análises de genômica comparativa de fungos do gênero *Trametes*;
9. Avaliar o crescimento do isolado *T. villosa* CCMB561 quando submetido ao crescimento em meios de cultura contendo lignina como principal fonte de carbono;
10. Realizar testes bioquímicos para avaliar a produção das enzimas Lacase (Lac) e Mangânes Peroxidase (MnP) quando o isolado *T. villosa* CCMB561 é submetido ao crescimento em fermentação submersa utilizando meios indutores e não indutores para a produção de enzimas ligninolíticas;
11. Desenhar primers para a quantificação relativa dos genes codificadores da Lacase (Lac), Lignina Peroxidase (LiP) e Manganês Peroxidase (MnP) por qPCR;
12. Sequenciar o transcriptoma de *T. villosa* CCMB561 cultivado em diferentes fontes de carbono para avaliar a expressão diferencial e redes de co-expressão dos genes envolvidos na degradação da lignocelulose.

4. CAPÍTULOS

4.1 CAPÍTULO I: Hybrid Assembly Improves Genome Quality and Completeness of *Trametes villosa* CCMB561 and Reveals a Huge Potential for Lignocellulose Breakdown.

Este capítulo foi publicado na revista ***Journal of Fungi*** 2022
DOI: https://doi.org/10.3390/jof8020142

Journal of
Fungi

MDPI

Article

Hybrid Assembly Improves Genome Quality and Completeness of *Trametes villosa* CCMB561 and Reveals a Huge Potential for Lignocellulose Breakdown

Luiz Marcelo Ribeiro Tomé [1], Felipe Ferreira da Silva [1], Paula Luize Camargos Fonseca [2], Thairine Mendes-Pereira [1], Vasco Ariston de Carvalho Azevedo [3], Bertram Brenig [4], Fernanda Badotti [5] and Aristóteles Góes-Neto [1,*]

[1] Molecular and Computational Biology of Fungi Laboratory, Department of Microbiology, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte 31270-901, MG, Brazil; marcelofsa_rt@hotmail.com (L.M.R.T.); felselva@gmail.com (F.F.d.S.); thairinemp@gmail.com (T.M.-P.)
[2] Departamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte 31270-901, MG, Brazil; camargos.paulaluize@gmail.com
[3] Laboratório de Genética Celular e Molecular, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte 31270-901, MG, Brazil; vascoariston@gmail.com
[4] Institute of Veterinary Medicine, Burckhardtweg, University of Göttingen, 37073 Göttingen, Germany; bbrenig@gwdg.de
[5] Department of Chemistry, Centro Federal de Educação Tecnológica de Minas Gerais, Belo Horizonte 30421-169, MG, Brazil; fbadotti@outlook.com
* Correspondence: arigoesneto@icb.ufmg.br; Tel.: +55-31-994130996

Abstract: *Trametes villosa* is a wood-decaying fungus with great potential to be used in the bioconversion of agro-industrial residues and to obtain high-value-added products, such as biofuels. Nonetheless, the lack of high-quality genomic data hampers studies investigating genetic mechanisms and metabolic pathways in *T. villosa*, hindering its application in industry. Herein, applying a hybrid assembly pipeline using short reads (Illumina HiSeq) and long reads (Oxford Nanopore MinION), we obtained a high-quality genome for the *T. villosa* CCMB561 and investigated its genetic potential for lignocellulose breakdown. The new genome possesses 143 contigs, N50 of 1,009,271 bp, a total length of 46,748,415 bp, 14,540 protein-coding genes, 22 secondary metabolite gene clusters, and 426 genes encoding Carbohydrate-Active enzymes. Our CAZome annotation and comparative genomic analyses of nine *Trametes* spp. genomes revealed *T. villosa* CCMB561 as the species with the highest number of genes encoding lignin-modifying enzymes and a wide array of genes encoding proteins for the breakdown of cellulose, hemicellulose, and pectin. These results bring to light the potential of this isolate to be applied in the bioconversion of lignocellulose and will support future studies on the expression, regulation, and evolution of genes, proteins, and metabolic pathways regarding the bioconversion of lignocellulosic residues.

Keywords: *Trametes villosa* CCMB561; genome assembly; comparative genomics; lignocellulosic biomass; CAZymes

Citation: Tomé, L.M.R.; da Silva, F.F.; Fonseca, P.L.C.; Mendes-Pereira, T.; Azevedo, V.A.d.C.; Brenig, B.; Badotti, F.; Góes-Neto, A. Hybrid Assembly Improves Genome Quality and Completeness of *Trametes villosa* CCMB561 and Reveals a Huge Potential for Lignocellulose Breakdown. *J. Fungi* **2022**, *8*, 142. https://doi.org/10.3390/jof8020142

Academic Editor: Philippe Silar

Received: 17 December 2021
Accepted: 27 January 2022
Published: 30 January 2022

1. Introduction

Lignocellulosic biomass (LB), including agro-industrial residues, municipal solid wastes, and forest litter is one of the largest stocks of carbon and energy on Earth [1]. These are a sustainable, renewable, and abundant source of raw material, which can be (bio-) converted into high value-added products, such as bio-based chemicals, polymers, and fuels [1,2]. LB is composed mainly of cellulose (40–50%), hemicellulose (25–30%), and lignin (15–25%), as well as other compounds in lower proportions, such as pectin, proteins, extractables, and ash [3]. Although lignocellulosic materials are inexpensive, abundant, and considered a valuable feedstock for biorefineries (bio-based industry), a major challenge for their use is the degradation of lignin, which is a highly recalcitrant polymer, and the consecutive depolymerization of cellulose and hemicellulose (polysaccharides) to fermentable sugars (oligosaccharides and monosaccharides) [3–5]. Therefore, biomass needs to undergo physical, chemical, and/or biological pretreatment [4].

White-rot fungi (WRF) act efficiently in the biodegradation of plant biomass, possessing the metabolic machinery for the breakdown of all plant cell wall polymers (PCW) [6–12]. The depolymerization of the lignocellulose by WRF is carried out through the production and secretion of hydrolytic and oxidative enzymes belonging to glycoside hydrolases (GH), carbohydrate esterases (CE), pectate lyases (PL), and auxiliary activity oxidoreductases (AA) families, according to the classification in the Carbohydrate Active Enzymes database (CAZymes database) [13–15]. In addition to the PCW breakdown, WRF enzymes can also be applied in different industrial sectors, such as food, pulp and paper, textile, pharmaceutical, and biofuel, and be used in the bioremediation of polycyclic aromatic hydrocarbons (PAHs) and other xenobiotics [3–5,13,16,17].

Fungi of the genus *Trametes* (Polyporales, Basidiomycota) are classified as white-rot fungi and have the ability to simultaneously degrade all polymers of the lignocellulose [6–8,18,19]. In nature, they act as one of the major decomposers of wood and plant leaf litter, and play a central role in the carbon cycle, soil formation, and, consequently, in the maintenance of forest ecosystems [13,20,21]. To date, the genomes of the species *T. versicolor*, *T. coccinea*, *T. polyzona*, *T. hirsuta*, *T. cinnabarina*, *T. sanguinea*, and *T. pubescens*, among others, have already been sequenced, and research related to these genomes, as well as studies of transcriptomics and proteomics, described a set of mechanisms, genes, and metabolites involved in wood decay. Such knowledge has provided a huge aid toward the improvement in the production and industrial application of CAZymes [6,7,9,22–28].

Trametes villosa is an important species of the genus *Trametes*, and its potential to produce laccases, peroxidases, and cellulases in different lignocellulosic substrates has already been demonstrated at the small-scale biochemical level [29–32]. Regarding the strain *T. villosa* CCMB561, biochemical assays have shown its potential to produce ligninolytic enzymes using sugarcane bagasse as a substrate in different temperatures (from 20 °C to 33 °C) and pHs (from 4.60 to 9.38) [29,32]. Furthermore, Silva et al. (2014) obtained a reduction up to 63% in the lignin content of different agro-industrial wastes (sugarcane bagasse, sisal fiber, and coconut shell) using the enzyme Manganese Peroxidase (MnP) produced by *T. villosa* CCMB561 [32]. Despite the potential of this isolate for biotechnological applications, so far, no genomic study has been carried out to comprehensively understand the genetic repertoire and mechanisms involved in the complex breakdown of all plant cell wall components (lignin, hemicellulose, cellulose, and pectin).

Indeed, the genome of *T. villosa* CCMB561 has previously been sequenced and a public, preliminary draft version is available in the NCBI database (GenBank accession: GCA_002964805.1) [24]. Nevertheless, the draft genome has a high fragmentation rate (10,327 contigs), high duplication of single-copy ortholog genes, and a genome size larger than expected for the genus [24]. The low quality of these data limits downstream analysis, such as the identification, characterization, and understanding of the expression and regulation of genes and proteins.

Therefore, in order to obtain a high-quality genome and then contribute to further studies aimed at understanding the genetic bases of lignocellulose breakdown by the WRF, we have sequenced the genome of the isolate *T. villosa* CCMB561 jointly using second- (HiSeq 2500—Illumina) and third-generation (MinION—Oxford Nanopore) sequencing platforms. Different assembly strategies were tested and are described in this study. Coding regions, transfer RNAs (tRNAs), transposable elements, and CAZymes genes were annotated. Furthermore, comparative and phylogenomic analysis including genomes of other isolates of the genus *Trametes* were performed. Our findings revealed that using a hybrid assembly approach (a combination of short- and long-read sequencing from distinct platforms), it was possible to acquire a genome with much better completeness and contiguity when compared to the draft genome available. The new assembled genome presented 143 contigs, a size of 46.748 Mb, 14,540 proteins-encoding genes, and 22 secondary metabolite gene clusters (SMGCs). In addition, a wide array of genes encoding lignocellulose-modifying enzymes was identified, revealing a huge potential of the isolate *T. villosa* CCMB561 to act in the degradation of all lignocellulose polymers, making it a high-potential strain to be industrially used.

2. Materials and Methods

2.1. Fungal Strain and Extraction of Genomic DNA

The fungal strain *T. villosa* CCMB561 was isolated from field-collected basidiomata growing on a decaying tree branch (unidentified angiosperm) in the semiarid region of northeastern Brazil (Serra das Candeias, Quijingue, Bahia, Brazil; Lat: 39°04′30′′ W and Long: 10°55′16′′ S). Dehydrated basidiomata were deposited in the HUEFS herbarium (HUEFS108280), and the culture derived from the basidiomata tissue was preserved in sterile distilled water and deposited in the Culture Collection of Microorganisms of Bahia (CCMB, Feira de Santana, Bahia, Brazil) under access code CCMB561. The isolate was grown on Malt Extract Agar (2% Malt Extract, 2% dextrose and 2% Agar) at 28 ± 2 °C for seven days. Then, the total DNA was extracted using the ZymoBIOMICS^{TM} DNA Miniprep Kit (Zymo Research, Irvine, CA, USA). The DNA sample was analyzed qualitatively by agarose gel electrophoresis 1%, and quantitatively by a Nanodrop 1000 ND spectrophotometer (Thermo Scientific, Waltham, MA, USA) and Qubit fluorometer (Invitrogen, Waltham, MA, USA). For species-level certification of the extracted DNA, the internal transcribed region (ITS1-5.8S-ITS2) was amplified and sequenced using the ITS 6 (5′-TTCCCGCTTCACTCGCAGT-3′) and ITS 8 (5′-AGTCGTAACAAGGTTTCCGTAGGTG-3′) primers [33]. Amplification reaction, purification, and sequencing of the amplicons were carried out according to the methods described by Tomé et al. 2019 [34].

2.2. MinION Library Preparation and Sequencing

We fragmented genomic DNA [8 µg] to approximately 8 Kbp using the Covaris g-TUBE (Covaris, Woburn, MA, USA). After fragmentation, 1200 ng of DNA was purified using the AMPureXP beads (Beckman Coulter Inc., Brea, CA, USA), not adopting the DNA repair step. The sequencing library was prepared using the Ligation Sequencing Kit 1D (SQK-LSK108), the Native Barcoding Kit 1D (EXP-NBD103), and the Library Loading Bead Kit (EXP-LLB001), following the recommendations of Oxford Nanopore Technologies. The library was sequenced for 48 h in the flowcell FLO-MIN106 (ID: FAK07371) using the MinKNOW program with the real-time base calling function enabled. Porechop software (https://github.com/rrwick/Porechop, accessed on 15 January 2020) was used to demultiplex the libraries and remove the adapters (Figure S1).

2.3. Illumina Library Preparation and Sequencing

The sequencing library was prepared from genomic DNA [1 µg] using the NEBNext Fast DNA Fragmentation and Library Preparation Kit (New England Biolabs, Ipswich, MA, USA) following the manufacturer's recommendations. The library quality was assessed using the Agilent 2100 Bioanalyzer equipment, and the paired-end DNA sequencing was carried out in the Illumina HiSeq 2500 platform. After sequencing, the raw reads quality was assessed using the FastQC v0.11.5 software (https://github.com/s-andrews/FastQC, accessed on 15 January 2020). Adapter sequences and bases with low quality (Phred score <20) were trimmed using BBDuk software (https://sourceforge.net/projects/bbmap/, accessed on 15 January 2020) (Figure S1). Genome features such as size, heterozygosity, and repetitiveness were assessed prior to genome assembly using Jellyfish and GenomeScope 2.0 [35,36].

2.4. De Novo Genome Assembly and Assessment

The genome assembly was carried out using different approaches and software (Figure S1). MinION long-reads were assembled using (i) Flye [37], (ii) Canu [38], (iii) Racon [39], and the (iv) CANU-smartdenovo pipeline with default parameters [40] while Illumina short reads were assembled using the (v) MaSuRCA software with default parameters [41]. Hybrid assemblies using Illumina and MinION reads were performed using the software (vi) MaSuRCA, (vii) SPAdes [42], and the assembly workflow (viii) MaSuRCA-Purge_dups [41,43]. Genome quality and completeness for each assembly were evaluated using QUAST v4.6.0 [44] and BUSCO v4 (Benchmarking Universal Single-Copy Orthologs) [45]. BUSCO analyses were performed using the database basidiomycota_odb10.

2.5. Genome Annotation and Gene Ontology Analyses

Genome annotation was performed using the MAKER2 v2.31.9 software [46–48] and the following ab initio gene prediction software: SNAP [49], Augustus [50], and GeneMark [51] (Figure S2). Low- and high-complexity repetitive genomic regions were masked using the RepeatMasker [52], the Repbase database, and the RepeatRunner software [48]. The identification of gene regions and the prediction of proteins were performed through the alignment of ESTs (Expressed Sequence Tags) and proteins of the genus *Trametes* (obtained from NCBI until September 2020) using the BLAST algorithm and the Exonerate program. After the annotation by evidence, the software SNAP, Augustus, and GeneMark were used in the further annotation steps. The annotation metrics, such as the number of genes, exons, and introns, were obtained using GAG software (Genome Annotation Generator) [53]. The assignment of gene function was carried out with the support of tools provided by MAKER [47], the makeblastdb application, the UniProt database (uniprot_sprot.fasta), and the blastp algorithm. Gene ontology analyses were carried out in the web server GoFeat (Gene Ontology Functional Enrichment Annotation Tool), with the support of the following databases: Uniprot, NCBI protein, KEGG, InterPro, Pfam, EMBL, and Gene Ontology [54]. The prediction of transfer RNA (tRNA) was performed using the software tRNAscan-SE [55]. Secondary metabolite gene clusters (SMGCs) were predicted using the online tool antiSMASH 6.0.1 [56].

2.6. Repeat Annotation

Transposable elements (TE) were identified de novo using the RepeatModeler package (repeatmasker.org/RepeatModeler, accessed on 15 September 2020) with the support of RepeatMasker [52], RECON [57], RepeatScout [58], TRF [59], and RMBlast. The obtained TE library (consensus sequences) was filtered by removing all sequences <100 bp and those showing significant hits with proteins not identified as TE using blastx and the UniProt database [7]. The classification and number of occurrences of TE were assessed using the RepeatMasker tool.

2.7. Comparative Genomics and Phylogenomics

The genome of the fungus *T. villosa* CCMB561 was compared with the following seven genomes publicly available at the National Center for Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.gov/, accessed on 15 March 2020): *T. coccinea* (GCA_002092935.1), *T. sanguinea* (GCA_008973685.1), *T. cinnabarina* (GCA_000765035.1), *T. hirsuta* (GCA_001302255.2), *T. polyzona* (GCA_001939255.1), *T. pubescens* (GCA_001895945.1), and *T. versicolor* (GCF_000271585.1); and two publicly available at the Joint Genome Institute (JGI, https://genome.jgi.doe.gov/portal/, accessed on 15 March 2020): *T. ljubarskyi* (CIRM1659) and *T. elegans* (CIRM1663, synonym: *Artolenzites elegans*). The completeness and the main metrics of the retrieved genomes were assessed using BUSCO and QUAST, respectively. In order to standardize and improve the accuracy of the comparative analyses, all genomes were reannotated, and the transposable elements were identified using the methods described in Sections 2.5 and 2.6, respectively. Phylogenomic analyses were carried out using the script BUSCO_phylogenomics (https://github.com/jamiemcg/BUSCO_phylogenomics, accessed on 15 October 2020), in which single-copy ortholog genes were aligned using MUSCLE. The alignment was trimmed using trimAl [60], and the estimation of the best-fit model was performed using ModelFinder. The maximum likelihood phylogenetic tree was generated using the IQ-TREE software, adopting the supermatrix method [61]. The consensus tree was constructed considering 1000 bootstrap replicates and visualized in FigTree v1.4.3. The species *Polyporus brumalis* (Polyporales, Basidiomycota—GCA_001792895.1) was used in the phylogenomic analyses as the outgroup. Network and correlation analyses were performed based on the genome length, number of genes, TE coverage, GC content, and number of tRNAs, using the PAST 4.04 software. The Bray–Curtis dissimilarity index (edge cutoff: 50%) was adopted for the network analyses while the correlation analyses were conducted using Pearson's correlation.

2.8. CAZy Annotation and Potential for Lignocellulose Degradation

The Carbohydrate Active enzymes (CAZymes) of the *Trametes* species were functionally annotated using the dbCAN2 web server (http://bcb.unl.edu/dbCAN2/, accessed on 15 November 2020) with the integration of the following automated annotation tools/databases: (i) HMMER, (ii) DIAMOND, and (iii) Hotpep [62]. The dbCAN outputs were manually curated. Proteins identified by two or three tools (HMMER, DIAMOND, and HOTPEP) were considered correctly classified while those identified by only one tool were subjected to blastp analyses (protein–protein BLAST) to confirm the dbCAN classification. After annotation and manual curation, proteins related to cellulose, hemicellulose, lignin, and pectin degradation were counted and heat maps were generated using the pheatmap package (1.0.12) in R software (R 4.0.3).

3. Results and Discussion

3.1. Illumina and MinION Sequencing

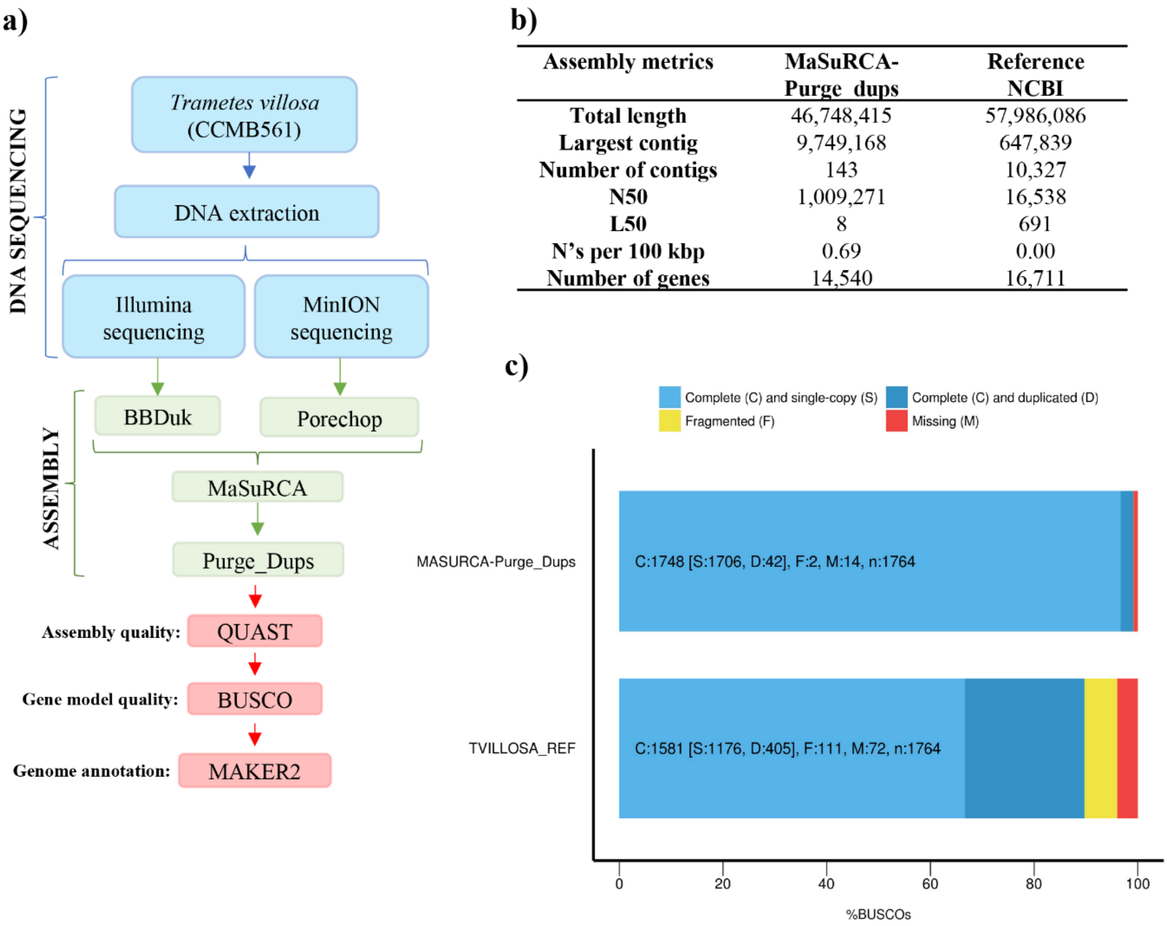
After adapter removal and quality trimming, 48,347,940 short reads were obtained through the Illumina sequencing (read length about 150 pb), corresponding to approximately 14 Gb (Table 1). Using these data and the GenomeScope software, the genome of the isolate *T. villosa* CCMB561 was estimated to have 44,895,640 bp in length, a homozygosity rate of 97.4%, and heterozygosity of 2.6% (Supplementary Data S1). The sequencing using the Oxford Nanopore platform generated 1,043,247 long reads, totaling 8.1 Gb. The long reads had an average size of 4.47 kb, N50 of 5.1 kb, and the longest read with 21,613 bp (Table 1). According to the estimated genome size of the CCMB561 strain, coverages of 129× and 93× were obtained through sequencing on the Illumina and Oxford Nanopore platforms, respectively. In previous studies, genomes with lengths similar to the estimated length of *T. villosa* CCMB561 were assembled with good contiguity and completeness using even smaller coverages than those obtained in this study. For example, the genome of the fungus *Leptosphaeria maculans* Nz-T4, which has a size of 43.42 Mb, was assembled in 288 contigs, using 56× long reads and 98× short reads [63]. In another study, the genome of the alga *Chlorella variabilis,* with a size of 46.67 Mb, was assembled into 302 contigs, using 56× long reads and 78× short reads [64]. Thus, the sequencing coverage obtained in this study was considered sufficient for the high-quality assembly of the genome of *T. villosa* CCMB561.

Table 1. Summary of the Illumina HiSeq and Oxford Nanopore MinION reads statistics after preprocessing step.

	Illumina	**MinION**
Total reads number	48,347,940	1,043,247
Total reads bases (bp)	5,798,237,268	4,189,223,607
Coverage	129×	93×
Longest read (bp)	151	21,613
Mean reads length (bp)	138	4476
GC content (%)	57.5	56

3.2. Genome Assembly and Assessment

In this study, eight assembly strategies were tested: one using exclusively short reads (Illumina HiSeq), four using only long reads (Oxford Nanopore MinION), and three based on the hybrid assembly, combining short and long reads (Table 2). The best result was obtained using the assembly workflow MaSuRCa-Purge_Dups (Hybrid assembly) (Table 2 and Figure 1a), which used MaSuRCa software to generate a primary assembly and the Purge_Dups program to identify and remove haplotypic duplications. The use of this workflow resulted in a genome with 143 contigs, a total length of 46,748,415 bp, the largest contig with 9,749,168 bp, and N50 of 1,009,271 bp (Table 2). This genome had the smallest difference according to the genome size estimated by GenomeScope2 (difference of 1,852,775 bp) and presented the best completeness index through the BUSCO analysis. The assembled genome presented 99.1% of the orthologous genes searched, of which 96.7% are single copies, 2.4% are duplicated, and 0.1% are fragmented. Finally, only 0.8% of the genes were not found (Table 3).



Assembly metrics	MaSuRCA-Purge_dups	Reference NCBI
Total length	46,748,415	57,986,086
Largest contig	9,749,168	647,839
Number of contigs	143	10,327
N50	1,009,271	16,538
L50	8	691
N's per 100 kbp	0.69	0.00
Number of genes	14,540	16,711

Figure 1. Overview of the newly assembled genome of *Trametes villosa* CCMB561. (**a**) Assembly workflow proposed as the best approach for genome assembly. (**b**) Summary evaluation of the genome assembled through MaSuRCa-Purge_Dups workflow and the reference genome of *Trametes villosa* deposited in the NCBI database (GCA_002964805.1). (**c**) BUSCO completeness assessment of the new genome and the reference of *Trametes villosa* previously deposited in the NCBI (GCA_002964805.1).

Table 2. Summary statistics for the assembled genomes of *Trametes villosa* CCMB561 using reads from Illumina HiSeq and Oxford Nanopore MinION.

	Assembly Short Reads (Illumina)	**Assembly Oxford Nanopore (MinION)**				**Hybrid Assembly (Illumina and Oxford Nanopore)**		
Assembly/ Software	MaSuRCa	CANU	CANU-smartdenovo	RACON	FLYE	SPADES	MaSuRCa	MaSuRCa-Purge_Dups
Number of contigs (≥0 bp)	4026	1836	337	1836	882	12,829	264	143
Number of contigs (≥500 bp)	3930	1836	337	1836	881	1940	264	143
Largest contig	470,636	1,594,329	1,660,310	1,605,280	1,891,910	1,207,893	4,772,416	9,749,168
Total length (≥500 bp)	58,820,861	63,704,316	42,774,667	63,971,542	49,876,064	65,406,907	62,711,988	46,748,415
GC (%)	59.40	59.36	59.39	59.41	59.35	59.39	59.39	59.45
N50	27,657	103,641	238,816	104,325	204,679	282,055	598,690	1,009,271
L50	503	115	43	114	55	69	21	8
# N's per 100 kbp	0.00	0.00	0.00	0.00	2.41	227.07	0.16	0.69

Table 3. Completeness assessment of *Trametes villosa* CCMB561 assemblies using BUSCO software.

	Complete (%)	**Single-Copy (%)**	**Duplicated (%)**	**Fragmented (%)**	**Missing (%)**
CANU	80.7	65.2	15.5	6.7	12.6
CANU-smartdenovo	76.2	73.8	2.4	8.6	15.2
FLYE	90.2	85.2	5.0	3.9	5.9
MaSuRCa (Hybrid)	99.0	64.0	35.0	0.2	0.8
MaSuRCa (Illumina)	97.4	64..3	33.1	0.9	1.7
MaSuRCa-Purge_Dups	99.1	96.7	2.4	0.1	0.8
RACON	88.2	70.0	18.2	4.5	7.3
SPADES	99.1	41.6	57.5	0.2	0.7

The assemblies in which the Purge_Dups software was not used showed a high degree of gene duplication and/or genome size greater than expected for the *Trametes* genus (~44 Mb), except when the CANU-smartdenovo pipeline was used (Tables 2 and 3). Previous studies have already demonstrated that, in order to facilitate the genome sequencing and assembly from dikaryotic fungi, dedikaryotization and thus the obtainment of a monokaryotic isolate is an essential step [27]. Hence, the high rate of gene duplication and genome size larger than expected may be related to the sequencing of the dikaryotic mycelium. Nevertheless, the use of Purge_Dups software [43] allowed us to remove duplications and increase genome contiguity without the need to obtain a monokaryotic isolate, which is laborious and time-consuming.

A high degree of fragmentation was detected when the genome was assembled using only Illumina sequencing data (Table 2). Conversely, when only long reads were used, the genome showed low fragmentation, but smaller completeness (Tables 2 and 3). This result is due to Illumina HiSeq sequencing generating reads with sizes between 100 and 250 bp, which leads to greater fragmentation of the genome. On the other hand, the MinION approach, despite generating long reads, which could exceed 2 Mb, presents a higher error rate (1D sequencing), which may imply a lower completeness rate [65].

Compared to the preliminary and draft genome of *T. villosa* CCMB561 deposited at the NCBI, herein, using the MaSuRCa-Purge_Dups assembly workflow (Figure 1a), we obtained a much higher-quality genome, with a significant reduction in the number of contigs (from 10,327 to 143) and improvement in the metrics N50, L50, and size of the largest contig (Figure 1b). Furthermore, based on the BUSCO analysis (Figure 1c), the newly assembled genome has better completeness indices and fewer duplicated, fragmented, and missing genes. Similarly, Maggiori et al. (2021) demonstrated that by using a hybrid assembly approach (HiSeq + MinION), it was possible to obtain a greater number of genomes from metagenomic samples, with longer contigs, more coding sequences, higher completeness, less contamination, and higher N50 [66].

3.3. Genome Annotation and Gene Ontology (GO) Analysis

The annotation results showed that *T. villosa* CCMB561 has 14,540 protein-coding genes, 86,516 exons, and 71,976 introns, which correspond to 66% of the genome (Table S1). Each gene had, on average, six exons and five introns with sizes of 265 bp and 112 bp, respectively, which agrees with other species of the genus (Table S1). In total, 274 transposable elements (TEs) were identified in the CCMB561 isolate, corresponding to 7.13% of the genome. The number of transport RNA (tRNAs) identified was 334.

According to gene ontology (GO) analysis, 8169 proteins of *T. villosa* CCMB561 were associated with GO terms, corresponding to 56.18% of the predicted sequences. Functionally annotated proteins were classified into three categories: (i) molecular function, (ii) cellular component, and (iii) biological process (Figure 2). In the “cellular component” category (Figure 2c), most proteins were associated with the terms “integral component of membrane” (2395), “nucleus” (638), and “cytoplasm” (319), which are terms related to the cell anatomy.

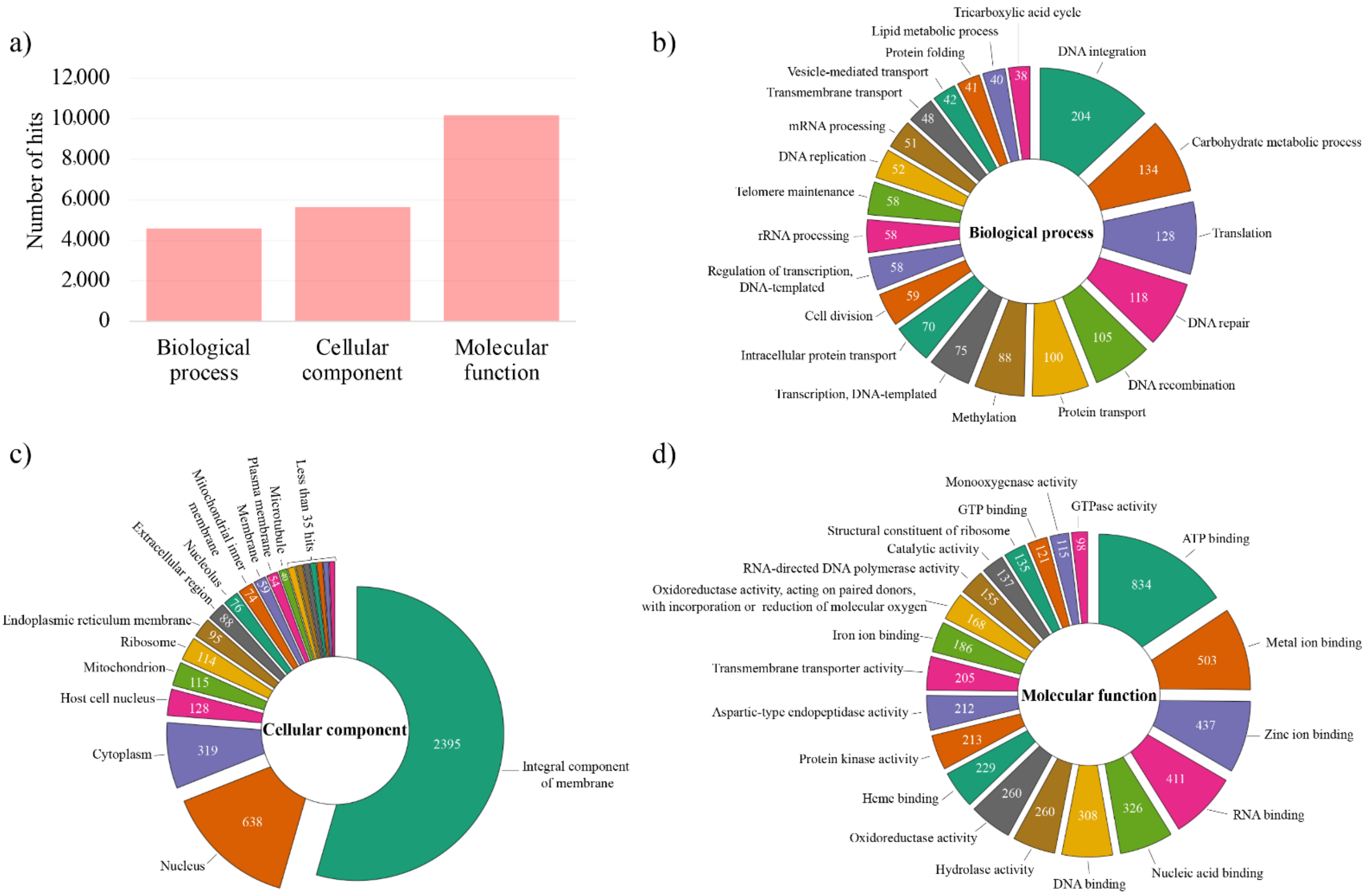


Figure 2. Gene Ontology (GO) functional annotation of *Trametes villosa* CCMB561 proteins. (**a**) Number of hits (GO terms) associated with the predicted proteins by GO categories (Biological process, Cellular component, and Molecular function), in which one protein can be associated with multiple GO terms. (**b**–**d**) The 20 most assigned terms per category in the GO enrichment analysis.

In the categories "biological processes" (Figure 2b) and "molecular function" (Figure 2d), terms related to the degradation of lignocellulosic biomass, such as "carbohydrate metabolic process" (GO:0005975), "hydrolase activity" (GO:0016787) "oxidoreductase activity" (GO:0016491), and "heme-binding" (GO:0020037) are among the most representative. This is an expected result since species from the genus *Trametes* act in the degradation of the main components of lignocellulosic biomass through the production and secretion of a large set of hydrolytic and oxidative enzymes [7,9,12,67]. Additionally, 205 proteins associated with "transmembrane transporter activity" (GO:0022857) have been identified (Figure 2d). This term could indicate enzymes that are secreted and have extracellular activity, such as lignocellulose-degrading enzymes [6].

