



UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE BIOLOGIA

DAMARIS BATISTÃO MARTIM

Investigação do metabolismo de compostos aromáticos relacionados à
lignina em um fitopatógeno *Xanthomonas*

Investigation of lignin-related aromatic compounds metabolism in a
Xanthomonas phytopathogen

CAMPINAS
2024

DAMARIS BATISTÃO MARTIM

**Investigação do metabolismo de compostos aromáticos relacionados à
lignina em um fitopatógeno *Xanthomonas***

**Investigation of lignin-related aromatic compounds metabolism in a
Xanthomonas phytopathogen**

*Tese apresentada ao Instituto de Biologia da
Universidade Estadual de Campinas como parte
dos requisitos exigidos para a obtenção do
Título de Doutora em Genética e Biologia
Molecular, na área de Genética de
Microorganismos.*

*Thesis presented to the Institute of Biology of the
University of Campinas in partial fulfillment of
the requirements for the degree of Doctor in
Genetics and Molecular Biology, in the area of
Microorganism Genetics.*

Orientadora: Dra. PRISCILA OLIVEIRA DE GIUSEPPE

ESTE ARQUIVO DIGITAL CORRESPONDE À
VERSÃO FINAL DA TESE DEFENDIDA PELA
ALUNA DAMARIS BATISTÃO MARTIM E
ORIENTADA PELA DR^a. PRISCILA OLIVEIRA DE
GIUSEPPE.

**CAMPINAS
2024**

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Biologia
Mara Janaina de Oliveira - CRB 8/6972

M362i Martim, Damaris Batistão, 1994-
Investigação do metabolismo de compostos aromáticos relacionados à lignina em um fitopatógeno *Xanthomonas* / Damaris Batistão Martim. – Campinas, SP : [s.n.], 2024.

Orientador: Priscila Oliveira de Giuseppe.
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Biologia.

1. Lignina. 2. Compostos aromáticos - Metabolismo. 3. *Xanthomonas*. 4. Transcriptoma. 5. Vias metabólicas. I. Giuseppe, Priscila Oliveira. II. Universidade Estadual de Campinas. Instituto de Biologia. III. Título.

Informações Complementares

Título em outro idioma: Investigation of lignin-related aromatic compounds metabolism in a *Xanthomonas* phytopathogen

Palavras-chave em inglês:

Lignin

Aromatic compounds - Metabolism

Xanthomonas

Transcriptome

Metabolic pathways

Área de concentração: Genética de Microorganismos

Titulação: Doutora em Genética e Biologia Molecular

Banca examinadora:

Priscila Oliveira de Giuseppe [Orientador]

Ljubica Tasic

Cristina Elisa Alvarez Martinez

Maxuel de Oliveira Andrade

Daniela Barretto Barbosa Trivella

Data de defesa: 22-02-2024

Programa de Pós-Graduação: Genética e Biologia Molecular

Identificação e informações acadêmicas do(a) aluno(a)

- ORCID do autor: <https://orcid.org/0000-0003-3170-0352>

- Currículo Lattes do autor: <http://lattes.cnpq.br/5347855819169430>

Campinas, 22 de fevereiro de 2024.

COMISSÃO EXAMINADORA

Prof.(a) Dra. Priscila Oliveira de Giuseppe

Prof.(a). Dra. Ljubica Tasic

Prof.(a) Dra. Cristina Elisa Alvarez Martinez

Prof.(a) Dr. Maxuel de Oliveira Andrade

Prof.(a) Dra. Daniela Barretto Barbosa Trivella

Os membros da Comissão Examinadora acima assinaram a Ata de Defesa, que se encontra no processo de vida acadêmica do aluno.

A Ata da defesa com as respectivas assinaturas dos membros encontra-se no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria do Programa de Pós-Graduação em Genética e Biologia Molecular da Unidade do Instituto de Biologia da Unicamp

AGRADECIMENTOS

Primeiramente agradeço a Deus, por todo cuidado, livramentos, superações e conquistas alcançadas ao longo desses anos.

Agradeço a minha orientadora Priscila Oliveira de Giuseppe pela excelente orientação, apoio, incentivo, dedicação, confiança e por todo o conhecimento e experiências construídos.

À equipe BCoolL, pelo companheirismo e por todas as discussões que contribuíram para a conclusão deste trabalho. Em especial, agradeço a Anna Ju e ao Augusto pela amizade, apoio, paciência e por todos os bons momentos que passamos ao longo desses 4 anos, compartilhando hipóteses, ideias e parcerias.

Ao Wagner, por todo apoio, incentivo, paciência, companheirismo e compreensão nos momentos mais difíceis.

A meus pais, Roseli e Pedro Gil por todo amor, carinho e incentivo.

A toda minha família e amigos, que mesmo distantes, sempre se fizeram presentes e me incentivaram na busca desse sonho.

Aos pesquisadores, funcionários e alunos do LNBR pela troca de experiências e agradável ambiente de trabalho. Em especial a turma do café lado B pelos momentos compartilhados.

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pelo apoio financeiro referente a bolsa de doutorado (Processo 2019/08590-8).

Ao Programa de Pós-Graduação em Genética e Biologia Molecular da UNICAMP.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) - Código de Financiamento 001.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Enfim, um agradecimento especial a todos que, de alguma forma, contribuíram para que eu concluísse este trabalho de maneira excepcional.

"Porque sou eu que conheço os planos
que tenho para vocês, diz o Senhor,
planos de fazê-los prosperar e não de lhes
causar dano, planos de dar-lhes
esperança e um futuro."

Jeremias 29:11

RESUMO

O aproveitamento sustentável dos resíduos de lignina para produção de compostos químicos de alto valor agregado tem emergido como uma estratégia promissora para se diminuir a dependência das indústrias fósseis. Nesse contexto, o conceito de se utilizar microrganismos para transformar a variedade de compostos resultantes da despolimerização da lignina em bioprodutos, formados em pontos de convergência do metabolismo microbiano, tem sido amplamente explorado nas últimas décadas. No entanto, o entendimento atual das vias metabólicas responsáveis pela bioconversão de compostos relacionados a lignina ainda é limitado a poucas espécies e elusivo para algumas dessas moléculas. Logo, a busca por bactérias capazes de metabolizar compostos relacionados a lignina é fundamental para identificação de novas enzimas ativas em aromáticos, bem como para o desenvolvimento de novas estratégias biotecnológicas visando a valorização da lignina. O presente trabalho teve como objetivo avaliar a compatibilidade dos produtos da despolimerização hidrotérmica da lignina com estratégias de bioconversão em diferentes gêneros bacterianos, além de identificar e validar novas vias do metabolismo de compostos aromáticos relacionados à lignina, usando como modelo de estudo um fitopatógeno *Xanthomonas*, gênero pouco explorado para esse propósito. Os resultados desse trabalho foram compilados em dois artigos científicos, sendo um já publicado e outro submetido para publicação, os quais são apresentados na forma de capítulos da tese. O primeiro capítulo descreve as análises sobre o efeito da despolimerização hidrotérmica de ligninas isoladas em termos de rendimento, composição e compatibilidade dos bio-óleos obtidos com a bioconversão microbiana. Embora o bio-óleo de alta massa molecular (pesado), não tenha promovido o crescimento bacteriano, o bio-óleo leve, rico em monômeros aromáticos, promoveu o crescimento de 9 bactérias, indicando sua compatibilidade com estratégias de conversão microbiana. Em sequência, o capítulo dois explora de forma detalhada as vias de bioconversão de compostos relacionados a lignina em *Xanthomonas citri* subsp. *citri* 306 (*X. citri* 306). *X. citri* 306 foi capaz de metabolizar diferentes compostos relacionados a lignina, incluindo os três principais monolignóis (álcool *p*-cumarílico (H), coniferílico (G) e sinapílico (S)). Utilizando abordagens como RNA-seq, nocauteamento gênico e ensaios bioquímicos, conseguimos identificar vias completas para o catabolismo desses três monolignóis, além de novas redutases e transportadores de efluxo provavelmente utilizados pela bactéria para contornar a toxicidade dos compostos aromáticos. Em suma, os dois capítulos fornecem conhecimento fundamental que pode servir como base para estudos futuros voltados para o desenvolvimento de estratégias biotecnológicas visando o aproveitamento de resíduos de lignina como matéria prima para uma produção mais sustentável de químicos.

ABSTRACT

The sustainable exploitation of lignin side-streams to produce high-value chemical compounds has emerged as a promising strategy to reduce the dependency of fossil industries. In this context, the concept of using microorganisms to transform the variety of compounds resulting from the lignin depolymerization into bioproducts, formed in convergence points of microbial metabolism, has been widely explored in recent decades. However, the current knowledge of metabolic pathways responsible for the bioconversion of lignin-related compounds is still constrained to a limited number of species and elusive for some of these compounds. Therefore, the search for bacteria capable of metabolizing lignin-related compounds is essential for identifying new enzymes active on aromatics, as well as for the development of new biotechnological strategies aimed at lignin valorization. The present study aimed to evaluate the compatibility of products from hydrothermal depolymerization of lignin with bioconversion strategies in different bacterial genera, in addition to identifying and validating new pathways for the metabolism of lignin-related aromatic compounds, using as a study model a plant pathogen *Xanthomonas*, a genus relatively unexplored for this purpose. The results of this work have been compiled into two scientific papers, one already published and the other submitted for publication, which are presented as chapters of the thesis. In the first chapter, we conducted analysis on the effect of hydrothermal depolymerization of isolated lignin in terms of yield, composition, and compatibility of the resulting bio-oils with microbial bioconversion. Although the heavy bio-oil with high molecular mass did not support bacterial growth, the light bio-oil, rich in aromatic monomers, promoted the growth of nine bacteria, indicating its compatibility with microbial conversion strategies. Subsequently, the second chapter extensively explores the bioconversion pathways of lignin-related compounds in *X. citri* subsp. *citri* 306 (*X. citri* 306). *X. citri* 306 was capable of metabolizing different lignin-related compounds, including the three main monolignols (*p*-coumaryl (H), coniferyl (G), and sinapyl alcohol (S)). Using approaches such as RNA-seq, gene knockout, and biochemical assays, we identified complete pathways for the metabolism of the three monolignols, in addition to new reductases and efflux transporters likely used by the bacterium to counterbalance the toxicity of aromatic compounds. In summary, these two chapters provide fundamental knowledge that can be used for further studies aimed at developing biotechnological strategies for utilizing lignin waste as a feedstock for a more sustainable production of chemicals.

SUMÁRIO

APRESENTAÇÃO.....	10
INTRODUÇÃO	11
Aplicação do afunilamento biológico para a valorização da lignina	16
<i>Xanthomonas citri</i> subsp. <i>citri</i> um alvo promissor e pouco explorado para o estudo do metabolismo de compostos relacionados a lignina.....	17
Capítulo 1	19
Exploring the compatibility between hydrothermal depolymerization of alkaline lignin from sugarcane bagasse and metabolization of the aromatics by bacteria.....	19
Abstract.....	20
1. Introduction.....	21
2. Material and methods.....	22
3. Results and discussion	26
4. Conclusions.....	35
References.....	37
Supplementary data.....	41
Capítulo 2	51
Unraveling aromatic metabolism in a plant pathogen reveals novel pathways for detoxification and metabolic convergence of lignin monomers	51
Abstract.....	52
Introduction.....	53
Results.....	54
Discussion	72
Material and Methods	76
References.....	87
Supplementary information	92
CONSIDERAÇÕES FINAIS.....	121
Referências.....	124
ANEXOS.....	127

APRESENTAÇÃO

Esta Tese está organizada em quatro seções principais: Introdução geral, Capítulo 1, Capítulo 2, e Considerações Finais. A Introdução Geral fornece uma visão sobre o tema de bioconversão microbiana de compostos relacionados a lignina e seu papel emergente na transformação de resíduos de lignina em produtos químicos de maior valor agregado.

Os dois capítulos seguintes foram escritos em formato de artigo na língua inglesa. O primeiro capítulo apresenta um artigo publicado na revista *International Journal of Biological Macromolecules* em novembro de 2022, elaborado em colaboração com o Professor George Jackson de Moraes Rocha, onde compartilho a posição de primeira-autora com a Dra. Fabrícia Farias de Menezes. Esse artigo pode ser dividido em duas partes. A primeira delas aborda a produção e a caracterização composicional de bio-óleos provenientes de diferentes condições de despolimerização hidrotérmica de uma lignina extraída do bagaço de cana-de-açúcar. Essa parte foi conduzida principalmente pela Dra. Fabrícia Menezes com o apoio técnico da equipe do Laboratório de Química Molecular do LNBR. Na segunda parte, nossa contribuição consistiu na avaliação da compatibilidade biológica desses bio-óleos com a conversão microbiana. Para isso, desenvolvemos um método de triagem de crescimento microbiano rápido e miniaturizado e o aplicamos para testar a capacidade de nove gêneros bacterianos diferentes, potencialmente capazes de metabolizar aromáticos, de crescer utilizando um bio-óleo de alta massa molecular (bio-óleo pesado) ou um bio-óleo rico em monômeros (bio-óleo leve) como fonte primária de carbono. Informações adicionais podem ser encontradas no material suplementar do artigo (<https://doi.org/10.1016/j.ijbiomac.2022.10.269>)

O segundo capítulo da tese refere-se a um manuscrito submetido para publicação na revista *Nature Communications*. Neste artigo, selecionamos *X. citri* 306 como organismo modelo para um estudo abrangente e detalhado de vias metabólicas de compostos aromáticos relacionados a lignina. *X. citri* 306 pertence a um gênero pouco explorado para o estudo de vias de bioconversão de aromáticos, mostrando-se como um nicho promissor para se desvendar novas estratégias metabólicas para a transformação desses compostos em outros produtos. Informações adicionais poderão ser encontradas nos dados suplementares do artigo após a publicação.

Por fim, a última seção da tese é composta por considerações finais sobre os dois capítulos aqui apresentados.

INTRODUÇÃO

A incorporação de fontes renováveis para produção de combustíveis, energia e produtos químicos de valor agregado tem emergido como uma solução viável e fundamental para reduzir a dependência da indústria fóssil, tornando nossos processos produtivos mais sustentáveis do ponto de vista social, econômico e ambiental. Nesse contexto, devido a seu baixo custo e alta disponibilidade, a biomassa lignocelulósica, um material altamente oxigenado, composto principalmente por três tipos de biopolímeros, celulose, hemicelulose e lignina, emerge como uma fonte rica de carbono orgânico renovável para a bioenergética e produção de químicos de interesse tecnológico^{1,2}.

A lignina, o componente mais subestimado de biomassas vegetais, é uma macromolécula aromática complexa encontrada abundantemente nas paredes celulares das plantas, e cuja função biológica primária é de conferir resistência física, rigidez e sustentação mecânica ao tecido vegetal, facilitando a absorção de água e transporte de nutrientes e formando uma barreira recalcitrante à degradação e ataques microbianos³. Na indústria de papel e celulose e de etanol celulósico, milhares de toneladas de lignina são produzidas anualmente como resíduo. Devido ao seu alto poder calorífico, esse subproduto é habitualmente utilizado para fins energéticos. No entanto, a riqueza de compostos aromáticos presente na lignina bem como sua abundância, fazem dela um resíduo valioso da indústria de biomassa, sendo uma promissora fonte de compostos aromáticos alternativa aos fósseis⁴. O crescente interesse na utilização da lignina e seus componentes para produção de substâncias como vanilina, polihidroxicumaratos e ácido adípico tem impulsionado diversas pesquisas que visam apoiar o desenvolvimento de tecnologias para uma indústria lignocelulósica com mínimo desperdício, contribuindo para a sustentabilidade econômica, social e ambiental de seus processos⁵.

A lignina é gerada pela polimerização de três principais álcoois aromáticos denominados de álcool *p*-cumarílico (H), álcool coniferílico (G), e álcool sinapílico (S), e suas proporções variam de espécie para espécie e de tecido para tecido (Figura 1). O teor de lignina encontrado para madeira macia, por exemplo, varia entre 25 a 30% (predominando unidades G), para madeira dura entre 16 e 24%, (unidades G e S) e para gramíneas entre 13 e 18% (contendo unidades G, S e H)². Além da heterogeneidade vinda das diferentes proporções das unidades de H, G e S, a abundância relativa das ligações que conectam esses blocos químicos também são variáveis dependendo do tipo e fonte de biomassa, e incluem ligações do tipo β -O-4' éter alquil-arílico e β -5' (fenil-cumarano), além de acilações nas unidades de hidroxila das cadeias laterais da lignina de tipo G, principalmente com *p*-cumaratos⁶.

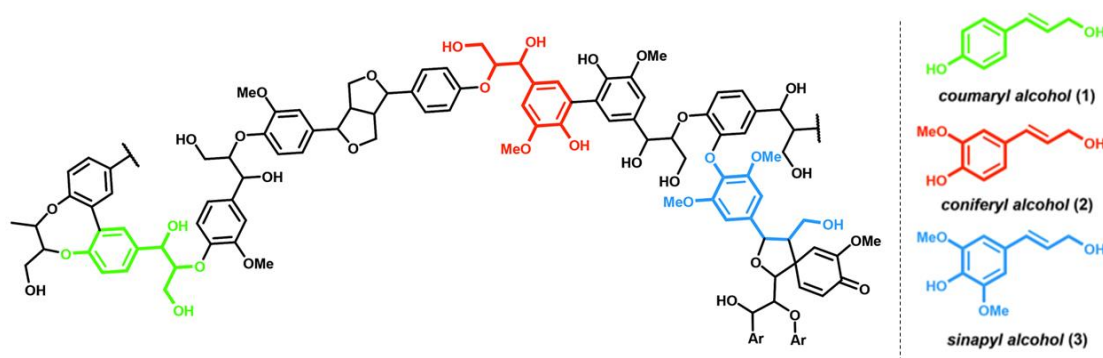


Figura 1. Estrutura molecular da lignina, destacando os principais blocos químicos utilizados para sua síntese. Adaptado de Kärkäs *et al.*⁷

A transformação eficaz da lignina em químicos de alto valor agregado demanda a integração hábil de três etapas-chave: (i) fracionamento da lignocelulose, (ii) fragmentação da lignina, e (iii) sua conversão em produtos químicos específicos desejados⁸. O fracionamento é essencial para reduzir a complexidade, heterogeneidade e a recalcitrância da biomassa lignocelulósica. Atualmente, existem diversas ligninas isoladas com características estruturais e químicas distintas, muitas das quais são produzidas com propósitos comerciais ou industriais, como a lignina kraft, lignosulfonatos e lignina alcalina⁸.

Para a fragmentação da lignina isolada, uma variedade de técnicas tem sido explorada, incluindo métodos redutores, oxidativos, catalisados por base e ácido, além de abordagens enzimáticas, solvolíticas e térmicas⁹. Cada método empregado apresenta suas vantagens e limitações que influenciam na natureza e rendimento final dos monômeros obtidos, enfatizando a necessidade de pesquisas direcionadas para a otimização de processos e a redução dos custos operacionais¹⁰.

Os métodos oxidativos ou catalisados por ácido ou base são técnicas químicas convencionais conhecidas por sua eficácia na conversão da lignina. Porém, essas abordagens demandam altas temperaturas e pressões, equipamentos caros e o uso de substâncias químicas tóxicas¹¹. Outras opções incluem as abordagens pirolíticas, mas estas também possuem desvantagens intrínsecas, como menor seletividade e necessidade de temperaturas mais elevadas, variando entre 300 e 600 °C¹². Por outro lado, a despolimerização hidrotérmica oferece algumas vantagens em comparação com os métodos convencionalmente utilizados. Nesse processo, emprega-se a técnica de hidrogenólise catalítica, em que a água atinge condições supercríticas (<374 °C; 22 MPa) exibindo propriedades únicas, como baixa viscosidade, alta solubilidade de compostos orgânicos, e o aumento na constante de dissociação,

que favorece a liberação de aromáticos bifuncionais¹⁰. Além disso, a água desempenha um papel essencial não só como solvente verde, mas também como um dos catalisadores, conferindo ao processo um caráter sustentável e ecologicamente amigável¹⁰.

No entanto, independentemente do método empregado, devido à heterogeneidade estrutural da lignina, sua fragmentação resulta em um amplo espectro de compostos aromáticos monoméricos, bem como dímeros ou oligômeros de lignina de alta massa molecular. Tal heterogeneidade química figura como um dos principais desafios para a utilização eficiente da lignina como matéria-prima para a produção de químicos de alto valor agregado, onde a pureza e altos rendimentos são de suma importância^{13,14}. Neste contexto, processos baseados em bactérias capazes de “afunilar” misturas complexas de compostos derivados de lignina em produtos alvo específicos emergem como um meio de se superar esse gargalo (Figura 2)¹⁴.

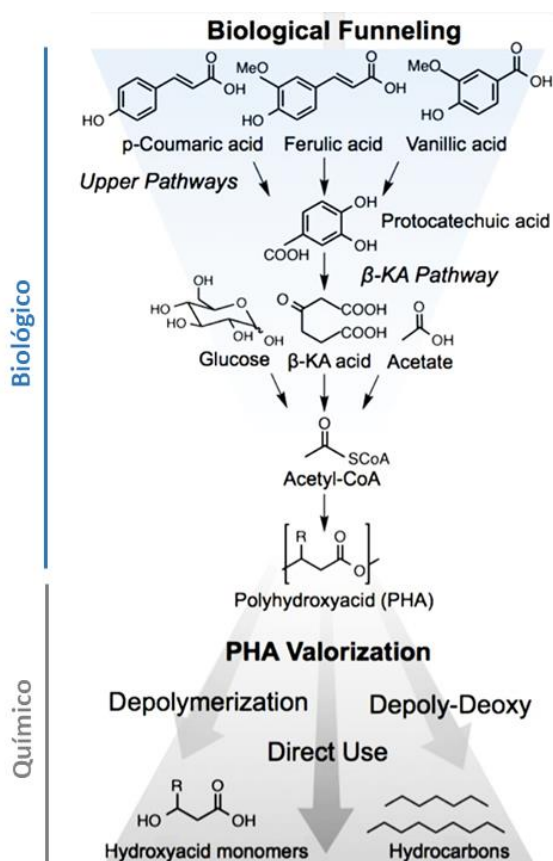


Figura 2. Esquema representativo do conceito de afunilamento biológico. Nesse esquema, alguns compostos modelo de lignina (*p*-cumarato, ferulato e vanilato) são convertidos ao intermediário químico central (protocatecuato) e subsequentemente ao acetil-coA que pode ser convertido a polihidroxialcanoato, o qual pode ser utilizado diretamente para produção do bioplástico PHA, ou convertido em outros componentes de valor agregado a partir de processos químicos. Adaptado de Linger *et al.*¹³.

Durante o catabolismo, os aromáticos derivados da lignina são progressivamente transformados através de uma série de reações enzimáticas conhecidas como vias periféricas ou vias de afunilamento^{13,15}, as quais convergem em dois possíveis intermediários químicos centrais, catecol ou protocatecoato, que seguem para vias de fissão de anel, onde os anéis aromáticos são clivados e os metabólitos subsequentes entram no metabolismo central do carbono¹⁶ (Figura 3). Tais vias, podem ser discretizadas de acordo com o tipo de monômero de lignina que ela processa: monômeros do tipo H, que derivam do álcool *p*-cumarílico (sem metoxilação no anel); do tipo G, que derivam do álcool coniferílico (1 metoxilação no anel); e do tipo S, que derivam do álcool sinapílico (2 metoxilações no anel).

A via de degradação do ácido *p*-cumárico é mais amplamente investigada para representar os monômeros do tipo H. Em geral, a degradação do ácido *p*-cumárico por bactérias pode ser dividida em três rotas: (1) Via de β -oxidação dependente de CoA; (2) Via de não- β -oxidação dependente de CoA e (3) Via independente de CoA, as quais convergem num intermediário comum, o ácido *p*-hidroxibenzóico, que posteriormente é transformado em protocatecoato (Figura 3). Este por sua vez pode ser convertido em catecol ou degradado à acetil-coA. A degradação do protocatecoato em bactérias é classificada em três vias: (1) Via de clivagem 3,4 (orto-clivagem); (2) Via de clivagem 4,5 (meta-clivagem) e (3) Via de clivagem 2,3^{17,18} (Figura 3).

A degradação de compostos do tipo G inclui a via de afunilamento do ácido ferúlico, um composto modelo padrão para esse tipo de unidade. A via de degradação do ácido ferúlico pode ser dividida em quatro vias distintas: (1) Via de descarboxilação não oxidativa; (2) Via de não- β -oxidação dependente de CoA; (3) Via de β -oxidação dependente de CoA e (4) Via de redução da cadeia lateral (Figura 3). Estas quatro vias envolvem diferentes intermediários e enzimas, no entanto todas apresentam como produto final o ácido vanílico, que é degradado via fissão do protocatecoato após uma etapa de desmetilação¹⁸.

A presença de dois grupos metóxi no anel aromático dos compostos do tipo S torna sua metabolização mais complicada. Nesta via, o composto modelo ácido sirínico é primeiramente *O*-desmetilado em 3-*O*-metilgalato (3OMG) que, posteriormente, pode ser convertido em três diferentes intermediários (Figura 3). Os três caminhos convergem a um intermediário da via de clivagem 4,5 do protocatecoato, o 4-oxalomesaconato, que ao final é integrado ao ciclo do ácido tricarboxílico (TCA), resultando em acetil-CoA (Figura 3)¹⁸.

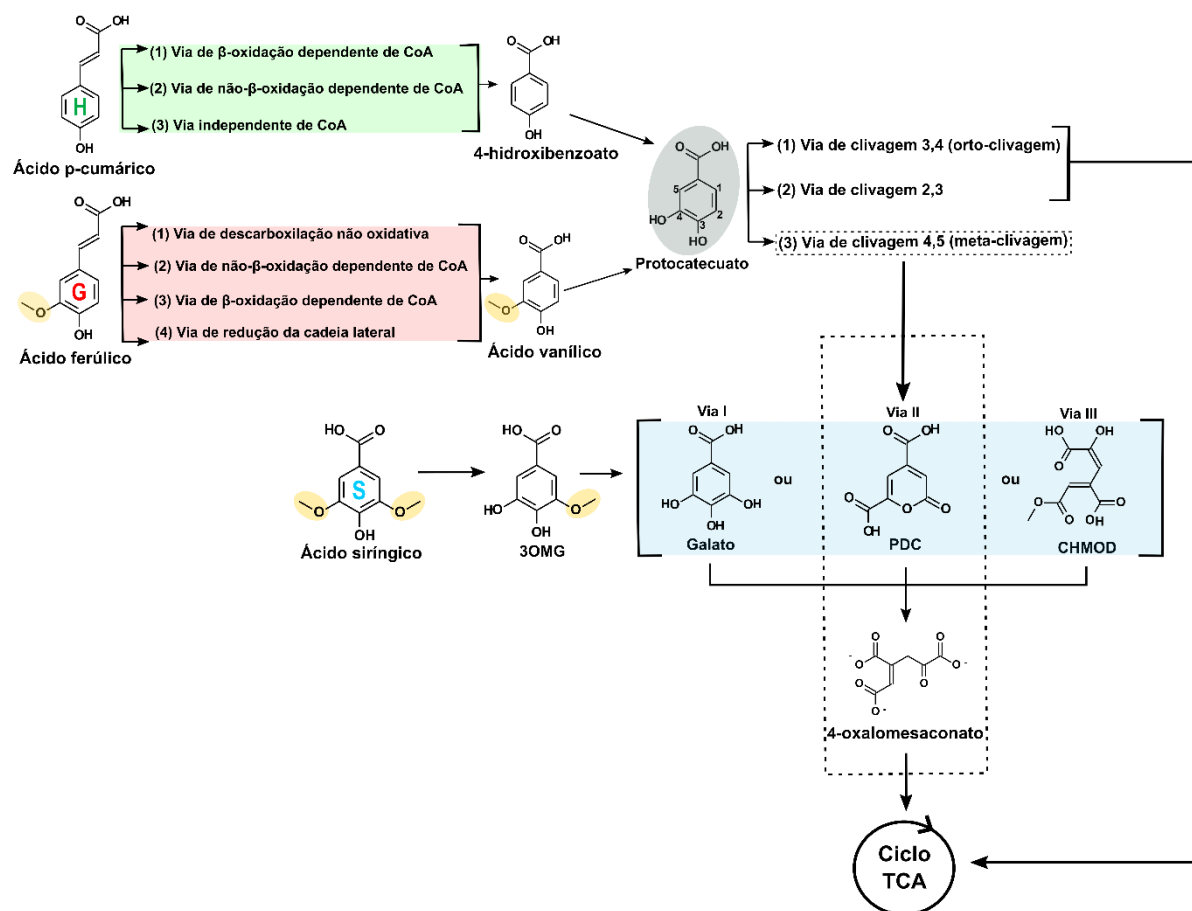


Figura 3. Esquema simplificado das vias catabólicas de compostos derivados dos três principais tipos (H,G,S) de unidades derivadas da lignina. 3OMG, 3-*O*-Metilgalato; PDC, 2-pirona-4,6-dicarboxilato; CHMOD, 4-carbóxi-2-hidróxi-6-metóxi-6-oxohexa-2,4-dienoato.

Apesar do interesse crescente na identificação das enzimas e mecanismos envolvidos na bioconversão microbiana dos compostos aromáticos derivados da lignina, a compreensão desses processos ainda é limitada. No "*eLignin Microbial Database*", um banco de dados manualmente curado que reúne informações disponíveis na literatura sobre enzimas envolvidas no metabolismo microbiano de compostos aromáticos¹⁹, mais de 170 espécies bacterianas foram identificadas como capazes de metabolizar compostos relacionados a lignina. No entanto, a maior parte do conhecimento atual sobre o metabolismo desses compostos baseia-se em estudos focados em organismos modelo comuns, como *Pseudomonas putida*, *Rhodococcus jostii* RHA1 e *Sphingobium* sp. SYK-6^{17,18}. Além disso, apenas cerca de 130 genes foram catalogados no *elignin database* com base em alguma evidência experimental que indique seu envolvimento na bioconversão de aromáticos, sendo que menos de 20% desses genes codificam enzimas com alguma caracterização bioquímica e/ou estrutural¹⁹.

Adicionalmente, a despolimerização da lignina pode gerar uma ampla diversidade de monômeros com diferentes grupos funcionais, tais como álcoois, ácidos, aldeídos e fenóis. Em geral, os estudos de identificação de vias para o afunilamento de derivados da lignina concentram-se na identificação das enzimas responsáveis pelo catabolismo dos compostos modelo padrão para cada unidade H (*p*-cumarato), G (ferulato) e S (siringato), e carecem de estudos e caracterização de enzimas responsáveis pela bioconversão dos compostos precursores desses ácidos, tais como álcoois e aldeídos¹⁹.

Juntos, esses fatos ressaltam a importância de expandir os estudos sobre enzimas e vias de bioconversão de aromáticos em uma variedade maior de organismos, contribuindo para uma compreensão mais abrangente da diversidade de estratégias disponíveis na natureza para bioconversão desses compostos. Essa base de conhecimento fundamental é de suma importância para apoiar o desenvolvimento e a engenharia de plataformas microbianas dedicadas a transformação de compostos derivados da lignina em bioprodutos, contribuindo para viabilizar o seu uso como matéria prima alternativa aos fósseis para a produção renovável de químicos de interesse industrial.

Aplicação do afunilamento biológico para a valorização da lignina

Um rápido levantamento da literatura recente voltada para o estudo e desenvolvimento de estratégias de bioconversão de lignina mostra uma quantidade substancial de artigos de pesquisa que se concentram em alguns organismos modelo, dentre os quais, *Sphingobium sp.* SYK-6 e *P. putida*. A compreensão de vias lignolíticas para compostos como β -aril éteres, bifenilos, pinosresinol e ácido ferúlico, por exemplo, é derivada principalmente de estudos com *Sphingobium sp.* SYK-6¹⁷.

O primeiro trabalho a demonstrar o conceito e a aplicação do afunilamento biológico para a valorização da lignina é relativamente recente e mostra a conversão de lignina (oriunda de pré-tratamento de biomassa) em polihidroxialcanoatos de cadeia média pela bactéria *Pseudomonas putida* KT2440¹³. Os polihidroxialcanoatos de cadeia média são aplicáveis na produção de bioplásticos, e podem também ser convertidos, por processos químicos, a hidrocarbonetos combustíveis, evidenciando não só a aplicação do afunilamento biológico, mas também a integração bem-sucedida deste, com métodos químicos para a valorização da lignina (Figura 2)¹³.

Alguns estudos recentes também mostraram que bactérias tais como *P. putida*, *Corynebacterium glutamicum* e *Amycolatopsis sp.* ATCC 39116 podem ser modificadas geneticamente para potencializar a bioconversão de compostos aromáticos derivados de lignina

em ácido *cis,cis* mucônico^{18,20}. Outros trabalhos têm explorado a integração de métodos de pré-tratamento, despolimerização da lignina com o afunilamento biológico dos monômeros fenólicos resultantes em produtos químicos de valor agregado como ácido 2-pirone-4,6-dicarboxílico (PDC)²¹.

Embora o conceito de afunilamento biológico seja relativamente novo, sua aplicação na utilização da lignina como matéria-prima para a produção de produtos químicos de base biológica mostra-se altamente promissora. Pesquisas futuras nessa direção são fortemente encorajadas, com o objetivo de aprimorar ainda mais sua relevância industrial, avaliar sua viabilidade para aplicações em biorrefinarias a longo prazo e impulsionar inovações que alinhem os três principais processos da biorrefinaria – fracionamento, despolimerização e transformação dos monômeros liberados da biomassa em bioprodutos²².

Xanthomonas citri subsp. citri um alvo promissor e pouco explorado para o estudo do metabolismo de compostos relacionados a lignina

Em 1978, Odier e Monties publicaram o primeiro trabalho sobre biodegradação de lignina de trigo por *Xanthomonas sp.*²³. Na década de 80, Kern demonstrou que um isolado bacteriano de *Xanthomonas sp.* era capaz de degradar dehidropolímeros de álcool coníferílico (DHP) e complexos lignocelulósicos derivados do milho²⁴. Três anos depois, Kern e Kirk²⁵ publicaram outro trabalho mostrando que *Xanthomonas sp. strain 99* era capaz de degradar ligninas com massa molecular na faixa de 600 a 1000 Da²⁵. Desde então, observa-se uma lacuna de quase 30 anos nas publicações sobre degradação de lignina por *Xanthomonas spp.* Porém, com o advento das ferramentas ômicas, estudos recentes sobre consórcios microbianos para degradação de biomassa vegetal voltaram a mostrar *Xanthomonas spp.* como bactérias capazes de degradar materiais lignocelulósicos²⁶.