In the genome of the CCMB561 strain, 147 cytochrome P450 (CYP) genes and one NADPH-cytochrome P450 reductase (CPR) gene were identified. Similar results were found by Sun et al. (2018), who identified in *T. versicolor* the presence of only one NADPH gene-cytochrome P450 reductase and multiple sequences belonging to genes of the cytochrome P450 family [68]. These genes are widely known for their importance in the degradation of lignin and organic pollutants (aromatic and xenobiotic compounds) [68]. Genes from the CYP family also play a role in the metabolism and adaptation of fungi to specific ecological niches [6,69]. Complementarily, Liu et al. (2019) described that the fungus *Trametes trogii* S0301 has 158 CYPs that may be related to a variety of metabolic functions [6].

3.4. Annotation of Secondary Metabolite Gene Clusters (SMGCs) and CAZymes of Trametes Villosa CCMB561

3.4.1. Secondary Metabolite Gene Clusters (SMGCs)

Fungi possess many gene clusters responsible for producing Secondary Metabolites (SMs), which have important ecological functions [70]. SMs are not essential for the normal growth of the organism but may act as defense compounds (e.g., against fungi and bacteria) and signaling molecules, being fundamental for ecological interactions and survival [71]. Because of their bioactive pharmacological properties, these molecules have been widely studied and tested in the healthcare industries to be used as antibiotics, antifungals, anti-inflammatory, and anticancer agents [71,72].

In *T. villosa* CCMB561, we identified 22 SM biosynthesis clusters, which comprise one non-ribosomal peptide synthetase (NRPS) cluster, one NRPS-like/betalactone cluster, three Type I Polyketide synthase (T1PKS) clusters, six NRPS-like clusters, and eleven Terpene clusters (Figure 3a and Supplementary Data S2). The NRPS-type cluster (Region 33.1/scf7180000000809) was identified with 100% of similarity with the basidioferrin compound cluster from *Gelatoporia subvermispora* (BGC0001527.1) (Polyporales, Basidiomycota). Basidioferrin is widely distributed in basidiomycetes and is part of the siderophore synthetases family, which are enzymes responsible for the biosynthesis of siderophores and iron metabolism [73]. Previous studies have reported that in most bacteria and fungi (pathogenic and non-pathogenic), the acquisition of high-affinity iron is mediated by siderophore-dependent pathways [74].

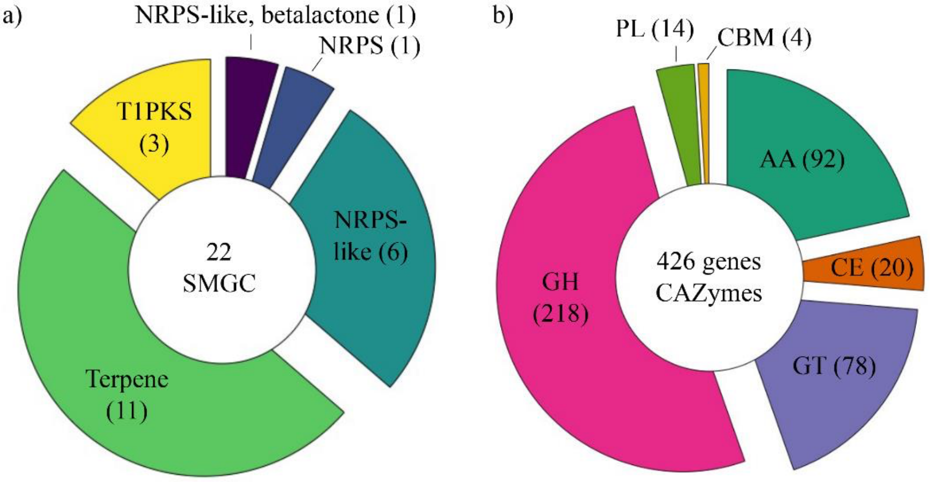


Figure 3. Annotation of Secondary metabolite gene clusters (SMGCs) and Carbohydrate-Active enzymes (CAZymes). (**a**) SMGCs identified in the genome of *Trametes villosa* CCMB561. (**b**) CAZymes identified in the genome of *Trametes villosa* CCMB561.

Most of the secondary metabolites biosynthesis clusters identified in *T. villosa* CCMB561 were assigned to the terpene type (11 clusters) (Figure 3a). Terpenoids have multiple biological activities and comprise sesquiterpenoids, diterpenoids, and triterpenoids. Their activities in inducing the apoptosis of human tumor cells, antibacterial, antimetastasis, and anti-HIV activity have already been demonstrated [72]. Using the MIBiG database (Minimum Information about a Biosynthetic Gene cluster), whole or partial genes of some known terpene clusters were identified. Our results showed that fragments of the clusters squalestatin S1 from *Aspergillus sp.* (BGC0001839.1), geosmin from *Streptomyces coelicolor* (BGC0001181.1), and koraiol from *Fusarium fujikuroi* (BGC0001642.1) were identified in the CCMB561 genome.

It is important to highlight that most clusters (21 clusters), despite being classified according to the type, had no similarity in the MIBiG database with the biosynthesis cluster of known compounds. Therefore, the CCMB561 isolate has the potential to produce a variety of secondary metabolites; however, these metabolites have not yet been identified or reported in the literature.

3.4.2. CAZome Annotation

The CAZome annotation results demonstrated that *T. villosa* CCMB561 possesses 426 genes encoding CAZymes, comprising 218 glycoside hydrolases (GH), 20 carbohydrate esterases (CE), 78 glycosyltransferases (GT), 14 polysaccharide lyases (PL), 4 Carbohydrate-binding modules (CBM), and 92 auxiliary activity enzymes (AA) (Figure 3b and Supplementary Data S3). Among the families that act in the depolymerization of lignocellulose, glycoside hydrolases (GHs) include glycosidases with activity in the hydrolysis of glycosidic bonds between two or more carbohydrates [14,15]. Carbohydrate esterases (CEs) act by removing ester-based modifications in polysaccharides, facilitating the action of GHs [14,15]. Polysaccharide Lyases (PLs) cleave glycosidic bonds from uronic acid-containing polysaccharides (e.g., pectin) [14,15]. Finally, the auxiliary activity families (AAs) include oxidative enzymes that act mainly in the depolymerization of lignin, helping the enzymes from GH, CE, and PL classes to gain access to carbohydrates from the plant cell wall [75].

The Carbohydrate-binding module (CBM) domain was detected in 33 genes encoding enzymes belonging to classes AA, GH, and CE (Supplementary Data S3). Most of the CBM domains (18 in total) found in the annotated genes belong to the Carbohydrate-binding module family 1 (CBM1) and were found in enzyme-encoding genes acting in the cellulose (AA9, GH3, GH5_5, GH6, and GH131 families) and hemicellulose breakdown (CE1, CE15, GH5_7, GH10, and GH74 families). CBM domains promote an enzyme association with the substrate and increase enzymatic hydrolysis and degradation of polysaccharides [76,77]. In addition, four dye-decolorizing peroxidases (DyPs) with the potential to oxidize lignin-like compounds and other phenolic polymers were identified in the genome of the isolate *T. villosa* CCMB561 [10].

Similar results of our CAZome annotation have been reported for the species *T. versicolor*, which has 424 genes encoding Carbohydrate-Active enzymes (CAZymes) [18]. *T. versicolor* is closely related to *T. villosa* and is one of the most common and widespread species of white-rot and basidiomata-forming fungi, showing great potential to act in lignocellulose breakdown [7,18].

3.5. Comparative Genomics and Phylogenomics of the Genus Trametes

The Maximum Likelihood phylogenetic matrix (RaxML) included 11 sequences with 798,158 amino acids from 1346 concatenate proteins of each genome. From these amino acids, 146,797 had distinct patterns, 152,144 were parsimoniously informative, 128,942 were parsimoniously non-informative, and 517,072 were constant characters. According to the tree topology (Figure 4a), *T. villosa*, *T. versicolor,* and *T. pubescens* were grouped in the same clade with a 100% bootstrap while the monophyletic clade formed by *T. coccinea*, *T. sanguinea,* and *T. cinnabarina* is the most phylogenetically distant from *Trametes villosa* CCMB561. These clustering patterns have already been reported in a previous phylogenetic reconstruction of Polyporales fungi, based on LSU and ITS ribosomal DNA markers [78]. In the Complex network analysis (Figure 4b), as well as in the phylogenomic analysis, *T. villosa*, *T. versicolor*, and *T. pubescens* were grouped together, suggesting these species have similar structural genomic characteristics, reinforcing the phylogenetic proximity.

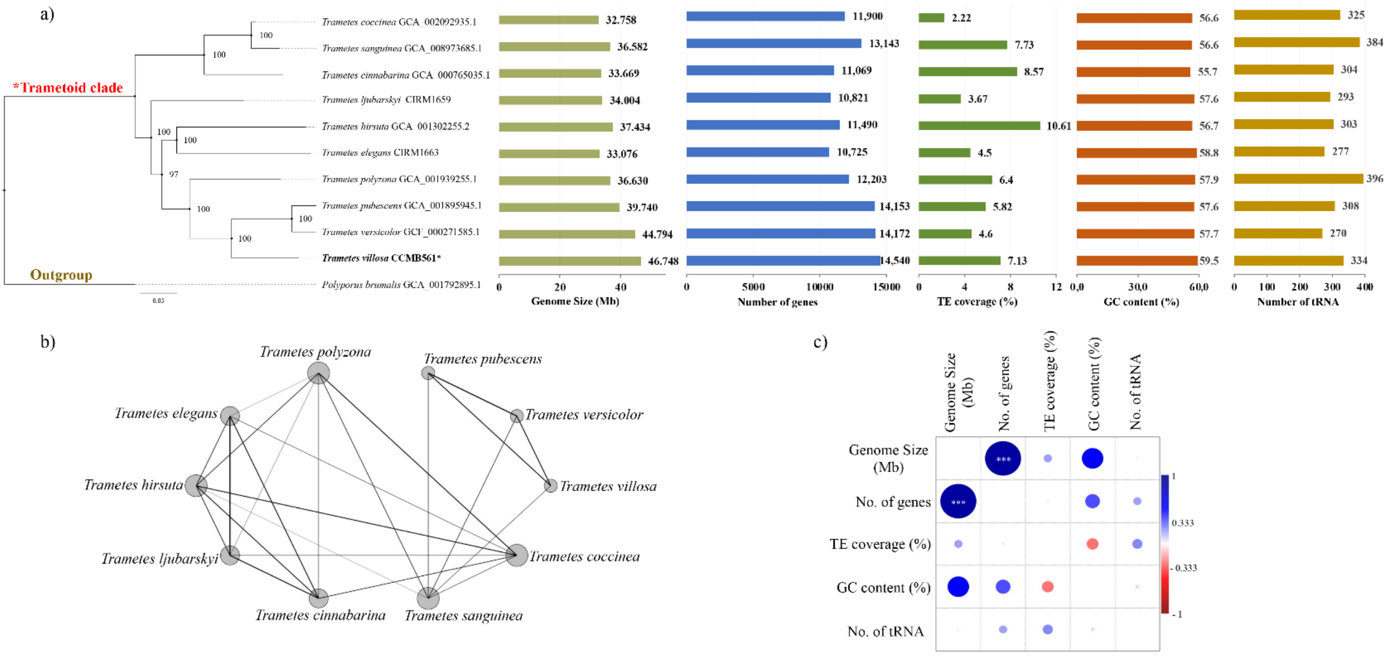


Figure 4. Comparative genomics results overview. (**a**) Maximum-likelihood phylogenomic tree constructed using the newly assembled genome of *Trametes villosa* CCMB561 (marked with *) and nine available genomes from the *Trametes* genus. Bootstrap values are expressed in percentage and the features of each genome are shown beside the phylogeny. (**b**) Network plot created using a matrix containing the values of genome size, number of genes, TE coverage, GC content, and number of tRNA of each genome. (**c**) Correlation analysis among the main metrics of the genome (statistically significant correlations are represented with *).

The fungal genome size can be impacted by different factors and is directly related to the fungal lifestyle, as well as adaptive and ecological needs [79–81]. Therefore, herein we evaluated the genome size of *Trametes* spp. and features that may impact the genome length, such as the number of genes, number and coverage of TEs, number of tRNAs, and GC content. Among the species analyzed, the genome sizes ranged from 32.758 Mb (*T. coccinea*) to 46.748 Mb (*T. villosa* CCMB561), and the number of genes from 10,725 (*T. elegans*) to 14,540 (*T. villosa* CCMB561) (Figure 4a). The average genome length for Basidiomycetes is 46.48 Mb, ranging from 9.82 (*Wallemia sebi*) to 130.65 Mb (*Dendrothele bispora*) [79]. The genomes of the *Trametes* spp. evaluated in this study had sizes within this range, and *T. villosa* CCMB561 was the species with a genome size closest to the mean described to the phylum Basidiomycota. This species also presented the highest number of genes, followed by the closely related *T. versicolor* and *T. pubescens*.

The coverage of TEs in the genomes ranged from 2.22% (*T. coccinea*) to 10.61% (*T. hirsuta*) (Figure 4a). Most of the classified TEs were LTR retrotransposons (long terminal repeats) belonging to Copia and/or Gypsy types (Table 4). Moreover, retrotransposons SINEs and LINEs, as well as DNA transposons and Helitrons were also identified. TEs can act in the modulation of genomes, through recombination and transposition, which can lead to chromosomal rearrangements and alter gene expression [80]. In Basidiomycota, the genome content corresponding to TEs can vary from 0.1 to 45.2% (average around 11%) and, usually, most of the TEs are LTR retrotransposons (Gypsy and Copia) [82]. A possible reason for the high copy number of TEs class I in Basidiomycota, including the studied species, is its transposition mechanism, which uses an intermediate RNA, resulting in the increased proliferative success of the TEs [82].

Table 4. Transposable elements (TE) identified in the *Trametes* species.

ID Fungo	Total No. TE	Total TE Coverage%	Retroelements				DNA Transposons	Helitron	Unclassified
			SINEs	LINEs	LTR Elements				
					Ty1/Copia	Gypsy/DIRS1			
Trametes cinnabarina	252	8.57	0	144	337	450	390	0	3079
Trametes coccinea	104	2.22	0	0	119	137	0	0	1187
Trametes elegans	129	4.50	0	0	210	451	14	0	1841
Trametes hirsuta	191	10.61	0	73	264	387	65	42	1949
Trametes ljubarskyi	172	3.67	0	0	317	263	89	0	2219
Trametes polyzona	349	6.41	0	113	416	912	105	69	5043
Trametes pubescens	303	5.82	19	106	107	160	258	30	3385
Trametes sanguinea	191	7.73	0	41	184	328	32	0	1855
Trametes versicolor	234	4.6	0	50	38	144	11	94	4165
Trametes villosa	274	7.13	17	97	186	503	74	70	4437

The GC content had no significant variation among the analyzed species and ranged from 55.7% (*T. cinnabarina*) to 59.5% (*T. villosa*) (Figure 4a). On the other hand, the number of tRNAs varied from 270 (*T. versicolor*) to 396 (*T. polyzona*), with a difference of more than 100 tRNAs among the species in this genus (Figure 4a). Transfer RNAs play a central role in protein biosynthesis and are involved in many biological functions in eukaryotic organisms, such as in the regulation of gene expression [83]. Therefore, this difference/expansion in the number of tRNAs could be related to specific evolutionary mechanisms and the lifestyle of each species [84].

Genome size and the number of genes were the only features that had a positive and statistically significant correlation ($p < 0.05$) in the genomes of *Trametes* spp. (Figure 4c). Interestingly, TE coverage was not significantly correlated with genome size. Similar results were described by Castanera et al. (2017), who described for some genera of the Agaricomycotina subphylum, including *Trametes* species, a high correlation between genome size and gene content while the correlation between the number of TEs and genome size was unclear [82].

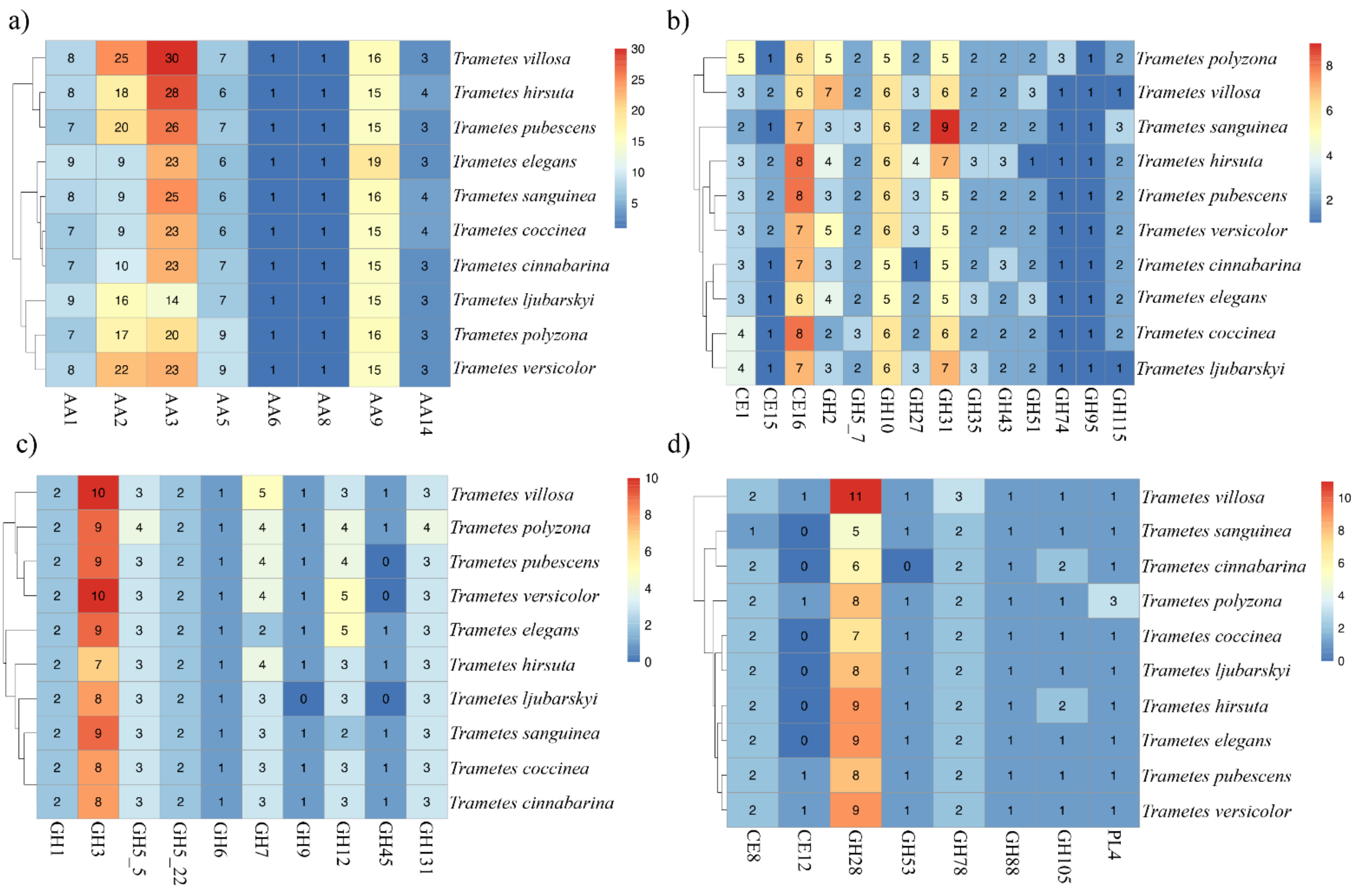
Overall, all the analyzed *Trametes* genomes displayed variability in relation to the analyzed features, with no clear pattern concerning the genome size, number of genes, number and coverage of TEs, and number of tRNAs. This variability may be related to the environment, selective pressures, and ecological and evolutionary factors to which each species is subjected to [84].

3.6. Potential for Lignocellulose Breakdown by Trametes spp.

Exploring the presence, abundance, and composition of oxidoreductases and carbohydrate-active enzymes (CAZymes) in wood-decay fungi provides important information on its nutritional preferences and adaptations, the metabolic pathways used, as well as the expansion and evolution of gene families related to lignocellulose breakdown. This information is important for an understanding of fungal biology and further application in the biotechnology industry. In this study, ten genomes of different *Trametes* species were evaluated for the presence of 40 gene families encoding enzymes that act in the breakdown of lignin, hemicellulose, cellulose, and pectin (Figure 5).

The lignin degradation is mainly performed by white-rot fungi and involves a series of enzymes classified as an Auxiliary Activity family (AA) [7,10,27,85]. In *Trametes* spp. genomes, many AAs were identified and have a central role in lignin modification (Figure 5a). The studied genomes harbor from 7 to 9 AA1-encoding genes, and in *T. villosa*, eight genes were recognized as AA1 (Figure 5a). The AA1 family encompasses multicopper oxidases (Laccases) that act directly on a wide range of aromatic and phenolic compounds, such as lignin [75,85,86]. These enzymes do not require cofactors for their activity, so they are of great interest for industrial applications [86].

The greatest number of lignin-modifying enzymes encoding genes was identified in families AA2 (9–25 genes) and AA3 (14–30 genes) (Figure 5a). The AA2 family includes Lignin Peroxidase (LiP), Manganese Peroxidase (MnP), and Versatile Peroxidase (VP). These enzymes are classified as class II peroxidases (PODs), since they use hydrogen peroxide (H_2O_2) as a cofactor for lignin breakdown [7,75,85,86]. The results displayed in Figure 5a also demonstrated an expansion in the number of AA2-encoding genes in *T. villosa* (25 copies), *T. hirsuta* (18 copies), *T. pubescens* (20 copies), *T. ljubarskyi* (16 copies), *T. polyzona* (17 copies), and *T. versicolor* (22 copies). This result could indicate a possible metabolic adaptation of these species to initially use lignin as the main carbon source for their growth. It is worth noting that *T. villosa* was the species with the largest number of AA2-encoding genes.



a)

	AA1	AA2	AA3	AA5	AA6	AA8	AA9	AA14
Trametes villosa	8	25	30	7	1	1	16	3
Trametes hirsuta	8	18	28	6	1	1	15	4
Trametes pubescens	7	20	26	7	1	1	15	3
Trametes elegans	9	9	23	6	1	1	19	3
Trametes sanguinea	8	9	25	6	1	1	16	4
Trametes coccinea	7	9	23	6	1	1	15	4
Trametes cinnabarina	7	10	23	7	1	1	15	3
Trametes ljubarskyi	9	16	14	7	1	1	15	3
Trametes polyzona	7	17	20	9	1	1	16	3
Trametes versicolor	8	22	23	9	1	1	15	3

b)

	CE1	CE15	CE16	GH2	GH5_7	GH10	GH27	GH31	GH35	GH43	GH51	GH74	GH95	GH115
Trametes polyzona	5	1	6	5	2	5	2	5	2	2	2	3	1	2
Trametes villosa	3	2	6	7	2	6	3	6	2	2	3	1	1	1
Trametes sanguinea	2	1	7	3	3	6	2	9	2	2	2	1	1	3
Trametes hirsuta	3	2	8	4	2	6	4	7	3	3	1	1	1	2
Trametes pubescens	3	2	8	3	2	6	3	5	2	2	2	1	1	2
Trametes versicolor	3	2	7	5	2	6	3	5	2	2	2	1	1	2
Trametes cinnabarina	3	1	7	3	2	5	1	5	2	3	2	1	1	2
Trametes elegans	3	1	6	4	2	5	2	5	3	2	3	1	1	2
Trametes coccinea	4	1	8	2	3	6	2	6	2	2	2	1	1	2
Trametes ljubarskyi	4	1	7	3	2	6	3	7	3	2	2	1	1	1

c)

	GH1	GH3	GH5_5	GH5_22	GH6	GH7	GH9	GH12	GH45	GH131
Trametes villosa	2	10	3	2	1	5	1	3	1	3
Trametes polyzona	2	9	4	2	1	4	1	4	1	4
Trametes pubescens	2	9	3	2	1	4	1	4	0	3
Trametes versicolor	2	10	3	2	1	4	1	5	0	3
Trametes elegans	2	9	3	2	1	2	1	5	1	3
Trametes hirsuta	2	7	3	2	1	4	1	3	1	3
Trametes ljubarskyi	2	8	3	2	1	3	0	3	0	3
Trametes sanguinea	2	9	3	2	1	3	1	2	1	3
Trametes coccinea	2	8	3	2	1	3	1	3	1	3
Trametes cinnabarina	2	8	3	2	1	3	1	3	1	3

d)

	CE8	CE12	GH28	GH53	GH78	GH88	GH105	PL4
Trametes villosa	2	1	11	1	3	1	1	1
Trametes sanguinea	1	0	5	1	2	1	1	1
Trametes cinnabarina	2	0	6	0	2	1	2	1
Trametes polyzona	2	1	8	1	2	1	1	3
Trametes coccinea	2	0	7	1	2	1	1	1
Trametes ljubarskyi	2	0	8	1	2	1	1	1
Trametes hirsuta	2	0	9	1	2	1	2	1
Trametes elegans	2	0	9	1	2	1	1	1
Trametes pubescens	2	1	8	1	2	1	1	1
Trametes versicolor	2	1	9	1	2	1	1	1

Figure 5. CAZyme-encoding genes involved in the degradation of lignocellulosic biomass. (**a**) Number of auxiliary redox enzyme-encoding genes. (**b**) Number of hemicellulose-degrading enzyme-encoding genes. (**c**) Number of cellulose breakdown enzyme-encoding genes. (**d**) Number of pectin-degrading enzyme-encoding genes.

In the metabolic pathway of lignin degradation, enzymes belonging to the AA3 and AA5 families play a fundamental role in the generation of H_2O_2 and activation of PODs [75,86]. AA3 are flavoproteins containing a flavin-adenine dinucleotide (FAD)-binding domain and are generally recognized as glucose-methanol-choline (GMC) oxidoreductases, which include Pyranose 2-oxidase, Alcohol oxidase, Glucose 1-oxidase, and Aryl-alcohol oxidase [75,86]. In our analyses, AA3 was the family with the highest number of identified genes, ranging from 14 to 30 copies per species. Additionally, *T. villosa* was the species with the highest number of AA3-encoding genes, possessing 30 copies.

The AA5 family is composed of copper radical oxidases and includes two described subfamilies: AA5_1 (glyoxal oxidase) and AA5_2 (galactose oxidases) [75,86]. The *Trametes* spp. genomes harbor 6 to 9 AA5-encoding genes, and in *T. villosa*, seven AA5 genes were identified. One AA6-encoding gene was conserved in all *Trametes* species. This enzymatic family includes 1,4-benzoquinone reductases, which are responsible for the intracellular cleavage of aromatic compounds and for the protection of fungal cells from reactive quinone compounds [75].

Regarding the families of hydrolytic enzymes with activity in the depolymerization of hemicellulose, CE16 (acetylesterases—six to eight copies), GH10 (endo-1,4-β-xylanases—five to six copies), and GH31 (alpha/beta-glucosidases/α-xylosidases—five to nine copies) were the families with the highest number of genes (Figure 5b). Carbohydrate Esterase family 16 (acetylesterase activity) plays a fundamental role in the deacetylation of hemicellulose units, allowing the activity of glycoside hydrolases [77]. The GH10 family includes endoxylanases (endo-β-1,4-endoxylanases) that act on the degradation of linear chains of β-1,4-linked D-xylose residues [3,18,77]. The CAZy family GH31 mainly includes enzymes with α-glucosidase activity [77]. Besides, other xylanases and xyloglucanases belonging to the families GH5_7 (β-1,4-endoxylanases/β-1,4-endoglucanases), GH51 (β-1,4-endoxylanases), GH43 (β-xylosidase), GH35 (β-galactosidases), and GH74 (endoglucanases) were also identified in the analyzed genomes (Figure 5b).

Xylan-type hemicellulose can also be degraded through the action of oxidative enzymes belonging to the AA14 family, which include copper-dependent lytic polysaccharide monooxygenases (LPMOs) [87]. In the evaluated *Trametes* spp. genomes, three to four genes encoding lytic xylan monooxygenase were identified (Figure 5a). Different families were annotated in the analyzed genomes responsible for the degradation of hemicellulose side chains, such as GH2 (β-galactosidases), GH27 (α-galactosidase), GH95 (α-L-fucosidase), GH115 (xylan α-1,2-glucuronidase), CE1 (acetyl xylan esterase, EC 3.1.1.72), and CE15 (glucuronoyl esterase) (Figure 5b) [77].

Cellulose is a linear polymer formed by residues of D-glucose linked by β-1,4-glycosidic bonds. This is the most abundant and the least complex polysaccharide of the plant cell wall and is degraded by three classes of enzymes, β-1,4-endoglucanases, cellobiohydrolases, and β-glucosidases [3,18,77]. Figure 5c exhibits the families of enzymes related to cellulose hydrolysis. The GH3 family (β-glycosidases) had the highest number of genes, ranging from 7 to 10 copies per genome. In general, families of β-1,4-endoglucanases (families GH5_5, GH5_22, GH9, GH12, GH45, GH131), exoglucanases/cellobiohydrolases (families GH6 and GH7), and β-glucosidases (GH1 and GH3 families) were identified in all fungi. The genes encoding the GH9 and GH45 families were absent in *T. ljubarskyi*. Likewise, in *T. pubescens* and *T. versicolor,* the GH45 gene family was not found (Figure 5c). Finally, 15 to 19 AA9-encoding genes (LPMOs) were identified in the *Trametes* spp. genomes (Figure 5a). AA9 genes are classified as copper-dependent lytic polysaccharide monooxygenases (LPMOs) that act in the oxidative depolymerization of crystalline cellulose [75].

Figure 5d displays the genes encoding enzymes with activity in pectin degradation. As evidenced, the GH28 family, which includes part of the glycosidic hydrolases, especially endo- and exo-polygalacturonases and endo- and exo-rhamnogalacturonases, had the largest number of genes, ranging from 5 to 11 copies per genome. Moreover, other enzymatic families were observed to be involved in pectin hydrolysis, but in a smaller proportion, such as GH78 (two to three genes), GH88 (one gene), GH105 (one to two genes), and GH53 (one gene), which include α-rhamnosidases, unsaturated glucuronyl hydrolases, unsaturated rhamnogalacturone hydrolases, and β-endogalactanases, respectively. In all genomes, pectinmethylesterases (CE8) were identified, and in most of them, enzymes belonging to the CE12 family (pectin acetylesterase) were not found. Finally, it was observed that all *Trametes* spp. genomes contain one gene encoding PL4 (rhamnogalacturonan endolyase), except for *T. polyzona*, which has three genes encoding this enzymatic family.

From the exploratory analysis of CAZymes, a set of genes encoding cellulases, hemicellulases, pectinases, and lignin-modifying enzymes were identified in the genomes of the *Trametes* species. These enzymes act synergistically, contributing to the breakdown of all polymers that make up the plant cell wall [3,7,18,27,77]. Among the analyzed genomes, *T. villosa* CCMB561 was the species with the highest number of genes encoding lignin-modifying enzymes (91 genes) and pectinases (21 genes) and the second with the highest number of genes encoding cellulases (31 genes) and hemicellulases (45 genes). It is also worth mentioning that *T. villosa* CCMB561 harbors all 40 searched genes related to the lignocellulose breakdown. Therefore, this isolate has great potential to be applied in the bioconversion of lignocellulosic biomass in the industry.

4. Conclusions

In this study, we demonstrated that through the hybrid assembly using short (Illumina HiSeq) and long reads (Oxford Nanopore MinION), and the assembly workflow MaSuRCA-Purge_dups, a high-quality genome for the isolate *T. villosa* CCMB561 was obtained. The contiguity and completeness of the genome assembled and presented in this study significantly increased when compared to the preliminary and draft version of this isolate previously sequenced using only short reads (Illumina HiSeq). The accurate annotation of the new genome, the comparative genomic analyses, associated with the functional annotation of the CAZymes-encoding genes demonstrated the genetic potential of the isolate *T. villosa* CCMB561 to act in the degradation of all components of lignocellulose. Among the analyzed genomes, *T. villosa* was the species with the highest number of genes encoding lignin-modifying enzymes. Lignin is the most recalcitrant polymer of the plant cell wall and, thus, its removal is considered the most limiting step for the conversion of lignocellulosic biomass. Taken together, data generated in this study provide support for future studies using genomics, transcriptomics, and proteomics tools. Still, they contribute to the understanding of the complex mechanisms involved in the expression, regulation, and evolution of genes and proteins associated with lignocellulose breakdown.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/jof8020142/s1, Figure S1: Workflow used for the genome sequencing and assembly of *Trametes villosa* CCMB561. Figure S2: Gene and protein annotation pipeline used for all *Trametes* species. Table S1: Gene annotation statistics generated by GAG software for all genomes from the *Trametes* genus analyzed in our study. Supplementary Data S1: GenomeScope results for the *Trametes villosa* CCMB561. Supplementary Data S2: Annotation of Secondary metabolite gene clusters (SMGCs) of *Trametes villosa* CCMB561. Supplementary Data S3: Annotation of Carbohydrate-active enzymes (CAZymes) for the isolate *Trametes villosa* CCMB561.

Author Contributions: Experimental design: L.M.R.T., F.B., V.A.d.C.A. and A.G.-N.; performed the experiments: L.M.R.T., P.L.C.F., T.M.-P. and B.B.; analyzed the data: L.M.R.T. and F.F.d.S.; wrote the paper: L.M.R.T., F.B., V.A.d.C.A. and A.G.-N. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Luiz Marcelo Ribeiro Tomé received a CNPq scholarship to conduct this study (Process number: 140156/2019-4) and Aristóteles Góes-Neto received a research grant for scientific productivity from CNPq (Process number: 310764/2016-5).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession PUDQ00000000. The version described in this paper is version PUDQ02000000 (BioSample ID: SAMN08579176, BioProject ID: PRJNA435407).

Acknowledgments: We would like to thank the Graduate Programs of Microbiology (http://www.microbiologia.icb.ufmg.br/pos/, accessed on 15 December 2021) and Bioinformatics (http://www.pgbioinfo.icb.ufmg.br/, accessed on 15 December 2021) of the Universidade Federal de Minas Gerais (UFMG) and the Department of Chemistry of the Centro Federal de Educação Tecnológica de Minas Gerais (CEFET-MG).

Conflicts of Interest: The authors declare that they have no competing interests and affiliations with or involvement in any organization or entity with any financial interest or non-financial interest that could be construed as a potential conflict of interest.

References

1. Srivastava, N.; Rawat, R.; Singh Oberoi, H.; Ramteke, P.W. A Review on Fuel Ethanol Production from Lignocellulosic Biomass. *Int. J. Green Energy* **2015**, *12*, 949–960. [CrossRef]
2. Isikgor, F.H.; Becer, C.R. Lignocellulosic biomass: A sustainable platform for the production of bio-based chemicals and polymers. *Polym. Chem.* **2015**, *6*, 4497–4559. [CrossRef]
3. Houfani, A.A.; Anders, N.; Spiess, A.C.; Baldrian, P.; Benallaoua, S. Insights from enzymatic degradation of cellulose and hemicellulose to fermentable sugars—A review. *Biomass Bioenergy* **2020**, *134*, 105481. [CrossRef]
4. Hernández-Beltrán, J.U.; Hernández-De Lira, I.O.; Cruz-Santos, M.M.; Saucedo-Luevanos, A.; Hernández-Terán, F.; Balagurusamy, N. Insight into Pretreatment Methods of Lignocellulosic Biomass to Increase Biogas Yield: Current State, Challenges, and Opportunities. *Appl. Sci.* **2019**, *9*, 3721. [CrossRef]
5. Bilal, M.; Nawaz, M.Z.; Iqbal, H.M.N.; Hou, J.; Mahboob, S.; Al-Ghanim, K.A.; Cheng, H. Engineering Ligninolytic Consortium for Bioconversion of Lignocelluloses to Ethanol and Chemicals. *Protein Pept. Lett.* **2018**, *25*, 108–119. [CrossRef]

6. Liu, Y.; Wu, Y.; Zhang, Y.; Yang, X.; Yang, E.; Xu, H.; Yang, Q.; Chagan, I.; Cui, X.; Chen, W.; et al. Lignin degradation potential and draft genome sequence of *Trametes trogii* S0301. *Biotechnol. Biofuels* **2019**, *12*, 256. [CrossRef]
7. Floudas, D.; Binder, M.; Riley, R.; Barry, K.; Blanchette, R.A.; Henrissat, B.; Martínez, A.T.; Otillar, R.; Spatafora, J.W.; Yadav, J.S.; et al. The Paleozoic Origin of Enzymatic Lignin Decomposition Reconstructed from 31 Fungal Genomes. *Science* **2012**, *336*, 1715–1719. [CrossRef]
8. Andriani, A.; Maharani, A.; Yanto, D.H.Y.; Pratiwi, H.; Astuti, D.; Nuryana, I.; Agustriana, E.; Anita, S.H.; Juanssilfero, A.B.; Perwitasari, U.; et al. Sequential production of ligninolytic, xylanolytic, and cellulolytic enzymes by *Trametes hirsuta* AA-017 under different biomass of Indonesian sorghum accessions-induced cultures. *Bioresour. Technol. Rep.* **2020**, *12*, 100562. [CrossRef]
9. Vasina, D.V.; Moiseenko, K.V.; Fedorova, T.V.; Tyazhelova, T.V. Lignin-degrading peroxidases in white-rot fungus *Trametes hirsuta* 072. Absolute expression quantification of full multigene family. *PLoS ONE* **2017**, *12*, e0173813. [CrossRef]
10. Mäkinen, M.; Kuuskeri, J.; Laine, P.; Smolander, O.-P.; Kovalchuk, A.; Zeng, Z.; Asiegbu, F.O.; Paulin, L.; Auvinen, P.; Lundell, T. Genome description of *Phlebia radiata* 79 with comparative genomics analysis on lignocellulose decomposition machinery of phlebioid fungi. *BMC Genom.* **2019**, *20*, 430. [CrossRef]
11. Atilano-Camino, M.M.; Álvarez-Valencia, L.H.; García-González, A.; García-Reyes, R.B. Improving laccase production from *Trametes versicolor* using lignocellulosic residues as cosubstrates and evaluation of enzymes for blue wastewater biodegradation. *J. Environ. Manag.* **2020**, *275*, 111231. [CrossRef] [PubMed]
12. Paës, G.; Navarro, D.; Benoit, Y.; Blanquet, S.; Chabbert, B.; Chaussepied, B.; Coutinho, P.M.; Durand, S.; Grigoriev, I.V.; Haon, M.; et al. Tracking of enzymatic biomass deconstruction by fungal secretomes highlights markers of lignocellulose recalcitrance. *Biotechnol. Biofuels* **2019**, *12*, 76. [CrossRef] [PubMed]
13. Espósito, E.; Azevedo, J.L. *De Fungos—Uma Introdução à Biologia, Bioquímica e Biotecnologia*, 2nd ed.; EDUCS, Editora da Universidade de Caxias do Sul: Caxias do Sul, Brasil, 2004; pp. 9–510.
14. Lombard, V.; Golaconda Ramulu, H.; Drula, E.; Coutinho, P.M.; Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **2014**, *42*, D490–D495. [CrossRef] [PubMed]
15. Cantarel, B.L.; Coutinho, P.M.; Rancurel, C.; Bernard, T.; Lombard, V.; Henrissat, B. The Carbohydrate-Active EnZymes database (CAZy): An expert resource for Glycogenomics. *Nucleic Acids Res.* **2009**, *37*, D233–D238. [CrossRef] [PubMed]
16. Kumar, A.; Chandra, R. Ligninolytic enzymes and its mechanisms for degradation of lignocellulosic waste in environment. *Heliyon* **2020**, *6*, e03170. [CrossRef] [PubMed]
17. Mendonça Maciel, M.J.; Castro e Silva, A.; Telles Ribeiro, H.C. Industrial and biotechnological applications of ligninolytic enzymes of the basidiomycota: A review. *Electron. J. Biotechnol.* **2010**, *13*. [CrossRef]
18. Sista Kameshwar, A.K.; Qin, W. Comparative study of genome-wide plant biomass-degrading CAZymes in white rot, brown rot and soft rot fungi. *Mycology* **2018**, *9*, 93–105. [CrossRef] [PubMed]
19. Daniel, G. Fungal and Bacterial Biodegradation: White Rots, Brown Rots, Soft Rots, and Bacteria. In: Deterioration and protection of sustainable biomaterials. *Am. Chem. Soc.* **2014**, 23–58. [CrossRef]
20. Giweta, M. Role of litter production and its decomposition, and factors affecting the processes in a tropical forest ecosystem: A review. *J. Ecol. Environ.* **2020**, *44*, 11. [CrossRef]
21. Hage, H.; Miyauchi, S.; Virágh, M.; Drula, E.; Min, B.; Chaduli, D.; Navarro, D.; Favel, A.; Norest, M.; Lesage-Meessen, L.; et al. Gene family expansions and transcriptome signatures uncover fungal adaptations to wood decay. *Environ. Microbiol.* **2021**, *23*, 5716–5732. [CrossRef]
22. Couturier, M.; Navarro, D.; Chevret, D.; Henrissat, B.; Piumi, F.; Ruiz-Dueñas, F.J.; Martinez, A.T.; Grigoriev, I.V.; Riley, R.; Lipzen, A.; et al. Enhanced degradation of softwood versus hardwood by the white-rot fungus *Pycnoporus coccineus*. *Biotechnol. Biofuels* **2015**, *8*, 216. [CrossRef] [PubMed]
23. Pavlov, A.R.; Tyazhelova, T.V.; Moiseenko, K.V.; Vasina, D.V.; Mosunova, O.V.; Fedorova, T.V.; Maloshenok, L.G.; Landesman, E.O.; Bruskin, S.A.; Psurtseva, N.V.; et al. Draft Genome Sequence of the Fungus *Trametes hirsuta* 072. *Genome Announc.* **2015**, *3*. [CrossRef] [PubMed]
24. Ferreira, D.S.S.; Kato, R.B.; Miranda, F.M.; da Costa Pinheiro, K.; Fonseca, P.L.C.; Tomé, L.M.R.; Vaz, A.B.M.; Badotti, F.; Ramos, R.T.J.; Brenig, B.; et al. Draft genome sequence of *Trametes villosa* (Sw.) Kreisel CCMB561, a tropical white-rot Basidiomycota from the semiarid region of Brazil. *Data Br.* **2018**, *18*, 1581–1587. [CrossRef] [PubMed]
25. Busk, P.K.; Lange, M.; Pilgaard, B.; Lange, L. Several Genes Encoding Enzymes with the Same Activity Are Necessary for Aerobic Fungal Degradation of Cellulose in Nature. *PLoS ONE* **2014**, *9*, e114138. [CrossRef]
26. Granchi, Z.; Peng, M.; Chi-A-Woeng, T.; de Vries, R.P.; Hildén, K.; Mäkelä, M.R. Genome Sequence of the Basidiomycete White-Rot Fungus *Trametes pubescens* FBCC735. *Genome Announc.* **2017**, *5*. [CrossRef]
27. Miyauchi, S.; Rancon, A.; Drula, E.; Hage, H.; Chaduli, D.; Favel, A.; Grisel, S.; Henrissat, B.; Herpoël-Gimbert, I.; Ruiz-Dueñas, F.J.; et al. Integrative visual omics of the white-rot fungus *Polyporus brumalis* exposes the biotechnological potential of its oxidative enzymes for delignifying raw plant biomass. *Biotechnol. Biofuels* **2018**, *11*, 201. [CrossRef]
28. Lin, W.; Jia, G.; Sun, H.; Sun, T.; Hou, D. Genome sequence of the fungus *Pycnoporus sanguineus*, which produces cinnabarinic acid and pH- and thermo- stable laccases. *Gene* **2020**, *742*, 144586. [CrossRef]
29. Carneiro, R.T.; Lopes, M.A.; Silva, M.L.; Santos, V.D.; Souza, V.B.; Sousa, A.O.; Pirovani, C.P.; Koblitz, M.G.; Benevides, R.G.; Góes-Neto, A. *Trametes villosa* Lignin Peroxidase (TvLiP): Genetic and Molecular Characterization. *J. Microbiol. Biotechnol.* **2017**, *27*, 179–188. [CrossRef]

30. Coniglio, R.O.; Díaz, G.V.; Fonseca, M.I.; Castrillo, M.L.; Piccinni, F.E.; Villalba, L.L.; Campos, E.; Zapata, P.D. Enzymatic hydrolysis of barley straw for biofuel industry using a novel strain of *Trametes villosa* from Paranaense rainforest. *Prep. Biochem. Biotechnol.* **2020**, *50*, 753–762. [CrossRef]
31. Yamanaka, R.; Soares, C.F.; Matheus, D.R.; Machado, K.M.G. Lignolytic enzymes produced by *Trametes villosa* ccb176 under different culture conditions. *Braz. J. Microbiol.* **2008**, *39*, 78–84. [CrossRef]
32. Silva, M.L.; de Souza, V.B.; da Silva Santos, V.; Kamida, H.M.; de Vasconcellos-Neto, J.R.; Góes-Neto, A.; Koblitz, M.G. Production of Manganese Peroxidase by *Trametes villosa* on Unexpensive Substrate and Its Application in the Removal of Lignin from Agricultural Wastes. *Adv. Biosci. Biotechnol.* **2014**, *5*, 1067–1077. [CrossRef]
33. White, T.J.; Bruns, T.; Lee, S.; Taylor, J. Amplification and Direct Sequencing of Fungal Ribosomal RNA Genes for Phylogenetics. In *PCR Protocols*; Elsevier: Amsterdam, The Netherlands, 1990; pp. 315–322.
34. Tomé, L.M.R.; Badotti, F.; Assis, G.B.N.; Fonseca, P.L.C.; da Silva, G.A.; da Silveira, R.M.B.; Costa-Rezende, D.H.; dos Santos, E.R.D.; de Carvalho Azevedo, V.A.; Figueiredo, H.C.P.; et al. Proteomic fingerprinting for the fast and accurate identification of species in the Polyporoid and Hymenochaetoid fungi clades. *J. Proteom.* **2019**, *203*, 103390. [CrossRef] [PubMed]
35. Marçais, G.; Kingsford, C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* **2011**, *27*, 764–770. [CrossRef] [PubMed]
36. Ranallo-Benavidez, T.R.; Jaron, K.S.; Schatz, M.C. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nat. Commun.* **2020**, *11*, 1432. [CrossRef]
37. Kolmogorov, M.; Yuan, J.; Lin, Y.; Pevzner, P.A. Assembly of long, error-prone reads using repeat graphs. *Nat. Biotechnol.* **2019**, *37*, 540–546. [CrossRef]
38. Koren, S.; Walenz, B.P.; Berlin, K.; Miller, J.R.; Bergman, N.H.; Phillippy, A.M. Canu: Scalable and accurate long-read assembly via adaptive k -mer weighting and repeat separation. *Genome Res.* **2017**, *27*, 722–736. [CrossRef]
39. Vaser, R.; Sović, I.; Nagarajan, N.; Šikić, M. Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res.* **2017**, *27*, 737–746. [CrossRef]
40. Schmidt, M.H.-W.; Vogel, A.; Denton, A.K.; Istace, B.; Wormit, A.; van de Geest, H.; Bolger, M.E.; Alseekh, S.; Maß, J.; Pfaff, C.; et al. De Novo Assembly of a New *Solanum pennellii* Accession Using Nanopore Sequencing. *Plant Cell* **2017**, *29*, 2336–2348. [CrossRef]
41. Zimin, A.V.; Marçais, G.; Puiu, D.; Roberts, M.; Salzberg, S.L.; Yorke, J.A. The MaSuRCA genome assembler. *Bioinformatics* **2013**, *29*, 2669–2677. [CrossRef]
42. Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.A.; Dvorkin, M.; Kulikov, A.S.; Lesin, V.M.; Nikolenko, S.I.; Pham, S.; Prjibelski, A.D.; et al. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *J. Comput. Biol.* **2012**, *19*, 455–477. [CrossRef]
43. Guan, D.; McCarthy, S.A.; Wood, J.; Howe, K.; Wang, Y.; Durbin, R. Identifying and removing haplotypic duplication in primary genome assemblies. *Bioinformatics* **2020**, *36*, 2896–2898. [CrossRef] [PubMed]
44. Gurevich, A.; Saveliev, V.; Vyahhi, N.; Tesler, G. QUAST: Quality assessment tool for genome assemblies. *Bioinformatics* **2013**, *29*, 1072–1075. [CrossRef] [PubMed]
45. Simão, F.A.; Waterhouse, R.M.; Ioannidis, P.; Kriventseva, E.V.; Zdobnov, E.M. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **2015**, *31*, 3210–3212. [CrossRef] [PubMed]
46. Cantarel, B.L.; Korf, I.; Robb, S.M.C.; Parra, G.; Ross, E.; Moore, B.; Holt, C.; Sanchez Alvarado, A.; Yandell, M. MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* **2007**, *18*, 188–196. [CrossRef] [PubMed]
47. Holt, C.; Yandell, M. MAKER2: An annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinform.* **2011**, *12*, 491. [CrossRef] [PubMed]
48. Campbell, M.S.; Holt, C.; Moore, B.; Yandell, M. Genome Annotation and Curation Using MAKER and MAKER-P. *Curr. Protoc. Bioinform.* **2014**, *48*, 4.11.1–4.11.39. [CrossRef] [PubMed]
49. Korf, I. Gene finding in novel genomes. *BMC Bioinform.* **2004**, *5*, 59. [CrossRef] [PubMed]
50. Stanke, M.; Steinkamp, R.; Waack, S.; Morgenstern, B. AUGUSTUS: A web server for gene finding in eukaryotes. *Nucleic Acids Res.* **2004**, *32*, W309–W312. [CrossRef] [PubMed]
51. Lomsadze, A. Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic Acids Res.* **2005**, *33*, 6494–6506. [CrossRef]
52. Tarailo-Graovac, M.; Chen, N. Using Repeat Masker to Identify Repetitive Elements in Genomic Sequences. *Curr. Protoc. Bioinform.* **2009**, *25*, 4.10.1–4.10.14. [CrossRef] [PubMed]
53. Geib, S.M.; Hall, B.; Derego, T.; Bremer, F.T.; Cannoles, K.; Sim, S.B. Genome Annotation Generator: A simple tool for generating and correcting WGS annotation tables for NCBI submission. *Gigascience* **2018**, *7*, giy018. [CrossRef] [PubMed]
54. Araujo, F.A.; Barh, D.; Silva, A.; Guimarães, L.; Ramos, R.T.J. GO FEAT: A rapid web-based functional annotation tool for genomic and transcriptomic data. *Sci. Rep.* **2018**, *8*, 1794. [CrossRef] [PubMed]
55. Lowe, T.M.; Eddy, S.R. tRNAscan-SE: A Program for Improved Detection of Transfer RNA Genes in Genomic Sequence. *Nucleic Acids Res.* **1997**, *25*, 955–964. [CrossRef] [PubMed]
56. Blin, K.; Shaw, S.; Kloosterman, A.M.; Charlop-Powers, Z.; van Wezel, G.P.; Medema, M.H.; Weber, T. antiSMASH 6.0: Improving cluster detection and comparison capabilities. *Nucleic Acids Res.* **2021**, *49*, W29–W35. [CrossRef] [PubMed]

57. Bao, Z. Automated De Novo Identification of Repeat Sequence Families in Sequenced Genomes. *Genome Res.* **2002**, *12*, 1269–1276. [CrossRef]
58. Price, A.L.; Jones, N.C.; Pevzner, P.A. De novo identification of repeat families in large genomes. *Bioinformatics* **2005**, *21*, i351–i358. [CrossRef]
59. Benson, G. Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Res.* **1999**, *27*, 573–580. [CrossRef]
60. Capella-Gutierrez, S.; Silla-Martinez, J.M.; Gabaldon, T. trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **2009**, *25*, 1972–1973. [CrossRef]
61. Nguyen, L.-T.; Schmidt, H.A.; von Haeseler, A.; Minh, B.Q. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol. Biol. Evol.* **2015**, *32*, 268–274. [CrossRef]
62. Zhang, H.; Yohe, T.; Huang, L.; Entwistle, S.; Wu, P.; Yang, Z.; Busk, P.K.; Xu, Y.; Yin, Y. dbCAN2: A meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* **2018**, *46*, W95–W101. [CrossRef]
63. Dutreux, F.; Da Silva, C.; D'Agata, L.; Couloux, A.; Gay, E.J.; Istace, B.; Lapalu, N.; Lemainque, A.; Linglin, J.; Noel, B.; et al. De novo assembly and annotation of three *Leptosphaeria* genomes using Oxford Nanopore MinION sequencing. *Sci. Data* **2018**, *5*, 180235. [CrossRef] [PubMed]
64. Minei, R.; Hoshina, R.; Ogura, A. De novo assembly of middle-sized genome using MinION and Illumina sequencers. *BMC Genom.* **2018**, *19*, 700. [CrossRef] [PubMed]
65. De Carvalho, L.M.; Borelli, G.; Camargo, A.P.; de Assis, M.A.; de Ferraz, S.M.F.; Fiamenghi, M.B.; José, J.; Mofatto, L.S.; Nagamatsu, S.T.; Persinoti, G.F.; et al. Bioinformatics applied to biotechnology: A review towards bioenergy research. *Biomass Bioenergy* **2019**, *123*, 195–224. [CrossRef]
66. Maggiori, C.; Raymond-Bouchard, I.; Brennan, L.; Touchette, D.; Whyte, L. MinION sequencing from sea ice cryoconites leads to de novo genome reconstruction from metagenomes. *Sci. Rep.* **2021**, *11*, 21041. [CrossRef]
67. Miyauchi, S.; Navarro, D.; Grigoriev, I.V.; Lipzen, A.; Riley, R.; Chevret, D.; Grisel, S.; Berrin, J.-G.; Henrissat, B.; Rosso, M.-N. Visual Comparative Omics of Fungi for Plant Biomass Deconstruction. *Front. Microbiol.* **2016**, *7*, 1335. [CrossRef]
68. Sun, X.; He, C.; Fang, Z.; Xiao, Y. Expression and characterization of NADPH-cytochrome P450 reductase from *Trametes versicolor* in *Escherichia coli*. *Sheng Wu Gong Cheng Xue Bao* **2018**, *34*, 1156–1168. [CrossRef]
69. Chen, W.; Lee, M.-K.; Jefcoate, C.; Kim, S.-C.; Chen, F.; Yu, J.-H. Fungal Cytochrome P450 Monooxygenases: Their Distribution, Structure, Functions, Family Expansion, and Evolutionary Origin. *Genome Biol. Evol.* **2014**, *6*, 1620–1634. [CrossRef]
70. Wang, M.; Ruan, R.; Li, H. The completed genome sequence of the pathogenic ascomycete fungus *Penicillium digitatum*. *Genomics* **2021**, *113*, 439–446. [CrossRef]
71. Osbourn, A. Secondary metabolic gene clusters: Evolutionary toolkits for chemical innovation. *Trends Genet.* **2010**, *26*, 449–457. [CrossRef]
72. Xiao, H.; Zhong, J.-J. Production of Useful Terpenoids by Higher-Fungus Cell Factory and Synthetic Biology Approaches. *Trends Biotechnol.* **2016**, *34*, 242–255. [CrossRef]
73. Brandenburger, E.; Gressler, M.; Leonhardt, R.; Lackner, G.; Habel, A.; Hertweck, C.; Brock, M.; Hoffmeister, D. A Highly Conserved Basidiomycete Peptide Synthetase Produces a Trimeric Hydroxamate Siderophore. *Appl. Environ. Microbiol.* **2017**, *83*, e01478-17. [CrossRef] [PubMed]
74. Miethke, M.; Marahiel, M.A. Siderophore-Based Iron Acquisition and Pathogen Control. *Microbiol. Mol. Biol. Rev.* **2007**, *71*, 413–451. [CrossRef] [PubMed]
75. Levasseur, A.; Drula, E.; Lombard, V.; Coutinho, P.M.; Henrissat, B. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol. Biofuels* **2013**, *6*, 41. [CrossRef] [PubMed]
76. Sidar, A.; Albuquerque, E.D.; Voshol, G.P.; Ram, A.F.J.; Vijgenboom, E.; Punt, P.J. Carbohydrate Binding Modules: Diversity of Domain Architecture in Amylases and Cellulases from Filamentous Microorganisms. *Front. Bioeng. Biotechnol.* **2020**, *8*, 871. [CrossRef]
77. Van den Brink, J.; de Vries, R.P. Fungal enzyme sets for plant polysaccharide degradation. *Appl. Microbiol. Biotechnol.* **2011**, *91*, 1477–1492. [CrossRef]
78. Justo, A.; Hibbett, D.S. Phylogenetic classification of *Trametes* (Basidiomycota, Polyporales) based on a five-marker dataset. *Taxon* **2011**, *60*, 1567–1583. [CrossRef]
79. Mohanta, T.K.; Bae, H. The diversity of fungal genome. *Biol. Proced. Online* **2015**, *17*, 8. [CrossRef]
80. Muszewska, A.; Steczkiewicz, K.; Stepniewska-Dziubinska, M.; Ginalski, K. Cut-and-Paste Transposons in Fungi with Diverse Lifestyles. *Genome Biol. Evol.* **2017**, *9*, 3463–3477. [CrossRef]
81. Muszewska, A.; Steczkiewicz, K.; Stepniewska-Dziubinska, M.; Ginalski, K. Transposable elements contribute to fungal genes and impact fungal lifestyle. *Sci. Rep.* **2019**, *9*, 4307. [CrossRef]
82. Castanera, R.; Borgognone, A.; Pisabarro, A.G.; Ramírez, L. Biology, dynamics, and applications of transposable elements in basidiomycete fungi. *Appl. Microbiol. Biotechnol.* **2017**, *101*, 1337–1350. [CrossRef]
83. Raina, M.; Ibba, M. tRNAs as regulators of biological processes. *Front. Genet.* **2014**, *5*, 171. [CrossRef] [PubMed]
84. Stajich, J.E. Fungal genomes and insights into the evolution of the kingdom. *Microbiol. Spectr.* **2017**, *5*, 1–25. [CrossRef] [PubMed]
85. Wong, D.W.S. Structure and Action Mechanism of Ligninolytic Enzymes. *Appl. Biochem. Biotechnol.* **2009**, *157*, 174–209. [CrossRef] [PubMed]

86. Dashtban, M.; Schraft, H.; Syed, T.A.; Qin, W. Fungal biodegradation and enzymatic modification of lignin. *Int. J. Biochem. Mol. Biol.* **2010**, *1*, 36–50. [PubMed]
87. Couturier, M.; Ladevèze, S.; Sulzenbacher, G.; Ciano, L.; Fanuel, M.; Moreau, C.; Villares, A.; Cathala, B.; Chaspoul, F.; Frandsen, K.E.; et al. Lytic xylan oxidases from wood-decay fungi unlock biomass degradation. *Nat. Chem. Biol.* **2018**, *14*, 306–310. [CrossRef] [PubMed]

4.1.1 Material suplementar capítulo I

Figure S1: Workflow used for the genome sequencing and assembly of *Trametes villosa* CCMB561.

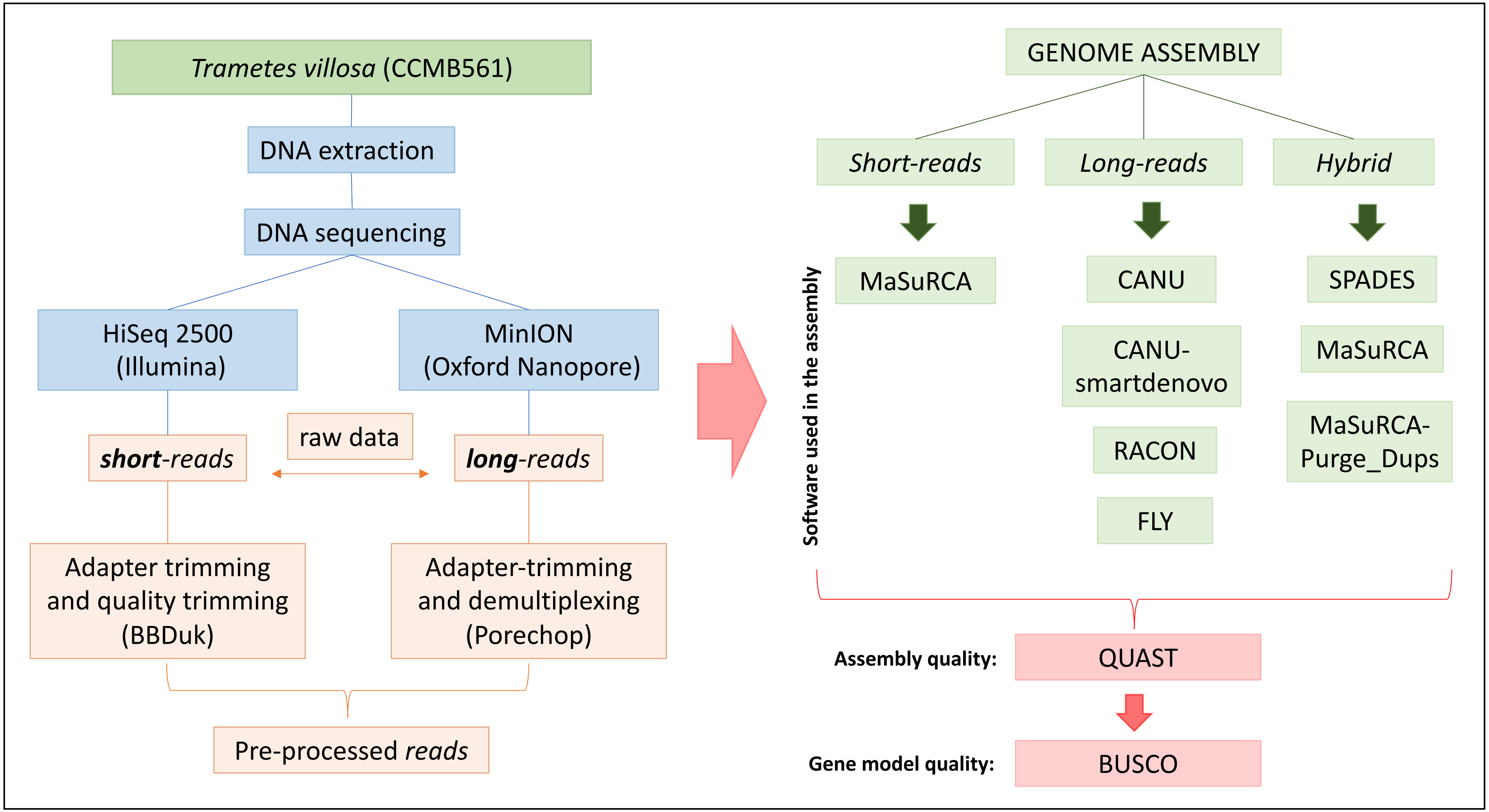
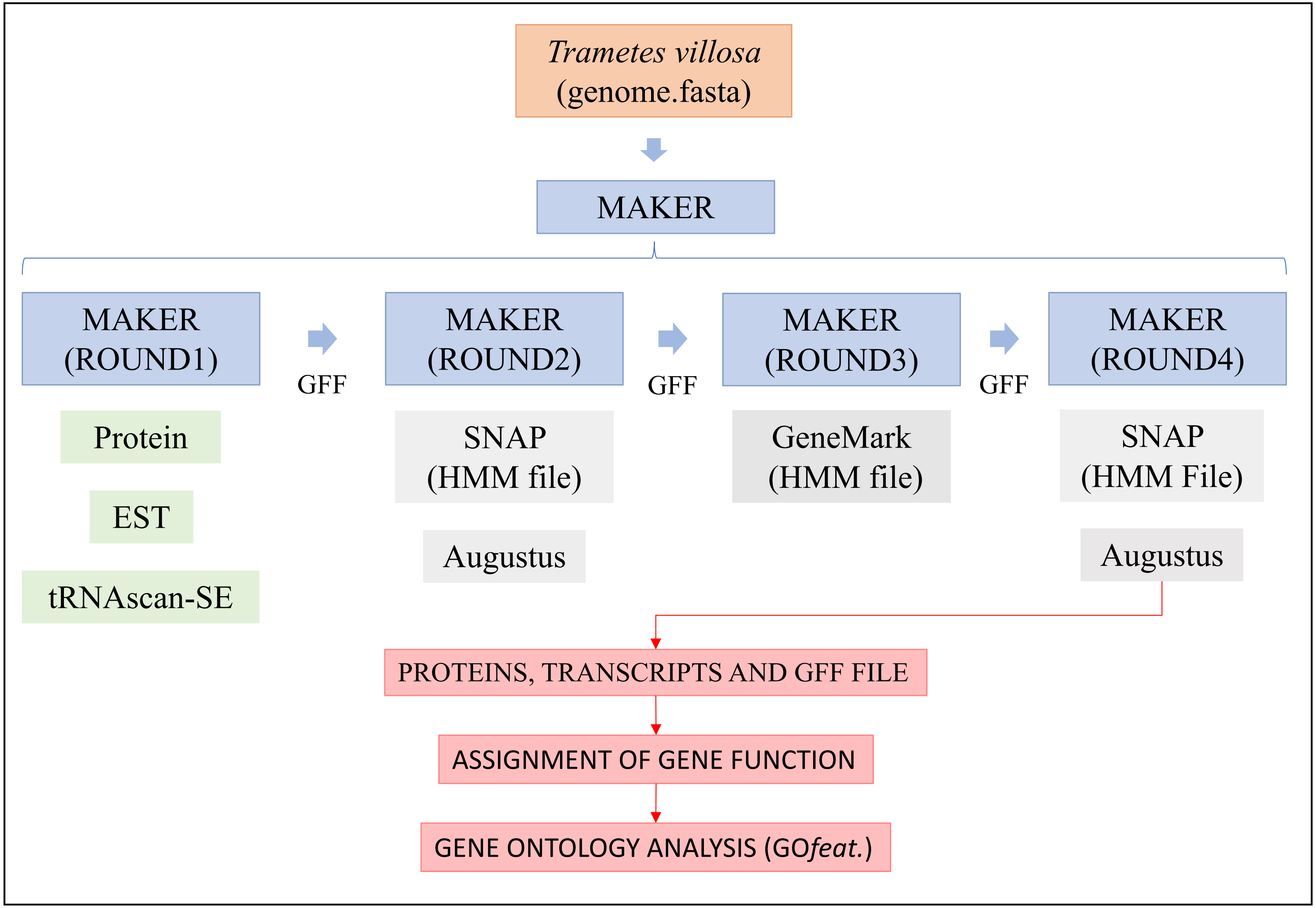
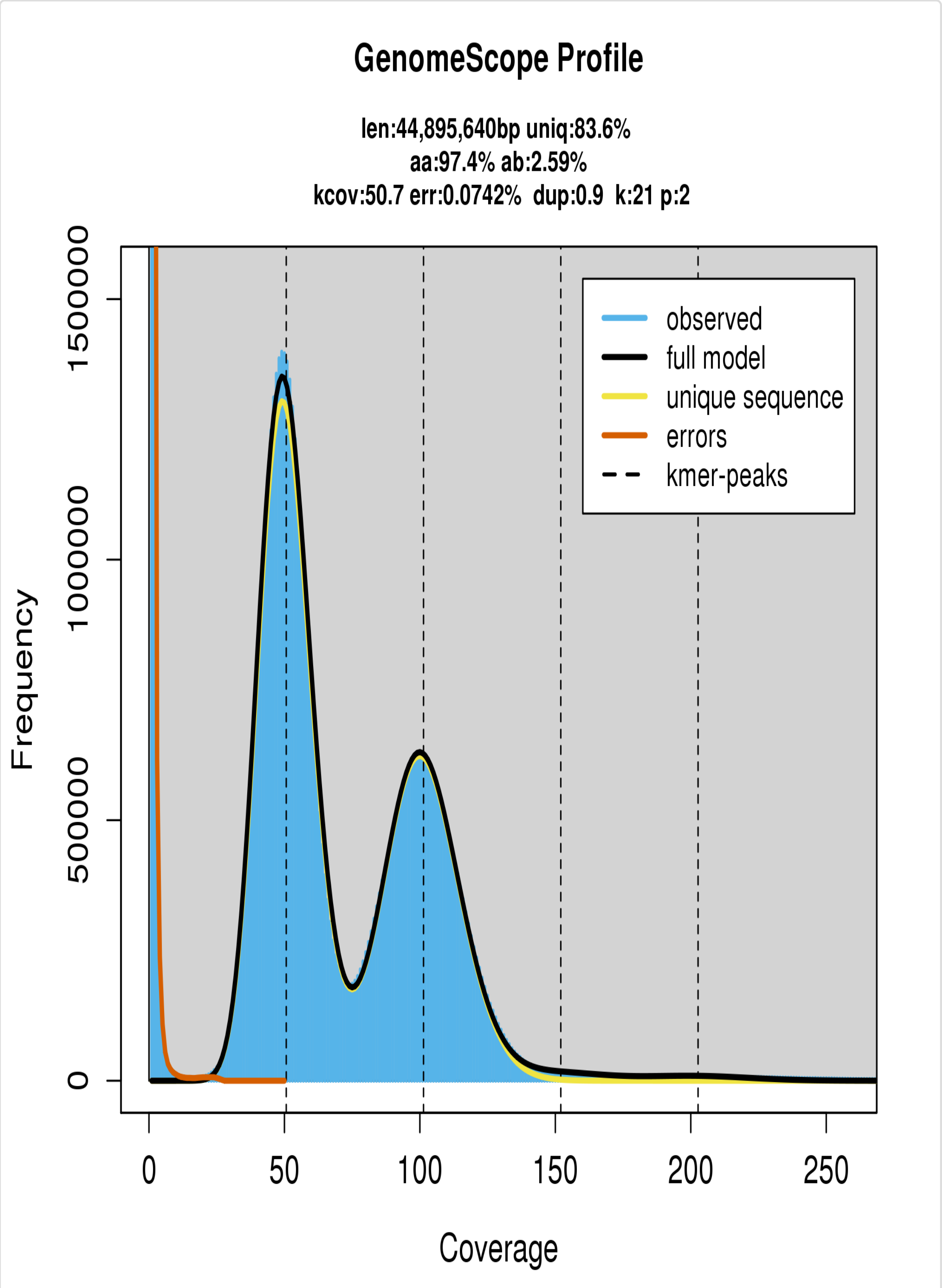
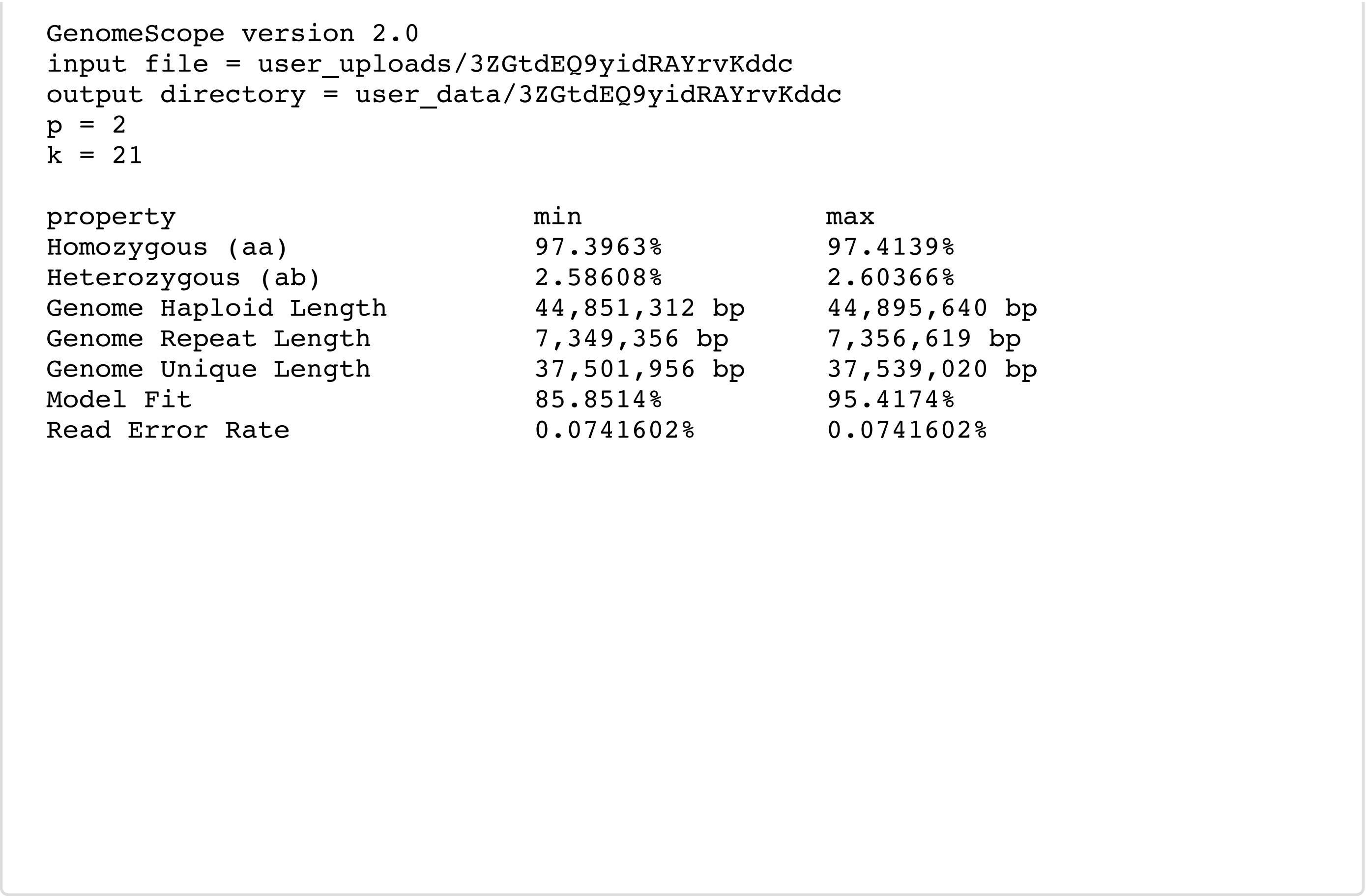


Figure S2: Gene and protein annotation pipeline used for all *Trametes* species.