Xanthomonas citri subsp. citri (*X. citri*) é uma bactéria Gram-negativa conhecida por sua habilidade em infectar e colonizar tecidos vegetais de plantas cítricas, bem como por sua versatilidade metabólica, refletida pelo seu vasto arsenal de enzimas ativas sobre carboidratos da parede celular vegetal^{27,28}. Além disso, o genoma de *X. citri* também apresenta genes relacionados ao metabolismo de compostos aromáticos; porém, os estudos sobre a capacidade de *X. citri* em catabolizar compostos derivados da lignina são escassos e o conhecimento sobre as proteínas envolvidas nesse processo ainda é bastante limitado em comparação com outros aspectos do seu metabolismo.

Segundo buscas no banco de dados BRENDA²⁹ e análises da literatura, no gênero *Xanthomonas*, apenas 5 enzimas dessa categoria apresentam dados funcionais: a hidroxilase

PobA (homóloga a XAC0356), que converte 4-hidroxibenzoato em protocatecuato³⁰, a transferase PcaIJ (homóloga a XAC0364/XAC0365), que transforma β -cetoadipato em β -cetoadipil-coA^{30,31} e o cluster hca (homólogo a XAC0881-83) composto por 3 enzimas envolvidas no metabolismo de ácidos hidroxicinâmicos³².

Além disso, a particularidade do habitat colonizado por essa bactéria, que difere dos modelos bacterianos tradicionalmente explorados – predominantemente bactérias de solos ou de efluentes industriais –, oferece uma perspectiva única de se investigar a existência de mecanismos adaptativos e estratégias metabólicas singulares em fitopatógenos. Deste modo, escolhemos *X. citri* 306 como organismo modelo para o estudo do metabolismo de aromáticos relacionados a lignina, visando descobrir, caracterizar e desvendar o papel biológico de novas enzimas ativas sobre compostos aromáticos, além de contribuir para ampliar a compreensão vigente sobre o metabolismo de aromáticos em patógenos de plantas.

Capítulo 1

Exploring the compatibility between hydrothermal depolymerization of alkaline lignin from sugarcane bagasse and metabolization of the aromatics by bacteria

Fabília Farias de Menezes^{a,1}, Damaris Batistão Martim^{a,b,1}, Liu Yi Ling^a, Aline Tieppo Nogueira Mulato^a, Elaine Crespim^c, Juliana Velasco de Castro Oliveira^a, Carlos Eduardo Driemeier^a, Priscila Oliveira de Giuseppe^{a,*}, George Jackson de Moraes Rocha^{a,*}

¹ Both authors contributed equally to this work and should be considered co-first authors.

^a Brazilian Biorenewable National Laboratory (LNBR), Brazilian Center for Research in Energy and Materials (CNPem), 13083-100, Campinas, SP, Brazil.

^b Graduate Program in Genetics and Molecular Biology, Biology Institute, State University of Campinas, 13083-970, Campinas, SP, Brazil.

^c Laboratory of Regulatory Systems Biology, Center for Nuclear Energy in Agriculture at the University of São Paulo (CENA/USP), 13416-000, Piracicaba, SP, Brazil.

*Corresponding authors. (P.O. de Giuseppe) priscila.giuseppe@lnbr.cnpem.br and (G.J. de Moraes Rocha) george.rocha@lnbr.cnpem.br

Revista: International Journal of Biological Macromolecules

Publicado em 4 de novembro de 2022

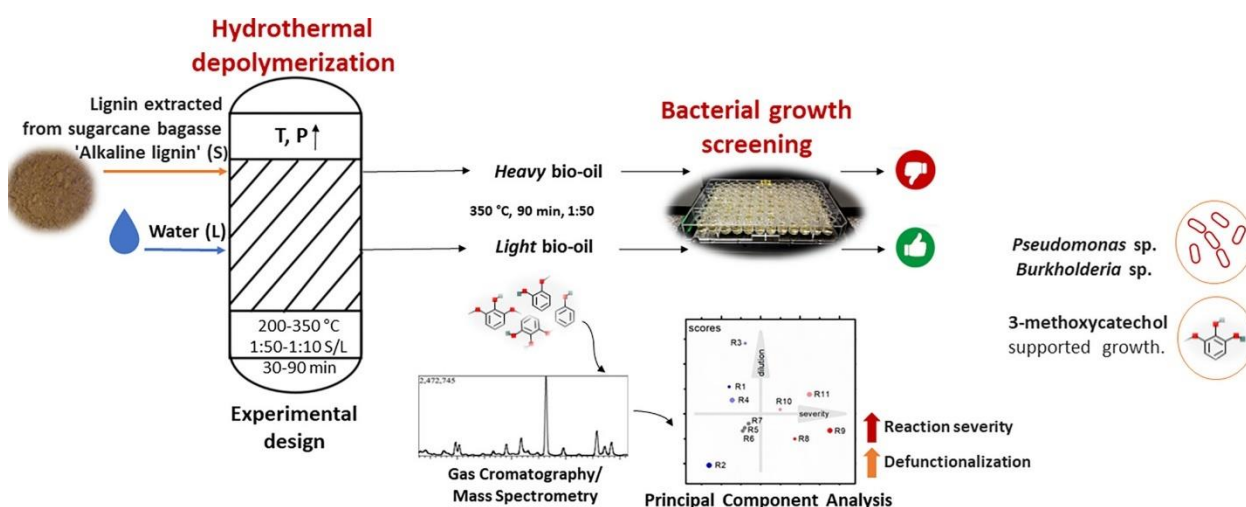
DOI: <https://doi.org/10.1016/j.ijbiomac.2022.10.269>

Abstract

Although hydrothermal treatments for biomass fractionation have been vastly studied, their effect on the depolymerization of isolated lignins in terms of yield, composition, and compatibility of the produced lignin bio-oils with bioconversion is still poorly investigated. In this study, we evaluated the hydrothermal depolymerization of an β -O-4'-rich lignin extracted from sugarcane bagasse by alkaline fractionation, investigating the influence of temperature (200-350 °C), time (30-90 min), and solid-liquid ratio (1:10-1:50 m.v⁻¹) on yield of bio-oils (up to 31 wt.%) rich in monomers (light bio-oils). Principal Components Analysis showed that the defunctionalization of the aromatic monomers was more pronounced in the most severe reaction conditions and that the abundance of more hydrophobic monomers increased in more diluted reactions. While the high-molecular-weight (heavy) bio-oil generated at 350 °C, 90 min, and 1:50 m.v⁻¹ failed to support bacterial growth, the corresponding light bio-oil rich in aromatic monomers promoted the growth of bacteria from 9 distinct species. The isolates *Pseudomonas* sp. LIM05 and *Burkholderia* sp. LIM09 showed the best growth performance and tolerance to lignin-derived aromatics, being the most promising for the future development of biological upgrading strategies tailored for this lignin stream.

Keywords: hydrothermal depolymerization; lignin-derived aromatics; aromatic-catabolizing bacteria.

Graphical abstract



1. Introduction

There is a growing recognition that lignin valorization is critical for the economic and environmental sustainability of biorefineries [1]. Processes to produce value-added chemicals from lignin have emerged as attractive approaches for lignin valorization and require three main steps: biomass fractionation, lignin depolymerization, and bio-oil upgrading [1,2]. Although a variety of methods have been developed individually for each step [2], few studies address them in combination [3,4], limiting our understanding of their compatibility and combined performance.

Sugarcane bagasse is a vast, inexpensive agroindustrial residue, and alkaline treatment allows the fractionation of bagasse into a lignin-rich stream and a cellulose-rich pulp highly digestible for bioconversion [5,6]. In this type of fractionation scheme, experimental studies in laboratory and pilot scales have demonstrated the separation of lignin with more than 43% of β -O-4' interunit linkages [5,6]. Such preservation of the labile β -O-4' linkages is a desirable feature for lignin depolymerization because milder depolymerization conditions are required. Therefore, it is worth examining how this type of alkaline lignin could be transformed into value-added chemicals through the combination of lignin depolymerization and upgrading.

Hydrothermal treatments for the depolymerization of isolated lignins are still poorly investigated, as they have been more commonly studied on whole lignocellulosic biomasses [7]. In a hydrothermal depolymerization process, wet feedstocks can be used without prior drying, and the reaction occurs in an environmentally friendly solvent [8]. Moreover, water at near supercritical conditions (<374 °C; 22 MPa) has special properties, such as increased dissociation constant, favoring the release of bifunctional aromatics [8]. Catalysts, homogeneous and heterogeneous, might be added to hydrothermal lignin depolymerization to enhance yield and selectivity. However, the overall benefits of catalysts are unclear considering the costs of catalyst, recovery process, and reactor corrosion, among other factors [9][10]. Therefore, uncatalyzed hydrothermal depolymerization remains an interesting approach, and the aqueous stream of depolymerized lignin may eventually be fed directly to biological upgrading [4]. Lignin depolymerization generates a wide spectrum of molecules, including the valuable monomeric aromatic compounds (MACs) as well as components of lower value such as high molecular weight lignin oligomers and low molecular weight alcohols, organic acids, and others. Transforming the diversity of MACs into chemical products with high yield requires the downstream upgrading of the complex mixtures [11,12], which can be performed by chemocatalytic or biocatalytic approaches [2].

Bio-based upgrading processes exploit the capacity some microorganisms have on funneling multiple lignin-derived compounds towards convergent metabolic intermediates that either have a high value *per se*, such as muconic acid, or can be further upgraded by chemical processes to other value-added products [13,14]. However, the diversity of lignin sub-structures available in nature has pressured the evolution of a diversity of pathways for their metabolism, which are not necessarily embedded in a single microbial species. Therefore, finding bacteria with metabolic capabilities as compatible as possible with the aromatic compounds present in a specific lignin stream is crucial to start the development of a microbial platform aiming at lignin aromatics upgrading via metabolic funneling.

In this study, we evaluated the influence of temperature, time, and solid-liquid (S/L) ratio on the yield and composition of light bio-oils (organic extracts from the aqueous phase) produced from the depolymerization of a β -O-4'-rich lignin, isolated by an alkaline process from sugarcane bagasse [5,6]. Gas Chromatography–Mass Spectrometry (GC-MS) studies were combined with a Principal Component Analysis (PCA) to provide a systematic evaluation on the effect of the hydrothermal reaction conditions on the chemical properties of the monomers released from this specific lignin. Moreover, we also developed a rapid and miniaturized microbial growth screening to evaluate whether the heavy or light bio-oil from the reaction condition leading to the highest monomer yield would support the growth of 11 aromatic-catabolizing bacteria from nine distinct genera. Besides presenting a systematic analysis of the influence of hydrothermal reaction conditions in the yield and composition of light bio-oils derived from a β -O-4'-rich lignin extracted from sugarcane bagasse, we also identified the bacteria *Pseudomonas* sp. LIM05 and *Burkholderia* sp. LIM09 as promising microorganisms for the future development of biological upgrading strategies tailored for this lignin stream.

2. Material and methods

2.1 Lignin depolymerization reactions

In our previous work [6], we produced four different lignins extracted from sugarcane bagasse using alkaline processes. From these lignins, we have chosen for the depolymerization reactions of this study the one with the highest amount (43%) of β -O-4' interunit linkages, produced by treating sugarcane bagasse with 1.5% *m.v*⁻¹ NaOH at 130 °C, 30 min, and a S/L ratio of 1:15 *m.v*⁻¹. This lignin-rich stream contained 81% of lignin, 14% of hemicelluloses, 3% of cellulose, and 2% of ashes. The depolymerization reactions of this alkaline lignin from sugarcane bagasse were performed according to a 2³ experimental design in triplicate at the

central point varying temperature (200 to 350 °C), time (30 to 90 min) and S/L ratio (1:10 to 1:50 $m.v^{-1}$).

The depolymerization reactions were performed in a high-pressure autoclave reactor, in batch mode (500 mL, model 4575A, Parr Instrument Company, up to 344 bar, 500 °C), with temperature, pressure, and agitation control (Parr 4848). All reactions occurred in an inert atmosphere (N₂). The air was initially purged three times with N₂ and then the reactor was pressurized to 30 bar N₂. The agitation was 500 rpm during the reactions. The reaction volume was set at 50% of the reactor volume. For the reactions at 1:50 $m.v^{-1}$, 245.1 mL of water and 4.9 g of alkaline lignin were added to the reactor. For 1:30 $m.v^{-1}$, 241.9 mL of water and 8.1 g of alkaline lignin were added to the reactor. For 1:10 $m.v^{-1}$: 227.3 mL of water and 22.7 g of alkaline lignin were added to the reactor. An experimental design was conducted to evaluate the effect of temperature, time, and S/L ratio on the mass yields of light bio-oil (*wt.%* of lignin in dry basis), resulting in 11 reactions named R1 to R11 (for details see Table 1 in the results section). After unloading the reactor, the aqueous fraction was filtered and sent to the liquid-liquid extraction step with ethyl acetate (Vetec PA ACS) (1:1, 3 times). The organic phase containing the aromatic monomers was dried with sodium sulfate (Synth PA ACS) and filtered. Afterward, the solvent was removed using a rotary evaporator. The yields of the resulting light bio-oils were calculated from equation (1):

$$\text{Light bio-oil yield (wt.\%)} = \left[\frac{\text{weight of light bio-oil}}{\text{weight of lignin}} \right] \times 100 \quad (1)$$

All light bio-oils obtained from the experimental planning described above were analyzed using a GCMS-QP2010 Ultra Shimadzu equipped with a CG2010Plus gas chromatograph and a selective mass detector Shimadzu QP-2010 (quadrupole). The column used was a RTX-5MS (30 m x 0.25 mm x 0.25 mm). The instrumental parameters used in the analysis of light bio-oils were the following: injection and detector temperature of 280 °C, split at 1:50, He_(gas) at 0.96 mL.min⁻¹, column oven temperature was 50 °C, and an oven temperature program was 50 °C (2 min), 120 °C (5 min), 280 °C (8 min) and 300 °C (2 min) with a heating rate of 10 °C.min⁻¹. Approximately 0.05 g of each light bio-oil sample was weighed and then diluted with ethyl acetate (HPLC grade) for injections into the GC-MS equipment. Compounds were identified by comparing their mass spectra to those in the NIST library. For the condition that generated the highest yield of light bio-oil (350 °C, 90 min, 1:50), a heavy bio-oil was additionally obtained by washing the solid residue of the hydrothermal reaction with acetone in excess (Synth PA ACS).

2.2 Analysis of the mean molecular properties of the compounds identified in the light bio-oils

The molecular properties of the compounds identified by GC-MS were obtained from the PubChem database [15]. The following computed molecular properties were used in the analysis: molecular weight (M_w), partition coefficient ($\log P$), hydrogen bond donor count (HB_D), hydrogen bond acceptor count (HB_A), and molecular complexity (C_{PLX}). For each reaction condition (R1-R11), weighted means of molecular properties were calculated using %area from GC-MS peaks as statistical weights. The resulting data matrix (rows: R1-R11; columns: M_w , $\log P$, HB_D , HB_A , and C_{PLX}) was submitted to principal component analysis (PCA) with columns normalized to unit variance. The scikit-learn Python package was employed for PCA [16].

2.3 Evaluation of bacterial growth using light and heavy bio-oils as the main carbon source

Both light and heavy bio-oils from the depolymerization reaction R11 (350 °C, 90 min, 1:50) were submitted to the biological conversion step to assess their capacity to support bacterial growth. Seven bacterial isolates and four model strains (Supplementary Tables S1 and S2) were cultured in Lysogeny Broth (LB) medium (10 g.L⁻¹ peptone, 5 g.L⁻¹ yeast extract, and 5 g.L⁻¹ NaCl), except *Xanthomonas citri* subsp. *citri* 306 (grew on LB without NaCl), overnight at 30 °C and 200 rpm. Cells were collected by centrifugation (6000 rpm, 5 min) and washed twice with XVM2m minimal medium (20 mmol.L⁻¹ NaCl, 10 mmol.L⁻¹ (NH₄)₂SO₄, 1 mmol.L⁻¹ CaCl₂, 0.01 mmol.L⁻¹ FeSO₄·7 H₂O, 5 mmol.L⁻¹ MgSO₄, 0.16 mmol.L⁻¹ KH₂PO₄, 0.32 mmol.L⁻¹ K₂HPO₄, 0.3 g.L⁻¹ casamino acids, pH 6.7) and inoculated for an initial OD_{600 nm} of 0.05 on XVM2m minimal media supplemented with 1 g.L⁻¹ (Test I) or 0.3 g.L⁻¹ (Test II) light bio-oil or 0.3 g.L⁻¹ heavy bio-oil. The experiment was performed in triplicate and the bacterial growth was monitored at 600 nm, for 45 h at 30 °C, in a SpectraMax® M3 Multi-mode microplate reader (Molecular Devices). The light bio-oil and heavy bio-oil used in the growth assays were solubilized on 2 mol.L⁻¹ NaOH solution, pH 12, at 40 g.L⁻¹, and then diluted to prepare the medium XVM2m with a final theoretical concentration of 1.0 g.L⁻¹ of the respective oils. After adjusting the pH to 6.7, the medium was filter-sterilized and used in the growth assays or diluted in XVM2m to produce a medium with 0.3 g.L⁻¹ of the respective oils. Specific growth rates (μ) were obtained using the Package Growth rates as described by (Petzoldt, 2020). The concentration of aromatic compounds in the medium was estimated by 280 nm absorbance, at pH 12, using the following equation 2 [17]:

$$Aromatics (g.L^{-1}) = (4.187 \times 10^{-2} * A_{280} - 3.279 \times 10^{-4}) * \text{dilution} \quad (2)$$

2.4 Evaluation of bacterial growth on molecular standards for aromatic monomers identified in light bio-oils

Pseudomonas sp. LIM05 and *Burkholderia* sp. LIM09 were cultured in LB medium, overnight at 30 °C and 200 rpm. After that, the harvested cells were washed twice with XVM2m minimal medium and inoculated for an initial OD_{600 nm} of 0.05 on XVM2m minimal media supplemented with individual 0.1, 0.5 and 1 g.L⁻¹ phenol (catalog number 185450, ≥ 99 % purity), guaiacol (catalog number G5502, ≥ 98% purity), syringol (catalog number D135550, ≥ 98.5 % purity), catechol (catalog number 135011, ≥ 99 % purity) and 3-methoxycatechol (catalog number M13203, ≥ 98.5 % purity) or a mix of them (each one representing 1/5 of the indicated concentration). The aromatic standards were purchased from Sigma-Aldrich. Growth was monitored at 600 nm for 45 h at 30 °C in an Infinite 200 PRO plate reader (Tecan) in triplicate.

2.5 Molecular identification of bacteria

Besides using model strains (Supplementary Table S2) we also employed bacteria that have been previously isolated from different niches (Supplementary Table S1) in LB medium, except for *Pseudomonas* sp. LIM05, which was isolated in liquid minimal medium Bushnell-Hass Broth [18] containing 0.5% (m.v⁻¹) of chopped sheets of newspaper as the sole carbon source. The strains were grown overnight in 10 mL of LB medium at 30 °C and 200 rpm. Subsequently, the cells were harvested by centrifugation at 4 °C for 5 min at 6000 × g, and DNA extraction from all the samples was performed using Quick-DNA Fungal/Bacterial Miniprep Kit (Zymo Research) following the manufacturer's instructions. The DNA was quantified using the Nanodrop 2000c Spectrophotometer (Thermo Fisher Scientific).

100 ng of purified DNA was used in a 50 µL Polymerase Chain Reaction (PCR). PCR reactions were performed using a final concentration of 2.5 U Platinum Pfx DNA Polymerase (Thermo Fischer Scientific), 1× PCR Buffer, 1 mmol.L⁻¹ MgCl₂, 10 mmol.L⁻¹ of each deoxynucleoside triphosphate (Thermo Fischer Scientific), and 30 pmol of each 16S bacterial primers V3-V5R (5' CCGTCAATTCMTTTRAGT 3') and V3-V5F (5' ACGGYTACCTTGTTACGACTT 3'). PCR conditions were 94 °C for 5 min, 30 cycles of 94 °C for 15 s, 55 °C for 30 seconds, and 68 °C for 60 s. A final extension cycle of 68 °C for 10 min was used. The amplified products were purified using the EasyPure Quick Gel Extraction kit (TransGen Biotech) and sequenced with both primers using BigDye terminator v3.1 cycle sequencing kit (Applied Biosystems).

For the bacterial isolate *Pseudomonas* sp. LIM05 the full-length 16S rRNA gene was amplified using the primers 20F (5′ GAGTTTGAT CCTGGCTCAG 3′) and 1500R (5′ GTTACCTTGTTACG ACTT 3′) and sequenced using 20F, 1500R, 520F (5′ CAGCAGCCG CGGTAATAC 3′), 520R (5′ GTATTACCGCGGCTGCTG 3′), 920F (5′ AAAC TCAAATGAATTGACGG 3′) and 920R (5′ CCGTCAATTCATTTGAGTTT 3′) [19]. All sequencing was performed in an ABI Prism 3500 xL Genetic Analyzer (Applied Biosystems) and the consensus sequences for each sample were assembled using Geneious v8.1.9 [20].

The comparison to the closest strain was performed using the EzBioCloud. All 16S rDNA sequences were deposited in NCBI under accession numbers ON550350 (*Pseudomonas* sp. LIM05); ON550351 (*Acinetobacter* sp. LIM08); ON550352 (*Burkholderia* sp. LIM09); ON550353 (*Kalamiella* sp. LIM10); ON550354 (*Bacillus* sp. LIM11); ON550355 (*Enterobacter* sp. LIM12) and ON550356 (*Pseudoxanthomonas* sp. LIM19).

3. Results and discussion

3.1 Hydrothermal reaction conditions affect the yield and molecular trends of light bio-oils

Evaluation of bio-oil yield (wt.%) is critical to analyze the potential of any lignin conversion technology. The observed yields of the light bio-oils (Table 1) obtained within the experimental design ranged from 9 to 31 wt.%. The triplicate depolymerization reactions at the central point (275 °C, 60 min, 1:30) showed the same numeric value (16 wt.%), indicating process repeatability. Although further increasing the severity and dilution of the process could have increased the bio-oil yields, the investigated conditions cover the most relevant range of subcritical (< 374 °C, 22 MPa) hydrothermal lignin depolymerization.

GC-MS analyses of the light bio-oils identified a set of aromatic monomers, including phenol, guaiacol, and syringol (Table 2). All chromatograms (Figures S1-S11) are reported in the supplementary material. The reactions at the central point (R5-R7) produced bio-oils with similar product compositions, also indicating good repeatability of the process and the compositional analysis. As an example of a molecular trend responding to variable hydrothermal conditions, the lower temperature (200 °C, R1-R4) favored the formation of vanillin, while the higher temperature (350 °C, R8-R11) favored the formation of 3-methoxycatechol.

Table 1. Experimental design 2^3 with central point of lignin hydrothermal depolymerization. The influence of temperature, time, and S/L ratio on light bio-oil yield (wt.%) is shown, highlighting in bold the reaction of maximum yield. *Triplicate reactions at the central point.

	T (°C)	t (min)	S/L ratio ($m.v^{-1}$)	Light bio-oil yield (wt.%)
R1	200	30	1:10	10
R2	200	90	1:10	10
R3	200	30	1:50	15
R4	200	90	1:50	18
R5*	275	60	1:30	16
R6*	275	60	1:30	16
R7*	275	60	1:30	16
R8	350	30	1:10	9
R9	350	90	1:10	10
R10	350	30	1:50	24
R11	350	90	1:50	31

Table 2. GC-MS analyses of light bio-oils from reaction conditions R1-R11. Numbers represent peak areas (expressed as % from total peak areas) of the monomeric aromatic compounds (MACs) identified by GC-MS. ^aCreosol co-eluted with catechol. HMF: hydroxymethylfurfural.

MACs	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11
Phenol	4	3	2	3	7	7	8	8	11	10	15
Guaiacol	7	6	4	5	14	14	13	11	7	17	16
Syringol	13	12	10	12	31	30	33	14	4	27	13
Vanillin	25	18	17	16	7	7	7	-	-	2	-
Syringaldehyde	9	9	7	7	3	4	3	-	-	-	-
Coumaran	16	4	28	15	1	-	-	-	-	-	-
4-vinylguaiacol	7	2	14	7	-	-	-	-	-	-	-
5-HMF	-	14	-	7	-	-	-	-	-	-	-
2-ethyl-phenol	-	1	-	1	4	4	5	4	4	6	8
Catechol/Creosol ^a	-	-	-	-	-	-	1	10	11	4	12
3-methoxy-catechol	-	-	-	-	-	-	-	23	12	7	11
Catechol	-	-	-	-	-	-	-	8	17	-	-
4-methyl-catechol	-	-	-	-	-	-	-	3	8	-	-

Vanillin is an aromatic aldehyde that was observed in the light bio-oils obtained from lignin depolymerization reactions at 200 °C with 16-25% of relative peak area in GC-MS analysis. In contrast, at 275 °C, vanillin relative peak area decreased to 7% and was not further observed in reactions performed at 350 °C, indicating higher thermal lability of the aldehyde

radical compared to methoxy or hydroxy groups appended at the aromatic ring. Similarly, Islam et al. (2018) observed that the yields of vanillin and syringaldehyde during the hydrothermal depolymerization of Kraft lignin decreased with an increase in temperature (200 to 300 °C), and those aromatic aldehyde compounds were not observed at 350 °C [21,22].

To have a broader vision of the effects of reaction conditions on the light bio-oil characteristics, the compositions originally expressed as GC-MS % areas (Table 2) were transformed into mean molecular properties of the detected compounds (Supplementary Table 3). The analysis of this matrix by PCA is shown in Figure 1. PC1 explained 66.3% of the variance, while PC2 did 30.5%, summing up 96.8% of the variance in this space of reduced dimensionality, allowing a comprehensive discussion of the main effects of the reaction conditions through two independent dimensions.

The relative positions of the reaction conditions (R1-R11) in the scores plot (Figure 1) indicate that PC1 is strongly associated with reaction severity since the reaction temperature increases along PC1. Furthermore, for the reactions at the highest temperature (350 °C), residence time also increases along PC1. Regarding PC2, it is strongly associated with reaction dilution. Comparing reaction pairs with dilution being the only difference (R1:R3, R2:R4, R8:R10, R9:R11), the more dilute reaction systems (lower S/L ratio) are found at higher PC2.

The loadings plot of Figure 1 reveals that mean molecular weight and complexity (M_w and C_{PLX}) are tightly associated with PC1 and, therefore, reaction severity. Considering the opposing directions of M_w and C_{PLX} (loadings plot) and severity (scores plot), it follows the mechanistic interpretation that higher reaction severity creates a trend for light bio-oils having compounds of lower M_w and C_{PLX} . For several fragments of lignin depolymerization, this trend means the defunctionalization of the aromatic monomers is more pronounced for the more severe reactions.

Regarding the dilution-PC2 axis, the loadings plot of Figure 1 shows an association with the octanol-water partition coefficient ($\log P$), which measures how hydrophobic a molecule is. A possible mechanistic explanation for this relationship can be rationalized based on the partition of compounds between the water and the heavy bio-oil phases during the water discharge from the depolymerization reactor. More hydrophilic compounds (lower $\log P$) would be almost fully in the aqueous phase, whichever the amount of water. On the other hand, more hydrophobic compounds would be partly in the heavy bio-oil phase, and the content of such compounds in water would increase if more water was present in the system. Thus, more dilute systems (lower S/L ratio) would have a higher share of compounds of higher hydrophobicity in the light bio-oil, in agreement with the PCA results of Figure 1.

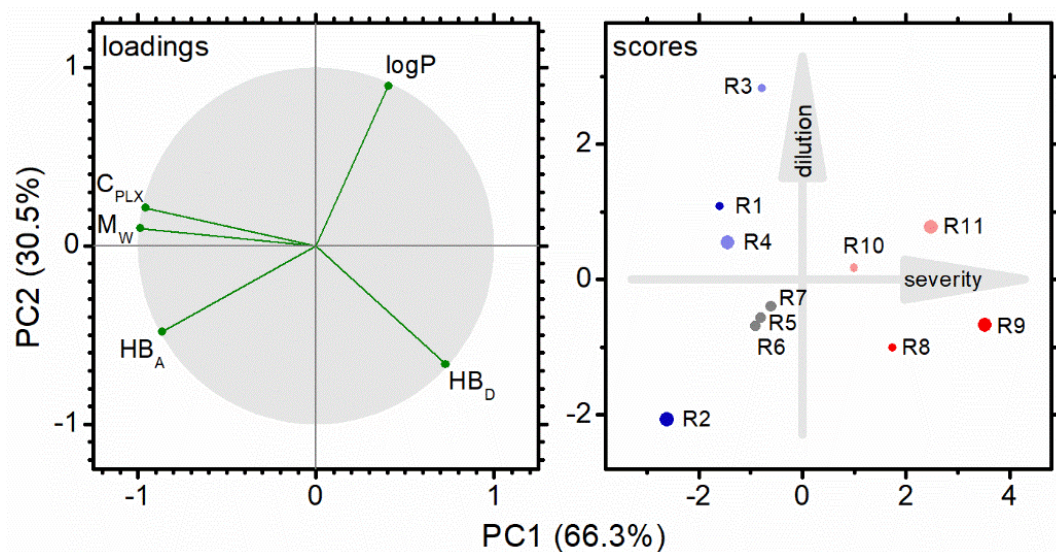


Figure 1. Principal Component Analysis (PCA) of the light bio-oil compositions. The loadings plot (left) shows the mean molecular properties: molecular weight (M_W), partition coefficient ($\log P$), hydrogen bond donor count (HB_D), hydrogen bond acceptor count (HB_A), and molecular complexity (C_{PLX}). The scores plot (right) shows the reaction conditions (R1-R11) of the experimental planning with variable reaction temperature (symbol color: blue < gray < red), residence time (symbol size: small < medium < large), and S/L ratio $m.v^{-1}$ (symbol tone: light < dark). The relative positions of the reaction conditions in the scores plot show PC1 strongly associated with severity (temperature and time) and PC2 associated with dilution (S/L ratio).

3.2 Bacterial growth screening reveals *Pseudomonas* sp. LIM05 and *Burkholderia* sp. LIM09 as the best bio-converters of R11 light bio-oil

From the previous analyses, we selected the reaction condition providing the highest light bio-oil yield to proceed with the microbial growth screening, seeking to identify bacteria capable of metabolizing the lignin-derived aromatics produced in this condition. In the screening, we also included samples of the corresponding heavy bio-oil, aiming to evaluate the capacity of these bacteria to use lignin oligomers as a carbon source for growth.

In this study, we used cell growth as an indicator of lignin oil bioconversion because it is an indirect measurement of how efficient a microorganism is at metabolically funneling aromatic compounds into the tricarboxylic acid cycle (TCA). From the TCA cycle, value-added chemicals such as polyhydroxyalkanoates (PHA) can be produced, for example [13]. Through the pathways leading to the TCA cycle, other value-added molecules can be targeted, such as muconic acid [14].

Eleven different bacterial species were tested for their capacity to grow on XVM2m minimal medium supplemented with either 1 g.L^{-1} light bio-oil or 0.3 g.L^{-1} heavy bio-oil from reaction R11 (Table 1). These bacteria were selected for their phylogenetic proximity to aromatic-metabolizing bacteria listed in the eLignin database [23] or found in the literature [24].

Four of them are model organisms and seven had been previously isolated from environmental samples and were classified at the genus level according to 16S rDNA phylogenetic analyses (Supplementary Table S1).

Under the condition supplemented with the heavy bio-oil, either no growth or minimal growth above the negative control (non-supplemented media) was observed for the tested bacteria, indicating their inefficiency to use heavy bio-oil as a carbon source (Supplementary Figure S12). Tests at higher heavy bio-oil concentrations were not possible due to its low solubility in water. The poor or no growth in the bio-oil enriched with heavier organic compounds indicates that either these bacteria do not have enzymatic machinery to efficiently break down these molecules into metabolizable monomers or that the tested conditions were insufficient to provide them with all nutritional requirements to decompose this material. The poor biodegradability of lignin oligomers has been observed previously for some aromatic-catabolizing bacteria, including *Pseudomonas putida* [25], in agreement with our results. The presence of β -O-4' bonds seem to be important for the enzymatic cleavage of lignin oligomers [25], so future studies aiming to elucidate the molecular structure of heavy bio-oil compounds derived from sugarcane lignin may provide insights about the extent to which their structures hinder their bioconversion.

In the light bio-oil condition, four bacterial isolates (*Pseudomonas* sp. LIM05, *Burkholderia* sp. LIM09, *Acinetobacter* sp. LIM08, and *Enterobacter* sp. LIM12) grew using 1 g.L⁻¹ light bio-oil as carbon source (Figure 2a-d). Nevertheless, no growth was observed for the other seven bacteria evaluated (Figure 2e-f, Table 3).

To verify whether the absence of growth for these bacteria was due to medium toxicity or due to their incapacity of metabolizing the available aromatic compounds, we performed a growth test decreasing the concentration of the light bio-oil R11 in the medium to 0.3 g.L⁻¹. Under this condition, the growth of five other bacteria (*Bacillus cereus* ATCC 14579, *Bacillus* sp. LIM11, *Kalamiella* sp. LIM10, *Agrobacterium tumefaciens* C58, and *Xanthomonas citri* subsp. *citri* 306) was observed (Supplementary Figure S13), although at lower specific growth rates (μ) compared to the best growers at 1 g.L⁻¹ light bio-oil (Table 3). Therefore, 9 out of the bacteria 11 tested were capable of using light bio-oil aromatics as a carbon source, but with distinct growth performance and toxicity tolerance.