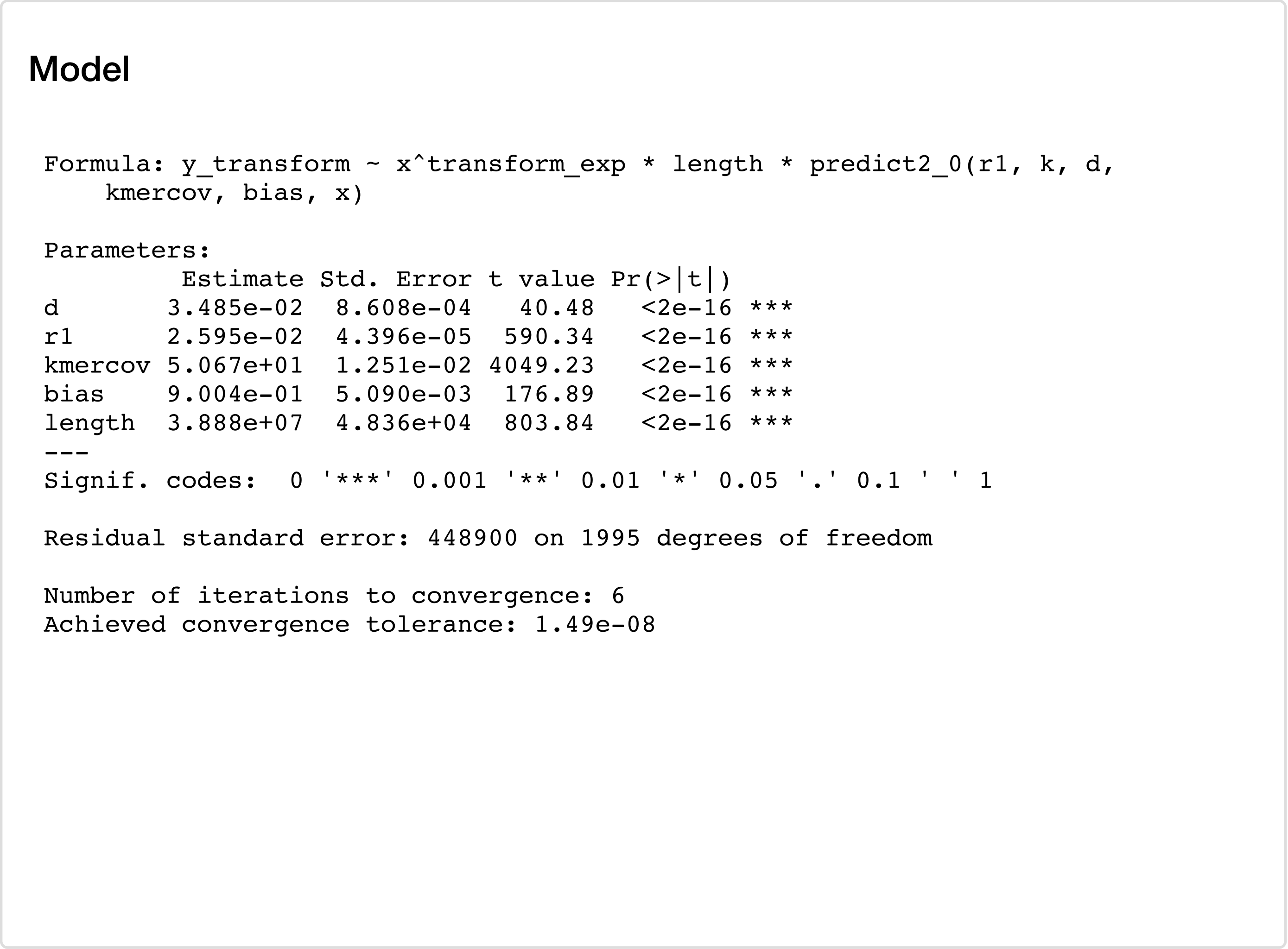
Data S1: GenomeScope results for the *Trametes villosa* CCMB561.





```
GenomeScope version 2.0
input file = user_uploads/3ZGtdEQ9yidRAYrvKddc
output directory = user_data/3ZGtdEQ9yidRAYrvKddc
p = 2
k = 21

property                      min               max
Homozygous (aa)               97.3963%          97.4139%
Heterozygous (ab)             2.58608%          2.60366%
Genome Haploid Length         44,851,312 bp     44,895,640 bp
Genome Repeat Length          7,349,356 bp      7,356,619 bp
Genome Unique Length          37,501,956 bp     37,539,020 bp
Model Fit                     85.8514%          95.4174%
Read Error Rate               0.0741602%        0.0741602%
```



Model

```
Formula: y_transform ~ x^transform_exp * length * predict2_0(r1, k, d,
    kmercov, bias, x)

Parameters:
         Estimate Std. Error t value Pr(>|t|)
d       3.485e-02  8.608e-04   40.48   <2e-16 ***
r1      2.595e-02  4.396e-05  590.34   <2e-16 ***
kmercov 5.067e+01  1.251e-02 4049.23   <2e-16 ***
bias    9.004e-01  5.090e-03  176.89   <2e-16 ***
length  3.888e+07  4.836e+04  803.84   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 448900 on 1995 degrees of freedom

Number of iterations to convergence: 6
Achieved convergence tolerance: 1.49e-08
```

Table S1: Gene annotation statistics generated by GAG software for all genomes from the *Trametes* genus analyzed in our study.

	Trametes villosa	*Trametes versicolor*	*Trametes pubescens*	*Trametes polyzona*	*Trametes elegans*	*Trametes hirsuta*	*Trametes ljubarskyi*	*Trametes cinnabarina*	*Trametes sanguinea*	*Trametes coccinea*
Number of genes	14,540	14,172	14,153	12,203	10,725	11,490	10,821	11,069	13,143	11,900
Number of exons	86,516	85,492	80,744	68,461	64,521	68,379	66,939	68,002	79,285	70,306
Number of introns	71,976	71,317	66,588	56,257	53,796	56,888	56,118	56,932	66,142	58,406
Mean gene length	2,121	2,092	1,903	1,779	1,933	2,018	2,003	1,971	1,934	1,920
Mean exon length	265	253	257	246	247	261	248	245	239	249
Mean intron length	112	115	95	88	91	96	92	93	99	94
Mean CDS length	1,572	1,514	1,455	1,379	1,487	1,551	1,535	1,504	1,443	1,469
% of genome covered by genes	66,0	66,2	67,8	59,3	62,7	61,9	63,7	64,8	69,5	69,7
% of genome covered by CDS	48,9	47,9	51,8	45,9	48,2	47,6	48,8	49,5	51,8	53,4
Mean exons per mRNA	6	6	6	6	6	6	6	6	6	6
Mean introns per mRNA	5	5	5	5	5	5	5	5	5	5

Data S2: Annotation of Secondary metabolite gene clusters (SMGCs) of *Trametes villosa* CCMB561.

SCAFFOLD	**Region**	**Type**	**From (nt)**	**To (nt)**	**Length (pb)**	**Most similar known cluster**	**Similarity**
scf7180000000754	Region 1.1	NRPS-like.betalactone	2.689.142	2.739.489	50.347		
	Region 1.2	terpene	6.146.859	6.160.834	13.975		
	Region 1.3	terpene	7.210.875	7.234.355	23.480		
scf7180000000756	Region 3.1	T1PKS	2.717.862	2.765.765	47.903		
	Region 3.2	terpene	3.769.644	3.791.600	21.956		
scf7180000000757	Region 4.1	NRPS-like	214.949	259.831	44.882		
scf7180000000766	Region 10.1	T1PKS	1.596.405	1.646.002	49.597		
scf7180000000770	Region 14.1	T1PKS	375.001	420.386	45.385		
scf7180000000772	Region 16.1	terpene	389.192	401.842	12.650		
	Region 16.2	terpene	521.292	547.243	25.951		
scf7180000000775	Region 18.1	terpene	289.272	310.238	20.966		
	Region 18.2	terpene	368.950	407.608	38.658		
scf7180000000778	Region 20.1	terpene	114.240	142.390	28.150		
	Region 20.2	terpene	518.993	538.358	19.365		
	Region 20.3	terpene	998.034	1.009.271	11.237		
scf7180000000783	Region 24.1	NRPS-like	241.039	284.086	43.047		
scf7180000000809	Region 33.1	NRPS	884.529	939.022	54.493	basidioferrin	100%
scf7180000000829	Region 47.1	NRPS-like	235.203	316.210	81.007		
scf7180000000841	Region 57.1	NRPS-like	390.868	436.736	45.868		
scf7180000000858	Region 67.1	terpene	170.272	192.007	21.735		
scf7180000000899	Region 84.1	NRPS-like	116.317	159.626	43.309		
	Region 84.2	NRPS-like	649.522	694.761	45.239		

4.2 CAPÍTULO II: Pure lignin induces overexpression of cytochrome P450 (CYP) encoding genes and brings insights into the lignocellulose depolymerization by *Trametes villosa*.

Pure lignin induces overexpression of cytochrome P450 (CYP) encoding genes and brings insights into the lignocellulose depolymerization by *Trametes villosa*

Luiz Marcelo Ribeiro Tomé [a+], Mariana Teixeira Dornelles Parise [a,b], Doglas Parise [a,b], Vasco Ariston de Carvalho Azevedo [b], Bertram Brenig [c], Fernanda Badotti [d], Aristóteles Góes-Neto [a*].

[a]Molecular and Computational Biology of Fungi Laboratory, Department of Microbiology, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte 31270-901, MG, Brazil; marcelofsa_rt@hotmail.com; m.dornelles19@gmail.com; doglasparise@gmail.com; arigoesneto@icb.ufmg.br.

[b] Laboratório de Genética Celular e Molecular, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte 31270-901, MG, Brazil; m.dornelles19@gmail.com; doglasparise@gmail.com; vascoariston@gmail.com.

[c] Institute of Veterinary Medicine, Burckhardtweg, University of Göttingen, 37073 Göttingen, Germany; bbrenig@gwdg.de.

[d] Department of Chemistry, Centro Federal de Educação Tecnológica de Minas Gerais, Belo Horizonte 30421-169, MG, Brazil; fbadotti@outlook.com.

+ First author
* Corresponding author (arigoesneto@icb.ufmg.br)

Highlights:

- *Trametes villosa* CCMB561 has the ability to grow in lignin as the sole carbon source;
- Lignin and sugarcane bagasse induce the expression and secretion of ligninolytic enzymes in *Trametes villosa* CCMB561;
- Pure lignin induces significant alterations in the global gene expression of *Trametes villosa* CCMB561;
- Cytochrome P450 was up-regulated in lignin medium and probably plays essential role in lignin breakdown;
- Specific clusters of CAZymes-encoding genes were up-regulated according to the carbon source.

ABSTRACT

Trametes villosa is a remarkable white-rot fungus (WRF) with potential to be applied in lignocellulose conversion to obtain chemical compounds and biofuels. Lignocellulose breakdown by WRF is carried out through the secretion of oxidative and hydrolytic enzymes. Despite the existing knowledge about this process, the complete molecular mechanisms involved in the regulation of this metabolic system have not yet been elucidated. Therefore, in order to understand the genes and metabolic pathways regulated during lignocellulose degradation, the strain *T. villosa* CCMB561 was cultured in media with different carbon sources (lignin, sugarcane bagasse, and malt extract). Subsequently, biochemical assays and differential gene expression analysis by qPCR and high-throughput RNA sequencing were carried out. Our results revealed the ability of *T. villosa* CCMB561 to grow on lignin (AL medium) as the unique carbon source. An overexpression of Cytochrome P450 was detected in this medium, which is associated with lignin O-demethylation pathway. Clusters of up-regulated CAZymes-encoding genes were identified in lignin and sugarcane bagasse, revealing that *T. villosa* CCMB561 acts simultaneously in the depolymerization of lignin, cellulose, hemicellulose, and pectin. Furthermore, genes encoding nitroreductases and homogentisate-1,2-dioxygenase that act in the degradation of organic pollutants were up-regulated in the lignin medium. Altogether, these findings provide new insights into the mechanisms of lignocellulose degradation by *T. villosa* and confirmed the ability of this fungal species to be applied in biorefineries and in the bioremediation of organic pollutants.

Keywords: Lignocellulose; CAZymes; transcriptomics; white-rot fungi, renewable energy; sustainability.

1 – Introduction

The lignocellulosic biomass (LB), one of the main sources of carbon and energy on Earth, has increased as a potential raw-material to obtain biofuels, biochemicals, and biomaterials [1–3]. LB is also a promising input for replacing non-renewable fossil fuels (coal, oil, and natural gas) and reducing greenhouse gas emissions [4]. The estimated annual world production of lignocellulosic wastes is 181.5 billion tons; however, only a small portion of that amount (8.2 billion tons) is annually used [5,6]. Lignocellulose is mainly composed of cellulose (40-50%), hemicellulose (25-30%) and lignin (15-25%); and the proportion of each biopolymer vary according to the plant species [7].

Cellulose is a natural and linear carbohydrate composed of D-glucose residues linked by β-1,4-glycosidic bonds [7–9]. Hemicellulose is chemically characterized as a branched heteropolymer composed of pentoses (D-xylose and D-arabinose) and/or hexoses (D-mannose, D-glucose and D-galactose). This polymer is often acetylated and their branches composed of L-arabinose, D-xylose, D-galactose, uronic acid, ferulic acid, and fucose monomers [7–9]. Lignin is a phenolic polymer composed of three phenylpropane units called monolignols, p-coumaryl, coniferyl, and sinapyl alcohols, linked by C-C, ester, and ether bonds [10]. This is the most recalcitrant and external compound of the plant cell wall, and is associated with cellulose and hemicellulose through hydrogen and covalent bonds, forming a heterogeneous structure [7,8,11,12].

The depolymerization of lignocellulosic biomass results in a plethora of marketable chemicals, such as biobased ethanol, biomethane, lactic acid, succinic acid, xylitol, sorbitol, phenols, vanillin, vanillic acid, among others, which have high-value for different economic sectors, such as in food, biofuels and pharmaceutical industries [3,13,14]. Nevertheless, the main challenge for lignocellulose to be widely used is the depolymerization of the highly recalcitrant lignin, which limits the saccharification of cellulose and hemicellulose to monomeric compounds (e.g., glucose and xylose) [3,7,13,15,16].

A select group of organisms known as wood-decay fungi has the potential to efficiently convert recalcitrant lignocellulosic biomass into monomers through the production of oxidative and hydrolytic enzymes [8,16–18]. Wood-decay fungi play a central role in carbon cycling in forest ecosystems and are mainly classified as white-rot or brown-rot fungi, depending on the type of degradation carried out [8,17,19].

Brown-rot fungi (BRF) depolymerize preferentially cellulose and hemicellulose, while the lignin fraction is modified only to some extent [17,20]. White-rot fungi (WRF) have the ability to breakdown all wood constituents simultaneous or selectively, with some species showing both degradation patterns [17,21]. WRF are considered the most effective lignocellulosic degraders due to their ability to depolymerize a large amount of lignin to CO_2 and H_2O, facilitating the access of cellulases and hemicellulases to the carbohydrates of plant cell wall [20].

White-rot fungi degrade lignin by secreting oxidative enzymes that are classified as oxidases and peroxidases, which belong to auxiliary activity oxidoreductases families (AA) [22]. Oxidases include laccases (AA1 family), and peroxidases include H_2O_2-dependent enzymes (AA2): Lignin Peroxidase (LiP), Manganese Peroxidase (MnP), and Versatile Peroxidase (VP). Cellulose is degraded by enzymes belonging to the glycoside hydrolases (GH) and AA families; hemicelluloses by enzymes belonging to the AA, carbohydrate esterases (CE), and GH families; and pectin by CE, GH, and pectate lyases (PL) families [9,22–24]. These enzyme families are classified according to the CAZy database (Carbohydrate Active Enzymes database) and besides acting in the degradation of biomass, they can also be applied to the bioremediation of organic pollutants and xenobiotics without producing toxic byproducts [23–25].

Trametes villosa is a WRF, and prior studies have demonstrated its amazing potential to effectively act in the degradation of lignin. Silva et al., (2014) [26] demonstrated that crude enzymatic extract produced from the isolate *T. villosa* CCMB561 can efficiently act in the delignification of agro-industrial residues. This activity was related to the high capacity of the fungi to produce Manganese Peroxidase (MnP), which was reached by optimizing the cultivation conditions. Similarly, Carneiro et al. [27] showed that *T. villosa* CCMB561 is able to produce Lignin Peroxidase (LiP) and Manganese Peroxidase (MnP) in culture media containing sugarcane bagasse as carbon source. Studies with other *T. villosa* isolates also demonstrated the potential of this fungal species to produce peroxidases and cellulases [28–30].

Recently, the complete genome of *T. villosa* CCMB561 was sequenced using the high-throughput sequencing platforms Illumina HiSeq and Oxford Nanopore MinION (hybrid assembly strategy) [31]. The obtained high-quality genome revealed an arsenal of genes encoding peroxidases, cellulases, hemicellulases, and pectinases that act synergistically in the degradation of all lignocellulose polymers. Furthermore, the comparative genomic analyses showed that *T. villosa* CCMB561 was the species with the highest number of genes encoding lignin-modifying enzymes when compared to other species of the same genus [31].

Therefore, to deeply understand the gene expression profile and metabolic pathways related to lignocellulose degradation in *T. villosa* CCMB561 we cultivated it in pure lignin, sugarcane bagasse, and malt extract. Subsequently, we performed biochemical assays, relative expression analysis by qPCR, transcriptome sequencing, differential gene expression analysis (DGE) and co-expression networks analysis. Taking together, the results described here will contribute to the improvement of enzyme production and use of agro-industrial residues for the production of chemical compounds and biofuels.

2 – Material and methods

2.1 – Evaluation of *Trametes villosa* CCMB561 growth and ligninolytic enzymes production

The isolate *T. villosa* CCMB561 was cultivated in six culture media containing different concentrations of alkali lignin (Sigma-Aldrich, St. Louis, Missouri, USA) with or without supplementation (Table 1). The capacity to produce ligninolytic enzymes was evaluated using the Remazol Brilliant Blue R dye (RBBR, Sigma-Aldrich, St. Louis, Missouri, USA), which was added to all media at final concentration of 0.02%. The lignin concentration was determined based on the amount of lignin present in sugarcane bagasse, according to Carneiro et al. [27]. The media were transferred to 90 x 15 mm Petri dishes and the fungal isolate inoculated and incubated at 28 ± 2°C for seven days. All the experiments were carried out with six replicates per culture medium (36 experiments in total), and the discoloration halo was measured on the third, fifth, and seventh days of incubation. Statistical analyses were performed using GraphPad Prism version 7.0 (GraphPad Software, San Diego, California, USA) applying analysis of variance (ANOVA) followed by the post-test of Bonferroni.

Table 1 – Different culture media compositions used to evaluate *T. villosa* CCMB561 growth capacity and ligninolytic enzymes production.

Culture media	Composition
AL20	2.4 g of agar, 0.48 g of alkali lignin and 120 mL of distilled water.
AL30	2.4 g of agar, 0.72 g of alkali lignin and 120 mL of distilled water.
ALG20	2.4 g of agar, 0.48 g of alkali lignin, 0.6 g of glucose and 120 mL of distilled water.
ALG30	2.4 g of agar, 0.72 g of alkali lignin, 0.6 g of glucose and 120 mL of distilled water.
ALSM	2.4 g of agar, 0.48 g of alkali lignin, 3.048 g of manganese sulfate and 120 mL of distilled water.
ALSA	2.4 g of agar, 0.48 g of alkali lignin, 0.6 g of ammonium sulfate and 120 mL of distilled water

2.2 – *Trametes villosa* CCMB561 growth conditions for enzymatic quantification, relative expression analysis by qPCR and transcriptome sequencing

The isolate *T. villosa* CCMB561 was inoculated into 50 mL Erlenmeyers flasks containing the following liquid culture media: AL (0.12 g of alkali lignin and 30 mL of water), BG (0.6 g of sugarcane bagasse and 30 mL of water) and MEA (0.015 of yeast extract, 0.06 g of malt extract and 30 mL of water). The AL media was selected based on the results obtained from the previously described experiment (section 2.1). The BG media was used as a model condition, since agro-industrial residues are a cheap and available source of lignocellulose to be used on an industrial scale, and MEA, a widely used media for fungal growth, worked as a control condition, since it is a non-inducing substrate. After fungal inoculation, the Erlenmeyers were incubated without rotation at 28 ± 2 °C. All experiments were performed with five replicates per culture media.

Biochemical assays, quantitative PCR, and transcriptome sequencing were carried out on the seventh day of growth. This choice was based on the findings of Miyauchi et al. [32] who described that common molecular functions of the fungus *Pycnoporus coccineus* (Current name: *Trametes coccinea*) were triggered in the initial phase of growth (day 3) while specific responses to the lignocellulosic substrates were identified at a later stage (day 7).

2.3 – Manganese Peroxidase and Laccase quantifications

On the seventh day of cultivation, an aliquot of 10 mL from the culture media (AL, BG, and MEA) was removed from each erlenmeyer flask and transferred to microtubes of 2 mL. The tubes were centrifuged at 15,300 g for 2 min at 4° C and the supernatant used to quantify the enzymes. The Laccase (Lac) activity was determined according to the method described by D'Agostini et al. using the 2,2'-azino-bis-(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS-1 mmol. 1-1, $C_{18}H_{24}N_6O_6S_4$, Sigma-Aldrich®) as substrate [33]. The Manganese Peroxidase (MnP) activity was determined through the oxidation of phenol red in the presence of hydrogen peroxide, following the method described by Kuwahara et al. [34]. Absorbances of 420 nm and 610 nm were used to evaluate the activity of Lac and MnP, respectively. The production of both enzymes was expressed in U/L applying the equation $U/L = \Delta Abs \times 10^6 / \varepsilon \times R \times t$, according to Almeida De Menezes et al. [35]. The statistical analyses were carried out using the software GraphPad Prism version 7.0 (GraphPad Software, San Diego, California, USA) applying analysis of variance (ANOVA) followed by Tukey's post-test.

2.4 – RNA extraction

On the seventh day of cultivation, the fungal mycelium was separated from the culture medium, transferred to 2 mL microtubes and macerated using zirconium beads in FastPrep (MP Biomedicals, Santa Ana, California, USA). Total RNA was extracted using the Fastzol reagent (Quatro G, Porto Alegre, Rio Grande do Sul, Brazil) according to the manufacturer’s recommendations. RNA samples were qualitatively analyzed by agarose gel electrophoresis 1%, and quantitatively by spectrophotometry Nanodrop 1000 ND (Thermo Scientific, Waltham, Massachusetts, USA) and automated capillary electrophoresis TapeStation System (Agilent Technologies, Santa Clara, California, USA).

2.5 – Relative expression of genes encoding ligninolytic enzymes by real time quantitative PCR (qPCR)

2.5.1 *Primers design and synthesis*

Specific primers (Table 2) for the amplification of Lacase (Lac), Lignin peroxidase (LiP), Manganese peroxidase (MnP) and Actin (actin1) were designed based on the genome and functional annotations of *T. villosa* CCMB561 [31]. The Geneious® software version 9.0.5 (Auckland, New Zealand) and the OligoAnalyzer™ tool (https://www.idtdna.com/) were used to design and analyze the primers, respectively. Primers were synthesized by the Integrated DNA Technologies (IDT, Coralville, Iowa, USA) and tested by conventional PCR and qPCR.

Table 2 – Primers used to evaluate gene expression of Laccase (Lac), Lignin Peroxidase (LiP) and Manganese Peroxidase (MnP) enzymes.

Primer ID	Sequence 5’- 3’
Lac_F	5’- ACCCCTTCCAYYTGCACGGB – 3’
Lac_R	5’- ATGTGGCAGTGGAGGAACCA – 3’
LiP_F	5’- TTCCACACBCCCGACCAGATC – 3’
LiP_R	5’- SGAAGAACTGCGTGTCCCAV – 3’
MnP_F	5’- TTCGACACSCAGTTCTTCAT – 3’
MnP_R	5’- TTGACSAAVGACTGCCACTCGC – 3’
actin1_F (constitutive gene)	5’- ACAACTCGATCTACAAGTGC – 3’
actin1_R (constitutive gene)	5’- CGACGATCTTGACCTTCATA- 3’

2.5.2 *Quantitative Polymerase Chain Reaction (qPCR)*

RNA samples were treated with Dnase I enzyme (New England BioLabs, Ipswich, Massachusetts, USA) and the complementary DNA (cDNA) synthesized using the iScript cDNA Synthesis Kit (Bio-Rad, Hercules, California, USA). Quantitative PCR (qPCR) was performed using the SsoAdvanced Universal SYBR Green Supermix kit (Bio-Rad, Hercules, California, USA) and the primers described in Table 2. The PCR reaction was carried out in the Applied Biosystems 7900HT Fast Real Time PCR System equipment (Thermo Fisher Scientific, Waltham, Massachusetts, USA) with the following amplification cycles: 50°C for 2 minutes, 95°C for 10 minutes and 40 cycles of 95°C for 15 seconds and 60° C for 1 minute, followed by a dissociation stage for recording the melting curve. For each growth condition, quantitative PCR was performed for five biological and two technical replicates.

2.5.3 *Relative gene expression quantification and statistical analyses*

The relative expression of the target genes was analyzed by the $2^{-\Delta\Delta Ct}$ method using the housekeeping gene encoding actin (actin1) as endogenous reference. Statistical analyses of gene expression were performed using the GraphPad Prism software version 7.00 (GraphPad Software, San Diego, California, USA) applying the analysis of variance (ANOVA) and the Tukey post-test.

2.6 – RNA library preparation and transcriptome sequencing

The library preparation was carried out using 1 µg of total RNA from each sample (section 2.4). Initially, the NEBNext Poly(A) mRNA Magnetic Isolation Module kit (New England BioLabs) was used for messenger RNA (mRNA) selection (enrichment) and ribosomal RNA (rRNA) depletion. Non-strand-specific library construction was performed using the NEBNext Ultra^{TM} II RNA Library Prep Kit for Illumina and the NEBNext Multiplex Oligos for Illumina. Single-end sequencing was performed on the NextSeq 500 Sequencing System platform. A total of nine libraries were sequenced, corresponding to the growth of the isolate *T. villosa* CCMB561 in the three different culture media (AL, BG, and MEA) in triplicate.

2.7 – Transcriptome analyses

The raw reads quality was assessed using the FastQC software (https://github.com/s-andrews/FastQC). The BBDuk tool was used to remove the adapters and trimming bases with Phred score < 20 (https://sourceforge.net/projects/bbmap/). The trimmed reads were mapped to the genome of *T. villosa* CCMB561 to verify the percentage of reads mapped using the Bowtie2 software [36]. Differential gene expression (DGE) analysis was performed according to the protocol described by Yalamanchili et al. [37]. Briefly, reads were mapped to the reference genome using the Tophat2 software [38] with the support of the Bowtie2 program. Subsequently, the HTSeq script was used to quantify the number of reads mapped to each gene [39]. The genes with counts less than ten were excluded, and the differential expression analysis was performed in the R software using the DESeq2 package for normalization, differential analysis of the high-dimensional count data, and visualization. The abundance of transcripts was also obtained using transcript expression levels as TPM (Transcript per Million) using the Kallisto software [40], since this metric normalizes the sequencing data by both gene length and sequencing depth, which improves the comparison between samples [41]. Co-expression networks were built using the R package WGCNA [42,43] and the count matrices (from each sample) generated with the HTSeq python script. The CAZymes encoding genes differentially expressed had their predicted proteins analyzed using WoLF PSORT and SignalP 6.0 software for subcellular location prediction and signal peptide identification, respectively [44,45].

3 – Results and Discussion

3.1 – Growth conditions analysis of *Trametes villosa* CCMB561

The RBBR discoloration test confirmed that *T. villosa* CCMB561 is able to grow in culture media containing only lignin as carbon source and to produce ligninolytic enzymes (AL20 and AL30) (Figure 1a and 1b). Studies with *Trametes versicolor*, a species phylogenetically close to *T. villosa* [31], demonstrated its ability to grow in medium with lignin as unique source of energy [46]. Nevertheless, a basal medium containing different nutrients, such as magnesium sulfate, ammonium nitrate, dipotassium hydrogen phosphate was used for supplementation. Herein, the AL20 and AL30 media had only distilled water and lignin (Table 1).

The discoloration halo measurement results revealed that lignin media supplemented with glucose (ALG20 and ALG30) and ammonium sulfate (ALSA) did not increase the fungal growth and size of the discoloration halo, when compared with media containing only lignin (AL20 and AL30). Still, medium supplemented with manganese sulfate (ALSM) inhibited the fungal growth and production of ligninolytic enzymes, which can be observed by the reduced halo of discoloration throughout the days (Figure 1a and 1b).

The supplementation of culture media with chemical compounds to stimulate fungal growth and enzyme activity is still a controversial issue, with some species requiring supplementation and others not [28,47,48]. Yamanaka et al. reported that the initial Mn^{2+} concentration did not induce the production and activity of the MnP enzyme, nor stimulate the activity of laccase in *T. villosa* CCB176 [28].

Therefore, these results suggested that the AL20 medium, which contain the lowest amount of lignin and no supplementation, induced the production of ligninolytic enzymes equal or better than the other conditions. This is a promising result, since lignin is the most recalcitrant biopolymer of the plant cell wall, and consequently, the main limiting factor for the wide use of plant biomass and lignin valorization [16,49]. Furthermore, the ability to growth on lignin without supplementation decreases the costs for ligninolytic enzymes production and reveals the potential of this isolate to grow in plant biomass residues.

3.2 – Quantification of Lac and MnP

From the screening performed by the RBBR dye discoloration assay (section 3.1), the AL20 culture medium was selected as the most suitable condition to investigate the production of ligninolytic enzymes. In addition, with the aim to evaluate the fungus response to other carbon sources, culture media composed of sugarcane bagasse (BG) and malt extract (MEA) were included in the analyses. Sugarcane bagasse is a largely available and cheap agro-industrial residue in Brazil. Furthermore, previous studies have demonstrated the ability of the isolate *T. villosa* CCMB561 to grow on this substrate [26,27]. The malt extract is a non-inducing medium since it is mainly constituted by easily metabolized sugars.

The medium containing only lignin (AL20) did not induce the production of laccase (Lac) at detectable levels until the seventh day of growth (Figure 1c). The highest production of Lac was observed in the BG medium (average of 141 U/L), followed by MEA (average of 106 U/L) (Figure 1c). A statistical difference in Lac production between BG and MEA conditions was observed ($p < 0.05$). The highest production of MnP was observed in the BG culture medium (average of 8 U/L), followed by AL20 (average of 7 U/L), with no statistical difference between these conditions (Figure 1c). In MEA, a smaller amount of enzyme was produced (5.7 U/L), which differed statistically ($p < 0.05$) from the amounts produced in AL and BG (Figure 1C). The MnP production in BG is supported by previous studies that detected this enzyme in culture media containing lignocellulosic substrates as carbon source [26,27].

Our results revealed the medium containing only lignin (AL) induced the MnP production; however, inhibited the secretion of Lac (Figure 1c). Previous studies with *Trametes togrii*, showed that supplementation of culture medium with lignocellulose (1%) increases the production of β-glucosidase, endoglucanase (EG), cellulase, xylanase, LiP, and MnP [50]. In contrast, the supplementation with glucose (1%) decreases the activity of these enzymes and increases the activity of laccase [50]. Thus, the absence of Lac in AL media may be related to the of lack of glucose, which is present in BG and MEA media.

3.3 – Relative gene expression analysis by qPCR

The dissociation curves of the target genes (Actin1, Lac, LiP and MnP) had homogeneous melting temperature for all the samples amplified (Supplementary figure 1), demonstrating that qPCR reactions were specific and without primer-dimer.

The relative expression of LiP-encoding gene by *T. villosa* CCMB561 was significantly increased ($p < 0.05$) when the fungus was cultivated in AL medium (Figure 1d). Regarding MnP encoding gene, there was an approximately 9-fold increase in the relative expression in AL and BG media, compared to MEA. Laccase gene expression was statistically increased in sugarcane bagasse and no significant difference was observed between AL and MEA (Figure 1d).

The results of relative expression by qPCR applying the $2^{-\Delta\Delta Ct}$ method showed higher expression of LiP and MnP in the AL condition (Figure 1d). In BG, higher expression of MnP and Lac encoding genes were observed. This result corroborates those obtained in the biochemical assays for MnP and Lac production, reinforcing the capacity of *T. villosa* CCMB561 to grown and produce ligninolytic enzymes in pure lignin and sugarcane bagasse. Besides, these results suggest that specific metabolic responses can be triggered according to the carbon source.

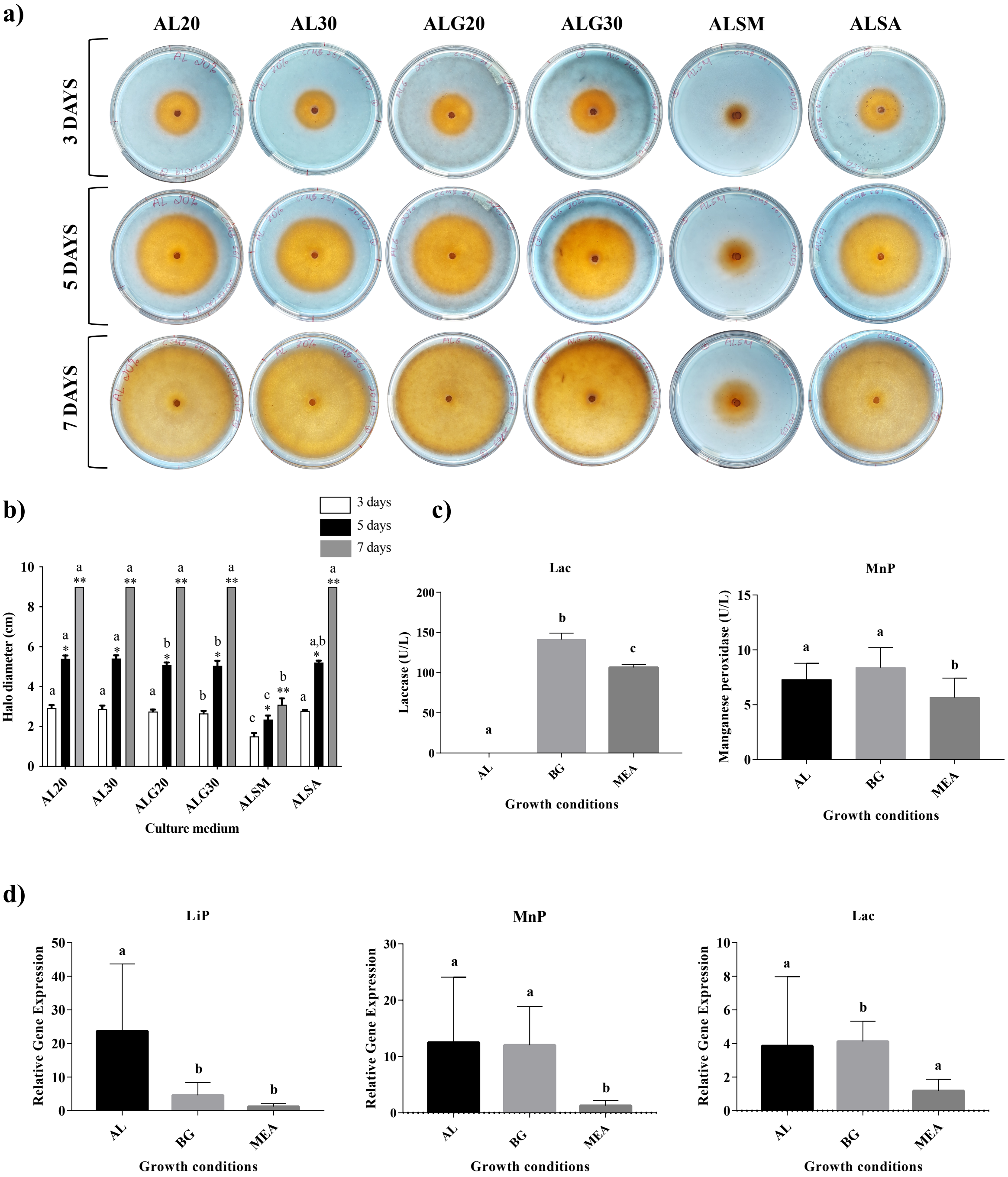


Figure 1 – Biochemical assays and relative gene expression analysis of ligninolytic enzymes of *Trametes villosa* CCMB561. a) RBBR dye discoloration halo on days 3, 5 and 7. b) Diameter of the discoloration halo in RBBR dye. Different letters (a, b, and c) indicate statistically significant difference ($p<0.05$) between groups, symbols (* and **) indicate a statistically significant difference between the days. c) Production of Laccase (Lac) and Manganese Peroxidase (MnP) enzymes by *T. villosa* CCMB561 in AL, BG and MEA media. d) Relative expression quantification of Lignin Peroxidase (LiP), MnP, and Lac encoding genes by *T. villosa* CCMB561 in AL, BG and MEA. Letters in the graphs from **c** and **d** indicate a statistically significant difference in enzymes production and gene relative expression among the culture media on the seventh day of growth.

3.4 – Transcriptome sequencing data and overall gene expression analysis

The RNA samples from *T. villosa* CCMB561 cultivated in AL, BG, and MEA were initially evaluated based on the total RNA capillary electrophoresis gel (TapeStation System). All samples, except for one AL replicate, showed the three characteristic bands corresponding to the 5.8S, 18S, and 28S ribosomal regions (Supplementary Figure 2). The presence of these bands confirms the RNA integrity. Samples for RNA-seq were selected based on quantification by Nanodrop (minimum of 1 μg per sample) and stability of the Cycle Threshold of the actin-encoding gene ($CT \leq 16$).

For high-quality differential gene expression (DGE) analysis using the Illumina platform, 10-30 million reads per sample are required [51]. Herein, the trimmed libraries had a total number of reads ranging from 36,609,047 to 43,947,185 (Table 3). An important parameter for assessing the overall accuracy of RNA sequencing is the rate of reads mapped to the reference genome or transcriptome. For the human genome, a mapping ratio of 70 to 90% is expected [52]. In this study, the sequences mapping rate in the reference genome of *T. villosa* CCMB561 ranged from 84.8 to 88.1% (Table 3).

In order to visualize the transcriptomes clustering from the different culture conditions tested, Principal Component Analysis (PCA) and correlation heatmap plots were generated (Figure 2a and 2b). The heatmap exhibited that the triplicates of each condition tested (AL, BG, and MEA) grouped together (white color) and differed from the other conditions. In the color scale, the bluer, the greater the difference in the expression profile among the samples (Figure 2a). The PCA graph corroborates the results depicted in the heatmap, showing that the triplicates of each condition differed less among them when compared with other coming from a different cultivation media (Figure 2b). It was also observed that the transcriptomes of the fungus cultivated in AL showed greater variation (PC1 with 60% of variation) than the transcriptomes from BG and MEA (PC2 with 22% of variation) (Figure 2b).

The MA-plot is a scatterplot of log2 fold change values (y-axis) versus the average of normalized counts (x-axis) (Figure 2c and 2d). From the analysis of this plot, it was possible to globally evaluate the differentially expressed genes in all analyzed samples. Figure 2c shows the MA-plot of the transcriptome of *T. villosa* CCMB561 cultivated in AL versus MEA media. In red, all genes differentially expressed considering an adjusted P-value < 0.01 are highlighted. Comparing AL and MEA, 2283 genes were up-regulated (17%) and 2260 genes were down-regulated (17%). Comparing BG versus MEA (Figure

2d), 1426 genes were up-regulated (11%) and 1401 down-regulated (11%). The most differentially expressed genes among the media tested had no assigned function and were named uncharacterized protein (Figure 2c and 2d).

Table 3 – *Trametes villosa* CCMB561 transcriptomes sequencing and mapping statistics from each culture condition tested.

Libraries	Number of reads	Mapping ratio (%)
AL1 (RF1)	41,686,793	84.9
AL2 (RF2)	40,526,137	88.1
AL3 (RF4)	43,947,185	86.1
BG1 (RF7)	38,570,847	84.8
BG2 (RF8)	36,609,047	85.2
BG3 (RF9)	43,335,879	84.8
MEA1 (RF12)	41,096,239	86.1
MEA2 (RF13)	40,709,103	85.6
MEA3 (RF14)	41,498,105	86

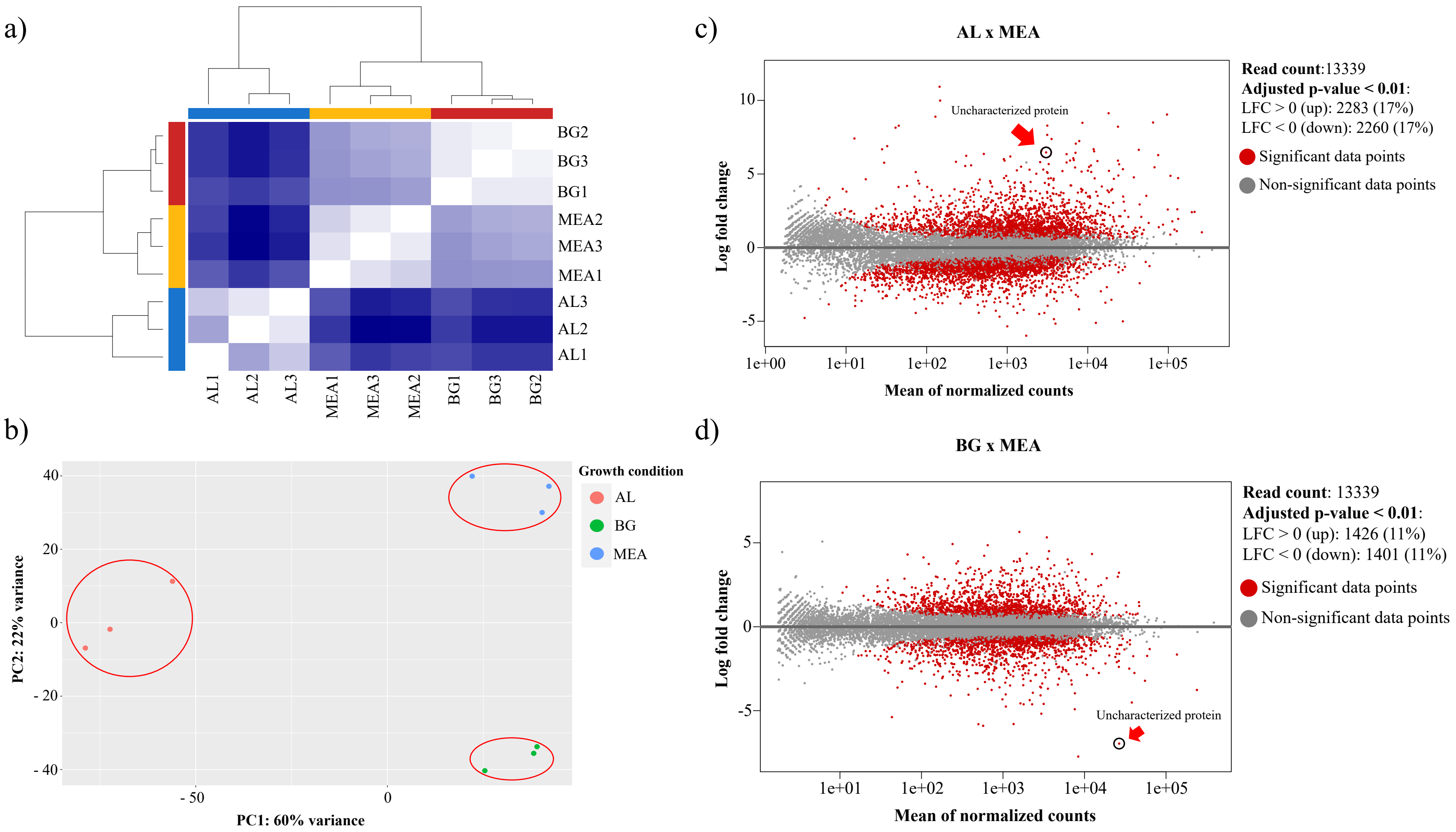


Figure 2 – Visualization of transcriptomes samples clustering and global gene expression. Heatmap clustering (a) and PCA (principal component analysis) plot (b) based on the expression profiles of the nine sequenced samples. MA plot showing genes differentially expressed in AL medium compared to MEA (c), and BG compared to MEA (d). In both MA plots (c-d) the most differentially expressed genes are highlighted (black circle) considering an adjusted P-value < 0.01.

3.5 – Differential gene expression (DGE) and co-expression network analyses

Comparing the transcriptomes of *T. villosa* CCMB561 sequenced in AL and MEA media (Figure 3a), 19 up-regulated genes were detected in AL. Among these, genes encoding Homogentisate 1,2-dioxygenase (n=1), Nitroreductase (n=2), Cytochrome P450 (n=3), Salicylate hydroxylase (n=1), Pirin (n=1), Inorganic phosphate transporter (n=1) and Heat shock protein (n=1). The Cytochrome P450 (CYP450) encoding gene (GENE 5632) had the highest value of log2 fold change (9.02), TPM (2886.41) and number of counts (192,208 reads). Two other genes encoding CYP450 (GENE 202 and 13169) are among the most differentially expressed (up-regulated) in AL medium (Figure 3b and 3c).

The weighted correlation network analysis (WGCNA) applying a minimum correlation of 0.8 generated a co-expression network only when the AL and MEA conditions were compared. The obtained network displays 38 genes and 68 connections (Figure 4). The Cytochrome P450-encoding gene (GENE 5632) showed the highest number of interactions and was co-expressed with 36 genes. The second gene with the highest number of interactions was the Nitroreductase encoding gene (GENE 6498).

Cytochrome P450 is a family of monooxygenases involved in secondary metabolite biosynthetic pathways, fungal adaptation, ecological roles, and organic matter decomposition [53–56]. Regarding lignin breakdown, for this polymer be oxidatively cleaved into ring-opened compounds, it must first be O-demethylated to diols. [53]. Cytochrome P450 are enzymes already reported catalyzing the O-demethylation reaction and also acting in the degradation of organic pollutants and xenobiotics [49,54–57]. Thus, the CYP450 encoding genes significantly expressed in the AL medium, may together be acting in the O-demethylation of lignin and formation of central intermediate compounds. This is an important and limiting step in lignin catabolism and to obtain bioproducts [49,56].

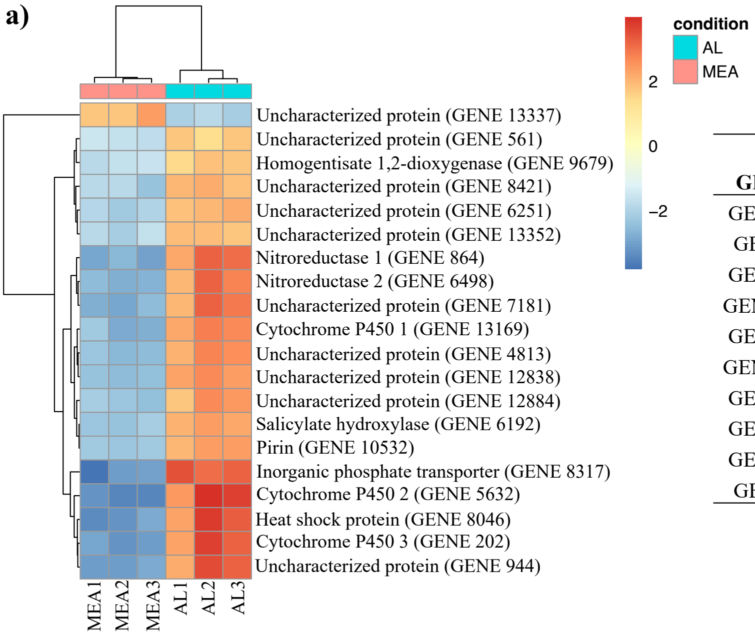
Nitroreductases are enzymes acting in the mineralization of nitroaromatic compounds, and can be used in the degradation of organic pollutants, such as 2,4,6-trinitrotoluene (TNT), which is a highly explosive compound used in the manufacture of bombs [57,58]. Nitroreductases are available for the degradation of nitroaromatics even under nitrogen-limited conditions [58]. In AL culture media, two genes encoding Nitroreductases (GENE 6498 and 864) are among the most differentially expressed (up-regulated) (Figure 3), and the gene 6498 was connected to other 22 genes in the co-expression network analysis (Figure 4).

Interestingly, the genes encoding both, CYP450 and Nitroreductase (GENE 6498) are being co-expressed with an RNA polymerase II transcription factor B (GENE 12516), which is a component of the initiation complex for RNA polymerase II and mRNA transcription [59]. The CYP450 (GENE 5632) is still co-expressed with a DNA helicase encoding gene (GENE 4528), which acts opening the DNA double strand and in the maintenance of nuclear and mitochondrial DNA (Figure 4) [60].

One homogentisate 1,2-dioxygenase (HGT) encoding gene is also among the most expressed in AL medium (*up-regulated*) (Figure 3). HGT was up-regulated in *Phanerochaete chrysosporium* in the presence of vanillin, one of the main intermediates found during lignin depolymerization [61,62]. Another study demonstrated that *Exophiala lecanii-corni* expressed homogentisate-1,2-dioxygenase in a culture medium containing ethylbenzene as carbon source (volatile organic compounds, VOCs) [63]. Thus, the high-expression of HGT in *T. villosa* CCMB561 may be an indication that the isolate is acting in the lignin depolymerization and forming bioproducts, such as vanillin.

In AL medium, two genes encoding Heat-shock proteins (HSPs, GENE 8034 and 11244) were co-expressed with CYP450 and nitroreductases genes (GENE 5632 and 6498) (Figure 4). The 8046 gene, which encodes an HSP, is among the 20 most significantly expressed genes in AL (Figure 3). Heat-Shock proteins play a key role in the physiological regulation of fungi in biotic and/or abiotic stress. HSPs are also involved in different biological functions such as transcription, translation, and post-translational modifications [64].

In BG medium, genes-encoding Lactamase_B domain-containg protein (n=1), Olipeptide transporter 1 (n=1), Lignin Peroxidase (n=1), Aspartic proteinase (n=1) and Cysteine proteinase (n=1) are among the most differentially expressed (up-regulated) (Figure 5). The gene encoding one beta-lactamase had the highest log2fold change value (5.65) (Figure 5b). Interestingly, among the differentially expressed genes, 11 were down-regulated in BG medium, such as the Catalase and Pyranose 2-oxidase (P2O) that presented a log2fold change value of -7.73 and -3.21, respectively. They displayed significantly higher TPM values and number of counts in the MEA culture medium (Figure 5b and 5c). P2O acts on the oxidation of carbohydrates such as D-glucose to generate 2-keto-sugars and H_2O_2 [22,65]. Catalase acts on the decomposition of H_2O_2, a compound that, in excess, may be responsible for the inactivation of Pyranose-2-oxidase [66]. Furthermore, one lignin peroxidase encoding gene (LiP) was identified among the most significantly expressed genes in BG (Figure 5). LiPs are H_2O_2-dependent peroxidases that act in the oxidation of phenolic aromatic compounds [22]. The other genes differentially expressed in this condition are part of the normal fungi metabolism and are not directly related with biomass depolymerization.



b)

GENE ID	Function	log2fold change	TPM (AL)	TPM (MEA)
GENE 9679	Homogentisate 1,2-dioxygenase	4.39	49.03	4.37
GENE 864	Nitroreductase 1	7.65	242.3	2.11
GENE 6498	Nitroreductase 2	7.06	609.26	7.64
GENE 13169	Cytochrome P450 1	6.64	1293.17	21.40
GENE 6192	Salicylate hydroxylase	6.01	35.18	0.82
GENE 10532	Pirin	5.80	174.90	5.31
GENE 8317	Inorganic phosphate transporter	8.49	2177.75	9.70
GENE 5632	Cytochrome P450 2	9.03	2886.41	8.90
GENE 8046	Heat Shock Protein	8.47	1066.51	5.88
GENE 202	Cytochrome P450 3	8.21	835.93	5.20

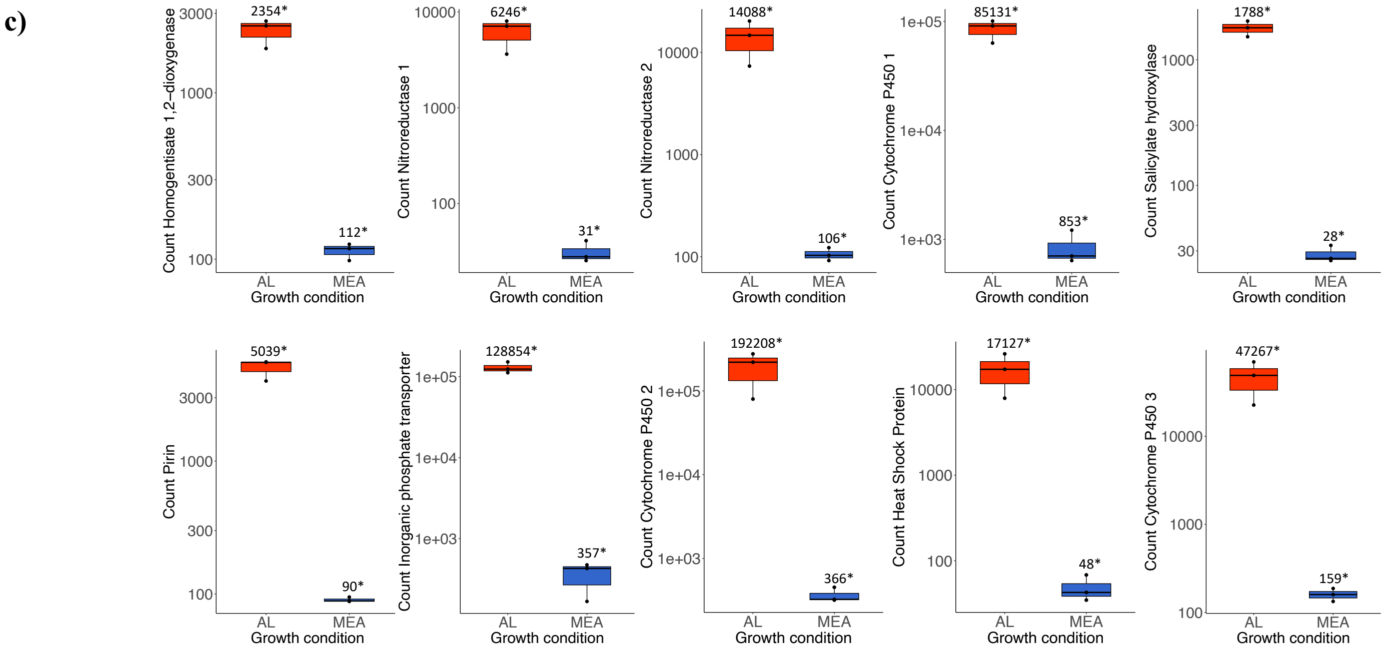


Figure 3 – Genes differentially expressed comparing AL x MEA culture media considering the adjusted P-value < 0.01. a) Heatmap based on the normalized count data (log2 scale/rlogTransformation) considering the amount in which each gene deviates in a specific sample from the average of the gene in all samples. b) Table containing the value of log2fold change and Transcripts Per Million (TPM) for genes that had an assigned function based on the GoFeat and blastp analyses using the NCBI and Uniprot databases. c) Graphs of absolute count number of the differentially expressed genes. The average number of reads mapped in each gene is highlighted with an asterisk.

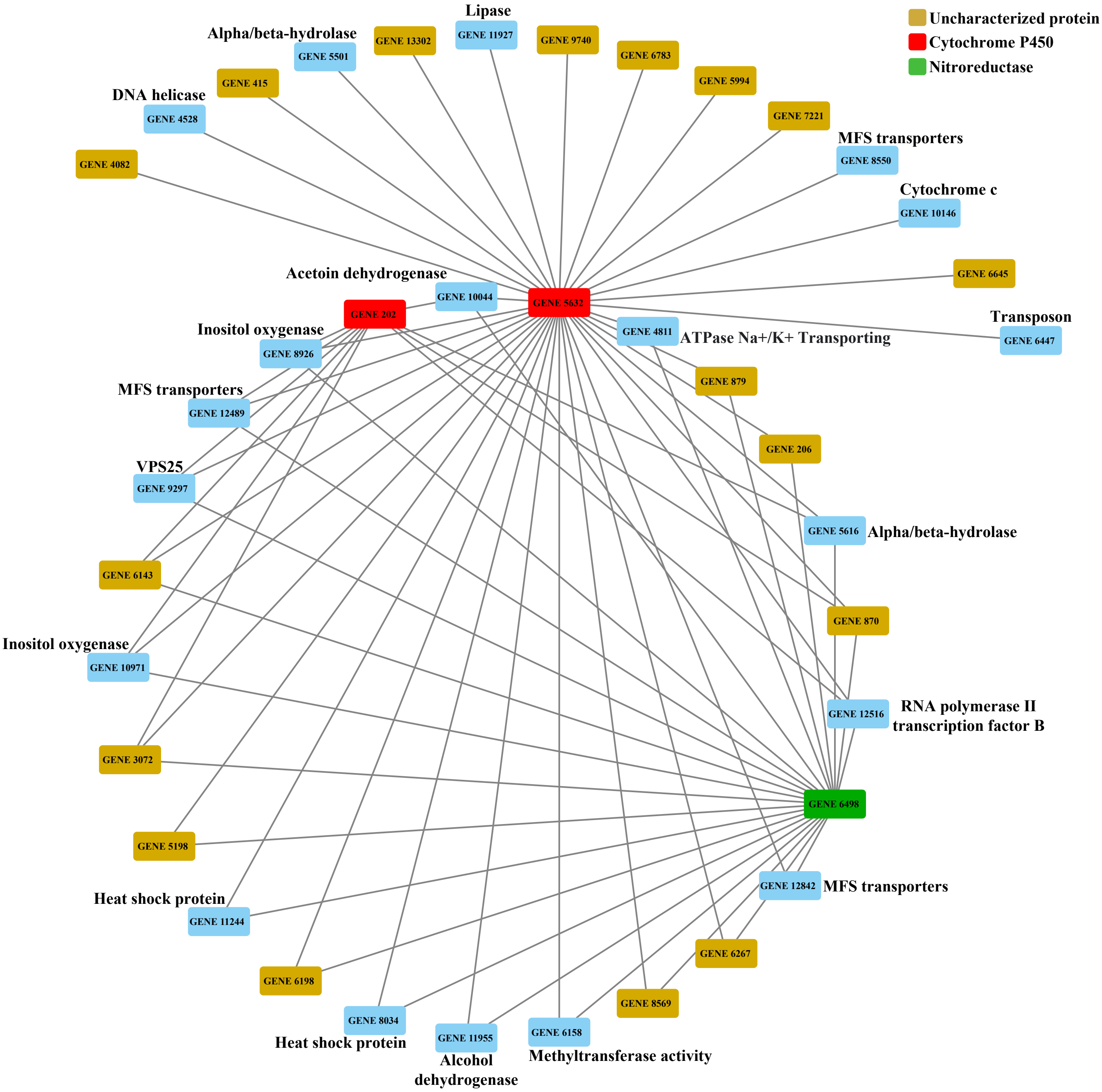
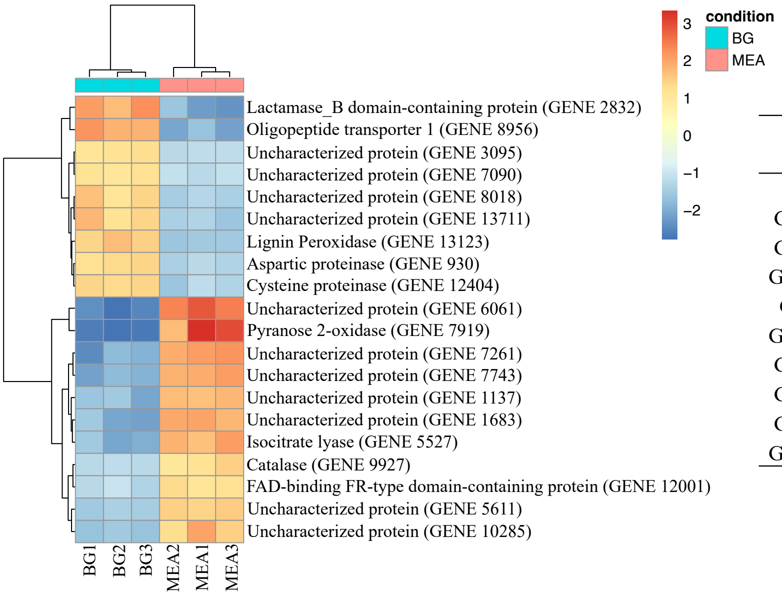


Figure 4 – Co-expression gene network based on transcriptomes of *Trametes villosa* CCMB561 cultivated in AL and MEA media. A minimum correlation value of 0.8 was considered.

a)



b)

GENE ID	Function	log2fold change	TPM (BG)	TPM (MEA)
GENE 2832	Lactamase_B domain-containing protein	5.65	122.62	2.29
GENE 8956	Oligopeptide transporter 1	5.32	102.03	2.40
GENE 13123	Lignin Peroxidase	4.92	18.61	0.46
GENE 930	Aspartic proteinase	3.90	14.87	0.97
GENE 12404	Cysteine proteinase	3.66	224.44	17.32
GENE 7919	Pyranose 2-oxidase	-7.73	1.75	382.58
GENE 5527	Isocitrate lyase	-4.91	9.96	317.65
GENE 9927	Catalase	-3.21	251.77	2386.52
GENE 12001	FAD-binding	-3.22	9.09	88.71

c)

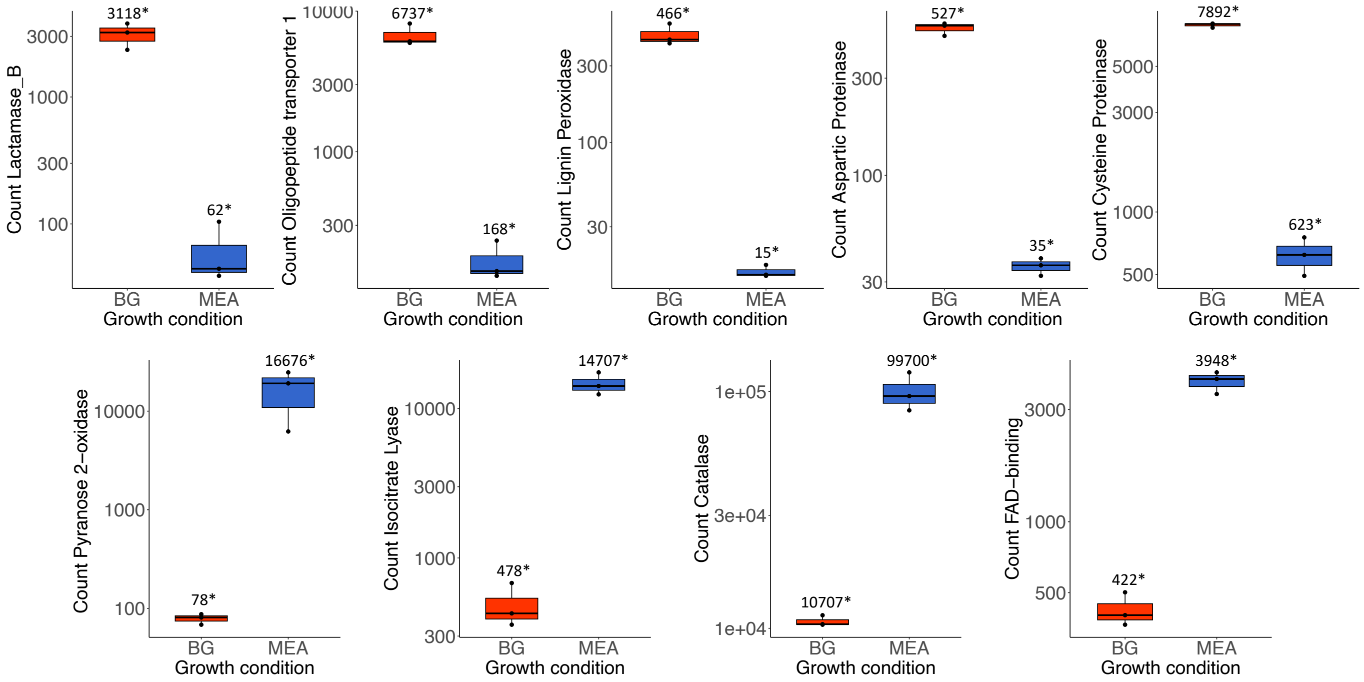


Figure 5 – Genes differentially expressed comparing BG x MEA culture media considering the adjusted P-value < 0.01. a) Heatmap based on the normalized count data (log2 scale/rlogTransformation) considering the amount in which each gene deviates in a specific sample from the average of the gene in all samples. B) Table containing the value of log2fold change and Transcripts Per Million (TPM) for the genes that had an assigned function based on the GoFeat and blastp analyses using the NCBI and Uniprot databases. C) Graphs of absolute count number of the differentially expressed genes. The average number of reads mapped in each gene is highlighted with an asterisk.

3.7 – Expression of genes encoding Carbohydrate-Active enzymes

Based on the study of Tomé et al. [31], 186 genes encoding Carbohydrate-Active enzymes were selected from the genome of *T. villosa* CCMB561, and their expression in the different carbon sources (AL, BG, and MEA) evaluated. Figure 6 shows the differential expression heatmaps of the CAZymes-encoding genes as mentioned.

In pure lignin (AL), genes encoding enzymes that act in the depolymerization of lignin (Glyoxal oxidase, Laccase, Lignin Peroxidase, and Manganese Peroxidase), cellulose (β-1,4-endoglucanases and Exoglucanases), hemicellulose (Endo-1,4-β-xylanases) and pectin (Endo-polygalacturonase) had its expression significantly increased (Table 4) and formed a cluster of 15 genes (Figure 6a). Among them, ten act on lignin depolymerization, and of these, five are LiP-encoding genes. Fourteen (14) genes displayed signal peptide and were identified as extracellular protein-coding ones. Two genes, one encoding for the endo-1,4-β-xylanases and other encoding for exoglucanases (GH6), exhibited the CBM1 domain (Table 4). Although the AL medium does not contain cellulose, hemicellulose and pectin, it induced the expression of genes that act in the degradation of all lignocellulose polymers.

In the BG and MEA media, the CAZymes-encoding genes had more homogeneous expression compared to the AL versus MEA media. In BG medium, a cluster of nine genes had statistically increased expression (adjusted P-value < 0.01) (Figure 6b). Among these, genes encoding proteins that act in the degradation of all lignocellulose constituents (Table 5), such as, LiP, MnP, Lytic polysaccharide monooxygenases, Alcohol oxidase, β-1,4-endoglucanases, β-glucosidases and Endo-polygalacturonase were identified. Seven genes displayed signal peptide and were classified as coding for extracellular proteins. In this cluster, the CBM domain was not identified.

In lignin, the results revealed that the fungus *T. villosa* CCMB561 simultaneously and differentially expressed genes encoding for enzymes that act in the depolymerization of all lignocellulose polymers (Table 4). Five genes encoding LiP were significantly increased in AL medium and two in BG (Table 4 and 5). This finding corroborates those obtained in the analysis of relative expression by qPCR, which demonstrated that the highest LiP expression was achieved in the AL medium. Interestingly, two genes encoding MnP had increased expression in both, AL and BG media, a result that also corroborates the qPCR data.