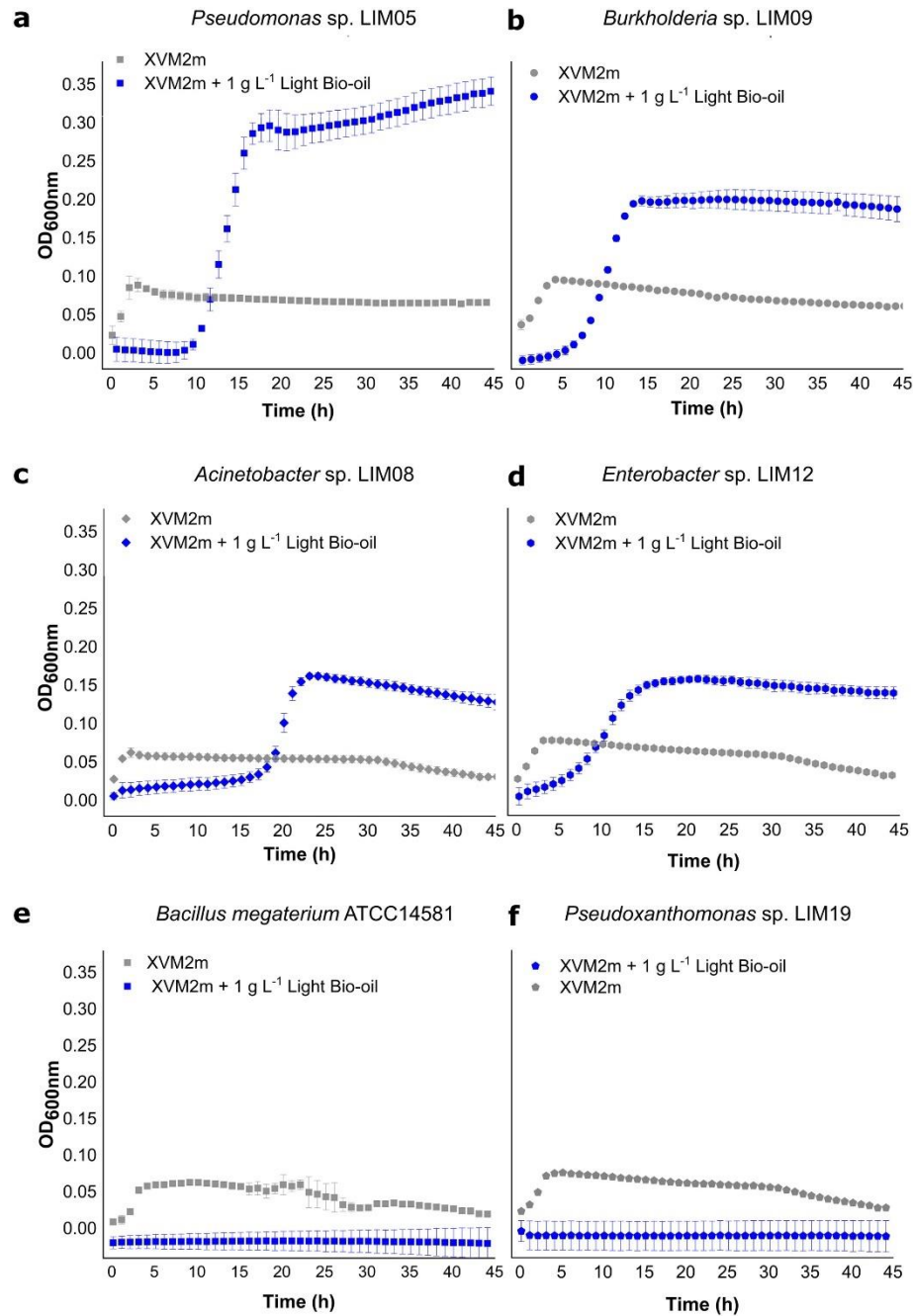


Figure 2. Growth profile of bacterial isolates in minimal medium supplemented with 1 g.L⁻¹ light bio-oil. In gray is the negative control curve (XVM2m medium without carbon source supplementation), in blue are the growth curves in XVM2m medium supplemented with 1 g.L⁻¹ light bio-oil. All curves are in the same x and y scales. Data are shown as mean ± SD of triplicates. *Bacillus megaterium* ATCC 14581 and *Pseudoxanthomonas* sp. are shown as representatives of “no growth” results observed for seven of the tested bacteria.

Table 3. Parameters of bacterial growth on XVM2m minimal medium supplemented with 1 g.L⁻¹ light bio-oil (Test I) and 0.3 g.L⁻¹ light bio-oil (Test II). The table shows the specific growth rates (μ) and maximum OD at 600 nm (OD_{600nm}Max) achieved for the culture of each microorganism. The symbol (–) represents no growth above the negative control (XVM2m without supplementation). OD_{600nm} Max is represented as the mean value with a standard deviation of three replicates (mean \pm SD). N.A.: not analyzed.

Microorganism	Test I (1 g.L ⁻¹)		Test II (0.3 g.L ⁻¹)	
	μ (h ⁻¹)	OD _{600nm} Max	μ (h ⁻¹)	OD _{600nm} Max
<i>Pseudomonas</i> sp. LIM05	1.08	0.30 \pm 0.02	N.A.	N.A.
<i>Burkholderia</i> sp. LIM09	0.65	0.20 \pm 0.01	N.A.	N.A.
<i>Acinetobacter</i> sp. LIM08	0.52	0.162 \pm 0.003	N.A.	N.A.
<i>Enterobacter</i> sp. LIM12	0.26	0.154 \pm 0.005	N.A.	N.A.
<i>Bacillus cereus</i> ATCC 14579	-	-	0.37	0.178 \pm 0.003
<i>Bacillus</i> sp. LIM11	-	-	0.19	0.20 \pm 0.01
<i>Kalamiella</i> sp. LIM10	-	-	0.18	0.152 \pm 0.008
<i>Agrobacterium tumefaciens</i> C58	-	-	0.08	0.174 \pm 0.007
<i>Xanthomonas citri</i> subsp. <i>citri</i> 306	-	-	0.05	0.136 \pm 0.003
<i>Bacillus megaterium</i> ATCC 14581	-	-	-	-
<i>Pseudoxanthomonas</i> sp. LIM19	-	-	-	-

According to the parameters μ and maximum OD_{600nm}, *Pseudomonas* sp. LIM05 and *Burkholderia* sp. LIM09 showed the best growth performance in light bio-oil R11 (Table 3). In addition, A₂₈₀ analysis indicated that the phenolic content in the medium supplemented with 1 g.L⁻¹ light bio-oil was reduced by approximately 43% and 35% in *Pseudomonas* sp. LIM05 and *Burkholderia* sp. LIM09 cultures after 45 h of growth, respectively. Therefore, these two isolates were identified as the most promising strains for the study and development of bioconversion strategies for light bio-oils akin to the bio-oil R11 produced in this study.

Bacteria from the genus *Pseudomonas* (Gammaproteobacteria), especially the strain *Pseudomonas putida* KT22440, have been one of the most studied and engineered chassis for the development of biological upgrading approaches aiming at lignin valorization [26]. Here, comparing representatives of several bacterial genera, we provide evidence that bacteria from the genus *Pseudomonas* are indeed a promising starting point for the development of biological upgrading strategies targeting hydrothermally depolymerized lignin, due to their natural capacity to metabolize the aromatic monomers generated in this process. On the other hand, our results also point to the genus *Burkholderia* (Betaproteobacteria) as a promising niche for the

study and development of lignin bioconversion strategies, since *Burkholderia* sp. LIM09 appeared as the second-best growth performer in our screening using light bio-oil. A previous review has highlighted the potential application of the genus *Burkholderia* in aromatics bioconversion [27] and here we provide the first evidence for the capacity of a *Burkholderia* specie to convert aromatic compounds derived from sugarcane bagasse lignin, adding to the few studies demonstrating this feature for other *Burkholderia* strains using other lignin sources or lignin model compounds [28,29].

To better investigate the capacity of *Pseudomonas* sp. LIM05 and *Burkholderia* sp. LIM09 to grow using, as a carbon source, selected aromatics identified in bio-oil R11, we assessed the growth of these isolates in a minimal medium supplemented with each individual compound or a mix of them in three different concentrations (Figure 3, Table 4). This assay is important to identify the individual contribution of each compound for bacterial growth or for toxicity, thus providing insights about the ideal composition of a bio-oil for a given bacterial strain and vice-versa.

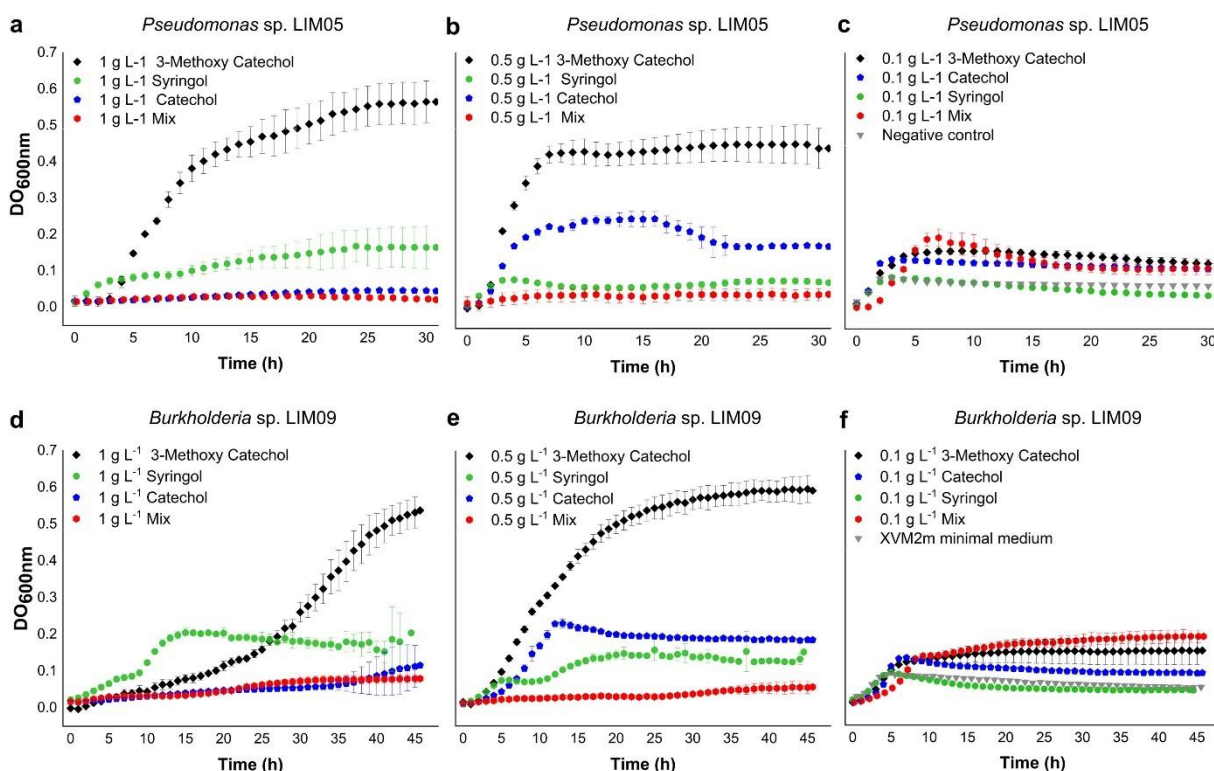


Figure 3. Growth curves of *Pseudomonas* sp. LIM05 and *Burkholderia* sp. LIM09 in minimal medium supplemented with different concentrations of representative aromatics identified in light bio-oil R11. *Pseudomonas* sp. LIM05 growth curves in XVM2m supplemented with aromatics at (a) 1 g.L⁻¹, (b) 0.5 g.L⁻¹, and (c) 0.1 g.L⁻¹. *Burkholderia* sp. LIM09 growth curves in XVM2m supplemented with aromatics at (d) 1 g.L⁻¹, (e) 0.5 g.L⁻¹, and (f) 0.1 g.L⁻¹. Black diamonds (3-methoxycatechol), Green circles (syringol), Blue pentagon (catechol), Red hexagon (mix of 3-methoxycatechol, syringol, catechol,

phenol, and guaiacol; each one representing 1/5 of the indicated concentration). For clarity purposes, the curves referring to phenol and guaiacol conditions were omitted since no growth was observed in the tested conditions. Data are presented as mean \pm SD of triplicates.

Table 4. Specific growth rates (μ) of *Pseudomonas* sp. LIM05 and *Burkholderia* sp. LIM09 on XVM2m supplemented with three concentrations of aromatics identified in light bio-oil. The bacteria were grown on XVM2m minimal medium supplemented with 3-methoxycatechol, catechol, syringol, guaiacol, phenol, and an equimass mix of them (1/5 each) in three different concentrations. The symbol (–) represents no growth above the negative control (XVM2m medium without supplementation).

Compounds	1 g.L ⁻¹	0.5 g.L ⁻¹	0.1 g.L ⁻¹
<i>Pseudomonas</i> sp. LIM05	μ (h ⁻¹)		
3-methoxycatechol	1.04	1.74	1.24
Catechol	-	0.95	1.20
Syringol	0.20	-	-
Guaiacol	-	-	-
Phenol	-	-	-
Mix	-	-	1.13
<i>Burkholderia</i> sp. LIM09			
3-methoxycatechol	1.10	0.77	0.57
Catechol	-	0.46	0.57
Syringol	0.36	0.54	-
Guaiacol	-	-	-
Phenol	-	-	-
Mix	-	-	0.36

In our analyses, the compound 3-methoxycatechol figured as the best growth supporter of both bacteria. In the literature, we found few studies showing evidence for enzymes that use 3-methoxycatechol as substrate [30–33], which indicates the existence of biotransformation reactions involving this compound. However, none of them show that 3-methoxycatechol can be converted into cell biomass. Therefore, we report, to the best of our knowledge, the first demonstration of 3-methoxycatechol as a growth supporter of aromatic-catabolizing bacteria. This finding implies the existence of a full pathway for 3-methoxycatechol uptake and funneling up to the central carbon metabolism (tricarboxylic acid cycle) in *Pseudomonas* sp. LIM05 and *Burkholderia* sp. LIM09, otherwise these bacteria would not be capable of producing cell biomass using this compound as a carbon source. Future studies will be required to elucidate the components of this pathway in both bacteria.

Our data also show that catechol was toxic at 1 g.L⁻¹ to *Pseudomonas* sp. LIM05 and *Burkholderia* sp. LIM09 but enabled their growth above the negative control when provided at 0.5 g.L⁻¹. Catechol toxicity has been reported for another *Pseudomonas* strain and microbial

communities [34,35], even though it is an intermediate in the catabolism of several aromatic compounds [36]. Syringol supported bacterial growth when provided at 1 g.L⁻¹ for *Pseudomonas* sp. LIM05 and at 1 g.L⁻¹ or 0.5 g.L⁻¹ for *Burkholderia* sp. LIM09, but only slightly higher than the unsupplemented XVM2m media. To date, no natural pathway has been reported for the syringol metabolism, but, through the structure-guided engineering of an aryl-*O*-demethylase, a recent study showed a rational way to enable its bioconversion [23]. For guaiacol and phenol, no growth was observed and the mix of all five compounds proved to be toxic at concentrations ≥ 0.5 g.L⁻¹, supporting modest growth at 0.1 g.L⁻¹. The metabolism of guaiacol and phenol has been reported for other bacterial strains [37]; therefore, the inclusion of heterologous pathways in *Pseudomonas* sp. LIM05 or *Burkholderia* sp. LIM09 might enable them to metabolize these compounds, further improving their capacity to convert bio-oils enriched with such molecules.

Comparing both isolates, *Pseudomonas* sp. LIM05 presented higher growth rates than *Burkholderia* sp. LIM09 on 3-methoxycatechol or catechol at 0.5 g.L⁻¹ or 0.1 g.L⁻¹ whereas *Burkholderia* sp. LIM09 had a better growth performance than *Pseudomonas* sp. LIM05 on syringol. Together, these results point *Pseudomonas* sp. LIM05 as an attractive chassis for light bio-oil bioconversion and *Burkholderia* sp. LIM09 as a potential source of enzymes to be further developed for an efficient syringol metabolism.

4. Conclusions

The highest light bio-oil yield (31 wt%) from the hydrothermal depolymerization of a β -O-4'-rich lignin isolated from sugarcane bagasse was obtained at 350 °C, 90 min and 1:50 S/L ratio, producing a complex mixture of aromatics including guaiacol, syringol, phenol, creosol, catechol, 3-methoxycatechol, and 2-ethyl-phenol. By exploring the role of hydrothermal depolymerization conditions in the light bio-oil composition, Principal Component Analysis revealed two independent dimensions of variation, corresponding to reaction severity (temperature and time) correlated to molecular defunctionalization and reaction dilution (S/L ratio) correlated to bio-oil hydrophobicity. Microbial growth assays indicated *Pseudomonas* and *Burkholderia* species as promising candidates for the development of upgrading strategies targeting a hydrothermally depolymerized lignin isolated from sugarcane bagasse with high preservation of β -O-4' linkages, revealing their ability to efficiently use 3-methoxycatechol as a carbon source for growth. Moreover, we also developed a rapid and miniaturized screen to identify and categorize bacteria according to their growth performance on minimum media supplemented with lignin bio-oil, using the efficient

production of cell biomass as an indicator of the efficient funneling of aromatic compounds towards the central carbon metabolism.

Acknowledgments

We thank Dr. Douglas Galante from Brazilian Synchrotron Light Laboratory (LNLS-CNPEN), Dr. Fabio Rodrigues from University of São Paulo (USP) and Dr. Benito Gomez-Silva from University of Antofagasta (Chile) for the collection of samples from the Atacama Desert that culminated in the isolation of *Pseudomonas* sp. LIM05. We also thank the Next Generation Sequencing facility and the Analytical Center from LNBR/CNPEN for all support on DNA sequencing and GC–MS analyses, respectively. This work was supported by the Sao ~ Paulo Research Foundation - FAPESP (grants 2019/22213-2 to GJMR, 2019/06921-7 to POG, 2019/08590-8 to DBM and 2022/00474-1 to JVCO) and the Brazilian Scientific and Technological Development Council - CNPq (300565/2022-4 to FFM).

Declaration of Competing Interests: The authors declare that they have no known competing financial interests or personal relationships that could influence the work reported in this paper.

CRedit authorship contribution statement: Wrote the paper: FFM, DBM, POG; Principal Component Analysis: LYL, CD; Lignin depolymerization reactions: FFM, GJMR; Designed the bacterial growth curve assays: DBM, POG; Performed the bacterial growth curve assays: DBM; Analyzed bacterial growth data: DBM, POG; Designed 16S DNA sequencing assays: JVCO, ATNM; Prepared samples for 16S DNA sequencing: ATNM; Analyzed 16S sequencing data: JVCO; Isolated *Pseudomonas* sp. LiM05: EC. All authors read, revised, and approved the final manuscript.

Data availability

No data was used for the research described in the article.

References

- [1] R. Rinaldi, R. Jastrzebski, M.T. Clough, J. Ralph, M. Kennema, P.C.A. Bruijninx, B.M. Weckhuysen, Paving the Way for Lignin Valorisation: Recent Advances in Bioengineering, Biorefining and Catalysis, *Angewandte Chemie International Edition*. 55 (2016) 8164–8215. <https://doi.org/10.1002/anie.201510351>.
- [2] W. Schutyser, T. Renders, S. van den Bosch, S.F. Koelewijn, G.T. Beckham, B.F. Sels, Chemicals from lignin: An interplay of lignocellulose fractionation, depolymerisation, and upgrading, *Chem Soc Rev*. 47 (2018) 852–908. <https://doi.org/10.1039/c7cs00566k>.
- [3] Z.-H. Liu, N. Hao, Y.-Y. Wang, C. Dou, F. Lin, R. Shen, R. Bura, D.B. Hodge, B.E. Dale, A.J. Ragauskas, B. Yang, J.S. Yuan, Transforming biorefinery designs with ‘Plug-In Processes of Lignin’ to enable economic waste valorization, *Nat Commun*. 12 (2021) 3912. <https://doi.org/10.1038/s41467-021-23920-4>.
- [4] A. Rodriguez, D. Salvachúa, R. Katahira, B.A. Black, N.S. Cleveland, M. Reed, H. Smith, E.E.K. Baidoo, J.D. Keasling, B.A. Simmons, G.T. Beckham, J.M. Gladden, Base-Catalyzed Depolymerization of Solid Lignin-Rich Streams Enables Microbial Conversion, *ACS Sustain Chem Eng*. 5 (2017) 8171–8180. <https://doi.org/10.1021/acssuschemeng.7b01818>.
- [5] V.M. Nascimento, S.C. Nakanishi, G.J.M. Rocha, S.C. Rabelo, M.T.B. Pimenta, C.E. v. Rossell, Effect of Anthraquinone on Alkaline Pretreatment and Enzymatic Kinetics of Sugarcane Bagasse Saccharification: Laboratory and Pilot Scale Approach, *ACS Sustain Chem Eng*. 4 (2016) 3609–3617. <https://doi.org/10.1021/acssuschemeng.5b01433>.
- [6] F.F. de Menezes, J. Rencoret, S.C. Nakanishi, V.M. Nascimento, V.F.N. Silva, A. Gutierrez, J.C. del Rio, G.J. Moraes Rocha, Alkaline pretreatment severity leads to different lignin applications in sugarcane biorefineries, *ACS Sustain Chem Eng*. (2017) [acssuschemeng.7b00265](https://doi.org/10.1021/acssuschemeng.7b00265). <https://doi.org/10.1021/acssuschemeng.7b00265>.
- [7] G.T. Beckham, Lignin Valorization, Royal Society of Chemistry, Cambridge, 2018. <https://doi.org/10.1039/9781788010351>.
- [8] J. Schuler, U. Hornung, A. Kruse, N. Dahmen, J. Sauer, Hydrothermal Liquefaction of Lignin, *J Biomater Nanobiotechnol*. 08 (2017) 96–108. <https://doi.org/10.4236/jbnb.2017.81007>.
- [9] A.K. Deepa, P.L. Dhepe, Solid acid catalyzed depolymerization of lignin into value added aromatic monomers, *RSC Adv*. 4 (2014) 12625–12629. <https://doi.org/10.1039/c3ra47818a>.
- [10] Z. Sun, B. Fridrich, A. de Santi, S. Elangovan, K. Barta, Bright Side of Lignin Depolymerization: Toward New Platform Chemicals, *Chem Rev*. 118 (2018) 614–678. <https://doi.org/10.1021/acs.chemrev.7b00588>.
- [11] S.S. Wong, R. Shu, J. Zhang, H. Liu, N. Yan, Downstream processing of lignin derived feedstock into end products, *Chem Soc Rev*. 49 (2020) 5510–5560. <https://doi.org/10.1039/D0CS00134A>.
- [12] I. Vural Gursel, J.W. Dijkstra, W.J.J. Huijgen, A. Ramirez, Techno-economic comparative assessment of novel lignin depolymerization routes to bio-based aromatics,

Biofuels, Bioproducts and Biorefining. 13 (2019) 1068–1084. <https://doi.org/10.1002/bbb.1999>.

[13] J.G. Linger, D.R. Vardon, M.T. Guarnieri, E.M. Karp, G.B. Hunsinger, M.A. Franden, C.W. Johnson, G. Chupka, T.J. Strathmann, P.T. Pienkos, G.T. Beckham, Lignin valorization through integrated biological funneling and chemical catalysis, *Proceedings of the National Academy of Sciences*. 111 (2014) 12013–12018. <https://doi.org/10.1073/pnas.1410657111>.

[14] M. Kohlstedt, S. Starck, N. Barton, J. Stolzenberger, M. Selzer, K. Mehlmann, R. Schneider, D. Pleissner, J. Rinkel, J.S. Dickschat, J. Venus, J. B.J.H. van Duuren, C. Wittmann, From lignin to nylon: Cascaded chemical and biochemical conversion using metabolically engineered *Pseudomonas putida*, *Metab Eng.* 47 (2018) 279–293. <https://doi.org/10.1016/j.ymben.2018.03.003>.

[15] S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B.A. Shoemaker, P.A. Thiessen, B. Yu, L. Zaslavsky, J. Zhang, E.E. Bolton, PubChem in 2021: new data content and improved web interfaces, *Nucleic Acids Res.* 49 (2021) D1388–D1395. <https://doi.org/10.1093/nar/gkaa971>.

[16] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, É. Duchesnay, Scikit-learn: Machine Learning in Python, *Journal of Machine Learning Research*. 12 (2011) 2825–2830. <https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.PCA.html> (accessed July 26, 2022).

[17] G.J.D.M. Rocha, V.M. Nascimento, A.R. Gonçalves, V.F.N. Silva, C. Martín, Influence of mixed sugarcane bagasse samples evaluated by elemental and physical–chemical composition, *Ind Crops Prod.* 64 (2015) 52–58. <https://doi.org/10.1016/j.indcrop.2014.11.003>.

[18] L.D. Bushnell, H.F. Haas, The Utilization of Certain Hydrocarbons by Microorganisms, *J Bacteriol.* 41 (1941) 653–673. <https://doi.org/10.1128/jb.41.5.653-673.1941>.

[19] P. Yukphan, W. Potacharoen, S. Tanasupawat, M. Tanticharoen, Y. Yamada, *Asaia krungthepensis* sp. nov., an acetic acid bacterium in the α -Proteobacteria, *Int J Syst Evol Microbiol.* 54 (2004) 313–316. <https://doi.org/10.1099/ijs.0.02734-0>.

[20] M. Kearse, R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Meintjes, A. Drummond, Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data, *Bioinformatics*. 28 (2012) 1647–1649. <https://doi.org/10.1093/bioinformatics/bts199>.

[21] M.N. Islam, G. Taki, M. Rana, J.-H. Park, Yield of Phenolic Monomers from Lignin Hydrothermolysis in Subcritical Water System, *Ind Eng Chem Res.* 57 (2018) 4779–4784. <https://doi.org/10.1021/acs.iecr.7b05062>.

[22] D. Alam, M.Y. Lui, A.K.L. Yuen, Z. Li, X. Liang, T. Maschmeyer, B.S. Haynes, A. Montoya, Substituted Aromatic Aldehyde Decomposition under Hydrothermal Conditions, *Energy and Fuels*. (2022). <https://doi.org/10.1021/acs.energyfuels.2c00361>.

- [23] D.P. Brink, K. Ravi, G. Lidén, M.F. Gorwa-Grauslund, Mapping the diversity of microbial lignin catabolism: experiences from the eLignin database, *Appl Microbiol Biotechnol.* 103 (2019) 3979–4002. <https://doi.org/10.1007/s00253-019-09692-4>.
- [24] H.W. Kern, Bacterial degradation of dehydropolymers of coniferyl alcohol, *Arch Microbiol.* 138 (1984) 18–25. <https://doi.org/10.1007/BF00425401>.
- [25] D. Salvachúa, A.Z. Werner, I. Pardo, M. Michalska, B.A. Black, B.S. Donohoe, S.J. Haugen, R. Katahira, S. Notonier, K.J. Ramirez, A. Amore, S.O. Purvine, E.M. Zink, P.E. Abraham, R.J. Giannone, S. Poudel, P.D. Laible, R.L. Hettich, G.T. Beckham, Outer membrane vesicles catabolize lignin-derived aromatic compounds in *Pseudomonas putida* KT2440, *Proceedings of the National Academy of Sciences.* 117 (2020) 9302–9310. <https://doi.org/10.1073/pnas.1921073117>.
- [26] M. Kumar, S. You, J. Beiyuan, G. Luo, J. Gupta, S. Kumar, L. Singh, S. Zhang, D.C.W. Tsang, Lignin valorization by bacterial genus *Pseudomonas*: State-of-the-art review and prospects, *Bioresour Technol.* 320 (2021) 124412. <https://doi.org/10.1016/j.biortech.2020.124412>.
- [27] R. Morya, D. Salvachúa, I.S. Thakur, Burkholderia: An Untapped but Promising Bacterial Genus for the Conversion of Aromatic Compounds, *Trends Biotechnol.* 38 (2020) 963–975. <https://doi.org/10.1016/j.tibtech.2020.02.008>.
- [28] H. Akita, Z. Kimura, M.Z. Mohd Yusoff, N. Nakashima, T. Hoshino, Isolation and characterization of Burkholderia sp. strain CCA53 exhibiting ligninolytic potential, *Springerplus.* 5 (2016) 596. <https://doi.org/10.1186/s40064-016-2237-y>.
- [29] R. Morya, M. Kumar, S.S. Singh, I.S. Thakur, Genomic analysis of Burkholderia sp. ISTR5 for biofunneling of lignin-derived compounds, *Biotechnol Biofuels.* 12 (2019) 277. <https://doi.org/10.1186/s13068-019-1606-5>.
- [30] M. Fujiwara, L.A. Golovleva, Y. Saeki, M. Nozaki, O. Hayaishi, Extradiol cleavage of 3 substituted catechols by an intradiol dioxygenase, pyrocatechase, from a pseudomonad, *Journal of Biological Chemistry.* 250 (1975) 4848–4855. [https://doi.org/10.1016/s0021-9258\(19\)41246-5](https://doi.org/10.1016/s0021-9258(19)41246-5).
- [31] N. Thorenoor, Y.-H. Kim, C. Lee, M.-H. Yu, K.-H. Engesser, A previously uncultured, paper mill Propionibacterium is able to degrade O-aryl alkyl ethers and various aromatic hydrocarbons, *Chemosphere.* 75 (2009) 1287–1293. <https://doi.org/10.1016/j.chemosphere.2009.03.032>.
- [32] B. Venkatesagowda, Enzymatic Kraft lignin demethylation and fungal O-demethylases like vanillate-O-demethylase and syringate O-demethylase catalyzed catechol-Fe³⁺ complexation method, *J Microbiol Methods.* 152 (2018) 126–134. <https://doi.org/10.1016/j.mimet.2018.07.021>.
- [33] S. Zhang, X. Wu, Y. Xiao, Conversion of lignin-derived 3-methoxycatechol to the natural product purpurogallin using bacterial P450 GcoAB and laccase CueO, *Appl Microbiol Biotechnol.* 106 (2022) 593–603. <https://doi.org/10.1007/s00253-021-11738-5>.
- [34] X. Chen, X. Hu, Q. Lu, Y. Yang, S. Linghu, X. Zhang, Study on the differences in sludge toxicity and microbial community structure caused by catechol, resorcinol and

hydroquinone with metagenomic analysis, *J Environ Manage.* 302 (2022) 114027. <https://doi.org/10.1016/j.jenvman.2021.114027>.

[35] K.W. George, A. Hay, Less is more: reduced catechol production permits *Pseudomonas putida* F1 to grow on styrene, *Microbiology (N Y)*. 158 (2012) 2781–2788. <https://doi.org/10.1099/mic.0.058230-0>.

[36] C.S. Harwood, R.E. Parales, The β -ketoadipate pathway and the biology of self-identity, *Annu Rev Microbiol.* 50 (1996) 553–590. <https://doi.org/10.1146/annurev.micro.50.1.553>.

[37] M.M. Machovina, S.J.B. Mallinson, B.C. Knott, A.W. Meyers, M. Garcia-Borràs, L. Bu, J.E. Gado, A. Oliver, G.P. Schmidt, D.J. Hitchen, M.F. Crowley, C.W. Johnson, E.L. Neidle, C.M. Payne, K.N. Houk, G.T. Beckham, J.E. McGeehan, J.L. DuBois, Enabling microbial syringol conversion through structure-guided protein engineering, *Proc Natl Acad Sci U S A.* 116 (2019) 13970–13976. <https://doi.org/10.1073/pnas.1820001116>.

Supplementary data

Exploring the compatibility between hydrothermal depolymerization of alkaline lignin from sugarcane bagasse and metabolization of the aromatics by bacteria

Fabírcia Farias de Menezes^{a,1}, Damaris Batistão Martim^{a,b,1}, Liu Yi Ling^a, Aline Tieppo Nogueira Mulato^a, Elaine Crespim^c, Juliana Velasco de Castro Oliveira^a, Carlos Eduardo Driemeier^a, Priscila Oliveira de Giuseppe^{a,*}, George Jackson de Moraes Rocha^{a,*}

¹ Both authors contributed equally to this work and should be considered co-first authors.

^a Brazilian Biorenewable National Laboratory (LNBR), Brazilian Center for Research in Energy and Materials (CNPem), 13083-100, Campinas, SP, Brazil.

^b Graduate Program in Genetics and Molecular Biology, Biology Institute, State University of Campinas, 13083-970, Campinas, SP, Brazil.

^c Laboratory of Regulatory Systems Biology, Center for Nuclear Energy in Agriculture at the University of São Paulo (CENA/USP), 13416-000, Piracicaba, SP, Brazil.

*Corresponding authors. (P.O. de Giuseppe) priscila.giuseppe@lnbr.cnpem.br and (G.J. de Moraes Rocha) george.rocha@lnbr.cnpem.br

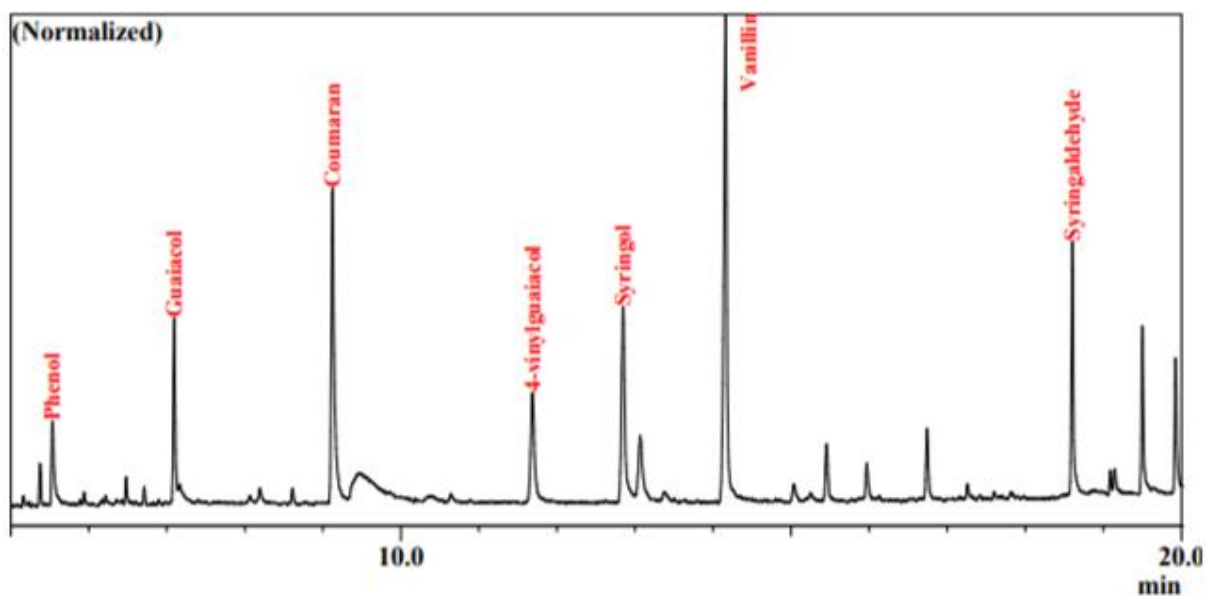


Figure S1. GC-MS chromatogram for the light bio-oil obtained at 200 °C, 30 min and 1:10 S/L (R1).