In the AL medium (Table 4), one Glyoxal oxidase encoding gene (AA5_2) exhibited significantly increased expression. This enzyme plays a central role in the generation of H_2O_2 [22,67]. In BG (Table 5), the gene with significantly increased expression was the Alcohol oxidase encoding gene (AA3), which is also responsible for hydrogen peroxide generation and activation of peroxidases (LiP, MnP and VP) [22].

Regarding hemicellulose degradation in AL medium (Table 4), one gene encoding Endo-1,4-β-xylanase (G10) had significantly increased expression. Endo-β-1,4-endoxylanases act on the degradation of linear chains of β-1,4-linked D-xylose residues [7,9]. In addition, the CMB1 domain was identified in this gene, which promotes the association of enzymes with the substrate and increases the enzymatic hydrolysis and degradation of polysaccharides [9,68]. In BG, a gene coding for Lytic polysaccharide monooxygenases (LPMOs) displayed increased expression. LPMOs oxidatively degrade xylan-type hemiceluloses [69]. In AL and BG media (Table 4 and 5), one Endo-polygalacturonase encoding gene (GH28), that acts on pectin hydrolysis, had its expression significantly increased in a similar way.

Regarding the cellulose depolymerization, a significantly increased expression of β-1,4-endoglucanases (GH12) and Exoglucanases (GH6) was identified in pure lignin. In sugarcane bagasse, a greater expression of β-1,4-endoglucanases and β-glucosidases were detected. β-1,4-endoglucanases act randomly in the internal amorphous region of the cellulose fiber, generating oligosaccharides. Exoglucanases act on reducing and non-reducing ends of oligosaccharides and release glucose and/or cellobiose units. Finally, β-glucosidases act by hydrolyzing the bond between the two glucose units of cellobiose, generating glucose monomers [7–9].

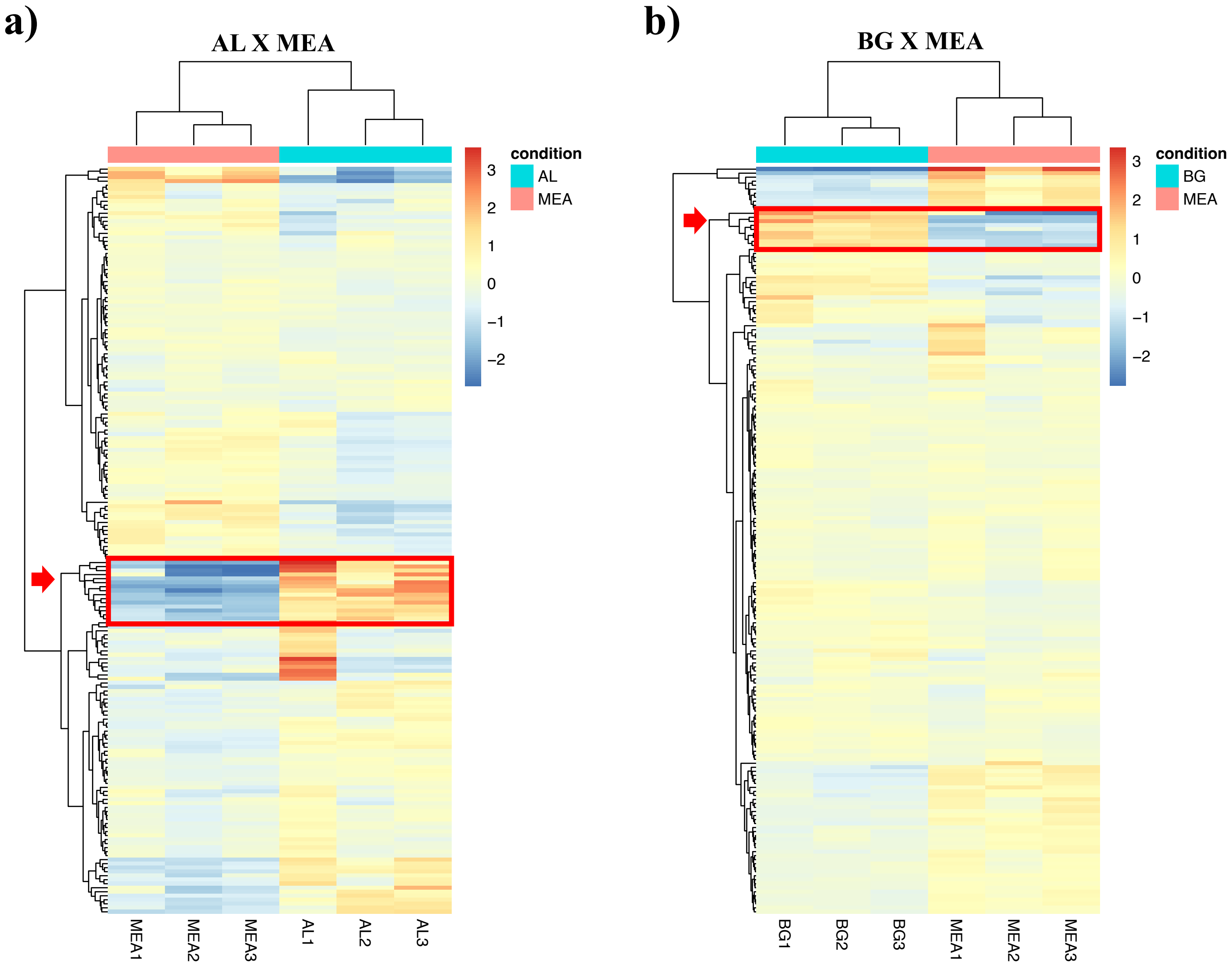


Figure 6 – Heatmap clustering of the differential expression of 186 genes encoding for enzymes belonging to the families of Auxiliary Activity oxidoreductases (AA) Carbohydrate Esterases (CE), Glycosyl Hydrolase (GH), Pectate Lyases (PL) of the *Trametes villosa* CCMB561 grown on different carbon sources. a) Differentially expressed genes comparing AL and MEA. b) Differentially expressed genes comparing BG and MEA.

Table 4 – The more differentially expressed CAZymes-encoding genes (cluster highlighted in red in Figure 6) of *Trametes villosa* CCMB561 cultivated in AL versus MEA.

GENE ID	Activity	Log2Fold change	padj	TPM AL (mean)	TPM MEA (mean)	Counts AL (mean)	Counts MEA (mean)	Signal Peptide (Sec/SPI)	Subcellular Localization	CBM Domain
GENE 7246	Lignin Peroxidase (AA2)	5.58	8.72e-06	23.27	0.82	698	15	YES (21/22)	Extracellular	NO
GENE 10648	Exoglucanases (GH6)	5.91	2.69e-20	242.26	6.02	11863	197	YES (20/21)	Extracellular	CBM1
GENE 5143	Lignin Peroxidase (AA2)	4.43	0.0007	55.73	3.95	2388	110	YES (21/22)	Extracellular	NO
GENE 4057	Manganese peroxidase (AA2)	3.28	0.03	224.62	35.20	8704	892	YES (21/22)	Extracellular	NO
GENE 882	Lignin Peroxidase (AA2)	4.12	1.36e-10	4.26	0.43	192	11	YES (21/22)	Extracellular	NO
GENE 7258	Lignin Peroxidase (AA2)	4.72	1.83e-13	41.62	2.32	1760	67	YES (26/27)	Extracellular	NO
GENE 13123	Lignin Peroxidase (AA2)	5.54	9.41e-35	88.08	0.46	590	13	YES (21/22)	Extracellular	NO
GENE 13714	Glyoxal oxidase (AA5_1)	5.29	2.93e-32	53.45	2.28	3689	94	YES (18/19)	Extracellular	NO
GENE 9985	β-1,4-endoglucanases (GH12)	4.39	7.58e-17	39.69	3.07	1192	57	YES (27/28)	Extracellular	NO
GENE 8352	Endo-1,4-β-xylanases (G10)	3.74	1.40e-26	36.82	4.79	1548	115	YES (19/20)	Extracellular	CBM1
GENE 13331	Laccase (AA1)	4.07	3.61e-25	40.48	3.60	2344	139	NO	Cytoplasmic	NO
GENE 6618	Manganese peroxidase (AA2)	3.04	1.83e-23	13.86	2.53	509	62	YES (21/22)	Extracellular	NO
GENE 3414	Laccase (AA1)	3.51	1.73e-16	329.58	48.50	24375	2130	YES (17/18)	Extracellular	NO
GENE 12198	β-1,4-endoglucanases (GH12)	2.85	1.32e-15	161.79	36.94	6591	908	YES (22/23)	Extracellular	NO
GENE 8952	Endo-polygalacturonase (GH28)	3.19	1.75e-12	12.76	2.00	534	58	YES (19/20)	Extracellular	NO

Table 5 – CAZymes-encoding genes more differentially expressed (cluster highlighted in red in Figure 6) of *Trametes villosa* CCMB561 cultivated in BG versus MEA.

GENE ID	Activity	Log2fold change	padj	TPM BG (mean)	TPM MEA (mean)	Counts BG (mean)	Counts MEA (mean)	Signal Peptide (Sec/SPI)	Subcellular Localization	CBM Domain
GENE 4057	Manganese peroxidase (AA2)	3.25	0.03	352.72	35.20	10388	1088	YES (21/22)	Extracellular	NO
GENE 13123	Lignin peroxidase (AA2)	4.92	8.37e-57	18.61	0.46	466	15	YES (21/22)	Extracellular	NO
GENE 14059	Manganese peroxidase (AA2)	4.50	2.69e-45	220.60	10.09	7024	310	YES (21/22)	Extracellular	NO
GENE 8313	Lytic polysaccharide monooxygenases (AA14)	2.61	1.11e-20	26.26	4.02	527	86	YES (19/20)	Extracellular	NO
GENE 11636	Alcohol oxidase (AA3)	2.16	1.08e-08	130.25	29.79	7154	1592	NO	Cytoplasmic	NO
GENE 907	β-glucosidases (GH1)	3.49	4.97e-45	76.46	7.35	3218	286	NO	Cytoplasmic	NO
GENE 12452	Lignin peroxidase (AA2)	3.17	5.05e-30	25.79	2.5	2086	231	YES (21/22)	Extracellular	NO
GENE 9985	β-1,4-endoglucanases (GH12)	2.75	1.71e-27	21.77	3.07	464	69	YES (27/28)	Extracellular	NO
GENE 8952	Endo-polygalacturonase (GH28)	3.15	8.08e-25	20.62	2.00	627	71	YES (19/20)	Extracellular	NO

5 – Conclusions

Our findings demonstrated that *Trametes villosa* CCMB561 is able to grow using lignin as the sole carbon and energy source, and such condition induces significative changes in the overall gene expression when compared to the other culture conditions evaluated. The overexpression of different genes encoding Cytochrome P450 monooxygenases was remarkable, demonstrating their putative key role in the O-demethylation of lignin for the subsequent depolymerization by oxidases (Lac) and peroxidases (LiP and MnP). Genes encoding nitroreductases and homogentisate-1,2-dioxygenase (HGT), which act in the detoxification of organic pollutants were also differentially expressed (up-regulated). Lignin and sugarcane bagasse also induced the expression of a specific cluster of CAZymes encoding genes, which revealed the ability of *T. villosa* CCMB561 to simultaneously act in the degradation of all lignocellulose polymers. Finally, our findings contribute for the comprehension of the complex genetic mechanisms of lignocellulose breakdown by *T. villosa* CCMB561. In agreement with other studies, we demonstrate here the great potential of this fungal strain to be used for the production of enzymes with wide biotechnological applications.

CRediT authorship contribution statement

Luiz Marcelo Ribeiro Tomé: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Data Curation, Funding acquisition, Writing - Original Draft, Writing - Review & Editing. **Mariana Teixeira Dornelles Parise:** Formal analysis. **Doglas Parise:** Formal analysis. **Vasco Ariston de Carvalho Azevedo:** Resources, Funding acquisition, Supervision. **Bertram Brenig:** Resources, Funding acquisition, Investigation. **Fernanda Badotti:** Supervision, Conceptualization, Resources, Funding acquisition, Writing - Review & Editing, Project administration. **Aristóteles Góes-Neto:** Supervision, Conceptualization, Resources, Funding acquisition, Writing - Review & Editing, Project administration.

Declaration of competing interest

The authors declare that they have no competing interests and affiliations with or involvement in any organization or entity with any financial interest or non-financial interest that could be construed as a potential conflict of interest.

Acknowledgments

We would like to thank the Graduate Programs of Microbiology and Bioinformatics of the Universidade Federal de Minas Gerais (UFMG) and the Department of Chemistry of the Centro Federal de Educação Tecnológica de Minas Gerais (CEFET-MG). This study was funded by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and FAPEMIG (Process number APQ-01550-18). Luiz Marcelo Ribeiro Tomé received a CNPq scholarship to conduct this study (Process number: 140156/2019-4) and Aristóteles Góes-Neto received a research grant for scientific productivity from CNPq (Process number: 310764/2016-5).

References

[1] V. Ashokkumar, R. Venkatkarthick, S. Jayashree, S. Chuetor, S. Dharmaraj, G. Kumar, W.H. Chen, C. Ngamcharussrivichai, Recent advances in lignocellulosic biomass for biofuels and value-added bioproducts - A critical review, Bioresour Technol. 344 (2022). https://doi.org/10.1016/j.biortech.2021.126195.

[2] N. Srivastava, R. Rawat, H. Singh Oberoi, P.W. Ramteke, A Review on Fuel Ethanol Production From Lignocellulosic Biomass, Int J Green Energy. 12 (2015) 949–960. https://doi.org/10.1080/15435075.2014.890104.

[3] F.H. Isikgor, C.R. Becer, Lignocellulosic biomass: a sustainable platform for the production of bio-based chemicals and polymers, Polym Chem. 6 (2015) 4497–4559. https://doi.org/10.1039/C5PY00263J.

[4] A. Devi, S. Bajar, H. Kour, R. Kothari, D. Pant, A. Singh, Lignocellulosic Biomass Valorization for Bioethanol Production: a Circular Bioeconomy Approach, Bioenergy Res. (2022). https://doi.org/10.1007/s12155-022-10401-9.

[5] N. Dahmen, I. Lewandowski, S. Zibek, A. Weidtmann, Integrated lignocellulosic value chains in a growing bioeconomy: Status quo and perspectives, GCB Bioenergy. 11 (2019) 107–117. https://doi.org/10.1111/gcbb.12586.

[6] B. Singh, J. Korstad, A. Guldhe, R. Kothari, Editorial: Emerging Feedstocks & Clean Technologies for Lignocellulosic Biofuel, Front Energy Res. 10 (2022). https://doi.org/10.3389/fenrg.2022.917081.

[7] A.A. Houfani, N. Anders, A.C. Spiess, P. Baldrian, S. Benallaoua, Insights from enzymatic degradation of cellulose and hemicellulose to fermentable sugars– a review, Biomass Bioenergy. 134 (2020) 105481. https://doi.org/10.1016/j.biombioe.2020.105481.

[8] E. Espósito, J.L. de Azevedo, Fungos - Uma Introdução à Biologia, Bioquímica e Biotecnologia, 2ª, EDUCS, Editora da Universidade de Caxias do Sul, Caxias do Sul, 2010.

[9] J. van den Brink, R.P. de Vries, Fungal enzyme sets for plant polysaccharide degradation, Appl Microbiol Biotechnol. 91 (2011) 1477–1492. https://doi.org/10.1007/s00253-011-3473-2.

[10] Z. Luo, S. Qin, S. Chen, Y. Hui, C. Zhao, Selective conversion of lignin to ethylbenzene, Green Chemistry. 22 (2020) 1842–1850. https://doi.org/10.1039/c9gc04246f.

[11] J. Zhang, Y.S. Choi, C.G. Yoo, T.H. Kim, R.C. Brown, B.H. Shanks, Cellulose-hemicellulose and cellulose-lignin interactions during fast pyrolysis, ACS Sustain Chem Eng. 3 (2015) 293–301. https://doi.org/10.1021/sc500664h.

[12] H. Nishimura, A. Kamiya, T. Nagata, M. Katahira, T. Watanabe, Direct evidence for α ether linkage between lignin and carbohydrates in wood cell walls, Sci Rep. 8 (2018). https://doi.org/10.1038/s41598-018-24328-9.

[13] M. Bilal, M.Z. Nawaz, H.M.N. Iqbal, J. Hou, S. Mahboob, K.A. Al-Ghanim, H. Cheng, Engineering Ligninolytic Consortium for Bioconversion of Lignocelluloses to Ethanol and Chemicals, Protein Pept Lett. 25 (2018) 108–119. https://doi.org/10.2174/0929866525666180122105835.

[14] F. Wang, D. Ouyang, Z. Zhou, S.J. Page, D. Liu, X. Zhao, Lignocellulosic biomass as sustainable feedstock and materials for power generation and energy storage, Journal of Energy Chemistry. 57 (2021) 247–280. https://doi.org/10.1016/j.jechem.2020.08.060.

[15] J.U. Hernández-Beltrán, I.O. Hernández-De Lira, M.M. Cruz-Santos, A. Saucedo-Luevanos, F. Hernández-Terán, N. Balagurusamy, Insight into Pretreatment Methods of Lignocellulosic Biomass to Increase Biogas Yield: Current State, Challenges, and Opportunities, Applied Sciences. 9 (2019) 3721. https://doi.org/10.3390/app9183721.

[16] D.W.S. Wong, Structure and Action Mechanism of Ligninolytic Enzymes, Appl Biochem Biotechnol. 157 (2009) 174–209. https://doi.org/10.1007/s12010-008-8279-z.

[17] A.K. Sista Kameshwar, W. Qin, Comparative study of genome-wide plant biomass-degrading CAZymes in white rot, brown rot and soft rot fungi, Mycology. 9 (2018) 93–105. https://doi.org/10.1080/21501203.2017.1419296.

[18] H. Hage, S. Miyauchi, M. Virágh, E. Drula, B. Min, D. Chaduli, D. Navarro, A. Favel, M. Norest, L. Lesage-Meessen, B. Bálint, Z. Merényi, L. Eugenio, E. Morin, A.T. Martínez, P. Baldrian, M. Štursová, M.J. Martínez, C. Novotny, J.K. Magnuson, J.W. Spatafora, S. Maurice, J. Pangilinan, W. Andreopoulos, K. LaButti, H. Hundley, H. Na, A. Kuo, K. Barry, A. Lipzen, B. Henrissat, R. Riley, S. Ahrendt, L.G. Nagy, I. v. Grigoriev, F. Martin, M. Rosso, Gene family expansions and transcriptome signatures uncover fungal adaptations to wood decay, Environ Microbiol. (2021) 1462-2920.15423. https://doi.org/10.1111/1462-2920.15423.

[19] C. Schmidt-Dannert, Biocatalytic portfolio of Basidiomycota, Curr Opin Chem Biol. 31 (2016) 40–49. https://doi.org/10.1016/j.cbpa.2016.01.002.

[20] D. Martinez, J. Challacombe, I. Morgenstern, D. Hibbett, M. Schmoll, C.P. Kubicek, P. Ferreira, F.J. Ruiz-Duenas, A.T. Martinez, P. Kersten, K.E. Hammel, A. vanden Wymelenberg, J. Gaskell, E. Lindquist, G. Sabat, S.S. Bondurant, L.F. Larrondo, P. Canessa, R. Vicuna, J. Yadav, H. Doddapaneni, V. Subramanian, A.G. Pisabarro, J.L. Lavín, J.A. Oguiza, E. Master, B. Henrissat, P.M. Coutinho, P. Harris, J.K. Magnuson, S.E. Baker, K. Bruno, W. Kenealy, P.J. Hoegger, U. Kü Es R, P. Ramaiya, S. Lucas, A. Salamov, H. Shapiro, H. Tu, C.L. Chee, M. Misra, G. Xie, S. Teter, D. Yaver, T. James, M. Mokrejs, M. Pospisek, I. v Grigoriev, T. Brettin, D. Rokhsar, R. Berka, D. Cullen, Genome, transcriptome, and secretome analysis of wood decay fungus Postia placenta supports unique mechanisms of lignocellulose conversion, 2009. http://www.jgi.doe.gov/Postia.

[21] A. Andriani, A. Maharani, D.H.Y. Yanto, H. Pratiwi, D. Astuti, I. Nuryana, E. Agustriana, S.H. Anita, A.B. Juanssilfero, U. Perwitasari, C.F. Pantouw, A.N. Nurhasanah, V.E. Windiastri, S. Nugroho, D. Widyajayantie, J. Sutiawan, Y. Sulistyowati, N. Rahmani, R.A. Ningrum, Yopi, Sequential production of ligninolytic, xylanolytic, and cellulolytic enzymes by Trametes hirsuta AA-017 under different biomass of Indonesian sorghum accessions-induced cultures, Bioresour Technol Rep. 12 (2020) 100562. https://doi.org/10.1016/j.biteb.2020.100562.

[22] A. Levasseur, E. Drula, V. Lombard, P.M. Coutinho, B. Henrissat, Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes, Biotechnol Biofuels. 6 (2013) 41. https://doi.org/10.1186/1754-6834-6-41.

[23] B.L. Cantarel, P.M. Coutinho, C. Rancurel, T. Bernard, V. Lombard, B. Henrissat, The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics, Nucleic Acids Res. 37 (2009) D233–D238. https://doi.org/10.1093/nar/gkn663.

[24] V. Lombard, H. Golaconda Ramulu, E. Drula, P.M. Coutinho, B. Henrissat, The carbohydrate-active enzymes database (CAZy) in 2013, Nucleic Acids Res. 42 (2014) D490–D495. https://doi.org/10.1093/nar/gkt1178.

[25] M. Brazkova, R. Koleva, G. Angelova, H. Yemendzhiev, Ligninolytic enzymes in Basidiomycetes and their application in xenobiotics degradation , BIO Web Conf. 45 (2022) 02009. https://doi.org/10.1051/bioconf/20224502009.

[26] M.L.C. Silva, V.B. de Souza, V. da S. Santos, H.M. Kamida, J.R.T. de Vasconcellos-Neto, A. Góes-Neto, M.G. Bello Koblitz, Production of Manganese Peroxidase by <i>Trametes villosa</i> on Unexpensive Substrate and Its Application in the Removal of Lignin from Agricultural Wastes, Advances in Bioscience and Biotechnology. 05 (2014) 1067–1077. https://doi.org/10.4236/abb.2014.514122.

[27] R.T. de O. Carneiro, M.A. Lopes, M.L.C. Silva, V. da S. Santos, V.B. de Souza, A.O. de Sousa, C.P. Pirovani, M.G.B. Koblitz, R.G. Benevides, A. Góes-Neto, Trametes villosa lignin peroxidase (TvLiP): Genetic and molecular characterization, J Microbiol Biotechnol. 27 (2017) 179–188. https://doi.org/10.4014/jmb.1606.06055.

[28] R. Yamanaka, C.F. Soares, D.R. Matheus, K.M.G. Machado, Lignolytic enzymes produced by Trametes villosa ccb176 under different culture conditions, Brazilian Journal of Microbiology. 39 (2008) 78–84. https://doi.org/10.1590/S1517-83822008000100019.

[29] R.O. Coniglio, M.I. Fonseca, G. v. Díaz, O. Ontañon, S. Ghio, E. Campos, P.D. Zapata, Optimization of cellobiohydrolase production and secretome analysis of Trametes villosa LBM 033 suitable for lignocellulosic bioconversion, Arab J Basic Appl Sci. 26 (2019) 182–192. https://doi.org/10.1080/25765299.2019.1598107.

[30] J.C. Matei, M. Soares, A.C.H. Bonato, M.P.A. de Freitas, C. V. Helm, W. V. Maroldi, W.L.E. Magalhães, C.W.I. Haminiuk, G.M. Maciel, Enzymatic delignification of sugar cane bagasse and rice husks and its effect in saccharification, Renew Energy. 157 (2020) 987–997. https://doi.org/10.1016/j.renene.2020.05.028.

[31] L.M.R. Tomé, F.F. da Silva, P.L.C. Fonseca, T. Mendes-Pereira, V.A. de C. Azevedo, B. Brenig, F. Badotti, A. Góes-Neto, Hybrid Assembly Improves Genome Quality and Completeness of Trametes villosa CCMB561 and Reveals a Huge Potential for Lignocellulose Breakdown, Journal of Fungi. 8 (2022). https://doi.org/10.3390/jof8020142.

[32] S. Miyauchi, D. Navarro, S. Grisel, D. Chevret, J.G. Berrin, M.N. Rosso, The integrative omics of white-rot fungus Pycnoporus coccineus reveals co-regulated CAZymes for orchestrated lignocellulose breakdown, PLoS One. 12 (2017). https://doi.org/10.1371/journal.pone.0175528.

[33] É.C. D'Agostini, T.R.D. Mantovani, J.S. do Valle, L.D. Paccola-Meirelles, N.B. Colauto, G.A. Linde, Low carbon/nitrogen ratio increases laccase production from basidiomycetes in solid substrate cultivation, Sci Agric. 68 (2011) 295–300. https://doi.org/10.1590/S0103-90162011000300004.

[34] M. Kuwahara, J.K. Glenn, M.A. Morgan, M.H. Gold, Separation and characterization of two extracelluar H 2 O 2 -dependent oxidases from ligninolytic cultures of Phanerochaete chrysosporium, FEBS Lett. 169 (1984) 247–250. https://doi.org/10.1016/0014-5793(84)80327-0.

[35] T. Almeida De Menezes, A. Simões Da, R. Bispo, M.G. Bello Koblitz, L. Porto De Souza Vandenberghe, H.M. Kamida, A. Góes-Neto, Production of Basidiomata and Ligninolytic Enzymes by the Lingzhi or Reishi Medicinal Mushroom, Ganoderma lucidum (Agaricomycetes), in Licuri (Syagrus coronata) Wastes in Brazil, 2016. www.begellhouse.com.

[36] B. Langmead, S.L. Salzberg, Fast gapped-read alignment with Bowtie 2, Nat Methods. 9 (2012) 357–359. https://doi.org/10.1038/nmeth.1923.

[37] H.K. Yalamanchili, Y. Wan, Z. Liu, Data Analysis Pipeline for RNA-seq Experiments: From Differential Expression to Cryptic Splicing, Curr Protoc Bioinformatics. 59 (2017). https://doi.org/10.1002/cpbi.33.

[38] D. Kim, G. Pertea, C. Trapnell, H. Pimentel, R. Kelley, S.L. Salzberg, TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions, Genome Biol. 14 (2013) 1–13.

[39] S. Anders, P.T. Pyl, W. Huber, HTSeq-A Python framework to work with high-throughput sequencing data, Bioinformatics. 31 (2015) 166–169. https://doi.org/10.1093/bioinformatics/btu638.

[40] N.L. Bray, H. Pimentel, P. Melsted, L. Pachter, Near-optimal probabilistic RNA-seq quantification, Nat Biotechnol. 34 (2016) 525–527. https://doi.org/10.1038/nbt.3519.

[41] Y. Zhao, M.C. Li, M.M. Konaté, L. Chen, B. Das, C. Karlovich, P.M. Williams, Y.A. Evrard, J.H. Doroshow, L.M. McShane, TPM, FPKM, or Normalized Counts? A Comparative Study of Quantification Measures for the Analysis of RNA-seq Data from the NCI Patient-Derived Models Repository, J Transl Med. 19 (2021). https://doi.org/10.1186/s12967-021-02936-w.

[42] P. Langfelder, S. Horvath, WGCNA: An R package for weighted correlation network analysis, BMC Bioinformatics. 9 (2008). https://doi.org/10.1186/1471-2105-9-559.

[43] B. Zhang, S. Horvath, A general framework for weighted gene co-expression network analysis, Stat Appl Genet Mol Biol. 4 (2005). https://doi.org/10.2202/1544-6115.1128.

[44] P. Horton, K.J. Park, T. Obayashi, N. Fujita, H. Harada, C.J. Adams-Collier, K. Nakai, WoLF PSORT: Protein localization predictor, Nucleic Acids Res. 35 (2007). https://doi.org/10.1093/nar/gkm259.

[45] F. Teufel, J.J. Almagro Armenteros, A.R. Johansen, M.H. Gíslason, S.I. Pihl, K.D. Tsirigos, O. Winther, S. Brunak, G. von Heijne, H. Nielsen, SignalP 6.0 predicts all five types of signal peptides using protein language models, Nat Biotechnol. 40 (2022) 1023–1025. https://doi.org/10.1038/s41587-021-01156-3.

[46] S. Gottlieb, W. Day, M. Pelczar Jr, The biological degradation of lignin. II. The adaptation of white-rot fungi to growth on lignin media., Phytopathology. 40 (1950) 926–935.

[47] H. Kamitsuji, Y. Honda, T. Watanabe, M. Kuwahara, Mn2+ is dispensable for the production of active MnP2 by Pleurotus ostreatus, Biochem Biophys Res Commun. 327 (2005) 871–876. https://doi.org/10.1016/j.bbrc.2004.12.084.

[48] C. Palma, A.T. Martínez, J.M. Lema, M.J. Martínez, Different fungal manganese-oxidizing peroxidases: a comparison between Bjerkandera sp. and Phanerochaete chrysosporium, J Biotechnol. 77 (2000) 235–245. https://doi.org/10.1016/S0168-1656(99)00218-7.

[49] E.S. Ellis, D.J. Hinchen, A. Bleem, L. Bu, S.J.B. Mallinson, M.D. Allen, B.R. Streit, M.M. Machovina, Q. V. Doolin, W.E. Michener, C.W. Johnson, B.C. Knott, G.T. Beckham, J.E. McGeehan, J.L. DuBois, Engineering a Cytochrome P450 for Demethylation of Lignin-Derived Aromatic Aldehydes, JACS Au. 1 (2021) 252–261. https://doi.org/10.1021/jacsau.0c00103.

[50] Y. Liu, Y. Wu, Y. Zhang, X. Yang, E. Yang, H. Xu, Q. Yang, I. Chagan, X. Cui, W. Chen, J. Yan, Lignin degradation potential and draft genome sequence of Trametes trogii S0301, Biotechnol Biofuels. 12 (2019) 256. https://doi.org/10.1186/s13068-019-1596-3.

[51] R. Stark, M. Grzelak, J. Hadfield, RNA sequencing: the teenage years, Nat Rev Genet. 20 (2019) 631–656. https://doi.org/10.1038/s41576-019-0150-2.

[52] A. Conesa, P. Madrigal, S. Tarazona, D. Gomez-Cabrero, A. Cervera, A. McPherson, M.W. Szcześniak, D.J. Gaffney, L.L. Elo, X. Zhang, A. Mortazavi, A survey of best practices for RNA-seq data analysis, Genome Biol. 17 (2016). https://doi.org/10.1186/s13059-016-0881-8.

[53] S.J.B. Mallinson, M.M. Machovina, R.L. Silveira, M. Garcia-Borràs, N. Gallup, C.W. Johnson, M.D. Allen, M.S. Skaf, M.F. Crowley, E.L. Neidle, K.N. Houk, G.T. Beckham, J.L. DuBois, J.E. McGeehan, A promiscuous cytochrome P450 aromatic O-demethylase for lignin bioconversion, Nat Commun. 9 (2018) 2487. https://doi.org/10.1038/s41467-018-04878-2.

[54] M.F. Khan, C.D. Murphy, Cytochrome P450 5208A3 is a promiscuous xenobiotic biotransforming enzyme in Cunninghamella elegans, Enzyme Microb Technol. 161 (2022) 110102. https://doi.org/10.1016/j.enzmictec.2022.110102.

[55] W. Chen, M.K. Lee, C. Jefcoate, S.C. Kim, F. Chen, J.H. Yu, Fungal cytochrome P450 monooxygenases: Their distribution, structure, functions, family expansion, and evolutionary origin, Genome Biol Evol. 6 (2014) 1620–1634. https://doi.org/10.1093/gbe/evu132.

[56] S.J.B. Mallinson, M.M. Machovina, R.L. Silveira, M. Garcia-Borràs, N. Gallup, C.W. Johnson, M.D. Allen, M.S. Skaf, M.F. Crowley, E.L. Neidle, K.N. Houk, G.T. Beckham, J.L. Dubois, J.E. McGeehan, A promiscuous cytochrome P450 aromatic O-demethylase for lignin bioconversion, Nat Commun. 9 (2018). https://doi.org/10.1038/s41467-018-04878-2.

[57] H. Harms, D. Schlosser, L.Y. Wick, Untapped potential: Exploiting fungi in bioremediation of hazardous chemicals, Nat Rev Microbiol. 9 (2011) 177–192. https://doi.org/10.1038/nrmicro2519.

[58] S. Rieble, D.K. Joshi, M.H. Gold, Aromatic Nitroreductase from the Basidiomycete Phanerochaete chrysosporium, Biochem Biophys Res Commun. 205 (1994) 298–304. https://doi.org/10.1006/bbrc.1994.2664.

[59] S. Buratowski, H. Zhou, Functional domains of transcription factor TFIIB., Proceedings of the National Academy of Sciences. 90 (1993) 5633–5637. https://doi.org/10.1073/pnas.90.12.5633.

[60] A.K. Byrd, K.D. Raney, Structure and function of Pif1 helicase, Biochem Soc Trans. 45 (2017) 1159–1171. https://doi.org/10.1042/BST20170096.

[61] T. Mori, G. Koyama, H. Kawagishi, H. Hirai, Effects of Homologous Expression of 1,4-Benzoquinone Reductase and Homogentisate 1,2-Dioxygenase Genes on Wood Decay in Hyper-Lignin-Degrading Fungus Phanerochaete sordida YK-624, Curr Microbiol. 73 (2016) 512–518. https://doi.org/10.1007/s00284-016-1089-6.

[62] M. Shimizu, N. Yuda, T. Nakamura, H. Tanaka, H. Wariishi, Metabolic regulation at the tricarboxylic acid and glyoxylate cycles of the lignin-degrading basidiomycetePhanerochaete chrysosporium against exogenous addition of vanillin, Proteomics. 5 (2005) 3919–3931. https://doi.org/10.1002/pmic.200401251.

[63] C.K. Gunsch, Q. Cheng, K.A. Kinney, P.J. Szaniszlo, C.P. Whitman, Identification of a homogentisate-1,2-dioxygenase gene in the fungus Exophiala lecanii-corni: analysis and implications, Appl Microbiol Biotechnol. 68 (2005) 405–411. https://doi.org/10.1007/s00253-005-1899-0.

[64] S. Tiwari, R. Thakur, J. Shankar, Role of Heat-Shock Proteins in Cellular Function and in the Biology of Fungi, Biotechnol Res Int. 2015 (2015) 1–11. https://doi.org/10.1155/2015/132635.

[65] T. Wongnate, P. Chaiyen, The substrate oxidation mechanism of pyranose 2-oxidase and other related enzymes in the glucose-methanol-choline superfamily, FEBS Journal. 280 (2013) 3009–3027. https://doi.org/10.1111/febs.12280.