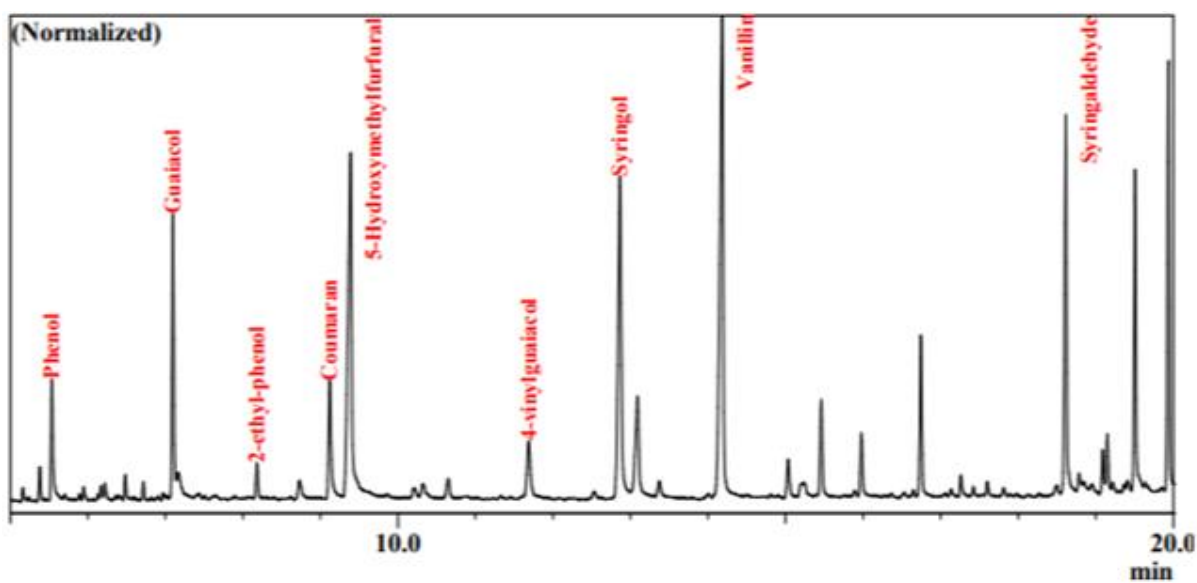


Figure S2. GC-MS chromatogram for the light bio-oil obtained at 200 °C, 90 min and 1:10 S/L (R2).

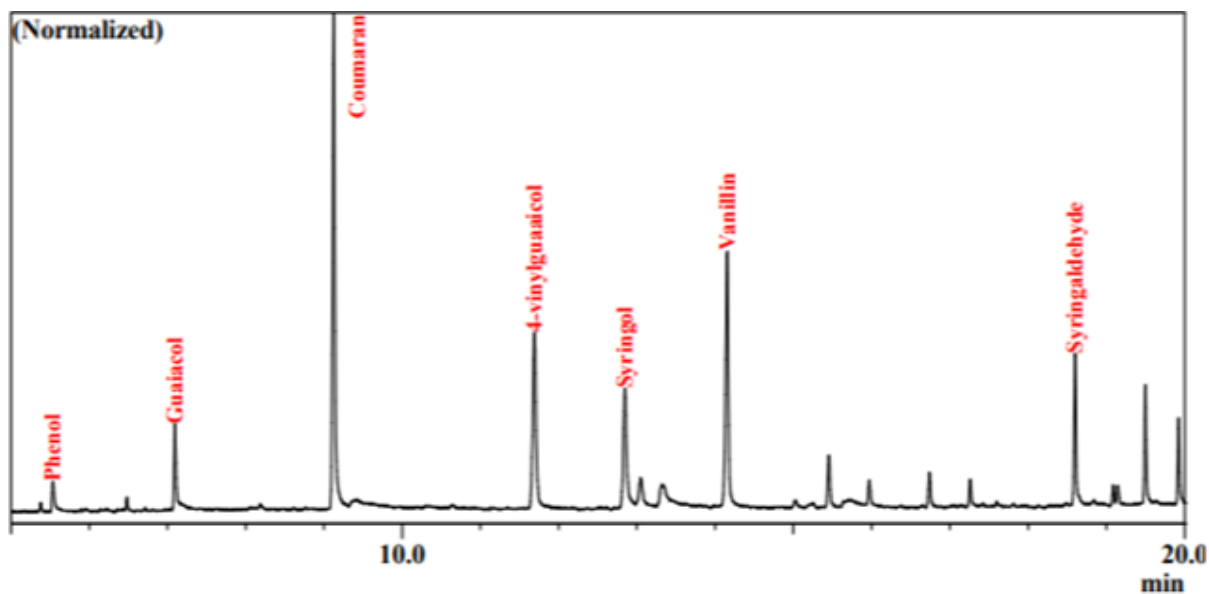


Figure S3. GC-MS chromatogram for the light bio-oil obtained at 200 °C, 30 min and 1:50 S/L (R3).

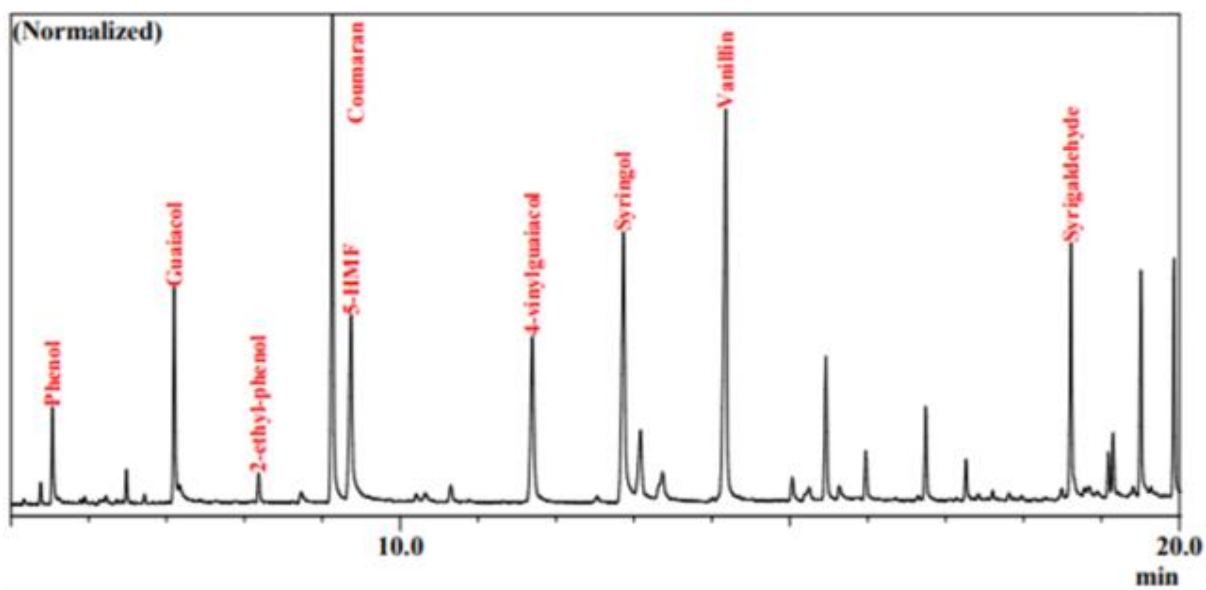


Figure S4. GC-MS chromatogram for the light bio-oil obtained at 200 °C, 90 min and 1:50 S/L (R4).

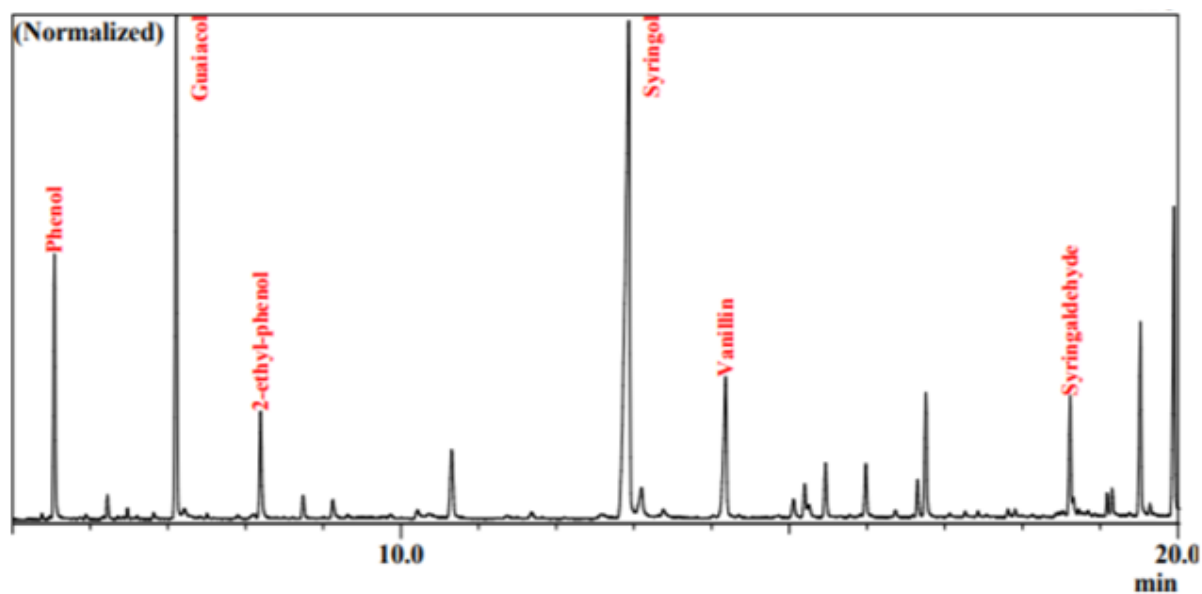


Figure S5. GC-MS chromatogram for the light bio-oil obtained at 275 °C, 60 min and 1:30 S/L (R5).

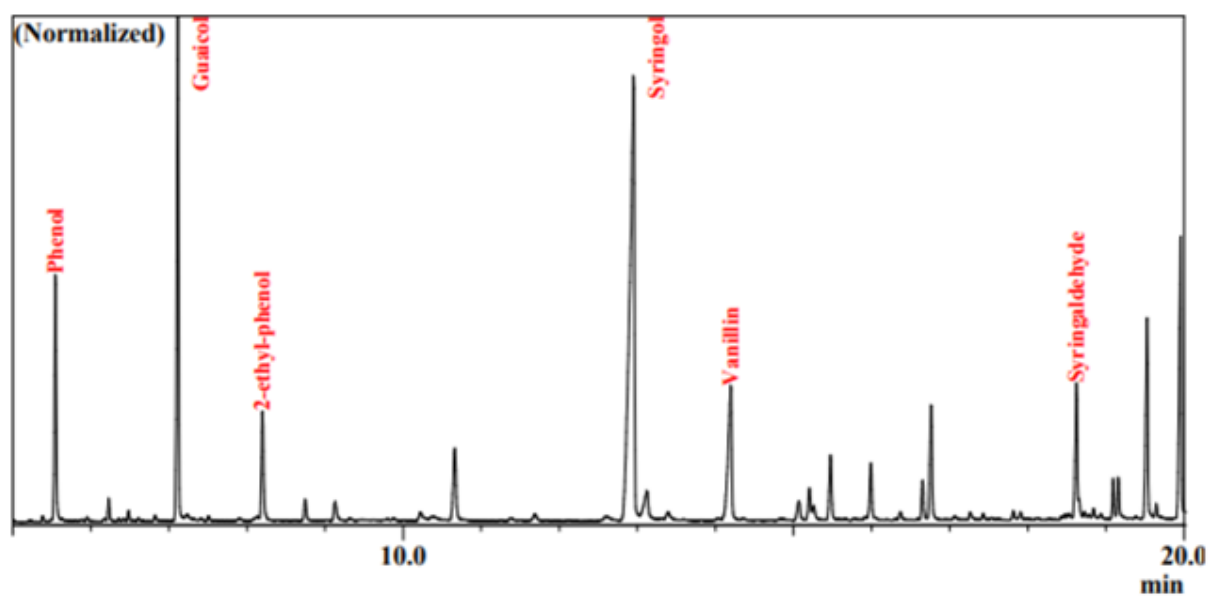


Figure S6. GC-MS chromatogram for the light bio-oil obtained at 275 °C, 60 min and 1:30 S/L (R6).

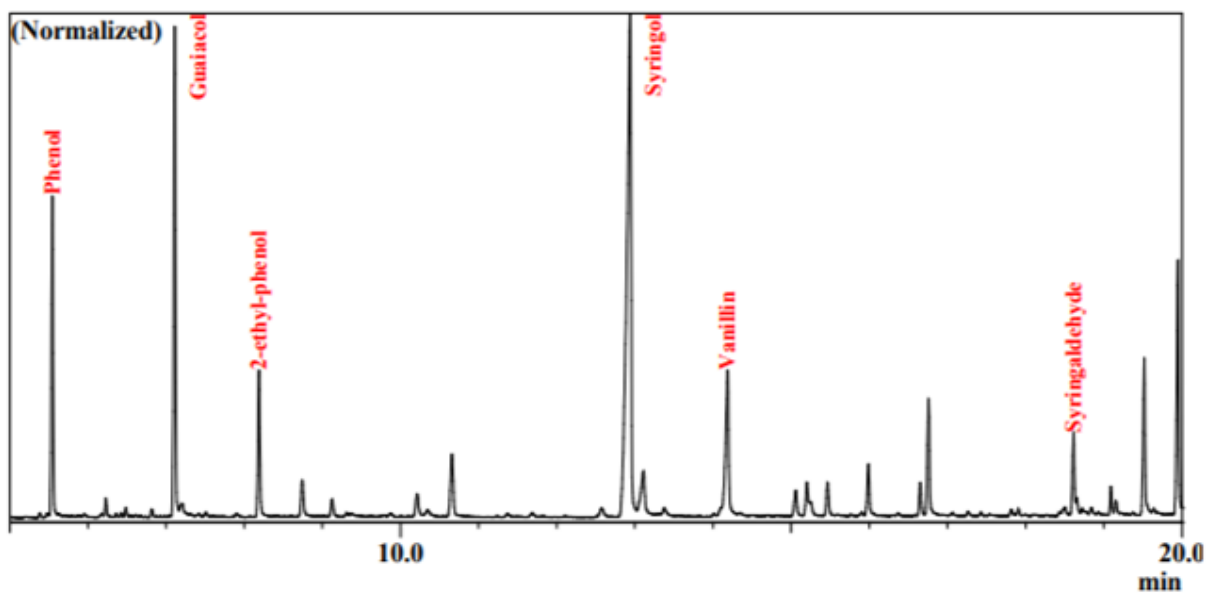


Figure S7. GC-MS chromatogram for the light bio-oil obtained at 275 °C, 60 min and 1:30 S/L (R7).

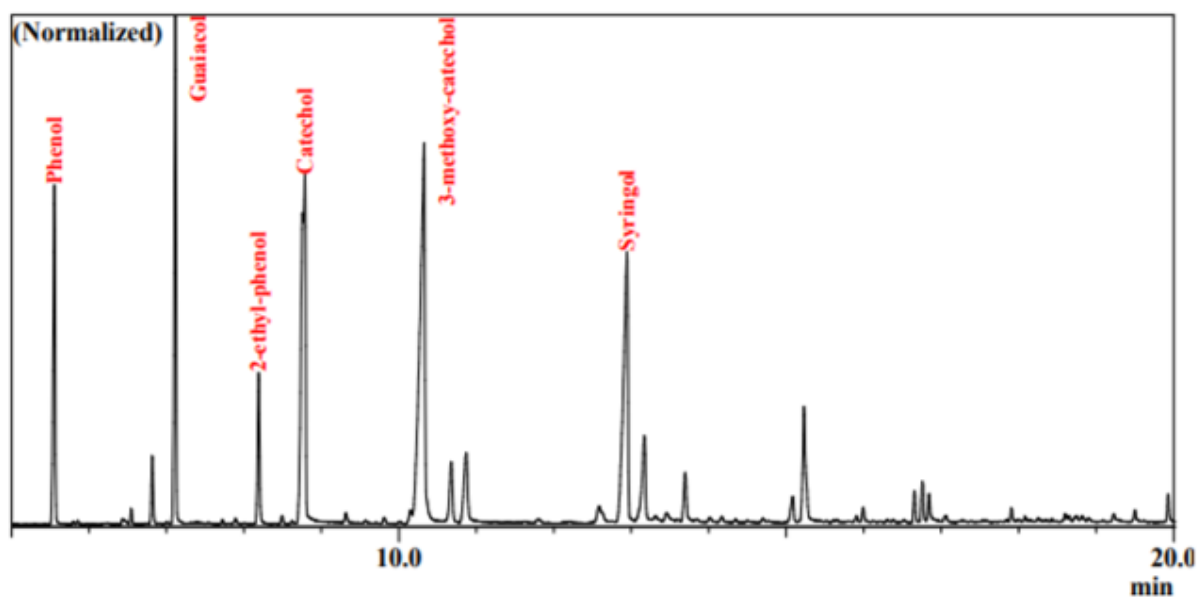


Figure S8. GC-MS chromatogram for the light bio-oil obtained at 350 °C, 30 min and 1:10 S/L (R8).

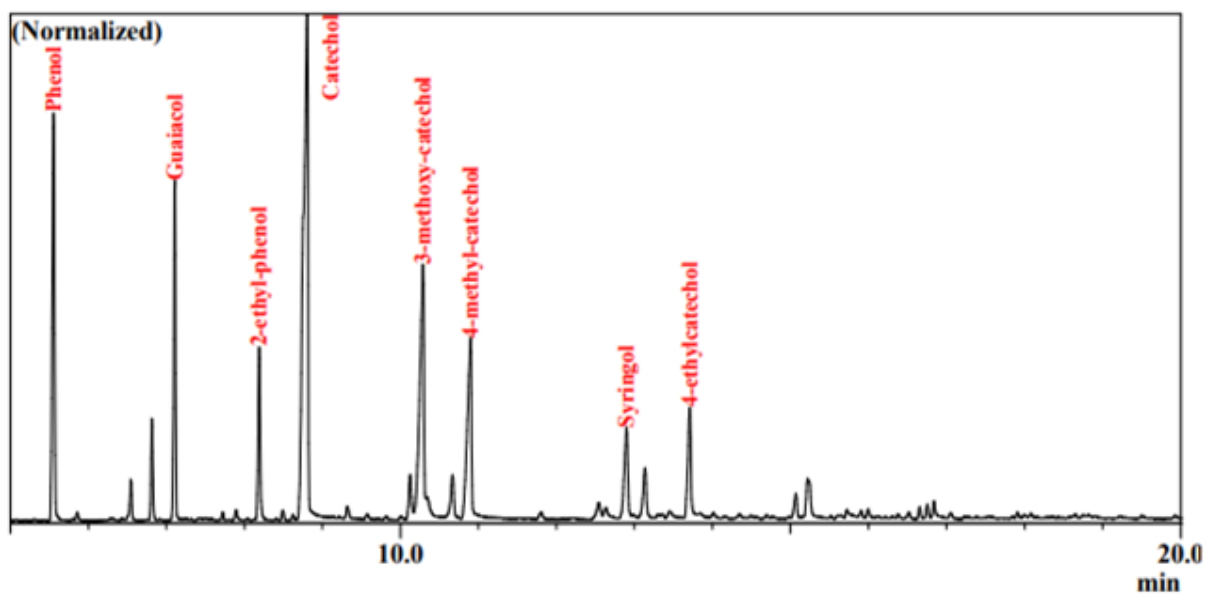


Figure S9. GC-MS chromatogram for the light bio-oil obtained at 350 °C, 90 min and 1:10 S/L (R9).

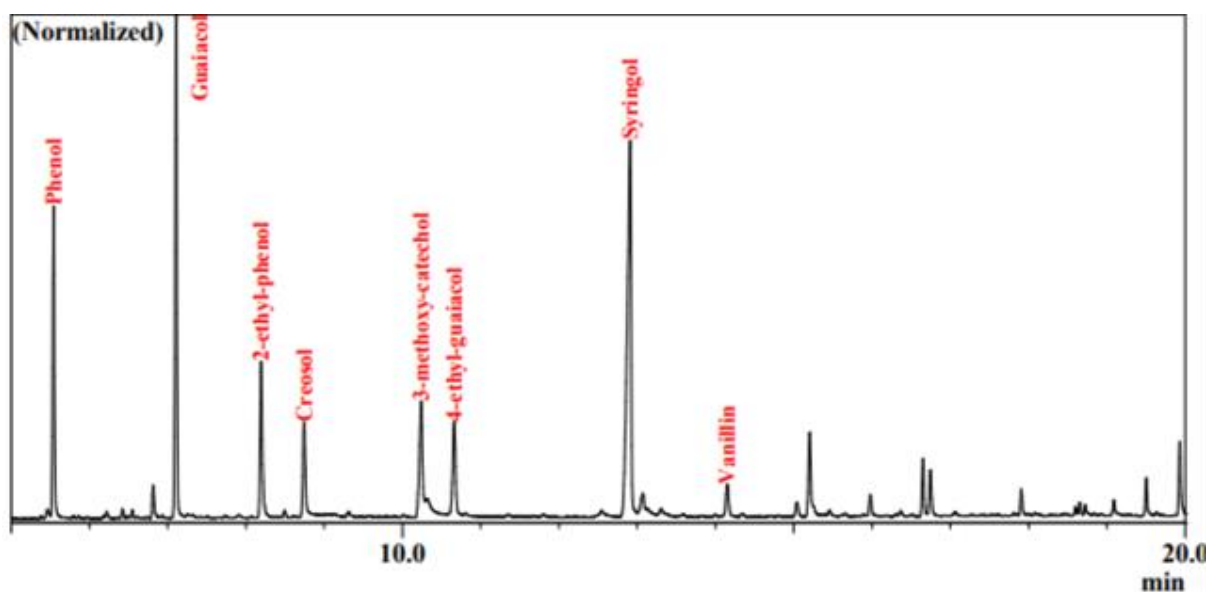


Figure S10. GC-MS chromatogram for the light bio-oil obtained at 350 °C, 30 min and 1:50 s/L (R10).

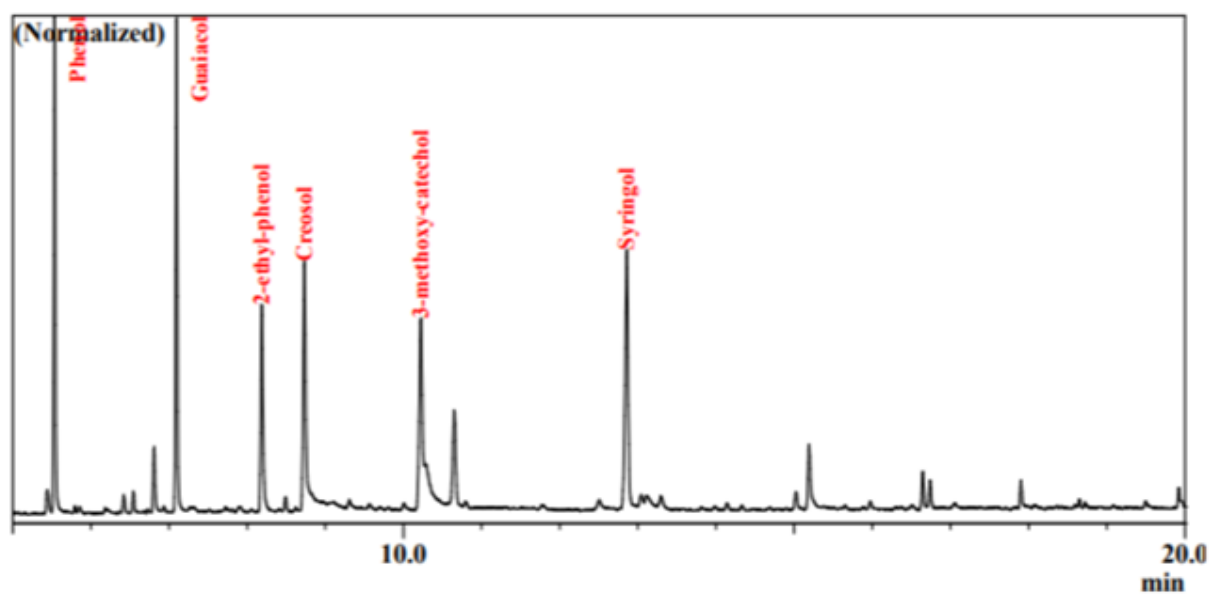


Figure S11. GC-MS chromatogram for the light bio-oil obtained at 350 °C, 90 min and 1:50 s/L (R11).

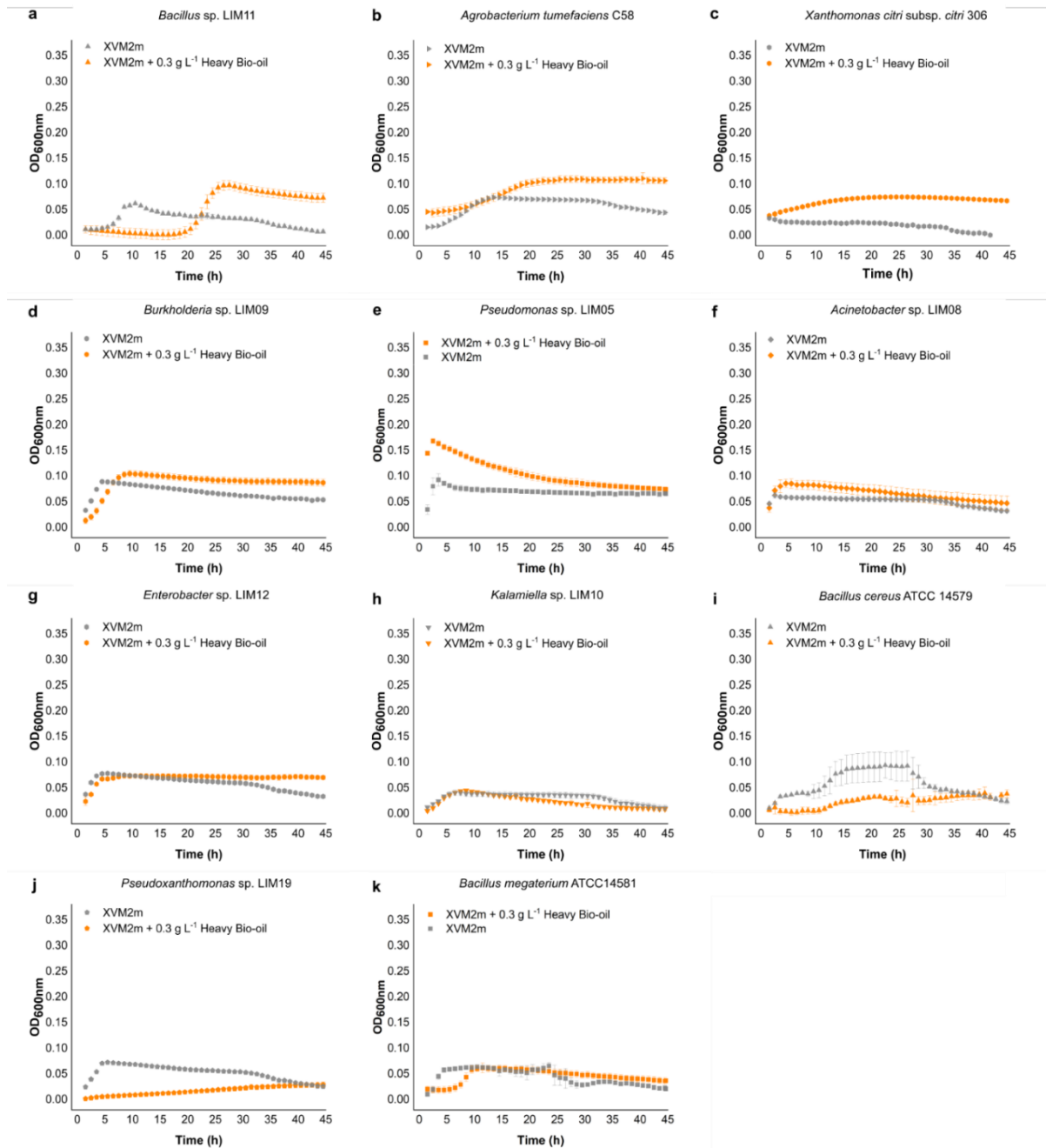


Figure S12. Growth curves of the bacterial isolates in 0.3 g.L⁻¹ heavy bio-oil. In gray is the negative control curve (XVM2m medium without carbon source supplementation), in orange are the growth curves with 0.3 g.L⁻¹ heavy bio-oil. Data shown as mean value of triplicate and standard deviation as error bar (mean \pm SD).

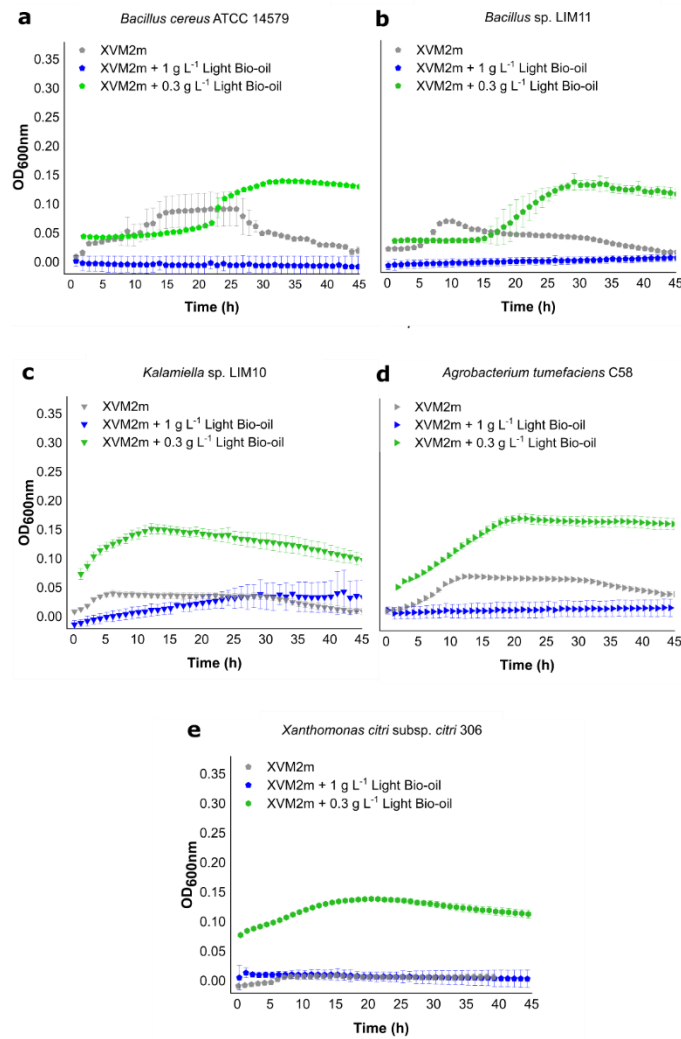


Figure S13. The decrease in light bio-oil concentration enabled the growth of 5 bacterial isolates sensitive to the 1 g.L⁻¹ supplementation in minimal medium. In gray is the negative control curve (XVM2m medium without carbon source supplementation), in blue are the growth curves in the minimal medium plus 1 g.L⁻¹ light bio-oil and in green is the growth profile in minimal medium plus 0.3 g.L⁻¹ of the light bio-oil. All curves are in the same x and y scales. Data shown as mean ± SD of triplicates.

Table S1. Taxonomy classification based on the 16S rDNA sequencing analysis of the bacteria isolated from environmental samples. Available in Supplementary file Table S1.xls. (<https://doi.org/10.1016/j.ijbiomac.2022.10.269>)

Name	Origin
<i>Pseudomonas</i> sp. LIM05	Water + soil from a pool in Puritama Hot Springs - Atacama Desert/Chile
<i>Acinetobacter</i> sp. LIM08	Sugarcane soil - São Paulo/BR
<i>Burkholderia</i> sp. LIM09	Sugarcane soil - Paraná/BR
<i>Kalamiella</i> sp. LIM10	Sugarcane soil - Goiás/BR
<i>Bacillus</i> sp. LIM11	Sugarcane filter cake composting - São Paulo/BR
<i>Enterobacter</i> sp. LIM12	Energy cane soil - São Paulo/BR
<i>Pseudoxanthomonas</i> sp. LIM19	Sugarcane filter cake composting - São Paulo/BR

Table S2. Model strains used in this research.

<i>Bacillus cereus</i> ATCC 14579
<i>Agrobacterium tumefaciens</i> C58
<i>Xanthomonas citri</i> subsp. <i>citri</i> 306
<i>Bacillus megaterium</i> ATCC 14581

Table S3. Mean molecular properties of the compounds in the light bio-oils generated under reaction conditions R1-R11. Molecular weight (M_w), partition coefficient ($\log P$), hydrogen bond donor count (HB_D), hydrogen bond acceptor count (HB_A), and molecular complexity (C_{PLX}) were calculated using computed molecular properties from the PubChem database and % peak area from GC-MS data as statistical weights.

Reaction conditions	Mean molecular properties				
	M_w	$\log P$	HB_D	HB_A	C_{PLX}
R1	144.1	1.29	0.87	2.54	120.0
R2	144.0	0.72	1.06	2.94	121.5
R3	141.2	1.55	0.71	2.22	119.2
R4	141.7	1.16	0.87	2.53	117.9
R5	140.7	1.11	1.11	2.67	106.3
R6	141.2	1.09	1.12	2.69	106.8
R7	140.4	1.14	1.10	2.63	104.4
R8	131.3	1.15	1.43	2.33	90.8
R9	122.7	1.23	1.50	2.01	81.2
R10	134.3	1.26	1.13	2.33	92.8
R11	127.7	1.36	1.15	2.01	85.3

Capítulo 2

Unraveling aromatic metabolism in a plant pathogen reveals novel pathways for detoxification and metabolic convergence of lignin monomers

Damaris B. Martim^{1,2}, Anna J. V. C. Brilhante^{1,2}, Augusto R. Lima^{1,2}, Lucia D. Wolf², Douglas zA. A. Paixão², Joaquim M. Junior², Fernanda M. Kashiwagi², Juliana A. Aricetti², Fabrícia F. de Menezes², Gabriela F. Persinoti², George J. M. Rocha² and Priscila O. Giuseppe^{2,*}

¹ Graduate Program in Genetics and Molecular Biology, Institute of Biology, University of Campinas (UNICAMP), Campinas, São Paulo, Brazil.

² Brazilian Biorenewables National Laboratory (LNBR), Brazilian Center for Research in Energy and Materials (CNPEM), Campinas, São Paulo, Brazil

* Corresponding authors. (P.O. de Giuseppe) priscila.giuseppe@lnbr.cnpem.br

Submetido à Revista *Nature Communications* em 8 de janeiro de 2024.

Abstract

Lignin, a complex aromatic macromolecule abundantly found in plant cell walls, stands out as a promising renewable carbon source for the production of bio-based chemicals. In recent decades, microbial convergent metabolism has emerged as an attractive approach for converting lignin side-streams into value-added chemicals, but our current knowledge of metabolic pathways enabling this process remains limited to a few microbial species and unavailable for some lignin-derived compounds. Here, we conducted an in-depth investigation into the enzymes and pathways employed by the model plant pathogen *Xanthomonas citri* subsp. *citri* 306 for metabolizing a wide range of lignin-related aromatic compounds. Based on an integrative approach, including RNA-seq, enzymology and gene knockout studies, we revealed peripheral pathways for the catabolism of the three main monolignols (*p*-coumaryl, coniferyl, and sinapyl alcohols), encoded by a previously unknown operon, designated as *molRKAB*. Our comprehensive analysis also demonstrated all the necessary enzymatic steps for funneling the monolignols into the tricarboxylic acid cycle, concurrently uncovering aryl aldehyde reductases and efflux transporters that likely protect the pathogen from aromatics toxicity. Together, these findings enhance the current understanding of the microbial metabolism and transcriptional responses to lignin-related aromatic compounds, shedding light on the diverse metabolic pathways available in nature to enable the engineering of microbial chassis dedicated to the upcycling of lignin and other aromatic-rich agro-industrial residues.

Keywords: lignin, aromatics, phenolic compounds, monolignol, RNA-seq, transcriptome, *Xanthomonas*, metabolic pathways, dehydrogenase, reductase

Introduction

The bioconversion of aromatic compounds has a central role in carbon cycling³³, plant-pathogen interactions^{34–36} and detoxification of organic pollutants³⁷. This process typically starts in upper pathways (also named peripheral pathways), which converge the structural diversity of aromatic compounds to fewer intermediate metabolites, which are further funneled to central carbon metabolites through a narrower range of “lower” pathways, also known as fission pathways^{19,38}. During this convergent process, industrially-relevant molecules are formed, offering opportunities for engineering microbial chassis to produce chemicals from aromatic-rich mixtures derived from complex and abundant wastes such as lignin and mixed-plastics^{14,39}.

Hundreds of microorganisms across diverse phyla are known to metabolize lignin-related monomers, but the metabolic pathways of only a few have been characterized so far^{18,19}. The molecular mechanisms related to the bioconversion of aromatic compounds have been mainly studied in some model organisms isolated from soil, such as *Pseudomonas putida* KT2440, or industrial wastewater, such as *Sphingobium* sp. SYK-6^{40,41}. Thus, our understanding of how microorganisms metabolize aromatic molecules in other ecological niches, such as during plant infection or in the gut of herbivores, remains limited. Moreover, how these molecules impact microbial behavior and physiology is still partially understood^{42,43}.

Some plant pathogens, such as *Xanthomonas* species, have a vast arsenal of enzymes to degrade components of the plant cell wall such as xyloglucan²⁷ and xylan⁴⁴, using the released carbohydrates as a source of carbon, energy, and stimuli²⁸. However, little is known about their capacity and molecular strategies to metabolize other compounds available in the plant cell wall, especially the phenolics related to lignin, a major plant cell wall component^{30–32,45}. One of the plant defense mechanisms against *Xanthomonas* infection is to increase the lignification of the plant cell wall, which implies an increased secretion of the lignin precursors, named monolignols, in the infection site⁴⁶. Currently, it is known only a microbial catabolic pathway for coniferyl alcohol^{19,47}, remaining to be determined if there are also pathways for the catabolism of the other main monolignols (*p*-coumaryl and sinapyl alcohols).

In this work, by combining RNA-seq analysis, biochemical characterization, and gene knockout studies, we investigated, in a genome-wide scale, the metabolism of lignin-related aromatics in the model phytopathogen *Xanthomonas citri* subsp. *citri* 306 (*X. citri* 306), providing insights on the influence of these compounds in physiological responses. Our data revealed complete pathways for the catabolism of the three main monolignols, reductive metabolic detours and efflux approaches to cope with aromatics toxicity. Moreover, this study

also showed that lignin-related compounds activate chemotaxis and flagellar-dependent motility in this phytopathogen, which might have an important role during the infection of the plant host. Thus, this work expands the current knowledge on the molecular diversity of metabolic pathways for the conversion of lignin-related aromatics compounds, providing insights about molecular mechanisms involved in plant-pathogen interactions and fundamental knowledge to guide the metabolic engineering of microbial chassis targeting the valorization of agro-industrial side-streams rich in aromatic compounds.

Results

The model plant pathogen uptakes and metabolizes a diverse range of lignin-related aromatics.

To investigate the range of aromatic compounds metabolized by *X. citri* 306, we evaluated its capacity to grow on minimal medium using, as the main carbon source, 21 different aromatic monomers, representative of the three types of lignin units, *p*-hydroxyphenyl (H), guaiacyl (G) and syringyl (S), besides two complex mixtures of compounds released from lignin by thermochemical processes (Low-molecular-mass (LM) Lignin⁴⁸ and Light Bio-oil⁴⁹) (Table 1). Under the tested conditions, only 4-hydroxybenzoate (4HBA) and the two lignin-derived complex mixtures supported the growth of *X. citri* 306 (Fig. 1a and Supplementary Fig. 1). GC-MS analysis identified the presence of 11 aromatic monomers in LM lignin, of which three (syringaldehyde, vanillin, and 4-hydroxybenzaldehyde) were completely depleted after 30 h of *X. citri* 306 growth (Supplementary Table 1). In the Light Bio-oil sample, 4-hydroxybenzaldehyde was completely depleted by *X. citri* 306, along with other partially depleted compounds, including homovanillyl alcohol and catechol derivatives (Supplementary Table 1).

To address if the medium used was insufficient to provide the necessary energy for both aromatics metabolism and cell growth in the other tested conditions, we supplemented it with 5 mmol L⁻¹ of glucose. Remarkably, the glucose supplementation allowed the growth of *X. citri* 306 in the presence of the lignin-related aromatics tested, except for those displaying severe toxicity effects when provided at 5 mmol L⁻¹, such as some aryl aldehydes (Table 1 and Supplementary Fig. 1). Overall, the presence of 5 mmol L⁻¹ of aromatic compounds resulted in a decreased growth rate when compared to the medium containing only glucose (XVM2m(G)), indicating that these compounds display some level of toxicity to *X. citri* 306, with the aryl

aldehydes being generally more toxic than the correspondent aryl alcohols and aryl acids (Table 1 and Supplementary Fig. 1).

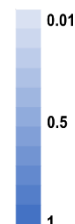
To determine if these aromatic compounds have been metabolized by *X. citri* 306, we performed HPLC analyses of the medium supernatant before and after bacterial growth on XVM2m(G) added by these compounds in two concentrations (millimolar and/or micromolar) (Table 1 and Supplementary Table 2). In aryl alcohol cultivations at the micromolar level, only the three monolignols (*p*-coumaryl, coniferyl, and sinapyl alcohol) were effectively consumed by *X. citri* 306 (Fig. 1b, Table 1 and Supplementary Table 2). For all the tested aldehydes, a depletion higher than 70% was observed at the millimolar condition (except for 4-hydroxybenzaldehyde, < 30%) (Table 1 and Supplementary Table 2).

Curiously, among the aryl acids tested, *X. citri* 306 showed a negligible capacity on consuming them when provided at the milli and micromolar level, except for 4-hydroxybenzoate. Together, these results indicate that *X. citri* 306 has pathways to metabolize aromatic compounds as complex as the monolignols and effectively metabolizes a more diverse range of aryl aldehydes compared to aryl alcohols and aryl acids provided at the extracellular environment (Table 1 and Supplementary Table 2).

Table 1. Summary of growth conditions and analysis of aromatics toxicity and depletion.

Growth assays in minimal medium (XVM2m) and XVM2m plus 5 mmol L⁻¹ glucose (XVM2m(G)) and or 5 mmol L⁻¹ of lignin-related aromatic compounds or 1 g L⁻¹ LM Lignin or 0.3 g L⁻¹ Light Bio-oil. LM Lignin is a fraction of low molar mass aromatics derived from the alkaline pretreatment of sugarcane bagasse. Light Bio-oil is the soluble fraction of aromatics resulted from the hydrothermal depolymerization of sugarcane bagasse lignin. The red gradient represents differences in growth parameters from low (blank) to high (red). HPLC analyses were performed using XVM2m(G) medium plus 2 mmol L⁻¹ aryl alcohols, 1 mmol L⁻¹ aldehydes, 5 mmol L⁻¹ acids, or 50 µmol L⁻¹ of either aryl alcohols or acids. The blue color bar represents the percentage of aromatic compound depleted by *X. citri* 306 after 15 hours of growth (as detailed in Supplementary Table 2). Conc. = concentration, N.A. = not evaluated. Data are shown as mean ± SD of at least three biological replicates.

Conditions	XVM2m	XVM2m(G)		Depletion (HPLC)	
	OD _{600max}	OD _{600max}	μ (h ⁻¹)	high conc.	low conc.
Aromatic-free control	0.03 ± 0.01	0.12 ± 0.01	0.44		N.A.
Aryl alcohols (mmol L⁻¹)	5	5	5	2	0.05
p-coumaryl	0.00	0.05 ± 0.00	0.08		
coniferyl	0.00	0.11 ± 0.01	0.27		
sinapyl	0.00	toxic	toxic	N.A.	
4-hydroxybenzyl	0.00	0.13 ± 0.01	0.20		
vanillyl	0.00	0.10 ± 0.02	0.18		
syringyl	0.00	0.10 ± 0.00	0.12		
benzyl	0.00	0.11 ± 0.01	0.09		
Aryl aldehydes (mmol L⁻¹)	5	5	5	1	0.05
p-coumaraldehyde	0.00	toxic	toxic		N.A.
coniferaldehyde	0.00	toxic	toxic		N.A.
sinapaldehyde	0.00	toxic	toxic		N.A.
4-hydroxybenzaldehyde	0.00	toxic	toxic		N.A.
vanillin	0.00	0.09 ± 0.00	0.08		N.A.
syringaldehyde	0.00	0.09 ± 0.00	0.07		N.A.
benzaldehyde	0.03 ± 0.00	0.11 ± 0.01	0.10		N.A.
Aryl acids (mmol L⁻¹)	5	5	5	5	0.05
p-coumarate	0.00	0.11 ± 0.00	0.17		
ferulate	0.02 ± 0.00	0.07 ± 0.00	0.14		
sinapate	0.02 ± 0.00	0.05 ± 0.00	0.17		
4-hydroxybenzoate	0.11 ± 0.00	N.A.	N.A.		
vanillate	0.00	0.08 ± 0.02	0.13		
syringate	0.02 ± 0.00	0.11 ± 0.02	0.25		
benzoate	0.00	0.10 ± 0.02	0.15		
Lignin-derived complex mixtures (g L⁻¹)					
LM Lignin (1.0)	0.10 ± 0.00	N.A.	N.A.	N.A.	N.A.
Light Bio-oil (0.3)	0.06 ± 0.00	N.A.	N.A.	N.A.	N.A.



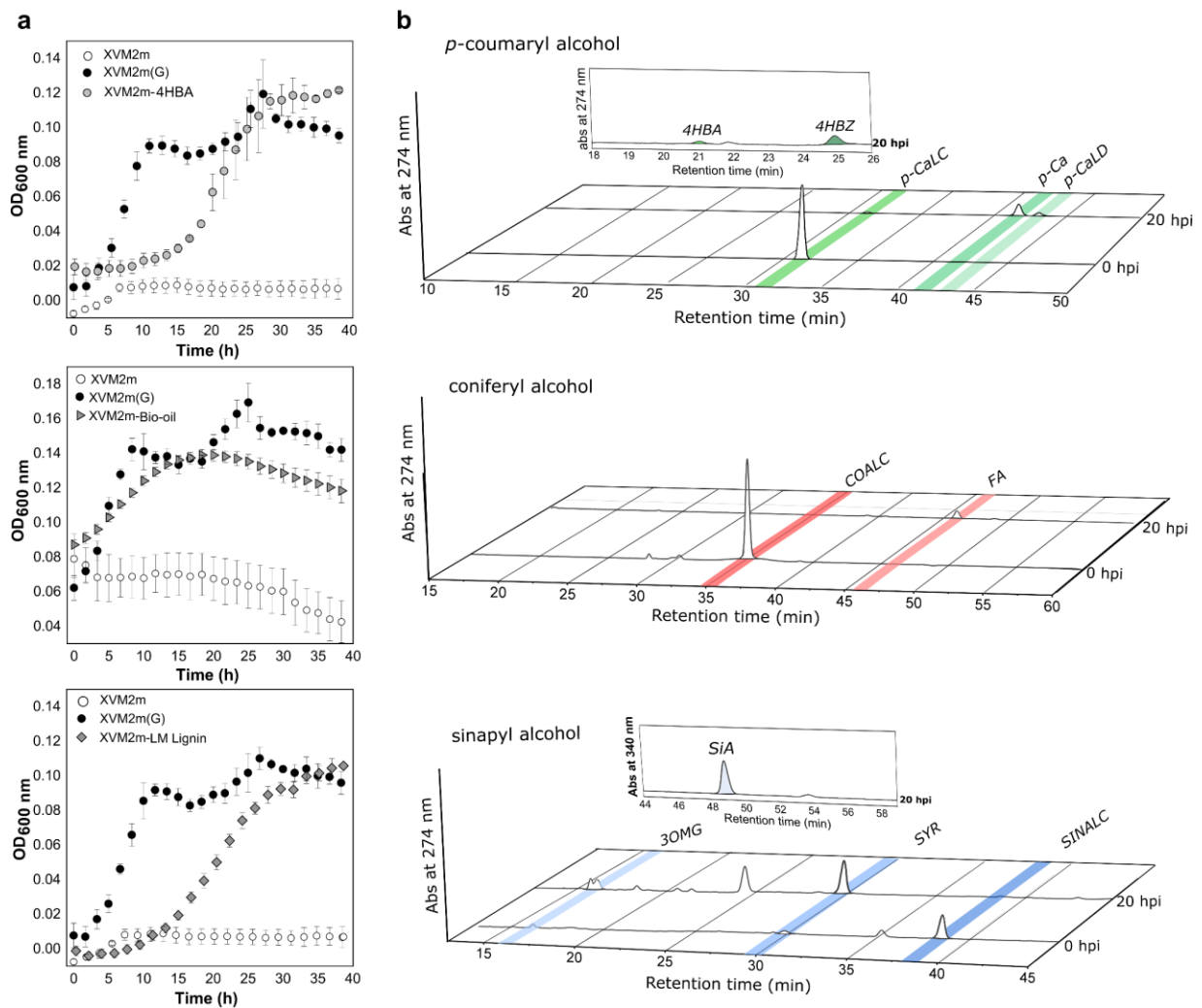


Fig. 1. *X. citri* 306 grows using 4-hydroxybenzoate and lignin-derived compounds as carbon sources and metabolizes the three main monolignols. a) Growth curves (hours) in XVM2m supplemented with 4-hydroxybenzoate (4HBA) and two complex mixtures of lignin-derived compounds. XVM2(G) – glucose-supplemented media. **(b)** HPLC chromatograms showing the depletion of the three monolignols and production of intermediate metabolites at 20 hours post inoculation (hpi). 4HBA = 4-hydroxybenzoate; 4HBZ = 4-hydroxybenzaldehyde; *p*-CaLC = *p*-coumaryl alcohol; *p*-Ca = *p*-coumarate; *p*-CaLD = *p*-coumaraldehyde; COALC = coniferyl alcohol; FA = ferulate; SiA = sinapate; 3OMG = 3-*O*-methylgalate; SYR = syringate; SINALC = sinapyl alcohol. Data are shown as mean \pm SD of at least three biological replicates.

Lignin-related compounds induces chemotaxis and flagellar-dependent motility

To investigate the transcriptional profile of *X. citri* 306 in response to lignin-related compounds and identify the metabolic pathways involved in their catabolism, RNA-seq studies were conducted. XVM2m(G) minimal medium was supplemented with six monomers (coniferyl alcohol, 4-hydroxybenzoate, 4-hydroxybenzaldehyde, vanillin, syringaldehyde, and

benzaldehyde), as well as three complex mixtures (LM lignin, Light Bio-oil, and aldehydes mix) (Supplementary Table 3). A total of 278 to 1464 differentially expressed genes (DEGs) in the aromatic-containing conditions compared to the control XVM2m(G) were identified (Supplementary Data 1), evidencing the importance of these compounds in modulating various physiological processes of *X. citri* 306 beyond the pathways related to their metabolism.

Notably, *X. citri* 306 demonstrates the ability to discern subtle structural variations among aromatic compounds, as shown by the incomplete overlap of upregulated genes in each condition (Fig. 2). For example, 4-hydroxybenzaldehyde activated the expression of around 400 genes that were not activated by 4-hydroxybenzoate, although they share a subset of about 200 genes activated by both (Fig. 2a). This discrepancy suggests that the simple oxidation of 4-hydroxybenzaldehyde into 4-hydroxybenzoate can elicit substantial alterations in the transcriptional response of *X. citri* 306. A similar trend is observed when comparing the same type of phenolic compound with distinct degrees of methoxylation (Fig. 2b). Additionally, analysis of the complex mixtures of lignin-derived compounds highlights only partial overlapping of the transcriptional response, indicating they present differences in terms of chemical composition that are sensed by *X. citri* 306 (Fig. 2c).

Although *X. citri* 306 is capable of recognizing subtle changes in the structure of lignin-related compounds, Gene Ontology (GO) enrichment analysis revealed that biological processes such as "signal transduction," "bacterial-type flagellum assembly," "bacterial flagellum-dependent cell motility," and "chemotaxis" were enriched and upregulated in all conditions featuring lignin-related aromatics (Fig. 2d, Supplementary Fig. 2 and Data 2). Within this context, we highlight the upregulation of chemotaxis and flagellar genes, including *cheAZY* (XAC1930-32), *motAB* (XAC3693-94), *fliC* (XAC1975), *flgL* (XAC1976), *flgG* (XAC1981), and *flgE* (XAC1983), suggesting that the sensing of lignin-related aromatics stimulates a motile state in *X. citri* 306. This observation is consistent with the results of the co-expression analysis, where a distinct co-expressed gene module (M1), predominantly composed of genes involved in flagellar assembly and chemotaxis, was identified (Fig. 2e). Significantly higher activity within this module becomes especially pronounced in the presence of 4-hydroxybenzaldehyde (4HBZ) and syringaldehyde (SYALD) while showing reduced activity under conditions involving complex lignin samples (LM Lignin and Bio-oil), which can be due to the comparatively lower concentration of aromatic compounds within these samples (Fig. 2e). Between the downregulated processes, translation and fatty acid biosynthesis were consistently enriched in most of the tested conditions (Supplementary Fig. 2 and Data 2).

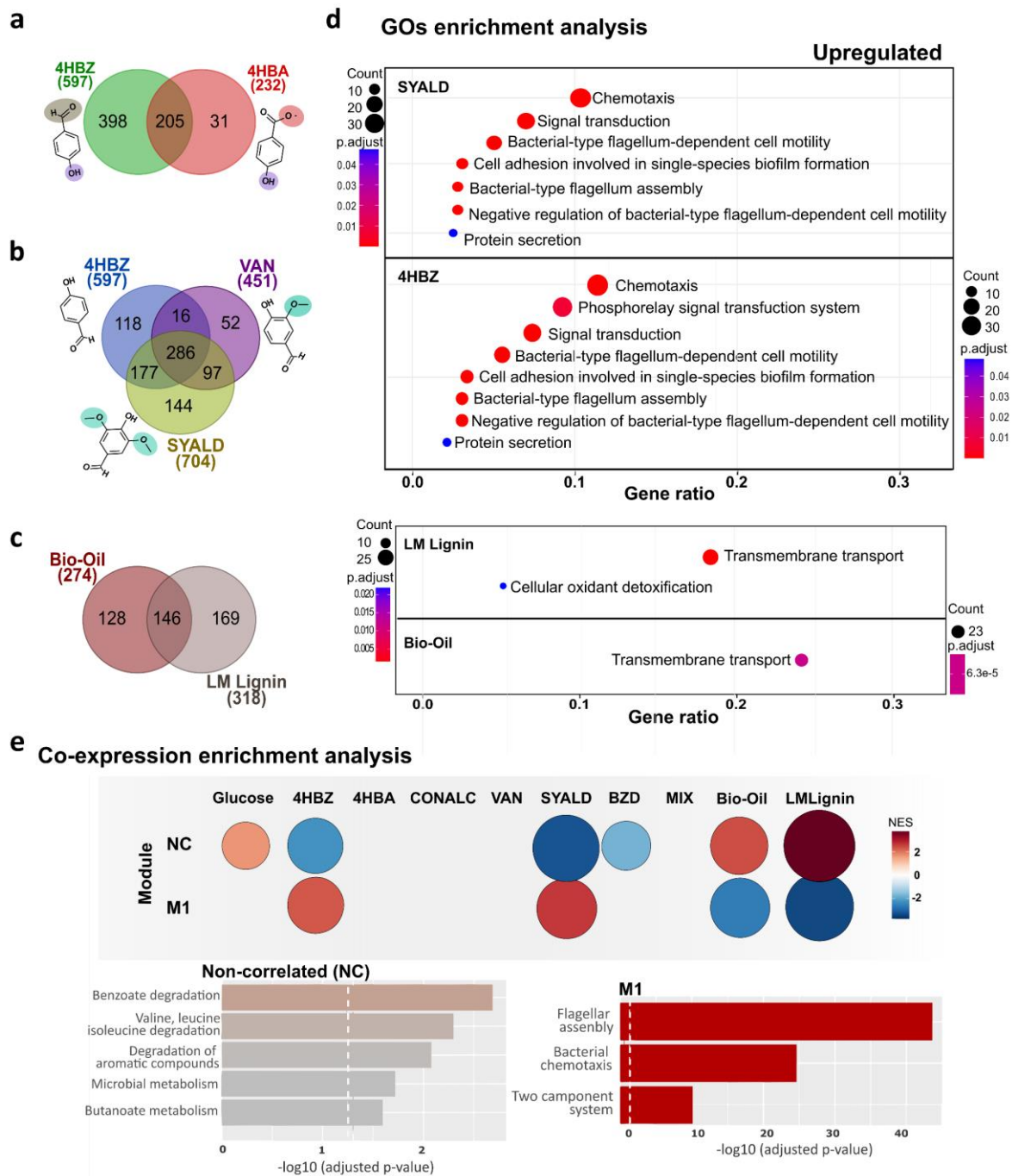


Fig. 2. Transcriptional responses triggered by lignin-related aromatic compounds. Venn diagrams comparing the distribution of unique and shared upregulated genes in conditions containing **a)** 4-hydroxybenzaldehyde (4HBZ) or 4-hydroxybenzoate (4HBA). **b)** 4HBZ, vanillin (VAN), or syringaldehyde (SYALD) and **c)** Low-molecular-mass lignin (LM Lignin) or Light Bio-oil (Bio-Oil) conditions. **d)** Gene ontology (GO) enrichment analysis of biochemical pathways considering the upregulated differentially expressed genes (DEGs) in conditions containing individual compounds (SYALD or 4HBZ), or complex mixtures of aromatics (LM Lignin or Bio-oil). The analysis of the other conditions is in [Supplementary Fig. 2](#). Circles' size and color represent the counts and adjusted *p*-values, respectively. Gene ratio corresponds to the number of DEGs related to a GO term divided by the total

number of annotated DEGs. The differential expression in each condition was compared to XVM2m(G) following the criteria \log_2 fold change ≥ 1 and p -adjusted ≤ 0.05 . **e)** Gene set enrichment analysis of co-expression gene modules identified using the CEMiTool package⁵⁰. The size and intensity of the circles correspond to the normalized enrichment score (NES) for the module in each condition, indicating biological functions enriched in each module. Positive NES reflects transcriptional activity above the median, whereas negative NES corresponds to transcriptional activity below the median in each condition. CONALC = coniferyl alcohol, BZD = benzaldehyde, MIX = aldehyde mixture.

The first steps of monolignols catabolism are performed by aryl alcohol and aryl aldehyde dehydrogenases

It has been known that the first step of coniferyl alcohol catabolism can be catalyzed by an NAD^+ dependent aryl alcohol dehydrogenase (ADH), generating coniferaldehyde, which is then converted to ferulate by a NAD^+ -dependent aryl aldehyde dehydrogenase (ALDH)^{51,52}. For *p*-coumaryl and sinapyl alcohols, this information is still missing, but considering their chemical similarity to coniferyl alcohol, we hypothesized that their catabolism might follow a similar pathway. Thus, to uncover the genes responsible for monolignols catabolism, we searched for ADH and ALDH genes upregulated in the presence of lignin-related compounds in *X. citri* 306. Based on their higher upregulation levels, genomic context, and the presence of common catalytic domains reported for dehydrogenases active on aromatics, we selected eight ADH genes and three ALDH genes for cloning, heterologous expression, and biochemical activity screening (Fig. 3).

The activity screening revealed several alcohol and aldehyde dehydrogenases active on aromatic compounds, with variable preferences in terms of substrate, co-substrate ($\text{NAD(H)}/\text{NADP(H)}$) and reaction direction (oxidation and/or reduction), which will be detailed in later sections (Fig. 3, Supplementary Tables 4 and 5).

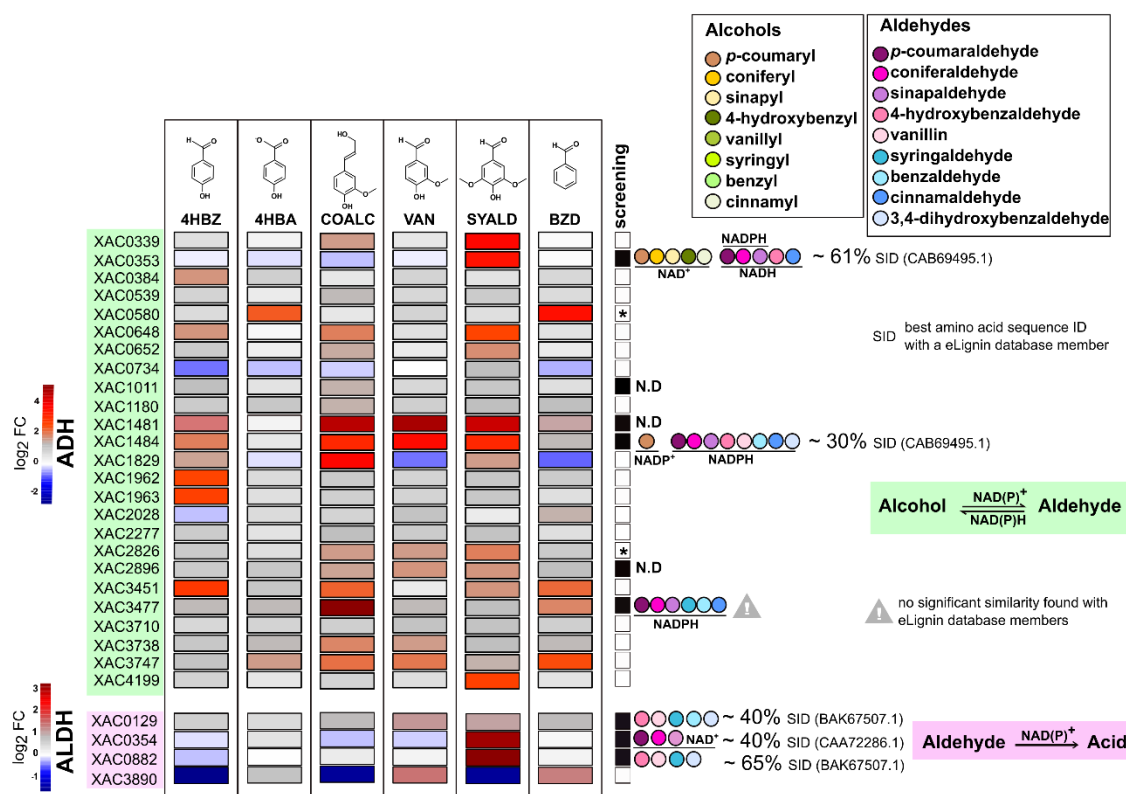


Fig. 3 Transcriptomic analysis and activity screening reveal novel ADH and ALDH enzymes. The heatmap presents transcription levels (as $\log_2\text{FC} = \log_2 \text{Fold Change}$), comparing each growth medium to the reference (XVM2M(G)) from at least biological triplicates. Genes were classified as up-regulated according to the following criteria: $\log_2\text{FC} \geq 1$, with an $p\text{-adjusted} \leq 0.05$. The black boxes indicate the enzymes that were selected for the activity screening, using purified enzymes as detailed in [Supplementary Data 3](#) and [Supplementary Tables 4 and 5](#) or whole cells assays (XAC0129 and XAC882) as detailed in [Supplementary Table 6](#). (*) indicates proteins that were not soluble in *E. coli*. 4HBZ = 4-hydroxybenzaldehyde; 4HBA = 4-hydroxybenzoate; COALC = coniferyl alcohol; VAN = vanillin; SYALD = syringaldehyde; BZD = benzaldehyde; N.D. = not determined. The ADH (green box) enzymes were subjected to screening for both direct and reverse reactions, using the corresponding substrates listed in the legend. Colored circles indicate the substrates and labels indicate the respective co-substrate in which the enzymes were active.

The screening revealed two genes that are clustered in the genome (XAC0353 and XAC0354) and encode enzymatic activities correspondent to the first steps of monolignols catabolism ([Fig. 4a](#)). XAC0353 gene product displayed a NAD^+ -dependent alcohol dehydrogenase activity over the three main monolignols, besides cinnamyl and 4-hydroxybenzyl alcohols, whereas XAC0354 gene product showed a NAD^+ -dependent aldehyde dehydrogenase activity on H-, G- and S-type hydroxycinnamic aldehydes, suggesting a role in the second step of monolignols catabolism in *X. citri* 306.

To confirm these results, we measured the specific activity of XAC0353 and XAC0354 in several aromatic substrates and confirmed the products by HPLC (Fig. 4b-c and Supplementary Fig. 3). As expected, XAC0353 converted aryl alcohols in their respective aldehydes using NAD^+ as co-substrate, showing the highest specific activity over the monolignols compared to benzyl alcohol derivatives (Fig. 4b). Accordingly, the aldehyde dehydrogenase encoded by XAC0354 transformed aryl aldehydes in their respective acids in a NAD^+ -dependent manner, showing the highest specific activity on coniferaldehyde (100%), followed by *p*-coumaraldehyde (51%) and sinapaldehyde (34%), compared to the other tested aromatics (Fig. 4c). Both enzymes were assayed in equivalent co-substrate concentrations, at the same temperature, under near-physiological pH, and displayed similar specific activity rates over their best substrates (monolignols and the respective aldehydes), which might be advantageous to avoid the accumulation of the intermediate aldehyde metabolites during the serial action of these enzymes.

To validate the *in vivo* involvement of these enzymes in monolignol catabolism, we evaluated the impact of XAC0353 or XAC0354 gene deletion on the bacterial capacity to metabolize monolignols, or hydroxycinnamic aldehydes, and excrete intermediate metabolites generated during this process. Growth assays in XVM2(G) containing $50 \mu\text{mol L}^{-1}$ monolignols revealed that XAC0353 deletion impaired monolignols consumption and decreased the excretion of hydroxycinnamic acids, which are intermediate metabolites of monolignols catabolism (Fig. 4d). Regarding sinapyl alcohol, although we did not detect its presence in the growth assays of WT and $\Delta\text{XAC0353}$ (*KO53*) strains after 15 h and 40 h incubation, likely due to instability issues⁵³, we noticed a significantly lower amount of sinapate being excreted by the *KO53* strain compared to the WT, which is consistent with the results observed for the other two monolignols (Fig. 4d). Together, these results indicate that the deletion of XAC0353 gene compromises the catabolism of the monolignols, leading to a lower excretion of the second intermediate metabolite of this process, whose conversion to the next metabolite might be a limiting factor in *X. citri* 306. They also point to the presence of a secondary gene related to the monolignols dehydrogenation, since the excretion of hydroxycinnamic acids was not completely abolished in the *KO53* strain.

The growth assays using $50 \mu\text{mol L}^{-1}$ of hydroxycinnamic aldehydes revealed that no hydroxycinnamic acids, the products of the enzymatic reaction catalyzed by XAC0354, were excreted by the *KO54* strain ($\Delta\text{XAC0354}$), supporting a crucial role for this gene in the catabolism of these aldehydes (Fig. 4e). On the other hand, a higher amount of excreted monolignols (*p*-coumaryl and coniferyl alcohol) was observed for the *KO54* strain compared to

the WT, indicating that the deletion of the XAC0354 gene triggered a metabolic shift toward reactions that reduce these aldehydes into their respective monolignols, which is consistent with the complete depletion of the hydroxycinnamic aldehydes by the WT and the *KO54* strain, although by preferentially following opposite metabolic directions.