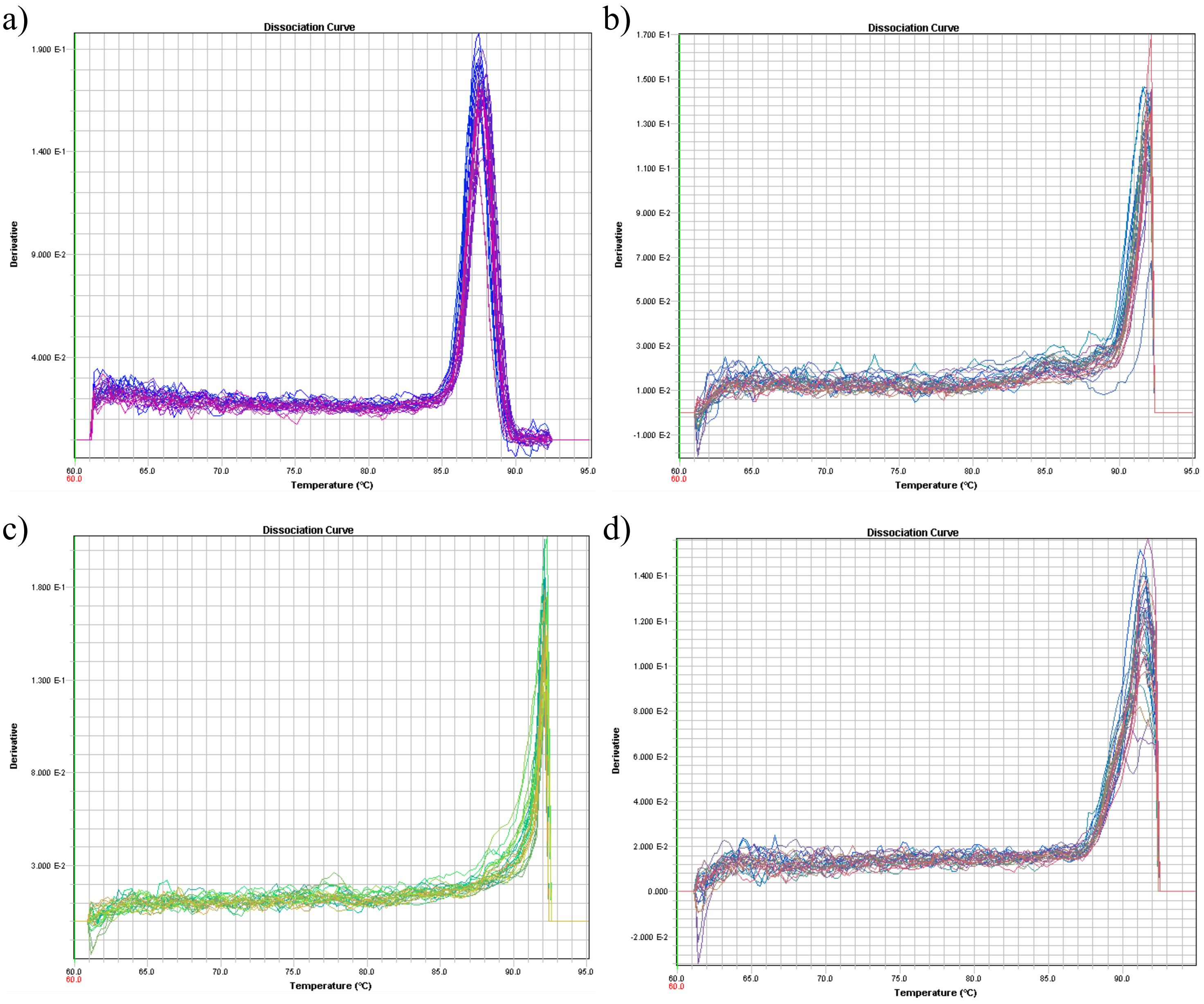
[66] G. Maria, M.D. Ene, I. Jipa, Modelling enzymatic oxidation of d-glucose with pyranose 2-oxidase in the presence of catalase, J Mol Catal B Enzym. 74 (2012) 209–218. https://doi.org/10.1016/j.molcatb.2011.10.007.

[67] M. Dashtban, H. Schraft, T.A. Syed, W. Qin, Fungal biodegradation and enzymatic modification of lignin., Int J Biochem Mol Biol. 1 (2010) 36–50. http://www.ncbi.nlm.nih.gov/pubmed/21968746.

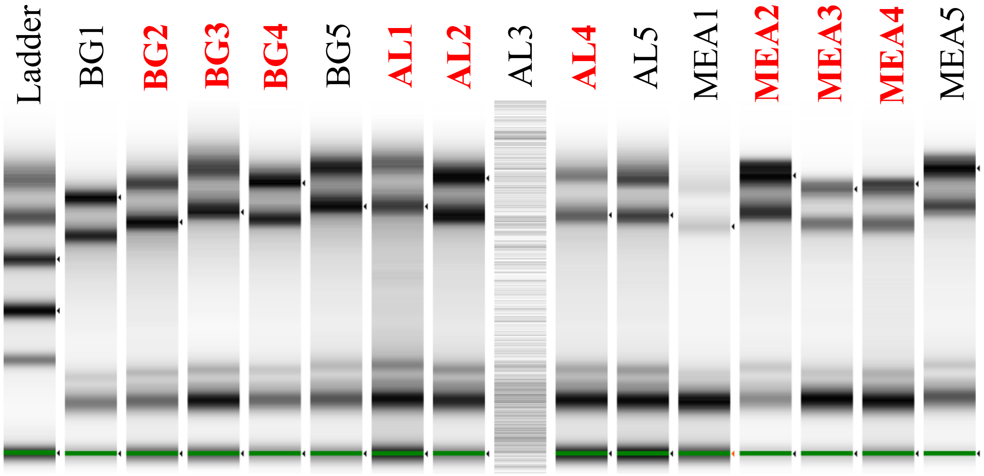
[68] A. Sidar, E.D. Albuquerque, G.P. Voshol, A.F.J. Ram, E. Vijgenboom, P.J. Punt, Carbohydrate Binding Modules: Diversity of Domain Architecture in Amylases and Cellulases From Filamentous Microorganisms, Front Bioeng Biotechnol. 8 (2020) 1–15. https://doi.org/10.3389/fbioe.2020.00871.

[69] M. Couturier, S. Ladevèze, G. Sulzenbacher, L. Ciano, M. Fanuel, C. Moreau, A. Villares, B. Cathala, F. Chaspoul, K.E. Frandsen, A. Labourel, I. Herpoël-Gimbert, S. Grisel, M. Haon, N. Lenfant, H. Rogniaux, D. Ropartz, G.J. Davies, M.-N. Rosso, P.H. Walton, B. Henrissat, J.-G. Berrin, Lytic xylan oxidases from wood-decay fungi unlock biomass degradation, Nat Chem Biol. 14 (2018) 306–310. https://doi.org/10.1038/nchembio.2558.

4.2.1 Material suplementar capítulo II



Supplementary Figure 1 – Dissociation curves (melting curve) obtained through real-time PCR of the actin (a), lignin peroxidase (b), manganese peroxidase (c) and laccase (d) genes.



Supplementary Figure 2 – RNA samples analyzed qualitatively by automated capillary electrophoresis TapeStation System (Agilent Technologies, Santa Clara, California, USA). The samples selected for transcriptome sequencing are highlighted in red.

5. DISCUSSÃO INTEGRADA

O crescimento populacional associado aos avanços tecnológicos tem aumentado a demanda energética global, gerando grandes desafios no âmbito social, econômico e ambiental (LIU, Z. *et al.*, 2022; SOLARTE-TORO *et al.*, 2022). Atualmente, a matriz energética mundial é baseada majoritariamente na utilização de combustíveis fósseis (petróleo, gás natural e carvão), que são fontes de energia não renováveis (EMPRESA DE PESQUISA ENERGÉTICA, 2022). Estas fontes de energia contribuem para a emissão de gases do efeito estufa e são um grande desafio à sustentabilidade. Neste contexto, há uma necessidade crescente para o desenvolvimento de tecnologias voltadas para produção e utilização de energias renováveis (LIU, Z. *et al.*, 2022).

A biomassa lignocelulósica surgiu como uma alternativa promissora frente aos combustíveis fósseis, por ser abundante, amplamente disponível e renovável (ISIKGOR; BECER, 2015). A biomassa é composta por polímeros presentes na parede celular vegetal, principalmente a celulose, hemicelulose e lignina (HOUFANI *et al.*, 2020). Resíduos agrícolas, industriais, florestais e sólidos urbanos são exemplos de biomassa que podem ser amplamente utilizadas para a obtenção de biocombustíveis e produtos químicos (SRIVASTAVA *et al.*, 2015). Apesar da grande relevância desses materiais, a despolimerização da lignocelulose envolve múltiplas etapas (pré-tratamento físico, químico e/ou biológico), o que torna esse processo um desafio para as biorrefinarias (BILAL *et al.*, 2018; HERNÁNDEZ-BELTRÁN *et al.*, 2019; HOUFANI *et al.*, 2020).

Pesquisas tem demonstrado que fungos, principalmente os classificados como da decomposição branca, possuem a capacidade de produzir enzimas hidrolíticas e oxidativas que atuam na despolimerização de todos os constituintes da lignocelulose (ANDRIANI *et al.*, 2020; ESPÓSITO; AZEVEDO, 2010; HAGE *et al.*, 2021; LIU, Y. *et al.*, 2019; MIYAUCHI *et al.*, 2018; SISTA KAMESHWAR; QIN, 2018a). Essas enzimas, além de serem utilizadas na conversão da biomassa, também podem ser aplicadas na biorremediação de poluentes orgânicos e xenobióticos (BRAZKOVA *et al.*, 2022).

O fungo *Trametes villosa* CCMB561 é classificado como um fungo da degradação branca, e pesquisas anteriores já demonstraram o potencial desse isolado em produzir lignina peroxidase e manganês peroxidase em substratos lignocelulósicos (CARNEIRO; LOPES; SILVA; SANTOS; DE SOUZA; *et al.*, 2017; SILVA *et al.*, 2014). Contudo, até o momento, nenhuma pesquisa utilizou este fungo para compreender a maquinaria metabólica envolvida na degradação de todos os constituintes da biomassa vegetal (celulose, hemicelulose, lignina e pectina).

Portanto, com objetivo de identificar os genes e vias metabólicas relacionados ao processo de degradação da celulose, hemicelulose, lignina e pectina, realizamos o sequenciamento, montagem e anotação do genoma do isolado *T. villosa* CCMB561, e, posteriormente, foram conduzidos ensaios bioquímicos de quantificação enzimática, e análises de expressão gênica por qPCR e sequenciamento do transcriptoma.

Nossos resultados demonstraram que por meio da montagem híbrida utilizando *short reads* (Illumina HiSeq) e *long reads* (Oxford Nanopore MinION), foi possível obter um genoma de alta qualidade para o isolado *T. villosa* CCMB561. Neste trabalho, utilizando o pipeline de montagem MaSuRCa-Purge_Dups, foi possível obter um genoma com redução no número de *contigs* (de 10.327 para 143) e melhora significativa das métricas N50, L50 e tamanho do maior *contig*, quando comparado com a versão *draft* do genoma previamente depositada no NCBI (FERREIRA *et al.*, 2018).

Com base na análise do BUSCO, o genoma montado apresentou melhores índices de completude e menor número de genes duplicados, fragmentados e não encontrados (ausentes). De forma similar, MAGGIORI et al., 2021 demonstraram que a partir da abordagem de montagem híbrida (HiSeq + MinION), foi possível obter um maior número de genomas de amostras de metagenômica, com *contigs* mais longos, com melhora significativa da métrica N50, maior número de sequências codificantes, maior completude e menor contaminação (MAGGIORI *et al.*, 2021). Desta forma, para genomas de fungos, a abordagem de montagem híbrida utilizando *short* e *long reads* é uma ferramenta promissora para obtenção de genomas completos e com qualidade.

A anotação do genoma obtido neste trabalho, associada as análises de genômica comparativa e anotação funcional dos genes que codificam CAZymes demonstraram o potencial genético do isolado *T. villosa* CCMB561 de atuar na degradação de todos os componentes da lignocelulose. Entre os genomas analisados, *T. villosa* foi a espécie com maior número de genes codificadores de enzimas que atuam na despolimerização da lignina.

Os resultados dos ensaios bioquímicos, expressão relativa por qPCR e sequenciamento do transcriptoma revelaram que *T. villosa* CCMB561 possui a capacidade de crescer em meio de cultura contendo apenas lignina como fonte de carbono, sem a necessidade de suplementação. Este resultado demonstrou que a adição de nutrientes no meio de cultura nem sempre é necessária para estimular o crescimento e atividade de determinadas enzimas. Por exemplo, pesquisas com o isolado *Trametes villosa* CCB176 demonstraram que a concentração inicial de Mn^{2+} não influencia no crescimento do fungo e na atividade das enzimas ligninolíticas (YAMANAKA *et al.*, 2008).

Uma super expressão de diferentes genes codificadores de citocromo P450 foi identificada no meio contendo apenas a lignina como fonte de carbono (AL). A expressão desse gene foi associada ao processo metabólico de O-desmetilação da lignina e subsequente despolimerização por oxidases e peroxidases. Citocromo P450 são enzimas que já foram descritas catalisando a reação de O-desmetilação e também atuam na degradação de poluentes orgânicos e xenobióticos (CHEN *et al.*, 2014; ELLIS *et al.*, 2021; HARMS; SCHLOSSER; WICK, 2011; KHAN; MURPHY, 2022; MALLINSON *et al.*, 2018).

Genes codificadores de nitroredutases e homogentisato-1,2-dioxigenase também foram identificados sendo expressos positivamente no meio AL. As nitroredutases atuam na mineralização de compostos nitroaromáticos (HARMS; SCHLOSSER; WICK, 2011; RIEBLE; JOSHI; GOLD, 1994). Homogentisato-1,2-dioxigenase atua no processo de degradação de compostos intermediários da despolimerização da lignina (vanilina) e na degradação de compostos orgânicos voláteis (VOCs) (GUNSCH *et al.*, 2005; MORI *et al.*, 2016; SHIMIZU *et al.*, 2005).

A lignina (meio AL) também induziu a expressão de um cluster de genes codificadores de CAZymes, que revelou a capacidade de *T. villosa* CCMB561 de atuar simultaneamente na degradação de todos os polímeros da lignocelulose. Esse resultado também foi corroborado pelo crescimento do fungo no bagaço da cana-de-açúcar, substrato no qual foram detectados genes diferencialmente expressos que atuam na degradação de todos os compostos da biomassa vegetal. Estudos prévios já demonstraram a capacidade de outras espécies do gênero *Trametes* em produzir de forma sequencial enzimas ligninolíticas, celulolíticas e hemicelulolíticas em substratos lignocelulósicos (ANDRIANI *et al.*, 2020).

Por fim, os resultados obtidos neste trabalho demonstraram os complexos mecanismos genéticos de degradação da lignocelulose por *T. villosa* CCMB561 e, em concordância com outros estudos, revelam o grande potencial deste fungo de ser utilizado na conversão da biomassa ligenocelulósica, e consequentemente, contribuir com o avanço na utilização de tecnologias sustentáveis.

6. CONCLUSÃO

Neste estudo, a montagem do genoma de *Trametes villosa* CCMB561 por meio da abordagem híbrida, utilizando o pipeline MaSuRCA-Purge_dups, se mostrou uma excelente opção para obtenção de um genoma com baixa taxa de fragmentação e duplicação gênica. As análises de anotação do CAZoma revelaram o potencial genético do isolado CCMB561 de atuar na despolimerização da lignina, celulose, hemicelulose e pectina.

O cultivo do fungo em diferentes substratos associado as análises do transcriptoma, demonstraram que genes codificadores de citocromo P450, nitroredutases, homogentisato 1,2-dioxigenases e de enzimas ligninolíticas são diferencialmente expressos (*up-regulated*) quando o fungo é cultivado em meio de cultura contendo a lignina como única fonte de carbono. Cluster de genes codificadores de CAZymes foram *up-regulated* em lignina e bagaço de cana, e revelaram que *T. villosa* CCMB561 atua na despolimerização simultânea da lignina, celulose, hemicelulose e pectina.

Por fim, as análises genômicas, de quantificação bioquímica, expressão relativa e transcriptômica, forneceram informações relevantes sobre os genes e transcritos relacionados ao processo de nutrição do fungo e degradação da lignocelulose, sendo estas informações importantes para a melhoria de bioprocessos, obtenção de produtos a partir da conversão de resíduos lignocelulósicos e degradação de poluentes orgânicos.

7. REFERÊNCIAS

AKYOL, Ç. *et al.* Biological pretreatment with Trametes versicolor to enhance methane production from lignocellulosic biomass: A metagenomic approach. *Industrial Crops and Products*, v. 140, p. 111659, nov. 2019.

ANDRIANI, A. *et al.* Sequential production of ligninolytic, xylanolytic, and cellulolytic enzymes by Trametes hirsuta AA-017 under different biomass of Indonesian sorghum accessions-induced cultures. *Bioresource Technology Reports*, v. 12, p. 100562, dez. 2020. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S2589014X20301833>.

ATILANO-CAMINO, M. M. *et al.* Improving laccase production from Trametes versicolor using lignocellulosic residues as cosubstrates and evaluation of enzymes for blue wastewater biodegradation. *Journal of Environmental Management*, v. 275, n. August, p. 111231, dez. 2020. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0301479720311555>.

BIELY, P. *et al.* Trichoderma reesei CE16 acetyl esterase and its role in enzymatic degradation of acetylated hemicellulose. *Biochimica et Biophysica Acta (BBA) - General Subjects*, v. 1840, n. 1, p. 516–525, jan. 2014.

BILAL, M. *et al.* Engineering Ligninolytic Consortium for Bioconversion of Lignocelluloses to Ethanol and Chemicals. *Protein & Peptide Letters*, v. 25, n. 2, p. 108–119, 2 maio 2018. Disponível em: <http://www.eurekaselect.com/159241/article>.

BRAZKOVA, M. *et al.* Ligninolytic enzymes in Basidiomycetes and their application in xenobiotics degradation . *BIO Web of Conferences*, v. 45, p. 02009, 2022.

BRETHAUER, S.; STUDER, M. H. Biochemical Conversion Processes of Lignocellulosic Biomass to Fuels and Chemicals – A Review. *CHIMIA International Journal for Chemistry*, v. 69, n. 10, p. 572–581, 28 out. 2015.

BUNDHOO, Z. M. A.; MUDHOO, A.; MOHEE, R. Promising Unconventional Pretreatments for Lignocellulosic Biomass. *Critical Reviews in Environmental Science and Technology*, v. 43, n. 20, p. 2140–2211, 5 jan. 2013.

CANTAREL, B. L. *et al.* The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Research*, v. 37, n. Database, p. D233–D238, 1 jan. 2009. Disponível em: <https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkn663>.

CARNEIRO, R. T. DE O.; LOPES, M. A.; SILVA, M. L. C.; SANTOS, V. DA S.; DE SOUZA, V. B.; *et al.* Trametes villosa lignin peroxidase (TvLiP): Genetic and molecular characterization. *Journal of Microbiology and Biotechnology*, v. 27, n. 1, p. 179–188, 1 jan. 2017.

CARNEIRO, R. T. DE O.; LOPES, M. A.; SILVA, M. L. C.; SANTOS, V. DA S.; SOUZA, V. B. DE; *et al.* Trametes villosa Lignin Peroxidase (TvLiP): Genetic and Molecular Characterization. *Journal of Microbiology and Biotechnology*, v. 27, n. 1, p. 179–188, 28 jan. 2017.

CARUS, M.; DAMMER, L. The Circular Bioeconomy—Concepts, Opportunities, and Limitations. *Industrial Biotechnology*, v. 14, n. 2, p. 83–91, abr. 2018.

CHEN, W. *et al.* Fungal cytochrome P450 monooxygenases: Their distribution, structure, functions, family expansion, and evolutionary origin. *Genome Biology and Evolution*, v. 6, n. 7, p. 1620–1634, 2014.

CONIGLIO, R. O. *et al.* Enzymatic hydrolysis of barley straw for biofuel industry using a novel strain of Trametes villosa from Paranaense rainforest. *Preparative Biochemistry & Biotechnology*, v. 50, n. 8, p. 753–762, 1 set. 2020.

CONIGLIO, R. O. *et al.* Optimization of cellobiohydrolase production and secretome analysis of Trametes villosa LBM 033 suitable for lignocellulosic bioconversion. *Arab Journal of Basic and Applied Sciences*, v. 26, n. 1, p. 182–192, 2 jan. 2019. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/25765299.2019.1598107>.

DASHTBAN, M. *et al.* Fungal biodegradation and enzymatic modification of lignin. *International journal of biochemistry and molecular biology*, v. 1, n. 1, p. 36–50, 2010. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/21968746>.

DE CARVALHO, L. M. *et al.* Bioinformatics applied to biotechnology: A review towards bioenergy research. *Biomass and Bioenergy*, v. 123, n. February, p. 195–224, abr. 2019.

DELSUC, F.; BRINKMANN, H.; PHILIPPE, H. Phylogenomics and the reconstruction of the tree of life. *Nature Reviews Genetics*, v. 6, n. 5, p. 361–375, maio 2005.

DÍAZ, G. V. *et al.* Secretomic analysis of cheap enzymatic cocktails of Aspergillus niger LBM 134 grown on cassava bagasse and sugarcane bagasse. *Mycologia*, v. 112, n. 4, p. 663–676, 3 jul. 2020.

DING, C.; WANG, X.; LI, M. Evaluation of six white-rot fungal pretreatments on corn stover for the production of cellulolytic and ligninolytic enzymes, reducing sugars, and ethanol. *Applied Microbiology and Biotechnology*, v. 103, n. 14, p. 5641–5652, 21 jul. 2019.

ELLIS, E. S. *et al.* Engineering a Cytochrome P450 for Demethylation of Lignin-Derived Aromatic Aldehydes. *JACS Au*, v. 1, n. 3, p. 252–261, 22 mar. 2021.

EMPRESA DE PESQUISA ENERGÉTICA. *Matriz Energética e Elétrica*.

ESPÓSITO, E.; AZEVEDO, J. L. DE. *Fungos - Uma Introdução à Biologia, Bioquímica e Biotecnologia*. 2ª ed. Caxias do Sul: EDUCS, Editora da Universidade de Caxias do Sul, 2010.

FERREIRA, D. S. S. *et al.* Draft genome sequence of Trametes villosa (Sw.) Kreisel CCMB561, a tropical white-rot Basidiomycota from the semiarid region of Brazil. *Data in Brief*, v. 18, p. 1581–1587, jun. 2018. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S235234091830427X>.

FLOUDAS, D. *et al.* The Paleozoic Origin of Enzymatic Lignin Decomposition Reconstructed from 31 Fungal Genomes. *Science*, v. 336, n. 6089, p. 1715–1719, 29 jun. 2012. Disponível em: <https://www.sciencemag.org/lookup/doi/10.1126/science.1221748>.

GAI, Y.-P. *et al.* Involvement of ligninolytic enzymes in degradation of wheat straw by Trametes trogii. *Journal of Applied Microbiology*, v. 117, n. 1, p. 85–95, jul. 2014.

GOFFEAU, A. *et al.* Life with 6000 Genes. *Science*, v. 274, n. 5287, p. 546–567, 25 out. 1996.

GUNSCH, C. K. *et al.* Identification of a homogentisate-1,2-dioxygenase gene in the fungus Exophiala lecanii-corni: analysis and implications. *Applied Microbiology and Biotechnology*, v. 68, n. 3, p. 405–411, 25 ago. 2005.

HAGE, H. *et al.* Gene family expansions and transcriptome signatures uncover fungal adaptations to wood decay. *Environmental Microbiology*, p. 1462- 2920.15423, 15 fev. 2021. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1111/1462-2920.15423>.

HARMS, H.; SCHLOSSER, D.; WICK, L. Y. *Untapped potential: Exploiting fungi in bioremediation of hazardous chemicals*. *Nature Reviews Microbiology*. [S.l: s.n.]. , mar. 2011

HERNÁNDEZ-BELTRÁN, J. U. *et al.* Insight into Pretreatment Methods of Lignocellulosic Biomass to Increase Biogas Yield: Current State, Challenges, and Opportunities. *Applied Sciences*, v. 9, n. 18, p. 3721, 6 set. 2019. Disponível em: <https://www.mdpi.com/2076-3417/9/18/3721>.

HOUFANI, A. A. *et al.* Insights from enzymatic degradation of cellulose and hemicellulose to fermentable sugars– a review. *Biomass and Bioenergy*, v. 134, n. January, p. 105481, mar. 2020. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0961953420300155>.

ISIKGOR, F. H.; BECER, C. R. Lignocellulosic biomass: a sustainable platform for the production of bio-based chemicals and polymers. *Polymer Chemistry*, v. 6, n. 25, p. 4497–4559, 2015. Disponível em: <http://xlink.rsc.org/?DOI=C5PY00263J>.

JAMES, T. Y. *et al.* Toward a Fully Resolved Fungal Tree of Life. *Annual review of microbiology*, v. 74, n. 1, p. 291–313, 8 set. 2020. Disponível em: <https://www.annualreviews.org/doi/10.1146/annurev-micro-022020-051835>.

JEYA, M. *et al.* Enhanced saccharification of alkali-treated rice straw by cellulase from Trametes hirsuta and statistical optimization of hydrolysis conditions by RSM. *Bioresource Technology*, v. 100, n. 21, p. 5155–5161, nov. 2009.

JIANG, J.-R.; CAI, L.; LIU, F. Oligotrophic fungi from a carbonate cave, with three new species of Cephalotrichum. *Mycology*, v. 8, n. 3, p. 164–177, 3 jul. 2017.

JÖNSSON, L. J.; MARTÍN, C. Pretreatment of lignocellulose: Formation of inhibitory by-products and strategies for minimizing their effects. *Bioresource Technology*, v. 199, p. 103–112, jan. 2016.

JUSTO, A.; HIBBETT, D. S. Phylogenetic classification of Trametes (Basidiomycota, Polyporales) based on a five-marker dataset. *TAXON*, v. 60, n. 6, p. 1567–1583, dez. 2011.

KHAN, M. F.; MURPHY, C. D. Cytochrome P450 5208A3 is a promiscuous xenobiotic biotransforming enzyme in Cunninghamella elegans. *Enzyme and Microbial Technology*, v. 161, p. 110102, nov. 2022.

KNEŽEVIĆ, A. *et al.* Induction of wheat straw delignification by Trametes species. *Scientific Reports*, v. 6, n. 1, p. 26529, 24 maio 2016.

KRAH, F.-S. *et al.* Evolutionary dynamics of host specialization in wood-decay fungi. *BMC Evolutionary Biology*, v. 18, n. 1, p. 119, 3 dez. 2018.

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. Disponível em: <http://biotechnologyforbiofuels.biomedcentral.com/articles/10.1186/1754-6834-6-41>.

LI, Y. *et al.* A genome-scale phylogeny of the kingdom Fungi. *Current Biology*, p. 1–13, fev. 2021.

LIU, X. *et al.* Protein Identification Using Top-Down Spectra. *Molecular & Cellular Proteomics*, v. 11, n. 6, p. M111.008524, jun. 2012.

LIU, Y. *et al.* Lignin degradation potential and draft genome sequence of Trametes trogii S0301. *Biotechnology for Biofuels*, v. 12, n. 1, p. 256, 30 dez. 2019. Disponível em: <https://doi.org/10.1186/s13068-019-1596-3>.

LIU, Z. *et al.* Assessing the effectiveness of biomass energy in mitigating CO2 emissions: Evidence from Top-10 biomass energy consumer countries. *Renewable Energy*, v. 191, p. 842–851, 1 maio 2022.

LOMBARD, V. *et al.* The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Research*, v. 42, n. D1, p. D490–D495, jan. 2014. Disponível em: <https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkt1178>.

LUO, Z. *et al.* Selective conversion of lignin to ethylbenzene. *Green Chemistry*, v. 22, n. 6, p. 1842–1850, 21 mar. 2020.

MACHADO, A. S. *et al.* The Secretome of Phanerochaete chrysosporium and Trametes versicolor Grown in Microcrystalline Cellulose and Use of the Enzymes for Hydrolysis of Lignocellulosic Materials. *Frontiers in Bioengineering and Biotechnology*, v. 8, 17 jul. 2020.

MAGGIORI, C. *et al.* MinION sequencing from sea ice cryoconites leads to de novo genome reconstruction from metagenomes. *Scientific Reports*, v. 11, n. 1, p. 21041, 26 dez. 2021. Disponível em: <https://www.nature.com/articles/s41598-021-00026-x>.

MÄKINEN, M. *et al.* Genome description of Phlebia radiata 79 with comparative genomics analysis on lignocellulose decomposition machinery of phlebioid fungi. *BMC Genomics*, v. 20, n. 1, p. 430, 28 dez. 2019.

MALLINSON, S. J. B. *et al.* A promiscuous cytochrome P450 aromatic O-demethylase for lignin bioconversion. *Nature Communications*, v. 9, n. 1, 1 dez. 2018.

MANAVALAN, T.; MANAVALAN, A.; HEESE, K. Characterization of Lignocellulolytic Enzymes from White-Rot Fungi. *Current Microbiology*, v. 70, n. 4, p. 485–498, 9 abr. 2015. Disponível em: <http://link.springer.com/10.1007/s00284-014-0743-0>.

MARTINEZ, D. *et al.* Genome sequence of the lignocellulose degrading fungus Phanerochaete chrysosporium strain RP78. *Nature Biotechnology*, v. 22, n. 6, p. 695–700, 1 jun. 2004.

MIYAUCHI, S. *et al.* Integrative visual omics of the white-rot fungus Polyporus brumalis exposes the biotechnological potential of its oxidative enzymes for delignifying raw plant biomass. *Biotechnology for Biofuels*, v. 11, n. 1, p. 201, 23 dez. 2018. Disponível em: <https://doi.org/10.1186/s13068-018-1198-5>.

MORI, T. *et al.* Effects of Homologous Expression of 1,4-Benzoquinone Reductase and Homogentisate 1,2-Dioxygenase Genes on Wood Decay in Hyper-Lignin-Degrading Fungus Phanerochaete sordida YK-624. *Current Microbiology*, v. 73, n. 4, p. 512–518, 30 out. 2016.

NIKOLAIVITS, E. *et al.* Functional and transcriptomic investigation of laccase activity in the presence of PCB29 identifies two novel enzymes and the multicopper oxidase repertoire of a marine-derived fungus. *Science of The Total Environment*, v. 775, p. 145818, jun. 2021.

NISHIMURA, H. *et al.* Direct evidence for α ether linkage between lignin and carbohydrates in wood cell walls. *Scientific Reports*, v. 8, n. 1, 1 dez. 2018.

PAËS, G. *et al.* Tracking of enzymatic biomass deconstruction by fungal secretomes highlights markers of lignocellulose recalcitrance. *Biotechnology for Biofuels*, v. 12, n. 1, p. 76, 1 dez. 2019.

QIAN, Y. *et al.* Enhancement of Cellulase Production in Trichoderma reesei via Disruption of Multiple Protease Genes Identified by Comparative Secretomics. *Frontiers in Microbiology*, v. 10, n. December, p. 1–12, 3 dez. 2019.

RAJA, H. A. *et al.* Fungal Identification Using Molecular Tools: A Primer for the Natural Products Research Community. *Journal of Natural Products*, v. 80, n. 3, p. 756–770, 24 mar. 2017.

RIEBLE, S.; JOSHI, D. K.; GOLD, M. H. Aromatic Nitroreductase from the Basidiomycete Phanerochaete chrysosporium. *Biochemical and Biophysical Research Communications*, v. 205, n. 1, p. 298–304, nov. 1994.

ROTH, J. C. G.; HOELTZ, M.; BENITEZ, L. B. Current approaches and trends in the production of microbial cellulases using residual lignocellulosic biomass: a bibliometric analysis of the last 10 years. *Archives of Microbiology*, v. 202, n. 5, p. 935–951, 11 jul. 2020.

SCHMIDT-DANNERT, C. Biocatalytic portfolio of Basidiomycota. *Current Opinion in Chemical Biology*, v. 31, p. 40–49, abr. 2016. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S136759311600003X>.

SHARMA, K. K. Fungal genome sequencing: basic biology to biotechnology. *Critical Reviews in Biotechnology*, v. 36, n. 4, p. 743–759, 3 jul. 2016.

SHERWOOD, J. The significance of biomass in a circular economy. *Bioresource Technology*, v. 300, p. 122755, mar. 2020.

SHIMIZU, M. *et al.* Metabolic regulation at the tricarboxylic acid and glyoxylate cycles of the lignin-degrading basidiomycetePhanerochaete chrysosporium against exogenous addition of vanillin. *PROTEOMICS*, v. 5, n. 15, p. 3919–3931, out. 2005.

SILVA, M. L. C. *et al.* Production of Manganese Peroxidase by <i>Trametes villosa</i> on Unexpensive Substrate and Its Application in the Removal of Lignin from Agricultural Wastes. *Advances in Bioscience and Biotechnology*, v. 05, n. 14, p. 1067–1077, 2014.

SISTA KAMESHWAR, A. K.; QIN, W. Comparative study of genome-wide plant biomass-degrading CAZymes in white rot, brown rot and soft rot fungi. *Mycology*, v. 9, n. 2, p. 93–105, 2018a. Disponível em: <https://doi.org/10.1080/21501203.2017.1419296>.

SISTA KAMESHWAR, A. K.; QIN, W. Understanding the structural and functional properties of carbohydrate esterases with a special focus on hemicellulose deacetylating acetyl xylan esterases. *Mycology*, v. 9, n. 4, p. 273–295, 2 out. 2018b.

SOLARTE-TORO, J. C. *et al.* Review of the impact of socio-economic conditions on the development and implementation of biorefineries. *Fuel*, v. 328, 15 nov. 2022.

SRIVASTAVA, N. *et al.* A Review on Fuel Ethanol Production From Lignocellulosic Biomass. *International Journal of Green Energy*, v. 12, n. 9, p. 949–960, 2 set. 2015. Disponível em: <http://www.tandfonline.com/doi/full/10.1080/15435075.2014.890104>.

STAJICH, J. E. Fungal genomes and insights into the evolution of the kingdom. *Microbiol Spectr.*, v. 5, n. 4, p. 1–25, 2017.

SUKUMARAN, R. K. *et al.* Cellulase production using biomass feed stock and its application in lignocellulose saccharification for bio-ethanol production. *Renewable Energy*, v. 34, n. 2, p. 421–424, fev. 2009.

TIŠMA, M. *et al.* Trametes versicolor in lignocellulose-based bioeconomy: State of the art, challenges and opportunities. *Bioresource Technology*, v. 330, p. 124997, jun. 2021.

TYERS, M.; MANN, M. From genomics to proteomics. *Nature*, v. 422, n. 6928, p. 193–197, mar. 2003.

VAN DEN BRINK, J.; DE VRIES, R. P. Fungal enzyme sets for plant polysaccharide degradation. *Applied Microbiology and Biotechnology*, v. 91, n. 6, p. 1477–1492, 23 set. 2011. Disponível em: <http://link.springer.com/10.1007/s00253-011-3473-2>.

VASINA, D. V. *et al.* Lignin-degrading peroxidases in white-rot fungus Trametes hirsuta 072. Absolute expression quantification of full multigene family. *PLOS ONE*, v. 12, n. 3, p. e0173813, 16 mar. 2017.

WALKER, G. M.; WHITE, N. A. Introduction to Fungal Physiology. *Fungi*. Chichester, UK: John Wiley & Sons, Ltd, 2005. p. 1–34.

WANG, F. *et al.* Lignocellulosic biomass as sustainable feedstock and materials for power generation and energy storage. *Journal of Energy Chemistry*, v. 57, p. 247–280, 1 jun. 2021.

WANG, Z. *et al.* Using evolutionary genomics, transcriptomics, and systems biology to reveal gene networks underlying fungal development. *Fungal Biology Reviews*, v. 32, n. 4, p. 249–264, set. 2018.

WILKEN, ST. E. *et al.* Linking ‘omics’ to function unlocks the biotech potential of non-model fungi. *Current Opinion in Systems Biology*, v. 14, p. 9–17, abr. 2019. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S2452310019300095>.