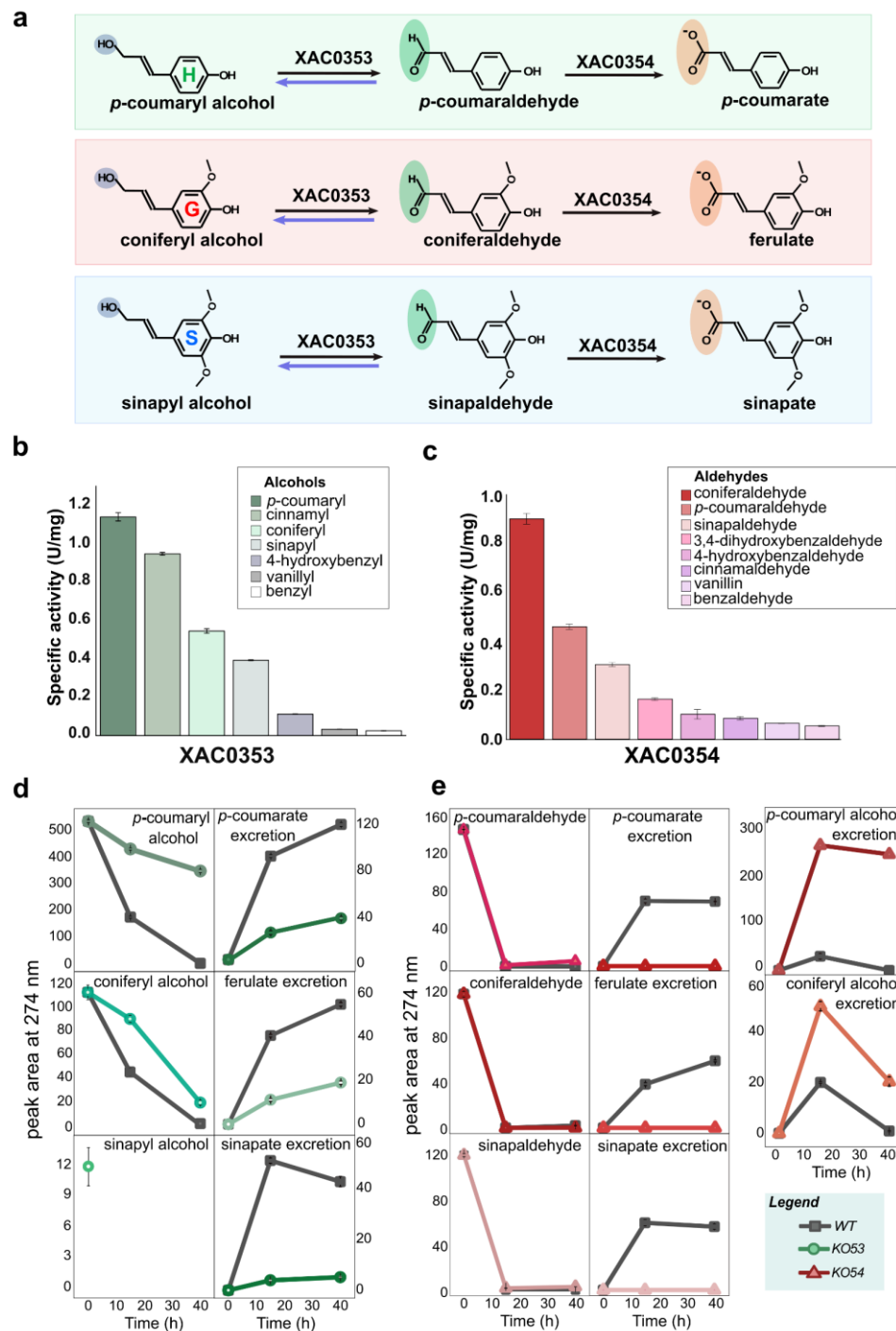


Fig. 4 XAC0353 and XAC0354 play a role in the first steps of monolignols catabolism. a) Schematic of the proposed *p*-coumaryl alcohol (green), coniferyl alcohol (red), and sinapyl alcohol (blue) bioconversion pathway in *X. citri* 306. Blue arrows indicate the reaction of aldehydes reduction (reductive pathways). **b-c)** Specific activity for the oxidation reactions of alcohols and aromatic

aldehydes catalyzed by enzymes encoded by XAC0353 and XAC0354, respectively. The activity was calculated based on NADH production, quantified by HPLC. **d)** HPLC analysis of the consumption of H, G, and S monolignols and the excretion of intermediate metabolites (hydroxycinnamic acids) by the *WT* (gray) and *KO53* knockout (shades of green) strains. Sinapyl alcohol was not detected during bacterial growth, likely due to instability issues. **e)** HPLC analysis of the consumption of H, G, and S hydroxycinnamic aldehydes and the excretion of metabolites (hydroxycinnamic acids or monolignols) by the *WT* (gray) and *KO54* knockout (shades of red) strains. Sinapyl alcohol excretion was not detected, which we attribute to its instability as previously reported⁵³. Data are shown as mean \pm SD of three biological replicates.

Syringaldehyde induces the catabolism of hydroxycinnamic acids

Although *X. citri* 306 did not utilize hydroxycinnamic acids (HCAs) when available in the growth medium ([Supplementary Table 2](#)), HPLC analysis indicated that they are produced as intermediate metabolites of monolignols catabolism, along with smaller intermediates resultant from their degradation such as 4-hydroxybenzoate and syringate ([Fig. 1b](#)). Thus, these findings suggest that *X. citri* 306 possesses the necessary enzymatic systems for metabolizing HCAs produced intracellularly.

Genome mining of *X. citri* 306 revealed a gene cluster (XAC0881-84) homologous to the *hca* gene cluster that plays a role in the degradation of *p*-coumarate, ferulate, and sinapate in *Xanthomonas campestris* pv. *campestris*³² ([Fig. 5 a,b](#)). In our RNA-seq data, this cluster was upregulated in the presence of syringaldehyde, an intermediate of the sinapate catabolism ([Fig. 5c](#)). The *hca* gene cluster from *X. citri* 306 harbors the XAC0882 gene, which encodes for an aldehyde dehydrogenase active on 4-hydroxybenzaldehyde, vanillin, and syringaldehyde, according to our activity screening results ([Fig. 3 and Supplementary Table 6](#)). This activity is compatible with the third step of HCAs catabolism, corroborating the involvement of the XAC0881-84 cluster in the metabolism of HCAs in *X. citri* 306. Besides XAC0882, another gene (XAC0129) also encodes for an aldehyde dehydrogenase active on aldehydes derived from HCAs catabolism ([Fig. 3 and Supplementary Table 6](#)), implying that there is a functional redundancy for this metabolic step in *X. citri* 306.

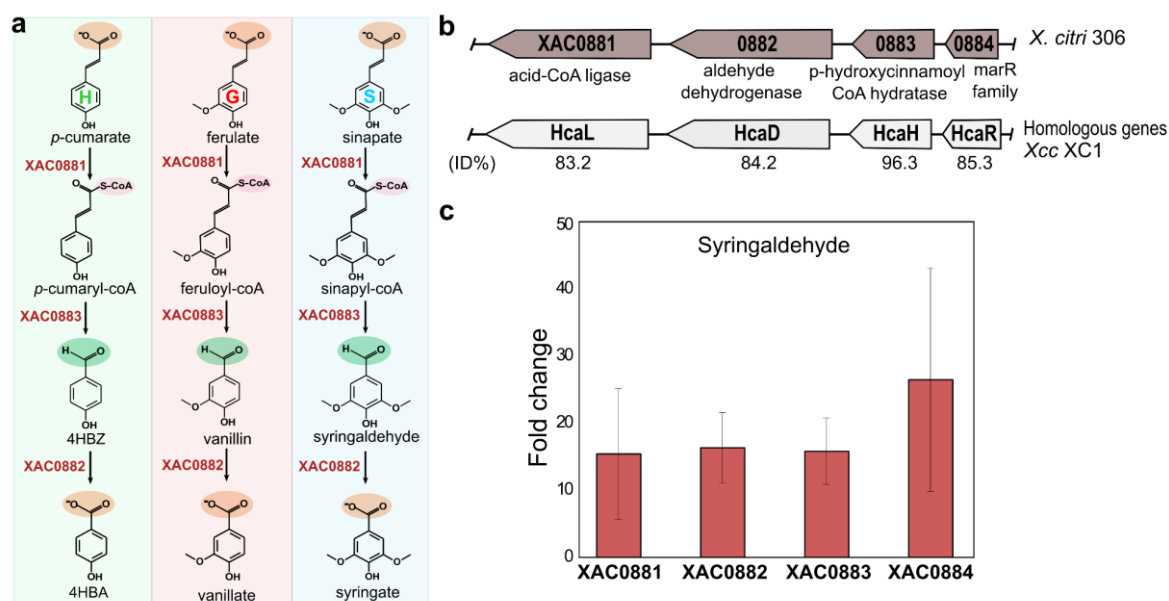


Fig. 5 Hydroxycinnamic acids catabolism is encoded by the *hca* gene cluster. **a)** Schematic representation of the degradation pathways proposed for *p*-coumarate (green), ferulate (red) and sinapate (blue) in *X. citri* 306. **b)** XAC0881-84 gene cluster in *X. citri* 306 (brown arrows), showing the amino acid sequence identity with the *hca* gene products (grey arrows), experimentally validated as responsible for HCAs degradation in *X. campestris* *pv. campestris* str. XC1 (*Xcc* XC1)³². **c)** Transcription level (Fold Change) of XAC0881-84 genes upregulated in the syringaldehyde condition compared to the glucose condition. 4HBZ = 4-hydroxybenzaldehyde, 4HBA = 4-hydroxybenzoate. Data are shown as mean ± SD from four biologically independent experiments. Genes were considered upregulated according to the following criteria: *p*-adjusted ≤ 0.05 and Fold Change > 2.

Novel reductive pathways for aryl aldehydes detoxification

Among the putative ADHs screened in the previous section, XAC1484 and XAC3477 displayed a NADPH-dependent reductase activity, catalyzing the reduction of aryl aldehydes in their respective alcohols, being active on derivatives of both cinnamyl and benzyl aldehydes (Fig. 3). Curiously, XAC0353 also displayed a similar activity, but in a NAD(P)H-dependent manner, indicating that it might also contribute to the reduction of aryl aldehydes depending on the intracellular balance of NAD⁺/NAD(P)H and aryl alcohol/aryl aldehyde substrates. Indeed, *X. citri* 306 growing on a mixture of aldehydes at millimolar level showed to be capable of converting, at least part of them, into aryl alcohols, which were then exported to the extracellular medium (Fig. 6a). In such condition, the genes XAC1484 and XAC3477 were upregulated, supporting that they have an important role in converting aryl aldehydes into their respective aryl alcohols (Supplementary Data 1). Moreover, XAC1484 is clustered with the genes XAC1482-83-85, which encodes for a putative Resistance-Nodulation-Division (RND)

multidrug efflux transporter that might be involved in the excretion of the aryl alcohols produced by XAC1484 when exposed to aryl aldehydes.

Since two metabolic directions, reductive or oxidative, are possible to the bioconversion of aryl aldehydes in *X. citri* 306, we argued how the concentration of these compounds affect their metabolic fate. To address this issue, we grew *X. citri* 306 in minimal medium supplemented with increasing concentrations of coniferaldehyde, used as a representative compound, and quantified its depletion as well as the production of intermediate metabolites excreted after 15 h of growth (Fig. 6b). Overall, the major fate of coniferyl aldehyde was its complete catabolism, but in lower coniferaldehyde concentrations, ferulate appeared as the major metabolite excreted to the extracellular medium whereas in higher coniferaldehyde concentrations, coniferyl alcohol assumed this role (Fig. 6b). Thus, these results indicate that increasing concentrations of aryl aldehydes activate reductive pathways, providing an additional mechanism to cope with toxicity, which proved to be higher for hydroxycinnamic aldehydes compared to their alcohol or acid counterparts (Table 1).

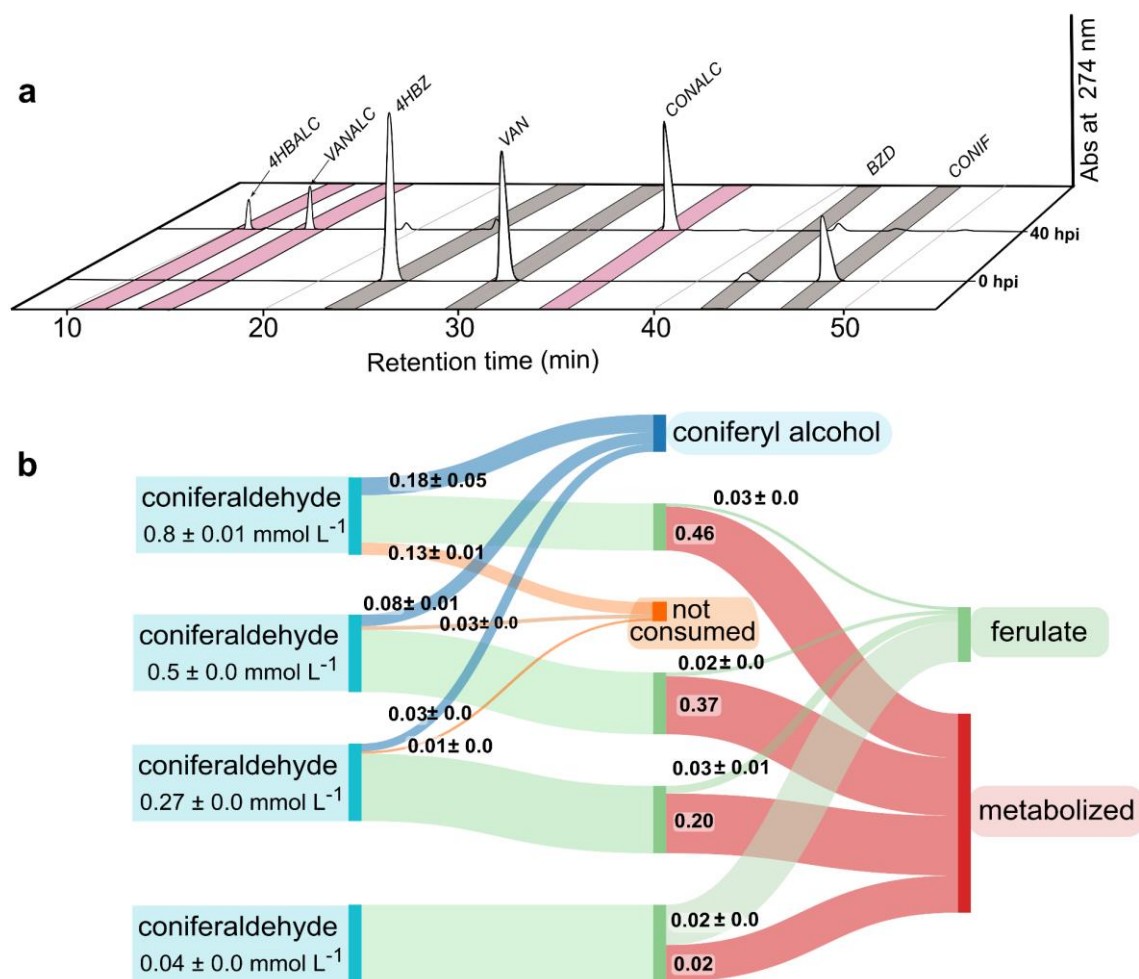


Fig. 6. Aryl aldehydes have multiple metabolic fates. a) HPLC chromatogram showing the consumption of a mixture of aldehydes (coniferaldehyde, benzaldehyde, vanillin, and 4-hydroxybenzaldehyde) and their respective alcohols excreted to the extracellular environment after 40 hpi. **b)** Sankey graph (<https://sankeymatic.com>) of HPLC quantification of coniferaldehyde consumption and excreted metabolites after 15 hpi. CONIF = coniferaldehyde, BZD = benzaldehyde, CONALC = coniferyl alcohol, VAN = vanillin, 4HBZ = 4-hydroxybenzaldehyde, VANALC = vanillyl alcohol, 4-HBALC = 4-hydroxybenzyl alcohol. Data are shown as mean ± SD of three biological replicates.

H and G-type compounds are metabolic funneled via protocatechuate ortho-cleavage

The final step in the funneling pathways of H- and G-type lignin units involves the formation of protocatechuate (PCA)^{18,19}. In the case of H-type monomers, it is known that 4-hydroxybenzoate (4HBA) is converted to PCA by the enzyme *p*-hydroxybenzoate hydroxylase (EC 1.14.13.2)⁵⁴. In *X. citri* 306 genome, we found a gene (XAC0356) encoding for a protein homologous to the *p*-hydroxybenzoate hydroxylase (PobA) from *Pseudomonas aeruginosa*⁵⁵

and from *X. campestris*³⁰. XAC0356 gene was upregulated in the 4HBA condition (Fig. 7a) and its deletion abolished *X. citri* 306 growth in a medium containing 4HBA as the primary carbon source, demonstrating that this gene encodes a functional PobA and is essential for the metabolism of 4HBA in *X. citri* 306 (Fig. 7b).

In the last step of the G-type units funneling pathway, vanillate needs to undergo *O*-demethylation to form PCA^{19,56}. Three main types of *O*-demethylase systems have been reported in the literature so far: tetrahydrofolate (THF)-dependent enzymes, Rieske-type oxygenases (ROs) and cytochromes P450 oxygenases (P450s)^{57,58}. In the *X. citri* 306 genome, we found genes homologous to ROs-type oxygenases, usually named VanA subunits (XAC0311 and XAC0363) clustered with genes encoding for putative VanB-like reductases (XAC0310 and XAC0362), as well as a predicted P450 gene (XAC3170, <24% sequence identity with GcoA)⁵⁹. In the vanillin condition, only XAC0362-63 were upregulated (Fig. 7c) and the deletion of this gene pair (but not of XAC0310-11) resulted in vanillate accumulation during bacterial growth (Fig. 7d). Enzymatic assays showed that XAC0362-63 display vanilate-*O*-demethylase activity, further supporting a role for these genes in converting vanillate into PCA (Fig. 7e and Supplementary Fig. 4).

Once formed, PCA is cleaved in one of three possible positions: 2,3-cleavage, 4,5-(meta-cleavage), or 3,4-(ortho-cleavage)¹⁹. In *X. citri* 306, conditions containing H and G-type lignin aromatic compounds induced the expression of a gene cluster (XAC0364-71) homologous to the *pcaIJFHGBDC* cluster previously found in *X. campestris*³⁰, indicating the presence of a 3,4-(ortho-cleavage) pathway (Fig. 7f and Supplementary Data 1).

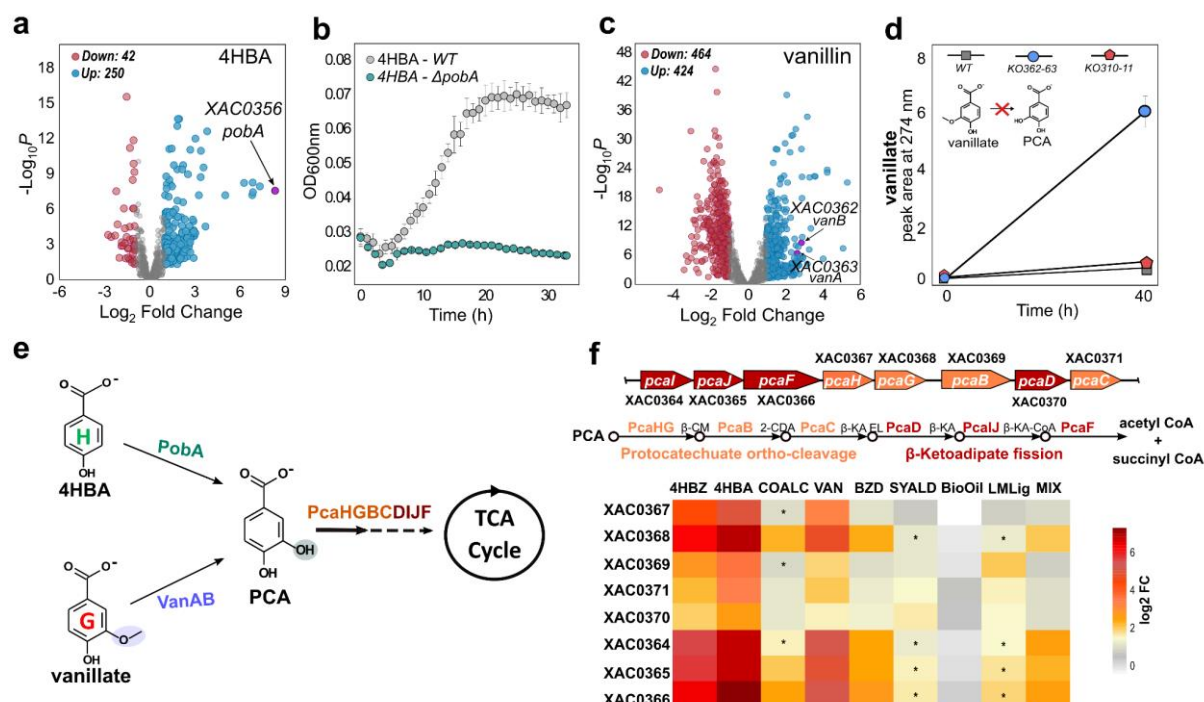


Fig. 7 H and G-type lignin monomers are funneled to the protocatechuate ortho-cleavage pathway.

a) Volcano plot of RNA-seq data highlighting the XAC0356 (*pobA* gene) upregulated in the 4HBA condition. **b)** Growth curve of *pobA* knockout strain (green circles) and *WT* strain (gray circles) on XVM2m containing only 4HBA as the primary carbon source. **c)** Volcano plot of RNA-seq data highlighting the XAC0362-63 upregulated in the vanillin condition. **d)** HPLC analysis demonstrating vanillate accumulation only by the Δ XAC0362-0363 (*KO362-63*) strain. **e)** Representative scheme of 4HBA and vanillate conversion into protocatechuate (PCA) and its ring cleavage pathway in *X. citri* 306. **f)** Schematic representation of the genomic context of *pca* genes, the corresponding metabolic steps, and a heat map of RNA-seq data (log2 Fold Change) showing the *pca* genes up-regulated under conditions containing H and G-type aromatic compounds and in benzaldehyde. COALC = coniferyl alcohol; VAN = vanillin; 4HBA = 4-hydroxybenzoate; 4HBZ = 4-hydroxybenzaldehyde; BZD = benzaldehyde (BZD); SYALD = syringaldehyde; PCA = protocatechuate; β -CM = β -carboxy-*cis,cis*-muconate; 2-CDA = γ -carboxymuconolactone; β -KA EL = β -keto adipate enol-lactone; β -KA = β -keto adipate; β -KACoA = β -keto adipyl-CoA. Genes were considered up-regulated according to the criteria, log2 Fold Change ≥ 1 , with adjusted *p*-value ≤ 0.05 . (*) indicate genes that do not fit the adjusted *p*-value ≤ 0.05 criterion.

Benzaldehyde metabolism generates dead-end products

In the RNA-seq data, we observed that benzaldehyde upregulates the gene *pobA* and the *pcaIJFHGBDC* cluster (Fig. 7f and Supplementary Data 1) and this intriguing result prompted us to investigate if its metabolism involves these genes. Our initial hypothesis was that

benzaldehyde would be oxidated to benzoate by XAC0129, benzoate would be sequentially hydroxylated by PobA to form 4-hydroxybenzoate and then PCA, followed by the PCA ortho-cleavage pathway.

To test this hypothesis, we performed activity assays and gene knockout studies targeting PobA (hydroxylase) and PcaHG (the first enzyme of the PCA ortho-cleavage pathway). PobA showed no detectable activity towards benzoate ([Supplementary Fig. 5a](#)) and the *X. citri* 306 Δ *pobA* strain exhibited neither the expected accumulation of benzoate or 4HBA nor significant differences in the excretion of intermediate metabolites of benzaldehyde metabolism, compared to the WT strain ([Supplementary Fig. 5b](#)). The Δ *pcaHG* strain accumulated the intermediate metabolite PCA in positive control conditions containing 4-hydroxybenzoate or vanillin, but the same was not observed in presence of benzaldehyde ([Supplementary Fig. 5c-e](#)). Together, these results indicate that benzaldehyde metabolism has no connection with PobA and the PCA ortho-cleavage pathway, despite its role as an inducer of *pobA* and *pca* genes expression. Since benzaldehyde was consumed by *X. citri* 306, and it is a substrate of aryl aldehyde reductases (XAC1484 and XAC3477) and dehydrogenase (XAC0129), we then hypothesized that benzaldehyde metabolism generates benzyl alcohol and benzoate as dead-end products, which was confirmed by quantitative analysis using HPLC ([Supplementary Fig. 5f](#)).

Transcriptome response to syringaldehyde reveals the presence of a complete pathway for S-type lignin monomers catabolism

In nature, S-type lignin-related aromatics catabolism can follow three different pathways (I,II,III), categorized according to the generated intermediates: gallate (I), 2-pyrone-4,6-dicarboxylate (PDC) (II), or 4-carboxy-2-hydroxy-6-methoxy-6-oxohexa-2,4-dienoate (III). In *X. citri* 306, we found genes (XAC0882, XAC0878, XAC4155, XAC4156, and XAC4157) that were upregulated in the presence of syringaldehyde ([Fig. 8a](#)) and are homologous to *desV*, *desB*, *ligU*, *ligJ*, and *ligK* involved in the catabolism of syringaldehyde via gallate in *Sphingobium* sp. SYK-6^{56,60} ([Fig. 8b](#), [Supplementary Table 7](#)).

The formation of gallate as an intermediate relies on a two-step *O*-demethylation process, catalyzed by tetrahydrofolate-dependent *O*-demethylases in *Sphingobium* sp. SYK-6⁶¹. However, no homologous to known tetrahydrofolate-dependent *O*-demethylases was detected within the genome of *X. citri* 306. Conversely, the two gene pairs XAC0362-63 and XAC0310-11, homologous to ROs-type *O*-demethylases, were upregulated in presence of syringaldehyde

(Fig. 8a), indicating a possible role in the *O*-demethylation of syringate (SYR) and 3-O-methylgallate (3OMG).

The growth tests conducted with the knockout strains of genes XAC0310-11 (*KO310-11*) and XAC0362-63 (*KO362-63*) in a minimal medium containing syringaldehyde revealed a higher accumulation of syringate by the *KO362-63* strain, but not by *KO310-11* (Fig. 8c). This observation suggests a significant role of XAC0362-63 in the first *O*-demethylation step of S-type lignin monomers catabolism. Whole-cell activity tests corroborated this finding, showing the production of 3OMG from syringate by *E. coli* cells co-expressing XAC0362-63, and confirmed the XAC0362-63 role in the second *O*-demethylation step that produces gallate from 3OMG (Fig. 8 d,e).

After the *O*-demethylation steps, gallate (GA) might be converted to 4-oxalomesaconate (OMA) by a gallate dioxygenase, which is likely encoded by the genes XAC0878 (*ligB*, β -chain) and XAC0879 (*ligA*, α -chain) according to their overexpression induced by syringaldehyde and homology to the respective N-terminal and C-terminal domains of gallate dioxygenases from *P. putida* KT2440 (GalA)⁶² and *Sphingobium* SYK-6 (DesB)⁶³ (Fig. 8b and Supplementary Fig. 6). Next, OMA is likely converted up to pyruvate and oxaloacetate by enzymes encoded by the XAC4155-56-57 genes, which are homologous to *ligK-ligU-ligJ* from *Sphingobium* sp. SYK-6⁵⁶ and were upregulated in the presence of syringaldehyde (Fig. 8 a,b).

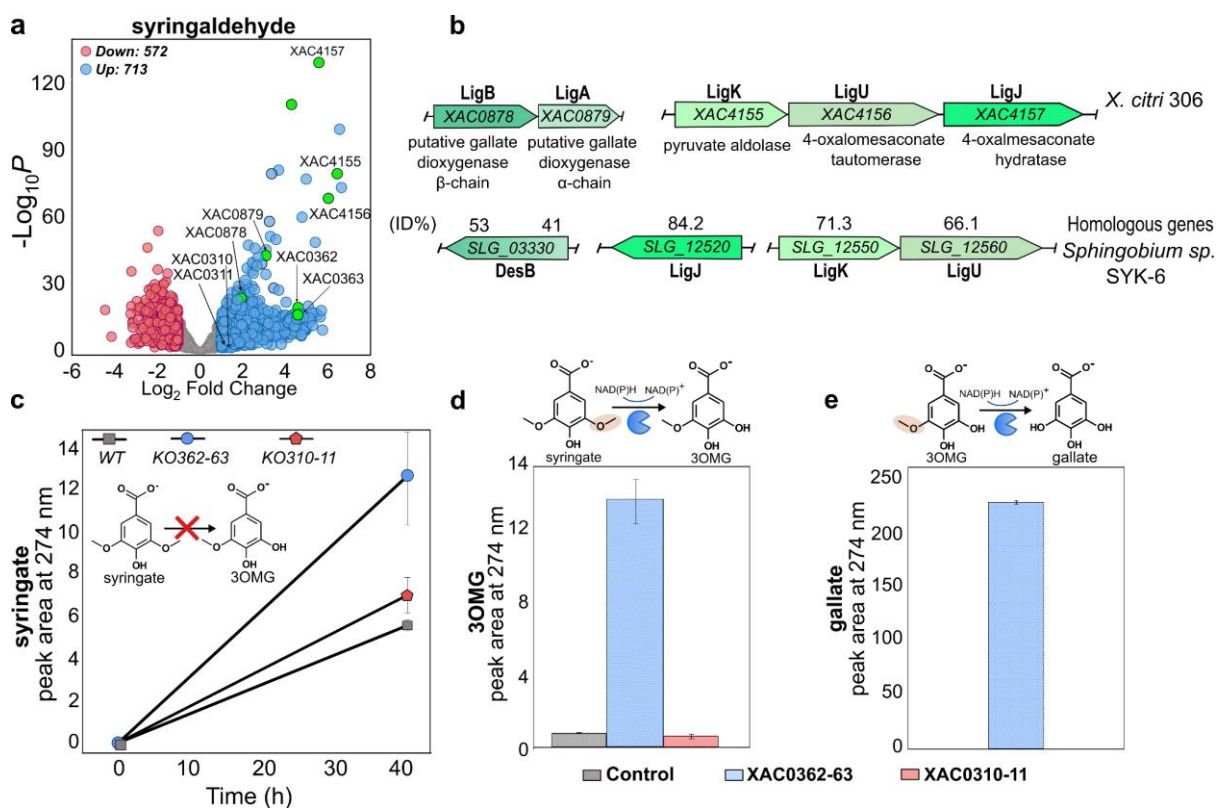


Fig. 8 Identification of enzymes involved in the metabolism of S-type lignin compounds. **a)** Volcano plot of RNA-seq data highlighting in green the genes upregulated related to the syringate catabolism. **b)** Schematic representation of the genomic context of XAC4155-57 gene cluster and XAC078-79 in *X. citri* 306, showing the amino acid sequence identity with homologous enzymes of the gallate fission pathway from *Sphingobium* sp. SYK-6⁵⁶. **c)** HPLC analysis demonstrating prominent syringate accumulation only by the KO362-63 strain. **d-e)** Syringate and 3-*O*-methylgallate (3OMG) whole-cell activity assays with *E. coli* BL21(DE3)- Δ slyD-pRARE2 cells expressing XAC0310-11 (red bars), XAC0362-63 (blue bars), and the negative control (empty vector – gray bars). Data are shown as mean \pm SD of three biological replicates.

Discussion

This study provides a comprehensive investigation of the metabolism of aromatic compounds in a model plant pathogen, *X. citri* 306, uncovering complete pathways for the catabolism of the three main lignin precursors (Fig. 9). So far, only the microbial catabolism of coniferyl alcohol has been reported¹⁹ and our study demonstrates the existence of catabolic pathways for the other two monolignols, *p*-coumaryl and sinapyl alcohols, adding new pieces in the puzzle of microbial metabolic pathways for lignin-related aromatic compounds.

According to our data, the catabolism of monolignols in *X. citri* 306 starts with their uptake, probably facilitated by the putative outer membrane transporter MolK (XAC0352), which belongs to a family (COG4313) that has been implicated in the uptake of hydrophobic molecules⁶⁴. Next, the monolignols are dehydrogenated to aldehydes and subsequently to hydroxycinnamic acids (HCAs) mainly by two NAD⁺-dependent enzymes: the aryl alcohol dehydrogenase MolA (XAC0353) followed by the aryl aldehyde dehydrogenase MolB (XAC0354).

The HCAs are then converted into their respective hydroxybenzylic aldehyde derivatives via the CoA-dependent non- β -oxidation pathway encoded by the gene cluster (XAC0881-84). In contrast to *X. campestris*, which uptakes and metabolizes HCAs³², *X. citri* 306 apparently metabolizes only HCAs produced intracellularly, likely due to the lack of transporters for their uptake. This adaptation correlates with studies showing that HCAs are commonly found in citrus fruits in their conjugated forms (ester- or glycoside-bond), with minimal concentrations in their free form^{65,66}. Part of the HCAs produced intracellularly is excreted by the cell, implying that HCAs deacetylation is probably a metabolic bottleneck to the flux towards the central carbon metabolism in *X. citri* 306.

Next, the H-G-S hydroxybenzyl aldehydes are converted to hydroxybenzoic acids by aryl aldehyde dehydrogenases, including XAC0882 and XAC0129. Then, 4HBA (H-type

subunit) undergoes hydroxylation by the PobA enzyme (XAC0356) while vanillate (G-subunit) is *O*-demethylated by an RO-type-oxygenase-reductase system VanAB (XAC0362-63), both steps resulting in the formation of PCA. Then, the PCA ring is ortho-cleaved and converted into β -ketoadipate by PcaHGBCD enzymes (XAC0367-68-69-71-70). Finally, PcaIJF enzymes (XAC0364-65-66) complete the conversions steps towards the tricarboxylic acid cycle (Fig. 9).