WONG, D. W. S. Structure and Action Mechanism of Ligninolytic Enzymes. *Applied Biochemistry and Biotechnology*, v. 157, n. 2, p. 174–209, 26 maio 2009. Disponível em: <http://link.springer.com/10.1007/s12010-008-8279-z>.

YAMANAKA, R. *et al.* Lignolytic enzymes produced by Trametes villosa ccb176 under different culture conditions. *Brazilian Journal of Microbiology*, v. 39, n. 1, p. 78–84, mar. 2008. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1517-83822008000100019&lng=en&nrm=iso&tlng=en>.

ZHANG, J. *et al.* Cellulose-hemicellulose and cellulose-lignin interactions during fast pyrolysis. *ACS Sustainable Chemistry and Engineering*, v. 3, n. 2, p. 293–301, 2 fev. 2015.

ATIVIDADES DESENVOLVIDAS NO PERÍODO DO DOUTORADO

- **Artigos completos publicados em periódicos (Total: 10)**

1 – Romero-Calle, D. X., Pedrosa-Silva, F., **Tomé, L. M. R.**, Sousa, T. J., Santos, L. T. S. D. O., Azevedo, V. A. D. C., ... & Góes-Neto, A. (2023). **Hybrid Genomic Analysis of Salmonella enterica Serovar Enteritidis SE3 Isolated from Polluted Soil in Brazil.** Microorganisms, 11(1), 111.

2 – de Menezes, T. A., Aburjaile, F. F., Quintanilha-Peixoto, G., **Tomé, L. M. R.**, Fonseca, P. L. C., Mendes-Pereira, T., ... & Góes-Neto, A. (2023). **Unraveling the Secrets of a Double-Life Fungus by Genomics: Ophiocordyceps australis CCMB661 Displays Molecular Machinery for Both Parasitic and Endophytic Lifestyles.** Journal of Fungi, 9(1), 110.

3 – **Tomé, L**. M. R., da Silva, F. F., Fonseca, P. L. C., Mendes-Pereira, T., Azevedo, V. A. D. C., Brenig, B., ... & Góes-Neto, A. (2022). **Hybrid Assembly Improves Genome Quality and Completeness of Trametes villosa CCMB561 and Reveals a Huge Potential for Lignocellulose Breakdown.** Journal of Fungi, 8(2), 142.

4 – Fonseca, M. S., Rodrigues, D. M., Sokolonski, A. R., Stanisic, D., **Tomé, L. M. R**., Góes-Neto, A., ... & Portela, R. D. (2022). **Activity of Fusarium oxysporum-Based Silver Nanoparticles on Candida spp. Oral Isolates.** Nanomaterials, 12(3), 501.

5 – Cordeiro, T. R. L., da Silva, S. B. G., da Cruz, M. O., de Lima, C. L. F., Lima, D. X., Góes-Neto, A., **Tomé, L. M** & de Azevedo Santiago, A. L. C. M. (2022). **Absidia pararepens Jurjević, M. Kolařík & Hubka (Mucorales, Mucoromycota) was isolated for the first time in South America.** Nova Hedwigia, 375-387.

6 – Quintanilha-Peixoto, G., Marone, M. P., Raya, F. T., José, J., Oliveira, A., Fonseca, P. L. C., **Tomé, L. M. R**... & Góes-Neto, A. (2022). **Phylogenomics and gene selection in Aspergillus welwitschiae: Possible implications in the pathogenicity in Agave sisalana.** Genomics, 114(6), 110517.

7 – Fonseca, P. L., De-Paula, R. B., Araújo, D. S., **Tomé, L. M. R.**, Mendes-Pereira, T., Rodrigues, W. F. C., ... & Góes-Neto, A. (2021). **Global characterization of fungal**

mitogenomes: new insights on genomic diversity and dynamism of coding genes and accessory elements. Frontiers in microbiology, 12.

8 – Araújo, D. S., De-Paula, R. B., **Tomé, L. M.**, Quintanilha-Peixoto, G., Salvador-Montoya, C. A., Del-Bem, L. E., ... & Góes-Neto, A. (2021). **Comparative mitogenomics of Agaricomycetes: Diversity, abundance, impact and coding potential of putative open-reading frames.** Mitochondrion, 58, 1-13.

9 – Dos Santos, A. B., **Ribeiro Tomé, L. M.**, De Carvalho, G. C., Camargos Fonseca, P. L., Pereira Figuiredo, H. C., Pereira, F. L., ... & Góes-Neto, A. (2020, August). Effect of the characteristics of municipal solid waste on biogas production in landfills. In Proceedings of the Institution of Civil Engineers-Waste and Resource Management (Vol. 173, No. 3, pp. 55-64). Thomas Telford Ltd.

10 – **Tomé, L. M. R.**, Badotti, F., Assis, G. B. N., Fonseca, P. L. C., da Silva, G. A., da Silveira, R. M. B., ... & Góes-Neto, A. (2019). **Proteomic fingerprinting for the fast and accurate identification of species in the Polyporoid and Hymenochaetoid fungi clades.** Journal of proteomics, 203, 103390.

- **Artigo aceito para publicação (Total: 1)**

1 – A polyphasic approach to the taxonomy of Backusella reveals two new species" (Mycological Progress)

- **Apresentação de Trabalhos (Total: 8)**

1 – **M. R. TOMÉ, LUIZ**; PEREIRA, T. M.; A. C. AZEVEDO, VASCO; BRENIG, B.; BADOTTI, F.; GÓES-NETO, A. Relative expression by qPCR, transcriptome sequencing, and differential gene expression (DGE) analysis reveal insights into lignocellulose degradation of *Trametes villosa* CCMB561. 2022. (Apresentação de Trabalho/Congresso).

2 – PEREIRA, T. M.; **M. R. TOMÉ, LUIZ**; KLOSS, T. G.; MAGALHAES, L. M. D.; A. C. AZEVEDO, VASCO; BRENIG, B.; GÓES-NETO, A. Whole-genome sequencing of the spider pathogen *Gibellula pulchra* (Ascomycota, Cordycipitaceae) reveals candidate genes related to parasitism and host manipulation mechanism. 2022. (Apresentação de Trabalho/Congresso).

3 – SANTOS, M.; **Luiz M. R. Tomé**; BADOTTI, F.; GÓES-NETO, A. Análise da produção de enzimas ligninolíticas de interesse industrial pelo fungo *Trametes villosa* CCMB561. 2020. (Apresentação de Trabalho/Outra).

4 – **TOMÉ, L. M. R.;** BADOTTI, F.; ASSIS, G. B.; FONSECA, P. L. C.; SILVA, G. A.; DA SILVEIRA, ROSA MARA BORGES; COSTA-REZENDE, DIOGO HENRIQUE; DRECHSLER-SANTOS, E. R.; AZEVEDO, VASCO ARISTON DE CARVALHO; FIGUEIREDO, H. C. P.; GÓES-NETO, A. Fast and Accurate Identification of Basidiomycotan fungi by Mass- Spectrometry-Based Proteomic Barcoding. 2019. (Apresentação de Trabalho/Simpósio).

5 – **TOMÉ, L. M. R**.; SANTOS, M.; FONSECA, P. L. C.; PEIXOTO, G. Q.; TOLEDO, A. T. Z.; ARAUJO, D. S.; GÓES-NETO, A.; BADOTTI, F. CARACTERIZÇÃO IN SILICO E FILOGENIA DE HEME-PEROXIDASES DE CLASSE II DE *Trametes versicolor*. 2019. (Apresentação de Trabalho/Simpósio).

6 – PEIXOTO, G. Q.; ARAUJO, D. S.; KATO, R. B.; FONSECA, P. L. C.; **TOMÉ, L. M. R.**; MIRANDA, F. M.; RAMOS, R. T. J.; BRENIG, B.; AZEVEDO, VASCO ARISTON DE CARVALHO; BADOTTI, F.; AGUIAR, E. R. G. R.; GÓES-NETO, A. The Death Is Red: Analysis of the Predicted Secretome of *Aspergillus welwitschiae*, with Emphasis in Pathogenicity and Carbohydrate Metabolism. 2019. (Apresentação de Trabalho/Conferência ou palestra).

7 – SILVA, V. O.; FONSECA, P. L. C.; **TOMÉ, L. M. R.**; GÓES-NETO, A. THE PAINTING I LIVE: FUNGI IDENTIFIED ON PICTORIAL SURFACE. 2019. (Apresentação de Trabalho/Conferência ou palestra).

8 – **TOMÉ, L. M. R.;** SIERRA, L. A. B.; FONSECA, P. L. C.; PEIXOTO, G. Q.; SANTOS, M.; PEREIRA, T. M; GÓES-NETO, A. O maravilhoso mundo dos fungos heróis e vilões. 2019. (Apresentação de Trabalho/Outra).

- **Prêmios (Total: 1)**

2020 – Relevância Acadêmica - Análise da produção de enzimas ligninolíticas de interesse industrial pelo fungo *Trametes villosa* CCMB561 - XXIX Semana de Iniciação Científica (UFMG), Universidade Federal de Minas Gerais.

- **Trabalho selecionado para prêmio e recebeu menção honrosa (Total: 1)**

2022 – Menção honrosa-Prêmio Milton Krieger: Expression by qPCR, transcriptome sequencing, and differential gene expression analysis reveal insights into lignocellulose degradation of *T. villosa* CCMB561, Sociedade Brasileira de Genética (SBG).

- **Participação em bancas de trabalho de conclusão de curso (Total: 1)**

1 – **TOMÉ, LUIZ M. R.**; COSTA, E. A. S.; SILVA, M. V. C. M.. Participação em banca de Ádila Souza; Maria Tatiana Feitosa; Milene Bianchi. Relação do gene STAG2 com o processo metastático do câncer de mama.. 2019. Trabalho de Conclusão de Curso (Graduação em BIomedicina) - Unidade de Ensino Superior de Feira de Santana.

- **Participação em eventos (Total: 11)**

1 – Genética 2022. Relative expression by qPCR, transcriptome sequencing, and differential gene expression (DGE) analysis reveals insights into lignocellulose degradation of Trametes villosa CCMB561. 2022. (Congresso).

2 – 1st Symposium of Fungal Diversity and Conservation in Cloud Forests. 2021. (Simpósio).

3 – 25th International Bioinformatic Workshop on Virus Evolution and Molecular Epidemiology (VEME light). 2021. (Outra).

4 – Bioinformatics Workshop on Virus Evolution and Molecular Epidemiology (COVEME). 2021. (Outra).

5 – VIII Simpósio de Microbiologia. 2021. (Simpósio).

6 – I Workshop Online de Bioinformática. 2020. (Outra).

7 – Jornada de Internacionalización del Programa de Bacteriología y Laboratorio Clínico.Herramientas proteómicas para la identificación rápida y precisa de hongos. 2020. (Outra).

8 – VII Simpósio de Microbiologia da UFMG - Conecta SIM. 2020. (Simpósio).

9 – #ExploraUFMGJovem. O maravilhoso mundo do fungos heróis e vilões. 2019. (Exposição).

10 – 1st SYMPOSIUM ON MASS SPECTROMETRY APPLIED TO PROTEOMICS AND STRUCTURAL BIOLOGY.Fast and Accurate Identification of Basidiomycotan fungi by Mass- Spectrometry-Based Proteomic Barcoding. 2019. (Simpósio).

11 – VI Simpósio de Microbiologia da UFMG - CONECTA SIM: Microbiologia Interligada.CARACTERIZÇÃO IN SILICO E FILOGENIA DE HEME-PEROXIDASES DE CLASSE II DE Trametes versicolor. 2019. (Simpósio).

- **Palestras/Aulas Ministradas (Total: 4)**

1 – M. R. TOMÉ, LUIZ. TRATAMENTO DAS READS, MONTAGEM POR REFERÊNCIA E CHAMADA DE SNP I. 2022. (Apresentação de Trabalho/Conferência ou palestra).

2 – M. R. TOMÉ, LUIZ. METAGENÔMICA E SUAS APLICAÇÕES NA SAÚDE PÚBLICA II. 2022. (Apresentação de Trabalho/Conferência ou palestra).

3 – M. R. TOMÉ, LUIZ. Candidíase Vulvovaginal (CVV). 2020. (Apresentação de Trabalho/Conferência ou palestra).

4 – M. R. TOMÉ, LUIZ. Herramientas proteómicas para la identificación rápida y precisa de hongos. 2020. (Apresentação de Trabalho/Conferência ou palestra).

- **Participação em cursos (Total: 6)**

1 – Análise de Dados em Linguagem R. (Carga horária: 20h). Escola Nacional de Administração Pública, ENAP, Brasil. 2022 - 2022

2 – Bioinformatics Workshop on Virus Evolution and Molecular Epidemiology. (Carga horária: 45h). Katholieke Universiteit Leuven, KU Leuven, Bélgica. 2021 - 2021

3 – I Workshop Online de Bioinformática (WOB20). (Carga horária: 30h). Universidade Federal de Minas Gerais, UFMG, Brasil. 2020 - 2020

4 – Filogenia: pequena-escala (Filogenética) e larga-escala (Filogenômica). (Carga horária: 15h). Universidade Federal de Minas Gerais, UFMG, Brasil. 2020 - 2020

5 – PROTEOMIC HIGH-THROUGHPUT DATA ANALYSIS. (Carga horária: 6h). Universidade Federal de Minas Gerais, UFMG, Brasil. 2019 - 2019

6 – Comunicação e Divulgação científica DivulgaMicro (American Society for Mi. (Carga horária: 8h). Universidade Federal de Minas Gerais, UFMG, Brasil. 2019 – 2019.

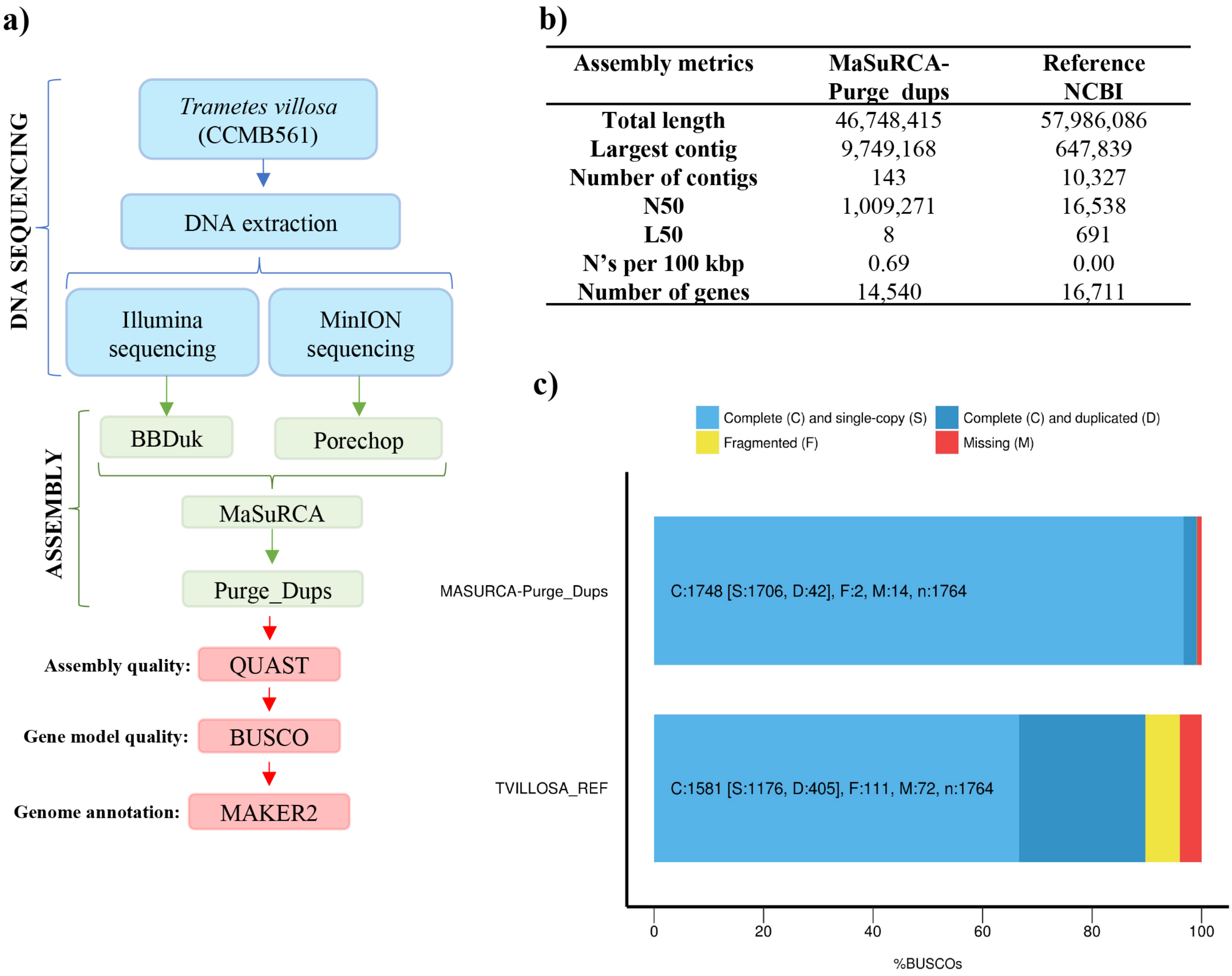
- **Orientação – Iniciação Científica**

  1. Mike Santos. Avaliação da produção de enzimas ligninolíticas de interesse industrial pelo fungo *Trametes villosa* CCMB561. Início: 2019. Iniciação científica (Graduando em Farmácia) - Universidade Federal de Minas Gerais. (Coorientador);

APÊNDICE

Figuras artigo capítulo I – ALTA RESOLUÇÃO

FIGURA 1 (ARTIGO CAPÍTULO 1)



Assembly metrics	MaSuRCA-Purge_dups	Reference NCBI
Total length	46,748,415	57,986,086
Largest contig	9,749,168	647,839
Number of contigs	143	10,327
N50	1,009,271	16,538
L50	8	691
N's per 100 kbp	0.69	0.00
Number of genes	14,540	16,711

FIGURA 2 (ARTIGO CAPÍTULO 1)

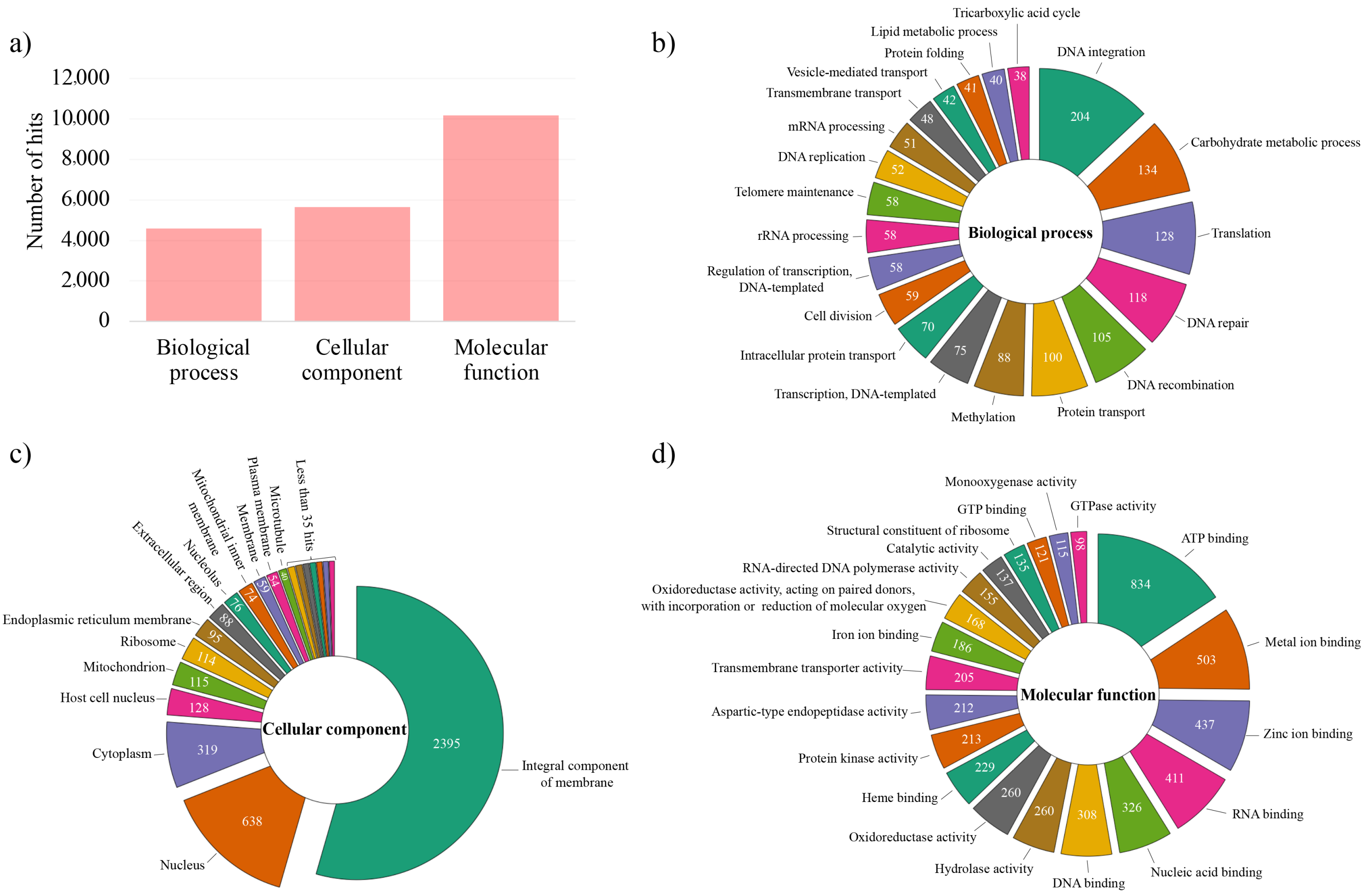


FIGURA 3 (ARTIGO CAPÍTULO 1)

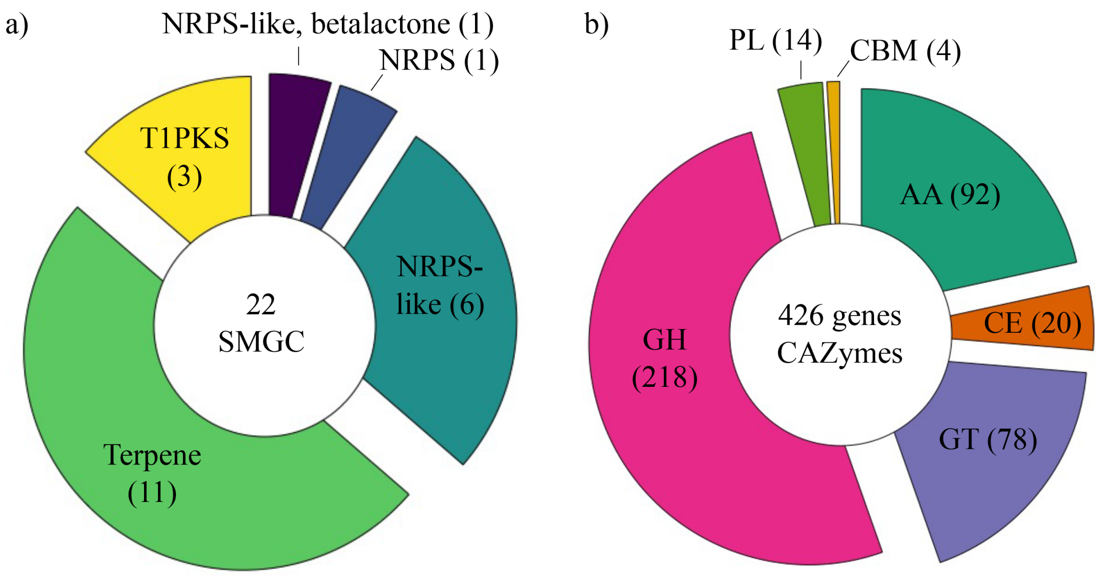
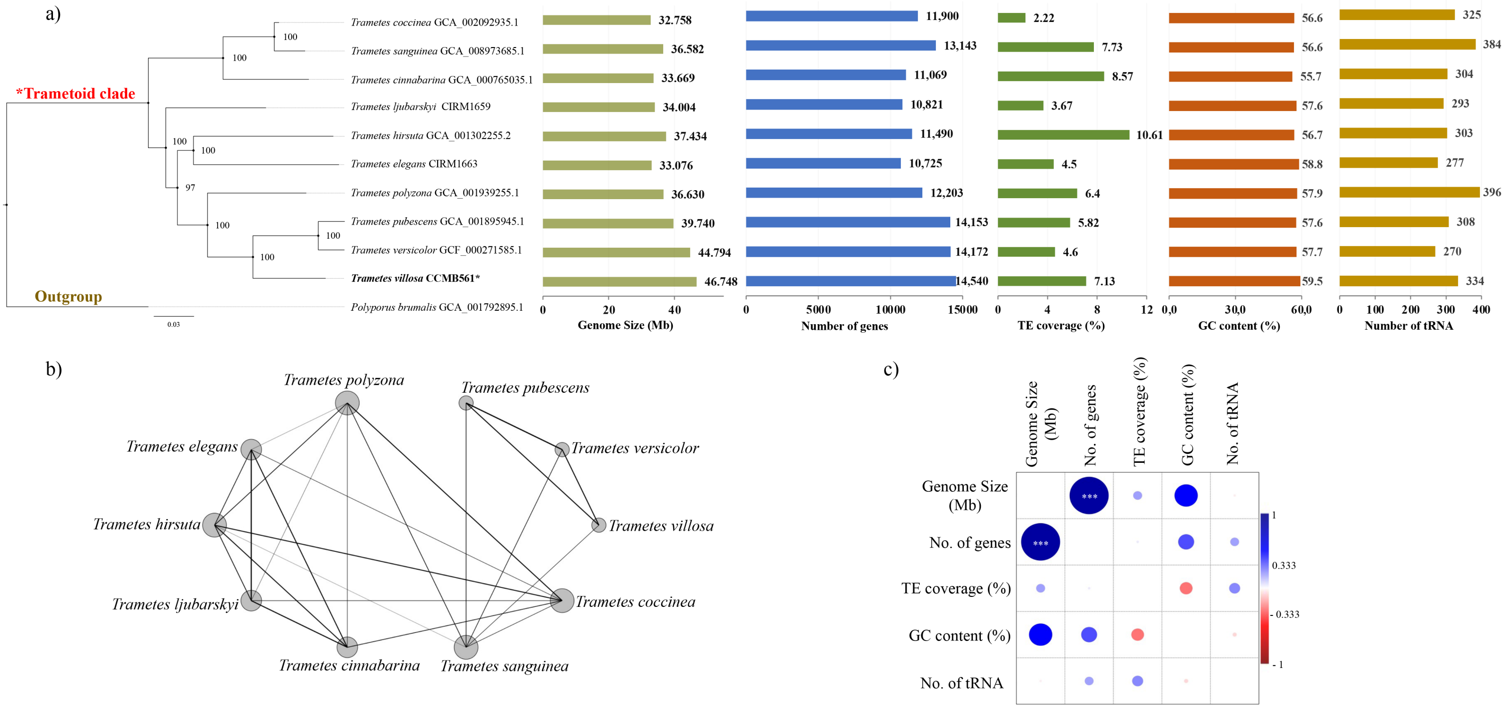
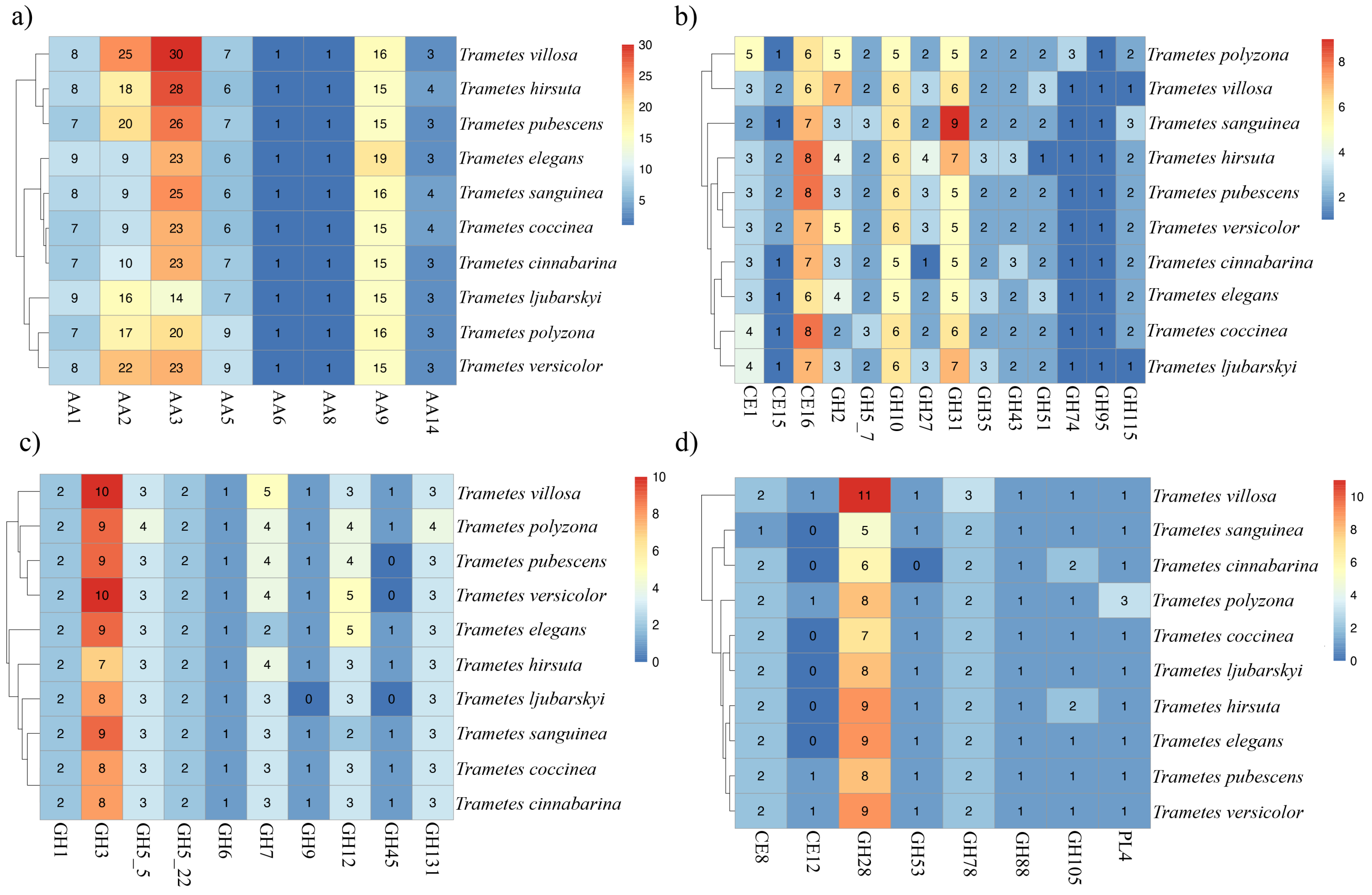


FIGURA 4 (ARTIGO CAPÍTULO 1)



Species	Genome Size (Mb)	Number of genes	TE coverage (%)	GC content (%)	Number of tRNA
Trametes coccinea GCA_002092935.1	32.758	11,900	2.22	56.6	325
Trametes sanguinea GCA_008973685.1	36.582	13,143	7.73	56.6	384
Trametes cinnabarina GCA_000765035.1	33.669	11,069	8.57	55.7	304
Trametes ljubarskyi CIRM1659	34.004	10,821	3.67	57.6	293
Trametes hirsuta GCA_001302255.2	37.434	11,490	10.61	56.7	303
Trametes elegans CIRM1663	33.076	10,725	4.5	58.8	277
Trametes polyzona GCA_001939255.1	36.630	12,203	6.4	57.9	396
Trametes pubescens GCA_001895945.1	39.740	14,153	5.82	57.6	308
Trametes versicolor GCF_000271585.1	44.794	14,172	4.6	57.7	270
Trametes villosa* CCMB561	46.748	14,540	7.13	59.5	334

FIGURA 5 (ARTIGO CAPÍTULO 1)



a)

	AA1	AA2	AA3	AA5	AA6	AA8	AA9	AA14
Trametes villosa	8	25	30	7	1	1	16	3
Trametes hirsuta	8	18	28	6	1	1	15	4
Trametes pubescens	7	20	26	7	1	1	15	3
Trametes elegans	9	9	23	6	1	1	19	3
Trametes sanguinea	8	9	25	6	1	1	16	4
Trametes coccinea	7	9	23	6	1	1	15	4
Trametes cinnabarina	7	10	23	7	1	1	15	3
Trametes ljubarskyi	9	16	14	7	1	1	15	3
Trametes polyzona	7	17	20	9	1	1	16	3
Trametes versicolor	8	22	23	9	1	1	15	3

b)

	CE1	CE15	CE16	GH2	GH5_7	GH10	GH27	GH31	GH35	GH43	GH51	GH74	GH95	GH115
Trametes polyzona	5	1	6	5	2	5	2	5	2	2	2	3	1	2
Trametes villosa	3	2	6	7	2	6	3	6	2	2	3	1	1	1
Trametes sanguinea	2	1	7	3	3	6	2	9	2	2	2	1	1	3
Trametes hirsuta	3	2	8	4	2	6	4	7	3	3	1	1	1	2
Trametes pubescens	3	2	8	3	2	6	3	5	2	2	2	1	1	2
Trametes versicolor	3	2	7	5	2	6	3	5	2	2	2	1	1	2
Trametes cinnabarina	3	1	7	3	2	5	1	5	2	3	2	1	1	2
Trametes elegans	3	1	6	4	2	5	2	5	3	2	3	1	1	2
Trametes coccinea	4	1	8	2	3	6	2	6	2	2	2	1	1	2
Trametes ljubarskyi	4	1	7	3	2	6	3	7	3	2	2	1	1	1

c)

	GH1	GH3	GH5_5	GH5_22	GH6	GH7	GH9	GH12	GH45	GH131
Trametes villosa	2	10	3	2	1	5	1	3	1	3
Trametes polyzona	2	9	4	2	1	4	1	4	1	4
Trametes pubescens	2	9	3	2	1	4	1	4	0	3
Trametes versicolor	2	10	3	2	1	4	1	5	0	3
Trametes elegans	2	9	3	2	1	2	1	5	1	3
Trametes hirsuta	2	7	3	2	1	4	1	3	1	3
Trametes ljubarskyi	2	8	3	2	1	3	0	3	0	3
Trametes sanguinea	2	9	3	2	1	3	1	2	1	3
Trametes coccinea	2	8	3	2	1	3	1	3	1	3
Trametes cinnabarina	2	8	3	2	1	3	1	3	1	3

d)

	CE8	CE12	GH28	GH53	GH78	GH88	GH105	PL4
Trametes villosa	2	1	11	1	3	1	1	1
Trametes sanguinea	1	0	5	1	2	1	1	1
Trametes cinnabarina	2	0	6	0	2	1	2	1
Trametes polyzona	2	1	8	1	2	1	1	3
Trametes coccinea	2	0	7	1	2	1	1	1
Trametes ljubarskyi	2	0	8	1	2	1	1	1
Trametes hirsuta	2	0	9	1	2	1	2	1
Trametes elegans	2	0	9	1	2	1	1	1
Trametes pubescens	2	1	8	1	2	1	1	1
Trametes versicolor	2	1	9	1	2	1	1	1