On another branch, syringate (S-type subunit) is *O*-demethylated to 3OMG and then to gallate by VanAB (XAC0362-63). Gallate is likely converted to 4-oxalomesaconate (OMA) by the action of LigAB (XAC0878-79) and follows a pathway up to pyruvate and oxaloacetate by enzymes encoded by the XAC4155-56-57 genes, which are homologous to *ligK-ligU-ligJ* from *Sphingobium* sp. SYK-6⁵⁶ (Fig. 8b). Alternatively, it has been reported that 3OMG can be ring-cleaved to form 4-carboxy-2-hydroxy-6-methoxy-6-oxohexa-2,4-dienoate (CHMOD) by the LigAB system from *Sphingobium* SKY-6⁶⁷. CHMOD is then transformed into 2-pyrone-4,6-dicarboxylic acid (PDC) spontaneously⁶⁸. However, further studies are needed to determine if the LigAB from *X. citri* 306 can convert 3OMG into CHMOD, resulting in the production of PDC.

Besides monolignols, *X. citri* 306 also uptakes their hydroxycinnamic and hydroxybenzilic aldehyde derivatives, driving them to two possible metabolic fates in rates that vary according to their concentration. At low micromolar levels, aryl aldehydes preferentially follow the catabolic pathways, but, as their concentration increases, part of them take reductive metabolic detours that leads to the production and excretion of aryl-alcohols. Two NADPH-dependent enzymes (XAC1484 and XAC3477) contribute to aryl aldehydes reduction and the putative XAC1482-83-85 efflux system probably facilitates the excretion of the aryl alcohols to the extracellular medium. XAC1484 belongs to the Short-chain dehydrogenases/reductases (SDR) family whereas XAC3477 belongs to the aldo/keto reductase (AKR) family, contrasting to the medium-chain dehydrogenase/reductase (MDR) family members reported to play a role in aryl aldehyde reduction in other bacterial species^{47,69–71}. XAC1484 and XAC3477 share less than 41% sequence identity to functionally characterized enzymes, according to BLAST searches at the Swiss-Prot database⁷², thus representing novel aryl aldehyde reductases with potential biotechnological applications in the production of value-added aryl alcohols^{73–75}.

In short, the metabolism of aromatic compounds in *X. citri* 306 converges to the central carbon metabolism but displays “escape routes” (efflux transporters), to cope with the overflow of aryl acids eventually produced during aromatics catabolism, and reductive pathways, which likely function as a complementary strategy to rapidly detoxify aryl aldehydes. The presence of the phenolic group seems to be mandatory for the complete catabolism of these compounds,

since benzaldehyde, a non-phenolic aromatic, was converted to dead-end products (benzyl alcohol and benzoate) (Fig. 9). Other adaptative responses to the presence of these aromatics in the model plant pathogen include the upregulation of genes associated with chemotaxis and flagellar assembly, which might benefit *X. citri* fitness during host colonization. In *X. campestris*, the catabolism of plant-derived phenolic compounds has been shown to be important for virulence^{30,31} and the induction of monolignols biosynthesis and plant cell wall lignification has been proved to be important host defense mechanisms^{46,76,77}. Thus, the capacity of *X. citri* 306 to sequester and metabolize lignin biosynthesis precursors during infection may locally disrupt plant defense mechanisms in the benefit of the bacterium, a feature shared with other plant pathogens, as suggested by the conservation of the *molRKAB* operon in other *Xanthomonas* phytopathogenic species and of the *molA* and *molB* genes in phytopathogens from other genera (Supplementary Fig. 7).

Therefore, this study substantially enriches the current understanding on the metabolism and transcriptional responses of a model plant pathogen to lignin-related compounds, providing fundamental knowledge to support not only the development of new approaches for plant diseases control but also for the improvement of microbial strains aiming the upcycling of aromatic-rich agroindustrial side-streams.

Fig. 9 Schematic representation of the monolignols bioconversion pathways in *X. citri* 306, including proposed transporters for their assimilation and an efflux system for the secretion of aromatic compounds. Black arrows represent the funneling pathways; red arrows indicate the aromatic ring fission pathways, and dark yellow arrows symbolize the reductive pathways. Enzymes identified by RNA-seq and biochemically and or genetically validated are highlighted in green, while candidates proposed through RNA-seq analyses, bioinformatics, and literature analyses are shown in orange. OM = outer membrane; IM = inner membrane; 4HBZ = 4-hydroxybenzaldehyde; 3OMG = 3-*O*-methylgallate; OMA = 4-oxalomesaconate.

Material and Methods

Bacterial strains and culture conditions

X. citri 306 was grown at 30 °C, 200 rpm, in LBON medium (10 g L⁻¹ bacto peptone and 5 g L⁻¹ yeast extract), minimal medium XVM2m (20 mmol L⁻¹ NaCl, 10 mmol L⁻¹ (NH₄)₂SO₄, 1 mmol L⁻¹ CaCl₂, 0.01 mmol L⁻¹ FeSO₄·7 H₂O, 5 mmol L⁻¹ MgSO₄, 0.16 mmol L⁻¹ KH₂PO₄, 0.32 mmol L⁻¹ K₂HPO₄, 0.03% m V⁻¹ casamino acids, pH 6.7) supplemented with different carbon sources, or NYG medium (5 g L⁻¹ peptone, 3 g L⁻¹ yeast extract and 20 g L⁻¹ glycerol) as better described in the next sections. *Escherichia coli* DH5αTM was used for DNA cloning, and *E. coli* BL21(DE3)-ΔslyD-pRARE2 or *E. coli* SHuffle® T7 Express lysY (New England Biolabs) were used for heterologous protein expression. *E. coli* strains were grown at 37 °C, 200 rpm in LB medium (10 g L⁻¹ bacto peptone, 5 g L⁻¹ yeast extract, 10 g L⁻¹ NaCl), 2xYT (16 g L⁻¹ peptone, 10 g L⁻¹ yeast extract and 5 g L⁻¹ NaCl), or in M9 minimal medium (6.78 g L⁻¹ Na₂HPO₄; 3 g L⁻¹ KH₂PO₄; 0.5 g L⁻¹ NaCl; 1 g L⁻¹ NH₄Cl; 2 mmol L⁻¹ MgSO₄; 100 μmol L⁻¹ CaCl₂). Bacterial growth was determined by measuring optical density at 600 nm (OD₆₀₀).

Preparation of lignin-derived complex samples

Low-molecular-mass lignin (LM lignin)

LM lignin was extracted from sugarcane bagasse by an alkaline process⁴⁸. Briefly, sugarcane bagasse was mixed with a 1.5% (m V⁻¹) NaOH solution at 1:20 (m V⁻¹) solid–liquid ratio, heated at 130 °C, for 1 h, under a stirring paddle of approximately 100 rpm, in a 25 L iron–carbon coated stainless steel reactor. Then, centrifugation at 1,100 rpm was performed to separate the cellulosic pulp (solid) and the liquor containing lignin (soluble fraction). The black liquor rich in lignin was collected, precipitated with sulfuric acid at pH 2, separated by filtration using a polypropylene filter in a Buchner funnel, and the liquid fraction was stored at -20 °C. The soluble LM lignin concentration was determined by UV spectroscopy at 280 nm, on 2 M NaOH solution, pH 12, using the following Eq. (1)⁷⁸.

$$\text{Aromatics (g L}^{-1}\text{)} = 4.187 \times 10^{-2} * (\text{Abs}_{280 \text{ nm}}) - (3.279 \times 10^{-4}) * \text{dilution} \quad (1)$$

Next, it was diluted in XVM2m, filter-sterilized, and the aromatics final concentration in the medium was adjusted to 1 g L⁻¹ for *X. citri* 306 cultivations.

Light Bio-oil

The Light Bio-oil (rich in aromatic monomers) was obtained from the hydrothermal depolymerization of sugarcane bagasse lignin, as described by Menezes et al.⁴⁹. Briefly, alkaline lignin from sugarcane bagasse⁷⁹ was depolymerized in a high-pressure autoclave reactor, in batch mode (500 mL, model 4575A, Parr Instrument Company, up to 344 bar, 500 °C), at 350 °C, for 90 min, 1:50 (m V⁻¹) solid-liquid ratio, in an inert atmosphere (N₂), and an agitation of 500 rpm. At the end of the reaction, the liquid fraction was filtered, and the Light Bio-oil was extracted using liquid-liquid extraction with ethyl acetate (Vetec PA ACS) (1:1, three times).

Light Bio-oil was solubilized on 1 mL 0.5 M NaOH solution, then diluted in minimal medium XVM2m for a final theoretical concentration of 1 g L⁻¹. After adjusting the pH to 6.7, the medium was filter-sterilized, and the concentration of aromatic compounds in the medium was estimated by 280 nm absorbance, at pH 12, according to the Eq. (1)⁷⁸. Then, this medium was diluted in XVM2m to adjust the aromatics concentration to 0.3 g L⁻¹.

***X. citri* 306 cultivations**

For growth curves analysis, *X. citri* 306 was cultured in LBON medium, 100 µg mL⁻¹ ampicillin, overnight at 30 °C and 200 rpm. Then, the harvested cells were washed once with XVM2m and inoculated for an initial OD₆₀₀ of 0.05 in XVM2m supplemented with chemical standards of aromatic compounds, complex mixtures of lignin-derived compounds, or glucose ([Supplementary Table 8](#)). Except for 4-hydroxybenzoate, LM lignin, Light Bio-oil, the other conditions were carried out in the presence of 5 mmol L⁻¹ glucose as an additional carbon source. The growth was monitored in 96 well plates incubated in a SpectraMax® M3 Multi-Mode Microplate Reader (Molecular Devices) or in an Infinite 200 PRO plate reader (Tecan), for 48 h, at 30 °C in biological triplicate. Specific growth rates (µ) were obtained using the Package Growth rates as described by Petzoldt⁸⁰.

For HPLC analysis, *X. citri* 306 was cultivated in 125 mL flasks containing 15 mL of medium as detailed in [Supplementary Table 9](#). The culture medium was sampled by removing 2 mL before and after specific post-inoculation time points ([Supplementary Table 9](#)), centrifuged at 6.500 rpm for 5 min to pellet the cells, and the supernatants were stored at -20 °C until the analysis.

Analysis of lignin-derived complex samples by Gas Chromatography coupled to Mass Spectrometry (GC-MS)

X. citri 306 was grown in 10 mL of LBON and prepared as described in the previous section. The OD₆₀₀ was adjusted to 0.1 in 15 mL of XVM2m supplemented with 1 g L⁻¹ LM lignin or 0.3 g L⁻¹ Light Bio-oil and the cells were incubated at 30 °C, 200 rpm. The culture medium was sampled by removing 2 mL in 0- and 30-hours post-inoculation (hpi), centrifuged at 6,500 rpm for 5 min, and the supernatants were utilized for further analyses. The experiment was conducted with four biological replicates.

For GC-MS analysis, 50 µL of 200 µg/mL trans-3-hydroxycinnamic acid (internal standard) was added to 1 mL of the samples, and the pH was adjusted to 2 with 5% HCl solution. Next, 500 µL of ethyl acetate was added, and the samples were vortexed for 5 min and centrifuged at 14,000 rpm, 10 min, at 5 °C. The organic fraction was separated, and the extraction process was repeated twice with 500 µL of ethyl acetate. The total organic extract was completely dried in a concentrator, and the extract was derivatized with 100 µL of pyridine and 100 µL of N,O Bis(trimethylsilyl)trifluoroacetamide (BSTFA) at 80 °C for 60 min. An 1 µL aliquot of derivatized sample was analyzed in a Gas chromatograph coupled to quadrupole mass spectrometry QP2010 Ultra (Shimadzu).

The chromatographic analysis was performed on a DB-35MS chromatographic column, 60 m x 0.25 mm x 0.25 µm (Agilent), with injector temperatures and ionization source at 250 °C. Helium was used as a carrier gas at a flow rate of 1 mL min⁻¹ in a run of 65 min under the following temperature program: 2 min of isotherm at 100 °C, followed by a ramp of 5 °C min⁻¹ until achieving 180 °C, remaining in this temperature for 10 min; followed by a ramp of 5 °C min⁻¹ to 340 °C, remaining for 5 min. The spectra were acquired in scan mode in 50-500 m/z mass range. Data were processed in GCMSsolution software, version 2.71 (Shimadzu), using NIST 2011 and in-house libraries.

High-Performance Liquid Chromatography (HPLC)

Aromatic acids in *X. citri* 306 cultures growth on NYG medium were extracted as described by Chen et al.³¹, with modifications. Briefly, 200 µL of ethyl acetate was added in 200 µL of sample with pH previously adjusted to 3.5. Then, the samples were vortexed for 15 min and centrifuged at 14,000 rpm, 10 min, at 5 °C. The organic fraction was separated, and the extraction process was repeat with more 200 µL of ethyl acetate. The total organic extract was dry in a concentrator for 30 min and dissolved in 200 µL 30% methanol. The aromatics resulting from growth in minimal medium (XVM2m or M9) were not extracted; the samples

were only diluted when necessary. For quantitative analysis the samples were diluted in methanol 1:1 (v/v). The peaks were quantified by measuring the area under the curve against their calibration standards prepared in methanol 50%.

All samples were homogenized by vortexing, filtered through a Millex 0.22 μm syringe filter and analyzed using an Ultimate 3000 high-efficiency liquid chromatograph (Thermo Scientific), equipped with a UV-visible detector set at 274 nm, or an Infinity II (Agilent) equipped with a DAD detector at 274, 260 e 340 nm, and an analytical column Acclaim C18 3 μm , 4.6 x 150 mm, at 30 °C. The eluent was composed by 0.3% acetic acid in water (A) and methanol (B), with an analysis flow of 0.5 mL min⁻¹ in the following gradient: T=0 to 40 min: 15% (B) and 85% (A) to 30% (B) and 70% (A); T= 40 to 55 min: 30% (B) and 70% (A); T= 55 to 55.1 min: 30% (B) and 70% (A) to 15% (B) and 85% (A); T= 55.1 at 60 min 15% (B) and 85% (A).

***X. citri* 306 genome mining**

For the initial prediction of the metabolic pathways related to the catabolism of aromatic compounds in *X. citri* 306, we used the BLASTp tool⁸¹ to search in the genome of this bacterium for proteins homologous to those deposited in the manually curated eLignin Database¹⁹. As a complementary tool, we also use the metabolic maps available in the KEGG database⁸² for this strain as a reference.

RNA extraction and sequencing

X. citri 306 cells grown at 30 °C, 200 rpm on 40 mL XVM2m minimal medium supplemented with different carbon sources ([Supplementary Table 3](#)) were collected at the middle exponential phase ($\text{OD}_{600} = 0.05\text{-}0.1$) from four biological replicates. Total RNA was extracted using the TRIzolTM/chloroform method⁸³. Samples were treated with RNase-free DNaseI (Invitrogen) and RNaseOUT (Invitrogen) for 30 min, at 37 °C, and purified with the RNeasy Mini Kit (Qiagen), following manufacturer's recommendations. RNA samples concentration was determined using Nanodrop 1000 (Thermo Scientific), and their integrity was evaluated in an Agilent 2100 Bioanalyzer (Agilent Technologies). The rRNA was depleted with Ribo-Zero Plus rRNA Depletion Kit (Illumina Inc.). Subsequently, these samples were used to synthesize cDNA libraries with TruSeq Stranded Total RNA kit (Illumina Inc.), according to the manufacturer's recommendations. The final RNA-Seq libraries were quantified via qPCR using the QIAseq Library Quant Assay kit (Qiagen), and library quality was verified using an Agilent 2100 Bioanalyzer (Agilent Technologies). Samples were pooled,

and the RNA sequencing was performed on an Illumina HiSeq 2500 platform (NGS facility LNBR-CNPEM, Campinas, Brazil).

RNA-seq data processing and analysis

The raw reads generated in the RNA-seq were filtered to remove adapters, primers and low-quality sequences using the fastp tool⁸⁴. Contaminant rRNA reads sequences were removed using sortmeRNA⁸⁵. High-quality reads were mapped in the *X. citri* 306 genome (GI: 21240769) using the Bowtie2 tool⁸⁶, allowing only two mismatches and unique alignments. Afterward, the Samtools program⁸⁷ was used to process the alignment files, which were inspected using the Integrative Genome viewer program⁸⁸. Reads mapped to the *X. citri* 306 genome were subjected to the featureCounts tool⁸⁹ to estimate the number of reads mapped to each transcript. Then, the data processed were summarized and plotted by multiQC package⁹⁰. Low count transcripts were removed, keeping only those that showed CPM (Counts Per Million) above 0.9, equivalent to a count of 10 to 15 reads per transcript. Differential expression analysis was carried out using edgeR package⁹¹ by pairwise comparisons between *X. citri* 306 grown in XVM2m supplemented with lignin-related aromatics (Supplementary Table 3) and XVM2m(G) containing 5 mmol L⁻¹ glucose (reference medium). Differentially expressed genes were defined using log₂ Fold Change ≥ 1 (upregulated genes) or ≤ -1 (downregulated genes), and a *p*-adjusted ≤ 0.05 as thresholds. Variance analyses were conducted utilizing PCA (Principal Component Analysis) to assess data integrity and comparability. According to this analysis, one non-concordant replica in the coniferyl alcohol and one from 4-hydroxybenzaldehyde condition were excluded, and the differential expression analysis for these conditions was done with only three biological replicates. Functional and pathway enrichment analyses were performed separately to predict the functions of differentially expressed genes. Modular gene co-expression analysis was performed using the CEMiTool package⁵⁰ and pathway enrichment analysis using the enrichKEGG function of the R clusterProfiler package⁹². One thousand simulations were performed for each condition, and the groups were considered significantly enriched when they presented *p* ≤ 0.05 . The Gene ontology (GO) enrichment analysis was performed using the clusterProfiler 3.14.3R/Bioconductor package and the categories were considered enriched based on hypergeometric test, implemented in the enrich function of the package⁹³. Venn diagram plots were generated using the InteractiVenn web-based tool⁹⁴ (<http://www.interactivenn.net/>). *de novo* reconstruction of *molRKAB* operon from RNA-seq data was performed using the Trinity software package⁹⁵ and comparative analysis with the reference genome (AE008923.1) and was

performed using pyGenomeViz package (<https://moshi4.github.io/pyGenomeViz/>), which was also employed for the analysis *molRKAB* operon conservation.

Gene cloning, protein expression and purification

The target open reading frames selected for biochemical assays based on the RNA-seq analysis were amplified by PCR using specific primers and cloned in the expression vector pET28a(+) or pET21b(+) using the In-Fusion® HD kit (Takara Bio) or by using restriction enzymes and DNA ligase following manufacturer's instructions ([Supplementary Table 10](#)). The constructs were confirmed by DNA sequencing and used to transform the *E. coli* strain BL21(DE3)-ΔSlyD-pRARE2 or BL21(DE3)-SHuffle®-lysY (for XAC0353 construct). The transformed cells were cultured in LB medium, 50 μg mL⁻¹ kanamycin or 100 μg mL⁻¹ ampicillin at 37 °C, 200 rpm until OD₆₀₀ ~ 0.8. Then, the protein expression was induced by adding 0.5 mmol L⁻¹ isopropyl-β-d-thiogalactopyranoside (IPTG) to the medium and incubating it at 20 °C for 16 h. Cells were harvested by centrifugation at 6,000 rpm, at 4 °C for 30 min, resuspended on buffer as detailed in the [Supplementary Table 11](#), and disrupted by sonication (pulses of 15 s with intervals of 30 s during 15 min, 30% amplitude). The cell extract was centrifuged at 16,000 rpm, at 4 °C for 30 min. The supernatant was incubated with 1-2 mL of Ni-NTA agarose resin (QIAGEN) for 2 h, at 4 °C, to purify the recombinant proteins by Ni⁺² affinity chromatography, except for XAC0353 and XAC0354 gene products, which were purified using a FPLC system and a Hitrap chelating column (Cytiva). The resin was transferred to a column and the target proteins were eluted as described in the [Supplementary Table 11](#). Fractions were examined by 13% sodium polyacrylamide-dodecyl sulfate gel electrophoresis (SDS-PAGE), stained with Coomassie Blue stain. Fractions containing the target protein were pooled and dialyzed twice as described in the [Supplementary Table 11](#). Proteins were concentrated using Amicon Ultra-15 10,000 MW (Millipore) when necessary. XAC0353 and XAC0354 were further purified by Size Exclusion Chromatography using a HiLoad 16/600 Superdex 200 pg column (Cytiva) ([Supplementary Table 11](#)). Protein concentration was estimated based on the absorbance of protein samples at 280 nm using the extinction coefficient calculated from their amino acid sequences using the ProtParam tool⁹⁶.

Enzyme activity screening

The activity of putative alcohol dehydrogenases was assessed by measuring the dehydrogenation of aryl alcohols in presence of NAD(P)⁺ or the reduction of aryl aldehydes in presence of NAD(P)H ([Supplementary Table 12](#)). Putative aldehyde dehydrogenases were evaluated for their activity over aryl aldehydes in presence of NAD(P)⁺ ([Supplementary Table](#)

12). The enzyme activity assays were performed in 100 μL of 50 mmol L^{-1} HEPES buffer, pH 7.5, containing 0.25 mmol L^{-1} NAD(P)^+ or NAD(P)H and specific substrates at a final concentration of 0.25 mmol L^{-1} (Supplementary Table 13). The PobA activity was assessed in 100 μL of 50 mM Tris/sulfate buffer, pH 7.5, containing 0.5 mmol L^{-1} NADPH, 60 $\mu\text{mol L}^{-1}$ FAD, and 0.5 mmol L^{-1} of either 4-hydroxybenzoate or benzoate. The reaction was initiated by adding 1 $\mu\text{mol L}^{-1}$ of the PobA enzyme. A control reaction was conducted to monitor the absorption decrease at 340 nm in the absence of aryl substrates.

NAD(P)H production or consumption was monitored by absorbance at 340 nm every 1 min for 30 or 60 min, at 30 $^{\circ}\text{C}$, using a 96-well plate on the Infinite® 200 PRO plate reader (TECAN). All assays were performed in triplicate, and a blank without enzyme was used as a negative control.

Due to non-additive effects caused by substrates and or products in the resultant absorbance value at 340 nm, dehydrogenases enzyme activity was defined here as $A = \frac{\alpha}{[E]}$, where α (arbitrary unit) is the slope of the linear portion of the time-course curves (initial absorbance rate) and $[E]$ is the enzyme concentration in the reaction ($\mu\text{mol L}^{-1}$).

Enzyme specific activity assay

The specific activity of XAC0353 and XAC0354 was accessed in 1 mL of 36 mmol L^{-1} HEPES buffer, pH 7.5, containing 0.25 mmol L^{-1} substrate, and 2 mmol L^{-1} NAD^+ , at 30 $^{\circ}\text{C}$ for 10 min. In reactions involving coniferyl alcohol, cinnamyl alcohol, *p*-coumaryl alcohol, and sinapyl alcohol, XAC0353 was used at a concentration of 0.1 $\mu\text{mol L}^{-1}$. For the remaining aryl alcohols listed in Supplementary Table 13, a concentration of 1.2 $\mu\text{mol L}^{-1}$ of XAC0353 was utilized. XAC0354 at 0.08 $\mu\text{mol L}^{-1}$ was used in reactions containing aryl aldehydes (see Supplementary Table 13). Reactions were stopped by heat at 100 $^{\circ}\text{C}$, 2 min, 650 rpm, and centrifuged (14,000 rpm, 4 $^{\circ}\text{C}$, 10 min). All assays were done in triplicate.

The supernatant was used to measure the NADH production by HPLC as described by Sporty et al.⁹⁷, with modifications. The samples were diluted, when necessary, homogenized, filtered through a Millex 0.22 μm syringe filter, and analyzed using a 1260 Infinity II high-performance liquid chromatograph (Agilent), equipped with DAD detector and a sampler refrigerated at 10 $^{\circ}\text{C}$. The analytical column was an Acclaim C18 3 μm , 4.6 x 150 mm, at 25 $^{\circ}\text{C}$, with flow 0.8 mL min^{-1} for 30 min. The UV absorption was monitored at 340 nm and 260 nm to NADH and NAD^+ , respectively. The eluent was composed of 50 mmol L^{-1} ammonium acetate (A) and 100% acetonitrile (B), in the proportion 97:3 (%), respectively, in an isocratic

method for 12 min. The calibration curve was prepared with NADH standard dissolved in 36 mmol L⁻¹ of HEPES buffer in a linear range from 0.0011 to 0.088 g L⁻¹, divided into 2 curves for better adjustment.

Whole-cell activity assays for dehydrogenases

Whole-cell assays were performed as described by García-Hidalgo et al.⁴⁷. In short, *E. coli* BL21(DE3)-ΔSlyD-pRARE2 transformed with expression vectors containing the genes XAC0882 or XAC0129 (or with the empty vector, used as a negative-control) were cultured in 5 mL of LB medium, overnight, at 30 °C and 200 rpm. Harvested cells were washed with M9 minimal medium and inoculated for an initial OD₆₀₀ of 0.8 in M9 medium containing 50 μg mL⁻¹ kanamycin or 100 μg mL⁻¹ ampicillin, 0.5 mmol L⁻¹ IPTG, 56 mmol L⁻¹ glucose, and 5 mmol L⁻¹ aryl aldehydes (Supplementary Table 13). The cells were incubated at 30 °C and 200 rpm for 20 h. Samples were taken at the final time, centrifuged at 6.500 rpm for 5 min to pellet the cells, and the supernatants were analysed as described in the HPLC section. All assays were done in biological duplicate.

Whole-cell activity assays for *O*-demethylases

O-demethylase whole-cell assays were performed as described by Lanfranchi et al.⁹⁸ with modifications. Briefly, the pairs XAC0310-XAC0311 and XAC0362-XAC0363 were coexpressed in *E. coli* BL21(DE3)-ΔslyD-pRARE2 in 50 mL of 2xYT medium supplemented with 1 mmol L⁻¹ L-cysteine, 0.1 mg mL⁻¹ FeCl₃, and 0.1 mg mL⁻¹ FeSO₄ during induction with 0.5 mmol L⁻¹ IPTG. After expression, the cells were centrifuged (2975 g, 4 °C, 10 min), washed with 20 mL of 50 mmol L⁻¹ Tris-HCl buffer, pH 7.5, and resuspended in 100 mmol L⁻¹ Tris-HCl, pH 7.5. The amount of buffer used for resuspension was adjusted to ensure that the OD₆₀₀ (~ 8.8) for the control cells (empty vector) and for the cells expressing the enzymes were the same, providing a theoretically equal number of cells across all samples. Reactions were prepared in 100 mmol L⁻¹ Tris-HCl, pH 7.5, containing 0.1 mmol L⁻¹ syringate, vanillate or 3OMG (final DMSO concentration 1%), 5 mmol L⁻¹ DTT, 0.1 mmol L⁻¹ FeSO₄, and 20% (v/v) of the cell stock in 2 mL Eppendorf tubes. The reactions were incubated at 30 °C, 150 rpm, for 21 h. After incubation, the reactions were stopped by heating at 90 °C for 10 min, then centrifuged (14.000 rpm, 23 °C, 10 min). The supernatants were frozen at -20 °C before being analyzed as described in the HPLC section. Stock solutions of substrates were prepared in 10 mmol L⁻¹ DMSO and stored at 4 °C. All assays were done in biological triplicate.

Gene deletion

Deletion mutants were generated using established methods⁹⁹ with some modifications. In brief, DNA fragments upstream and downstream of the target gene (~ 1 Kbp) were amplified by PCR from *X. citri* 306 genomic DNA using the primers listed in [Supplementary Tables 14 and 15](#). The PCR fragments were cloned into the pJET1.2/blunt vector using The CloneJET PCR Cloning Kit (Thermo Scientific). Next, they were digested with specific restriction enzymes for sequential cloning into the suicide vector (pNTPS138) using T4 DNA ligase (Thermo Scientific). Alternatively, the PCR fragments were directly cloned into the suicide vector pNTPS138 using the commercial In-Fusion® HD kit (Takara Bio) or NEBuilder® HiFi DNA Assembly Master Mix kit (New England Biolabs). The final constructions (~ 2 kbp) were confirmed by DNA sequencing.

The resultants recombinants plasmids were introduced into *X. citri* 306 by electroporation, and the colonies obtained were inoculated simultaneously on plates containing LBON, 100 µg mL⁻¹ ampicillin and 100 µg mL⁻¹ kanamycin or LBON, 100 µg mL⁻¹ ampicillin, 100 µg mL⁻¹ kanamycin and 5% sucrose to select colonies that suffered the first homologous recombination event. Positive colonies (sucrose-sensitive and kanamycin-resistant) were selected and grown in LBON, 100 µg mL⁻¹ ampicillin to allow the occurrence of the second recombination event. The cultures were diluted 1 x 10³ times in LBON and plated in LBON, 100 µg mL⁻¹ ampicillin and 5% sucrose. The resulting colonies were inoculated simultaneously on plates containing LBON, 100 µg mL⁻¹ ampicillin, 100 µg mL⁻¹ kanamycin and 5% sucrose or LBON, 100 µg mL⁻¹ ampicillin and 5% sucrose. Sucrose-resistant and kanamycin-sensitive colonies were selected, and the deletions were confirmed by PCR and DNA-sequencing.

Characterization of mutant strains

The growth curve assay of the *pobA* mutant strain was conducted in a minimal medium XVM2m or XVM2m(G) supplemented with 4-hydroxybenzoate and benzaldehyde. The cultures were performed in sealed 96-well plates and incubated in a SPARK® Multimode Microplate Reader (Tecan) at 30 °C with agitation. Each condition was replicated in at least three wells.

For HPLC analysis, all mutant strains generated in this work were cultivated in 125 mL flasks containing 15 mL of medium as detailed in [Supplementary Table 16](#). The culture medium was sampled by removing 2 mL before and after specific post-inoculation time points ([Supplementary Table 16](#)), centrifuged at 6.500 rpm for 5 min to pellet the cells, and the supernatants were stored at -20 °C until the analysis.

Acknowledgments

We are grateful to Renan H. S. Fernandes, for the preparation of LM lignin samples, to Mariany da Silva Costa and Matheus Coutinho Gazolla, for their contribution on HPLC data collection, to Angelica Luana Carrilo Barra for her contribution on XAC0129 purification and to Mario Tyago Murakami for reviewing the manuscript and offering valuable feedback. This research used facilities of the Brazilian Biorenewables National Laboratory (LNBR), part of the Brazilian Center for Research in Energy and Materials (CNPEM), a private non-profit organization under the supervision of the Brazilian Ministry for Science, Technology, and Innovations (MCTI). The High-Performance Sequencing (SEQ) and Biophysics of Macromolecules (BFM) staff is acknowledged for their assistance during the experiments (Proposal 27931 and 20230475). We also thank the Molecular Chemistry Laboratory-LNBR for all support on HPLC and GC–MS analyses. This research was supported by the São Paulo Research Foundation - FAPESP (grants 2019/06921-7 to P.O.G and scholarship 2019/08590-8 to D.B.M., 2022/01070-1 to A.J.V.C.B and 2021/07139-0 to A.R.L), Coordination for the Improvement of Higher Education Personnel – CAPES and by the Brazilian Scientific and Technological Development Council – CNPq.

Authors contributions

D.B.M performed *X. citri* 306 genome mining, *X. citri* 306 cultivations, RNA extraction, *molA* and *molB* gene knockout, $\Delta molA$, $\Delta molB$, $\Delta pobA$ and $\Delta pcaHG$ strains characterization, dehydrogenases cloning, activity screening, and whole cell assays. A.J.V.C.B performed XAC3477 purification, MolA and MolB cloning, expression, purification, and enzyme assays. A.R.L performed XAC0129 cloning and ROs-type *O*-demethylases cloning, whole cell assays, gene knockout and strains characterization. L.D.W designed and performed HPLC assays. D.B.M, D.A.A.P, J.M.J, G.F.P and P.O.G designed and performed RNA-seq assays and analyses. J.A.A designed and performed GC-MS assays. J.M.J and G.F.P performed XAC0351-54 operon conservation analysis. F.M.K performed *pobA* and *pcaHG* gene knockout. F.F.M, and G.J.M.R designed and performed lignin depolymerization reactions. P.O.G coordinated the work. D.B.M and P.O.G analyzed the results and wrote the manuscript. A.J.V.C.B, A.R.L., L.D.W and J.M.J contributed to the writing of the methods section and all authors revised the manuscript.

Data availability

RNA-seq data were deposited in the Gene Expression Omnibus database under accession number GSE252662. Source data are provided with this paper. Additional data that support the findings of this study are available from the corresponding authors on reasonable request.

References

1. Janusz, G. et al. Lignin degradation: microorganisms, enzymes involved, genomes analysis and evolution. *FEMS Microbiol Rev* 41, 941–962 (2017).
2. Khokhani, D. et al. Discovery of Plant Phenolic Compounds That Act as Type III Secretion System Inhibitors or Inducers of the Fire Blight Pathogen, *Erwinia amylovora*. *Appl Environ Microbiol* 79, 5424–5436 (2013).
3. BHATTACHARYA, A., SOOD, P. & CITOVSKEY, V. The roles of plant phenolics in defence and communication during *Agrobacterium* and *Rhizobium* infection. *Mol Plant Pathol* 11, 705–719 (2010).
4. Siqueira, J. O., Nair, M. G., Hammerschmidt, R., Safir, G. R. & Putnam, A. R. Significance of phenolic compounds in plant-soil-microbial systems. *CRC Crit Rev Plant Sci* 10, 63–121 (1991).
5. Cao, B., Nagarajan, K. & Loh, K.-C. Biodegradation of aromatic compounds: current status and opportunities for biomolecular approaches. *Appl Microbiol Biotechnol* 85, 207–228 (2009).
6. Fuchs, G., Boll, M. & Heider, J. Microbial degradation of aromatic compounds — from one strategy to four. *Nat Rev Microbiol* 9, 803–816 (2011).
7. Brink, D. P., Ravi, K., Lidén, G. & Gorwa-Grauslund, M. F. Mapping the diversity of microbial lignin catabolism: experiences from the eLignin database. *Appl Microbiol Biotechnol* 103, 3979–4002 (2019).
8. Sullivan, K. P. et al. Mixed plastics waste valorization through tandem chemical oxidation and biological funneling. *Science* (1979) 378, 207–211 (2022).
9. Beckham, G. T., Johnson, C. W., Karp, E. M., Salvachúa, D. & Vardon, D. R. Opportunities and challenges in biological lignin valorization. *Curr Opin Biotechnol* 42, 40–53 (2016).
10. Xu, Z., Lei, P., Zhai, R., Wen, Z. & Jin, M. Recent advances in lignin valorization with bacterial cultures: microorganisms, metabolic pathways, and bio-products. *Biotechnol Biofuels* 12, 32 (2019).
11. Jiménez, J. I., Miñambres, B., García, J. L. & Díaz, E. Genomic analysis of the aromatic catabolic pathways from *Pseudomonas putida* KT2440. *Environ Microbiol* 4, 824–841 (2002).
12. MASAI, E., KATAYAMA, Y. & FUKUDA, M. Genetic and Biochemical Investigations on Bacterial Catabolic Pathways for Lignin-Derived Aromatic Compounds. *Biosci Biotechnol Biochem* 71, 1–15 (2007).
13. Meyer, T. et al. Regulation of Hydroxycinnamic Acid Degradation Drives *Agrobacterium fabrum* Lifestyles. *Molecular Plant-Microbe Interactions*® 31, 814–822 (2018).
14. Broek A. V. & Vanderleyden J. The role of bacterial motility, chemotaxis, and attachment in bacteria plant interactions. *Mol Plant-Microbe Interact* 8, 800–810 (1995).
15. Vieira, P. S. et al. Xyloglucan processing machinery in *Xanthomonas* pathogens and its role in the transcriptional activation of virulence factors. *Nat Commun* 12, 4049 (2021).
16. Déjean, G. et al. The xylan utilization system of the plant pathogen *Xanthomonas campestris* pv *campestris* controls epiphytic life and reveals common features with oligotrophic bacteria and animal gut symbionts. *New Phytologist* 198, 899–915 (2013).

17. Giuseppe, P. O., Bonfim, I. M. & Murakami, M. T. Enzymatic systems for carbohydrate utilization and biosynthesis in *Xanthomonas* and their role in pathogenesis and tissue specificity. *Essays Biochem* 67, 455–470 (2023).
18. F. William & A. Mahadevan. Degradation of aromatic compounds by *Xanthomonas* species / Abbau aromatischer Verbindungen durch *Xanthomonas*-Arten. *Journal of Plant Diseases and Protection* 87, 738–744 (1980).
19. Chen, B. et al. The phytopathogen *Xanthomonas campestris* utilizes the divergently transcribed *pobA* / *pobR* locus for 4-hydroxybenzoic acid recognition and degradation to promote virulence. *Mol Microbiol* 114, 870–886 (2020).
20. Wang, J.-Y. et al. A functional 4-hydroxybenzoate degradation pathway in the phytopathogen *Xanthomonas campestris* is required for full pathogenicity. *Sci Rep* 5, 18456 (2015).
21. Chen, B. et al. The phytopathogen *Xanthomonas campestris* scavenges hydroxycinnamic acids in planta via the *hca* cluster to increase virulence on its host plant. *Phytopathology Research* 4, 12 (2022).
22. Daurelio, L. D. et al. Novel insights into the *Citrus sinensis* nonhost response suggest photosynthesis decline, abiotic stress networks and secondary metabolism modifications. *Functional Plant Biology* 42, 758 (2015).
23. García-Hidalgo, J. et al. Vanillin Production in *Pseudomonas* : Whole-Genome Sequencing of *Pseudomonas* sp. Strain 9.1 and Reannotation of *Pseudomonas putida* CalA as a Vanillin Reductase. *Appl Environ Microbiol* 86, (2020).
24. Rocha, G. J. de M., Nascimento, V. M. & da Silva, V. F. N. Enzymatic Bioremediation of Effluent from Sugarcane Bagasse Soda Delignification Process. *Waste Biomass Valorization* 5, 919–929 (2014).
25. de Menezes, F. F. et al. Exploring the compatibility between hydrothermal depolymerization of alkaline lignin from sugarcane bagasse and metabolization of the aromatics by bacteria. *Int J Biol Macromol* 223, 223–230 (2022).
26. Russo, P. S. T. et al. CEMiTool: a Bioconductor package for performing comprehensive modular co-expression analyses. *BMC Bioinformatics* 19, 56 (2018).
27. Achterholt, S., Priefert, H. & Steinbüchel, A. Purification and Characterization of the Coniferyl Aldehyde Dehydrogenase from *Pseudomonas* sp. Strain HR199 and Molecular Characterization of the Gene. *J Bacteriol* 180, 4387–4391 (1998).
28. Overhage, J., Steinbüchel, A. & Priefert, H. Biotransformation of Eugenol to Ferulic Acid by a Recombinant Strain of *Ralstonia eutropha* H16. *Appl Environ Microbiol* 68, 4315–4321 (2002).
29. SONG, Y.-N., SHIBUYA, M., EBIZUKA, Y. & SANKAWA, U. Identification of Plant Factors Inducing Virulence Gene Expression in *Agrobacterium tumefaciens*. *Chem Pharm Bull (Tokyo)* 39, 2347–2350 (1991).
30. Westphal, A. H., Tischler, D. & van Berkel, W. J. H. Natural diversity of FAD-dependent 4-hydroxybenzoate hydroxylases. *Arch Biochem Biophys* 702, 108820 (2021).
31. Entsch, B. & Van Berkel, W. J. H. Structure and mechanism of para-hydroxybenzoate hydroxylase. *The FASEB Journal* 9, 476–483 (1995).
32. Kamimura, N. et al. Bacterial catabolism of lignin-derived aromatics: New findings in a recent decade: Update on bacterial lignin catabolism. *Environ Microbiol Rep* 9, 679–705 (2017).

33. Venkatesagowda, B. & Dekker, R. F. H. Microbial demethylation of lignin: Evidence of enzymes participating in the removal of methyl/methoxyl groups. *Enzyme Microb Technol* 147, 109780 (2021).
34. Wolf, M. E., Hinchey, D. J., DuBois, J. L., McGeehan, J. E. & Eltis, L. D. Cytochromes P450 in the biocatalytic valorization of lignin. *Curr Opin Biotechnol* 73, 43–50 (2022).
35. Mallinson, S. J. B. et al. A promiscuous cytochrome P450 aromatic O-demethylase for lignin bioconversion. *Nat Commun* 9, 2487 (2018).
36. Kamimura, N. et al. A bacterial aromatic aldehyde dehydrogenase critical for the efficient catabolism of syringaldehyde. *Sci Rep* 7, 44422 (2017).
37. Abe, T., Masai, E., Miyauchi, K., Katayama, Y. & Fukuda, M. A Tetrahydrofolate-Dependent O -Demethylase, LigM, Is Crucial for Catabolism of Vanillate and Syringate in *Sphingomonas paucimobilis* SYK-6. *J Bacteriol* 187, 2030–2037 (2005).
38. Nogales, J., Canales, Á., Jiménez-Barbero, J., García, J. L. & Díaz, E. Molecular Characterization of the Gallate Dioxygenase from *Pseudomonas putida* KT2440. *Journal of Biological Chemistry* 280, 35382–35390 (2005).
39. Uchendu, S. N. et al. Identifying metabolic pathway intermediates that modulate the gallate dioxygenase (DesB) from *Sphingobium* sp. strain SYK-6. *Process Biochemistry* 102, 408–416 (2021).
40. Belchik, S. M., Schaeffer, S. M., Hasenoehl, S. & Xun, L. A β -barrel outer membrane protein facilitates cellular uptake of polychlorophenols in *Cupriavidus necator*. *Biodegradation* 21, 431–439 (2010).
41. Peleg, H., Naim, M., Rouseff, R. L. & Zehavi, U. Distribution of bound and free phenolic acids in oranges (*Citrus sinensis*) and Grapefruits (*Citrus paradisi*). *J Sci Food Agric* 57, 417–426 (1991).
42. Liu, S. et al. Review of phytochemical and nutritional characteristics and food applications of Citrus L. fruits. *Front Nutr* 9, (2022).
43. Masai, E., Katayama, Y., Nishikawa, S. & Fukuda, M. Characterization of *Sphingomonas paucimobilis* SYK-6 genes involved in degradation of lignin-related compounds. *J Ind Microbiol Biotechnol* 23, 364–373 (1999).
44. Sze, I. S. & Dagley, S. Degradation of substituted mandelic acids by meta fission reactions. *J Bacteriol* 169, 3833–3835 (1987).
45. Zhou, P., Khushk, I., Gao, Q. & Bao, J. Tolerance and transcriptional analysis of *Corynebacterium glutamicum* on biotransformation of toxic furaldehyde and benzaldehyde inhibitory compounds. *J Ind Microbiol Biotechnol* 46, 951–963 (2019).
46. Zhou, H., Xu, Z., Cai, C., Li, J. & Jin, M. Deciphering the metabolic distribution of vanillin in *Rhodococcus opacus* during lignin valorization. *Bioresour Technol* 347, 126348 (2022).
47. Weiland, F., Barton, N., Kohlstedt, M., Becker, J. & Wittmann, C. Systems metabolic engineering upgrades *Corynebacterium glutamicum* to high-efficiency cis, cis-muconic acid production from lignin-based aromatics. *Metab Eng* 75, 153–169 (2023).
48. Boutet, E., Lieberherr, D., Tognolli, M., Schneider, M. & Bairoch, A. UniProtKB/Swiss-Prot. in *Plant Bioinformatics* 89–112 (Humana Press, 2007). doi:10.1007/978-1-59745-535-0_4.
49. Chen, Z., Sun, X., Li, Y., Yan, Y. & Yuan, Q. Metabolic engineering of *Escherichia coli* for microbial synthesis of monolignols. *Metab Eng* 39, 102–109 (2017).

50. Lv, Y. et al. Improving bioconversion of eugenol to coniferyl alcohol by in situ eliminating harmful H₂O₂. *Bioresour Technol* 267, 578–583 (2018).
51. Chen, Z. et al. Establishing an Artificial Pathway for De Novo Biosynthesis of Vanillyl Alcohol in *Escherichia coli*. *ACS Synth Biol* 6, 1784–1792 (2017).
52. Abid, M., Khan, M. A. & Wahid, A. Screening and determination of phenolics in relation to resistance against citrus canker. *Pak. J. Phytopathol* 20, 109–116 (2008).
53. Giraldo – González, J. J. et al. Transcriptional changes involved in kumquat (*Fortunella* spp) defense response to *Xanthomonas citri* subsp. *citri* in early stages of infection. *Physiol Mol Plant Pathol* 116, 101729 (2021).
54. Rocha, G. J. de M., Nascimento, V. M., Gonçalves, A. R., Silva, V. F. N. & Martín, C. Influence of mixed sugarcane bagasse samples evaluated by elemental and physical–chemical composition. *Ind Crops Prod* 64, 52–58 (2015).
55. de Menezes, F. F. et al. Alkaline Pretreatment Severity Leads to Different Lignin Applications in Sugar Cane Biorefineries. *ACS Sustain Chem Eng* 5, 5702–5712 (2017).
56. Thomas Petzoldt. Estimation of Growth Rates with Package growthrates. Preprint at (2022).
57. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J Mol Biol* 215, 403–410 (1990).
58. Kanehisa, M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* 28, 27–30 (2000).
59. Chomczynski, P. & Sacchi, N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 162, 156–159 (1987).
60. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890 (2018).
61. Kopylova, E., Noé, L. & Touzet, H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* 28, 3211–3217 (2012).
62. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9, 357–359 (2012).
63. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079 (2009).
64. Thorvaldsdottir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform* 14, 178–192 (2013).
65. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930 (2014).
66. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 32, 3047–3048 (2016).
67. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. `edgeR`: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140 (2010).
68. Wu, T. et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation* 2, 100141 (2021).

69. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS* 16, 284–287 (2012).
70. Heberle, H., Meirelles, G. V., da Silva, F. R., Telles, G. P. & Minghim, R. InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinformatics* 16, 169 (2015).
71. Grabherr, M. G. et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 29, 644–652 (2011).
72. Gasteiger, E. et al. Protein Identification and Analysis Tools on the ExPASy Server. in *The Proteomics Protocols Handbook* 571–607 (Humana Press, 2005). doi:10.1385/1-59259-890-0:571.
73. Sporty, J. L. et al. Single sample extraction protocol for the quantification of NAD and NADH redox states in *Saccharomyces cerevisiae*. *J Sep Sci* 31, 3202–3211 (2008).
74. Lanfranchi, E., Trajković, M., Barta, K., de Vries, J. G. & Janssen, D. B. Exploring the Selective Demethylation of Aryl Methyl Ethers with a *Pseudomonas* Rieske Monooxygenase. *ChemBioChem* 20, 118–125 (2019).
75. Tófoli de Araújo, F. et al. Structural and Physiological Analyses of the Alkanesulphonate-Binding Protein (SsuA) of the Citrus Pathogen *Xanthomonas citri*. *PLoS One* 8, e80083 (2013).

Supplementary information

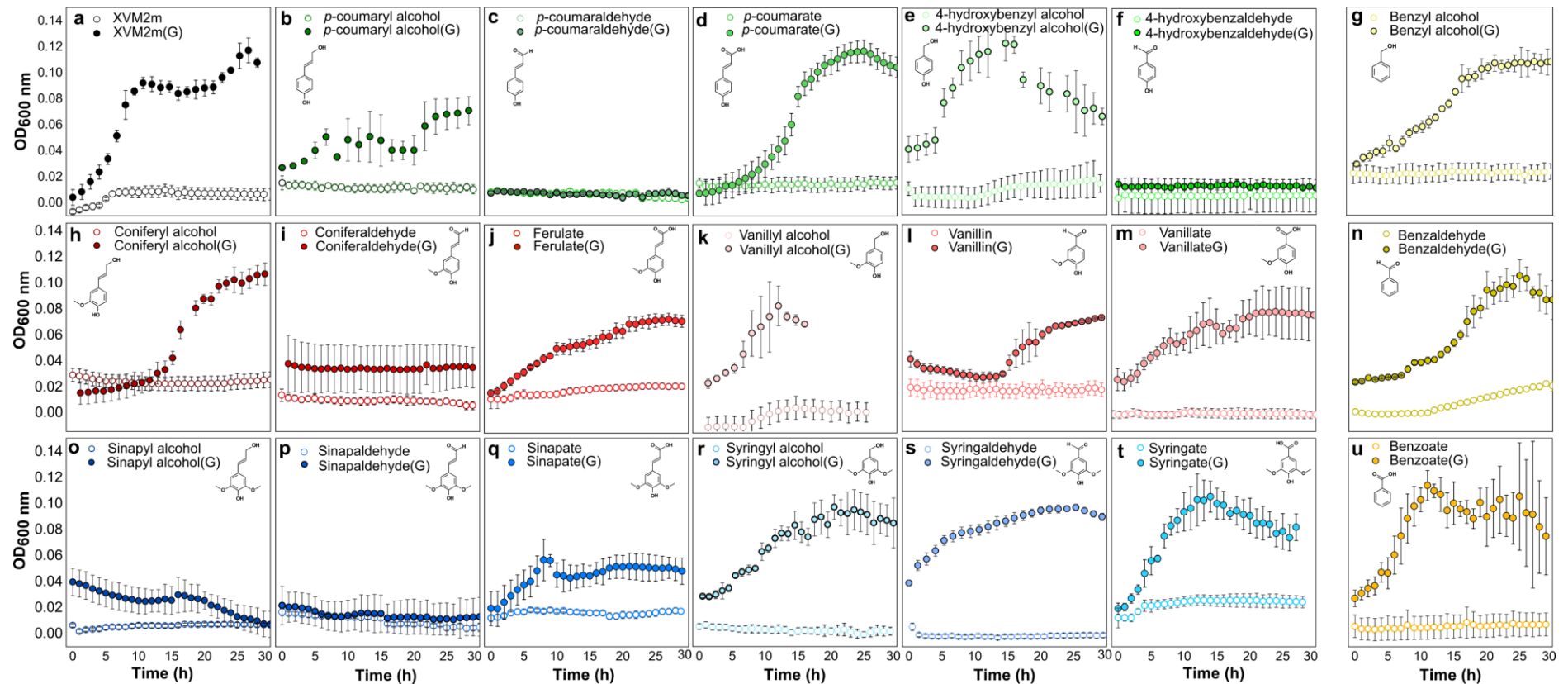
Unraveling aromatic metabolism in a plant pathogen reveals novel pathways for detoxification and metabolic convergence of lignin monomers

Damaris B. Martim^{1,2}, Anna J. V. C. Brilhante^{1,2}, Augusto R. Lima^{1,2}, Lucia D. Wolf², Douglas A. A. Paixão², Joaquim M. Junior², Fernanda M. Kashiwagi², Juliana A. Aricetti², Fabrícia F. de Menezes², Gabriela F. Persinoti², George J. M. Rocha² and Priscila O. Giuseppe^{2,*}

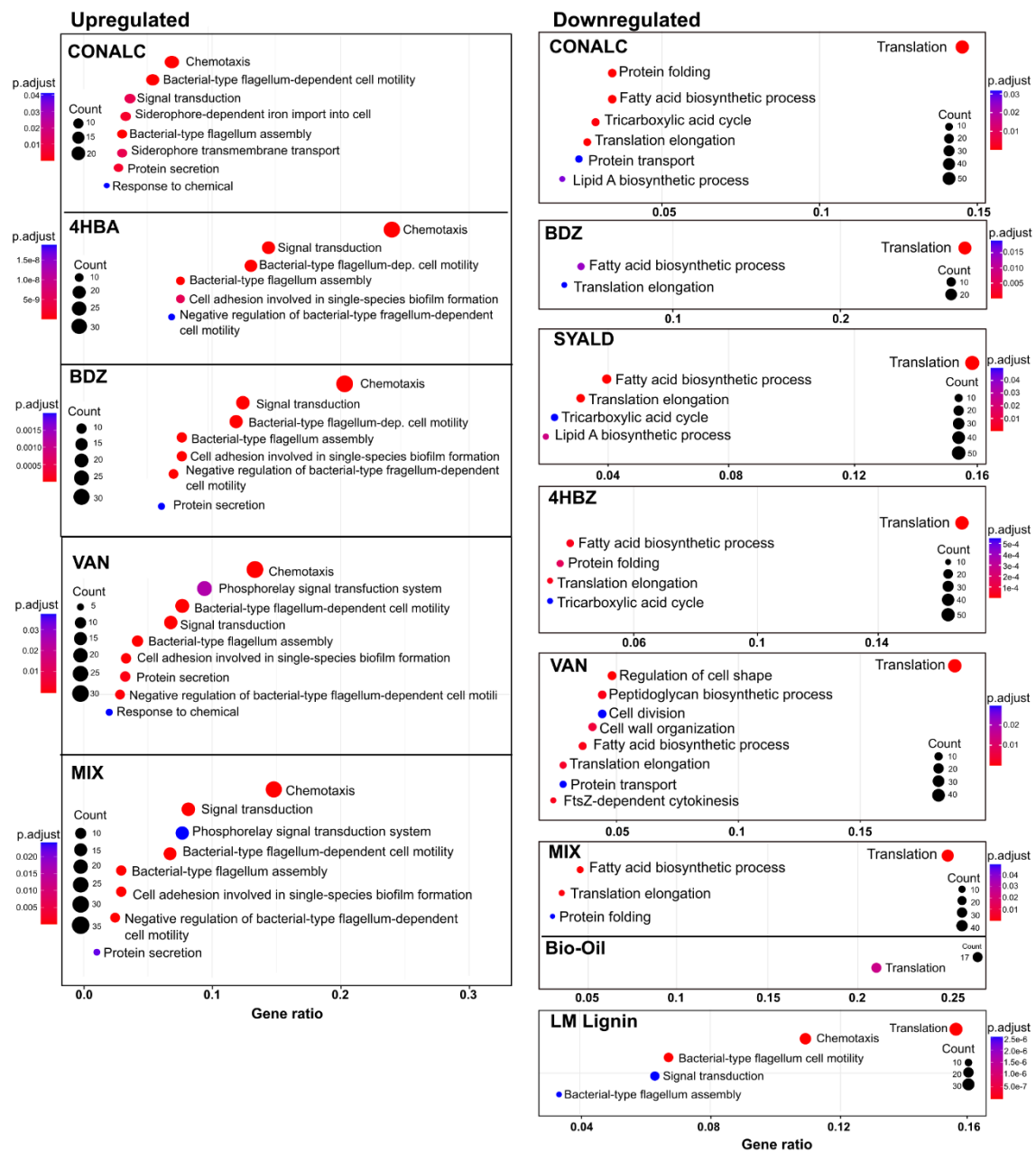
¹ Graduate Program in Genetics and Molecular Biology, Institute of Biology, University of Campinas (UNICAMP), Campinas, São Paulo, Brazil.

² Brazilian Biorenewables National Laboratory (LNBR), Brazilian Center for Research in Energy and Materials (CNPEM), Campinas, São Paulo, Brazil

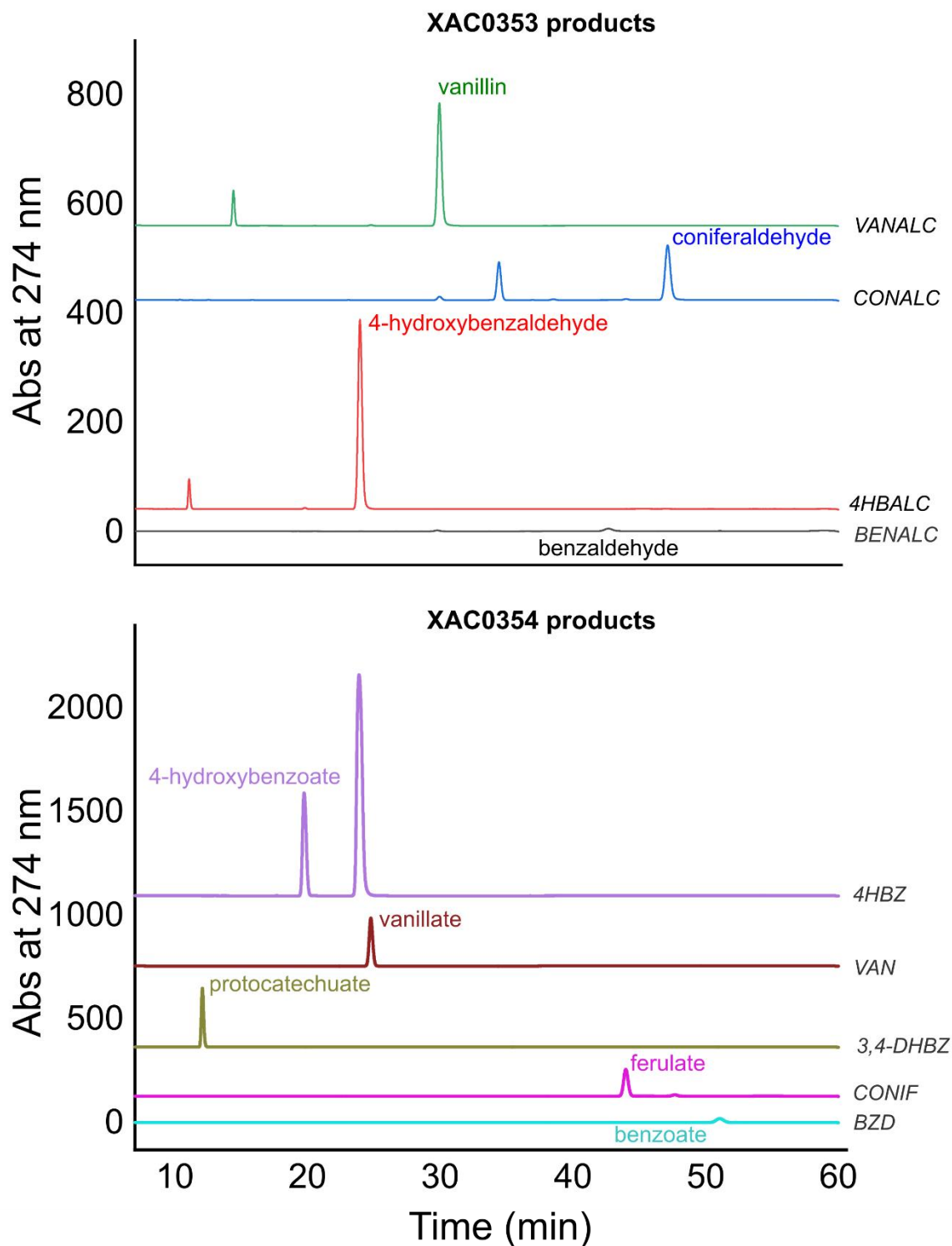
* Corresponding authors. (P.O. de Giuseppe) priscila.giuseppe@lnbr.cnpem.br



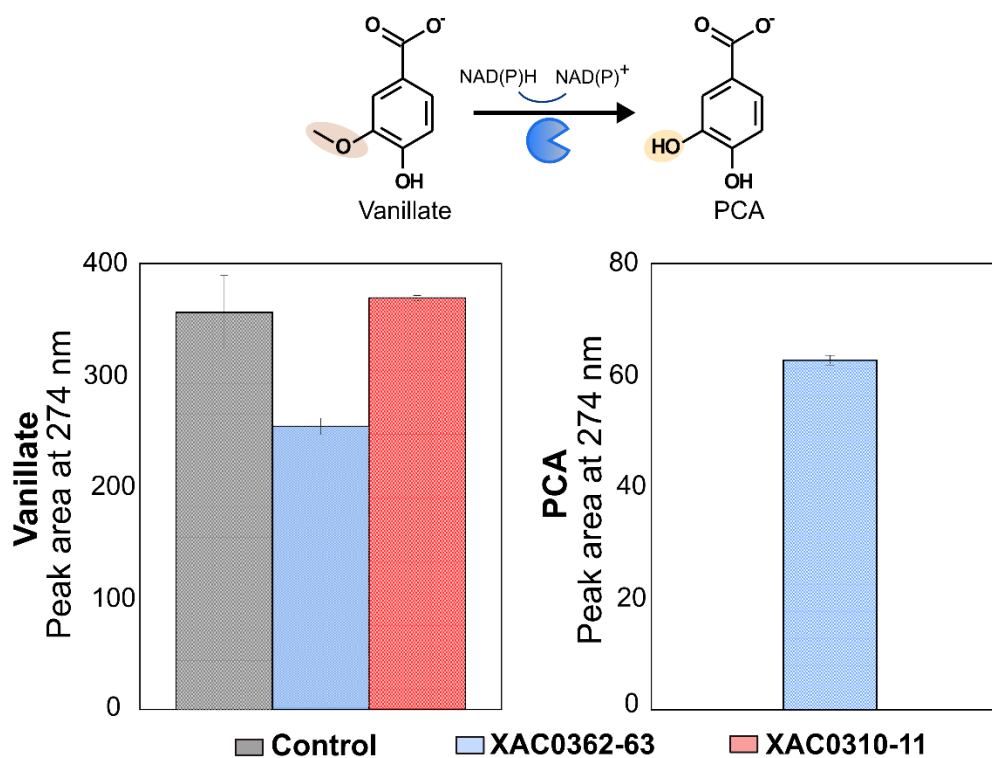
Supplementary Figure 1. Growth profile of *X. citri* 306 in XVM2m or XVM2m(G) supplemented with 5 mmol L⁻¹ lignin-related aromatic compounds. **a)** Growth curves with minimal medium XVM2m without carbon supplementation (empty circle) and XVM2m containing 5 mmol L⁻¹ glucose (XVM2m(G)) (black circles). **b-u)** Growth curves with XVM2m (empty circles) or XVM2m(G) (filled circles), both supplemented with 5 mmol L⁻¹ lignin-related aromatic compounds. Curves are color-coded according to the monomer type: H (green tones), G (red tones), S (blue tones), non-phenolic (yellow tones). All curves are in the same x and y scales. Data are shown as mean \pm SD for at least three biological replicates.



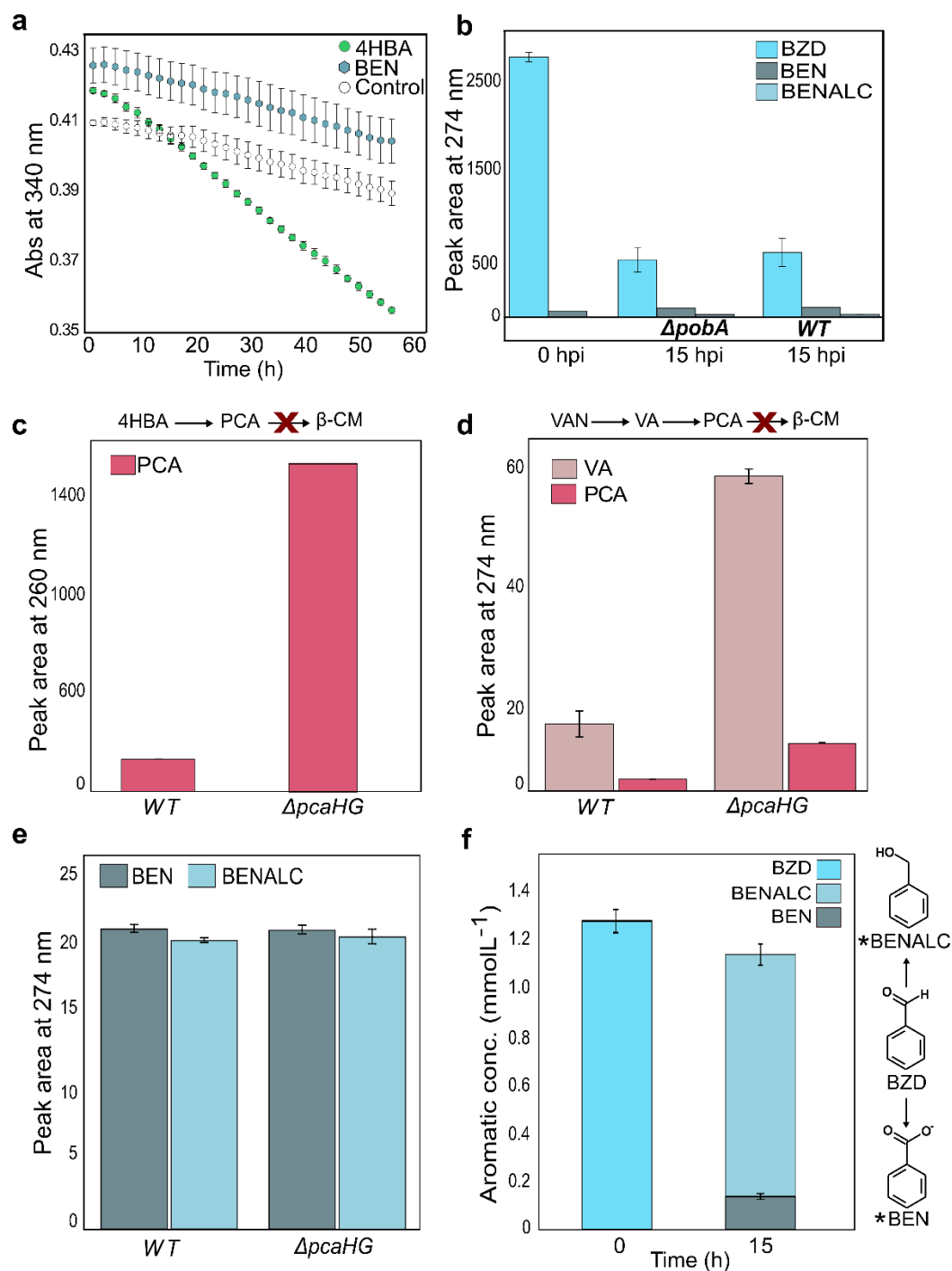
Supplementary Figure 2. Gene ontology (GO) enrichment analysis of biochemical pathways considering the differentially expressed genes (DEGs) in the tested conditions in comparison to the XVM2m(G) reference. Circles size and color represent the counts and adjusted *p*-values, respectively. Gene ratio corresponds to the number of DEGs related to a GO term divided by the total number of annotated DEGs. CONALC = coniferyl alcohol, 4HBA = 4-hydroxybenzoate, BDZ = benzaldehyde, VAN = vanillin, MIX = aldehydes mixture, SYALD = syringaldehyde, 4HBZ = 4-hydroxybenzaldehyde, Bio-Oil = Light lignin bio-oil, LM Lignin = Low-molecular-mass lignin.



Supplementary Figure 3. Representative HPLC chromatograms demonstrating the resulting products from reactions with select substrates of XAC0353 and XAC0354. Abbreviations represent the substrates used in each reaction: VANALC = vanillyl alcohol, CONALC = coniferyl alcohol, 4HBALC = 4-hydroxybenzyl alcohol, BENALC = benzyl alcohol, 4HBZ = 4-hydroxybenzaldehyde, VAN = vanillin, 3,4-DHBZ = 3,4-dihydroxybenzaldehyde, CONIF = coniferaldehyde, BZD = benzaldehyde. The peaks were assigned to the products (colored labels) by comparison with curves of analytical standards.



Supplementary Figure 4. Whole-cell activity assays against vanillate using *E. coli* BL21(DE3)- Δ slyD-pRARE2 cells expressing XAC0310-11 (red bars), XAC0362-63 (blue bars), and the **negative control** (empty vector – gray bars). PCA = protocatechuate. Data are shown as mean \pm SD of at three biological replicates.



Supplementary Figure 5. In vitro activity assay for the PobA enzyme and HPLC analysis of *X. citri* 306 wild-type strain (WT), and the knockouts for the *pobA* ($\Delta poba$) and *pcaHG* ($\Delta pcaHG$) genes. a) PobA enzyme activity against 4-hydroxybenzoate (4HBA) and benzoate (BEN). The control reaction was carried out in the absence of aryl substrates. **b)** HPLC analyses of *X. citri* 306 WT and $\Delta poba$ at 0 and 15 hours pos-inoculation (hpi) in XVM2m(G) plus 5 mmol L⁻¹ benzaldehyde (BZD). **c)** HPLC analysis highlighting the protocatechuete (PCA) accumulation in $\Delta pcaHG$ strain with 15 hours pos-inoculation (hpi) in XVM2m(G) plus 0.5 mmol L⁻¹ 4HBA. **d)** HPLC analysis highlighting the vanillate (VA) and PCA accumulation in $\Delta pcaHG$ strain with 15 hpi in XVM2m(G) plus 0.5 mmol L⁻¹ vanillin (VAN). **e)** HPLC analysis showing no difference in the WT and $\Delta pcaHG$ excreted metabolites with 15 hpi in XVM2m(G) plus 0.5 mmol L⁻¹ BZD. **f)** HPLC quantification of benzyl alcohol (BENALC) and BEN excreted by *X. citri* 306 with 15 hpi in XVM2m(G) plus 1.3 mmol L⁻¹ BZD. Data are shown as mean \pm SD at least three biological replicates. * represents reaction products.

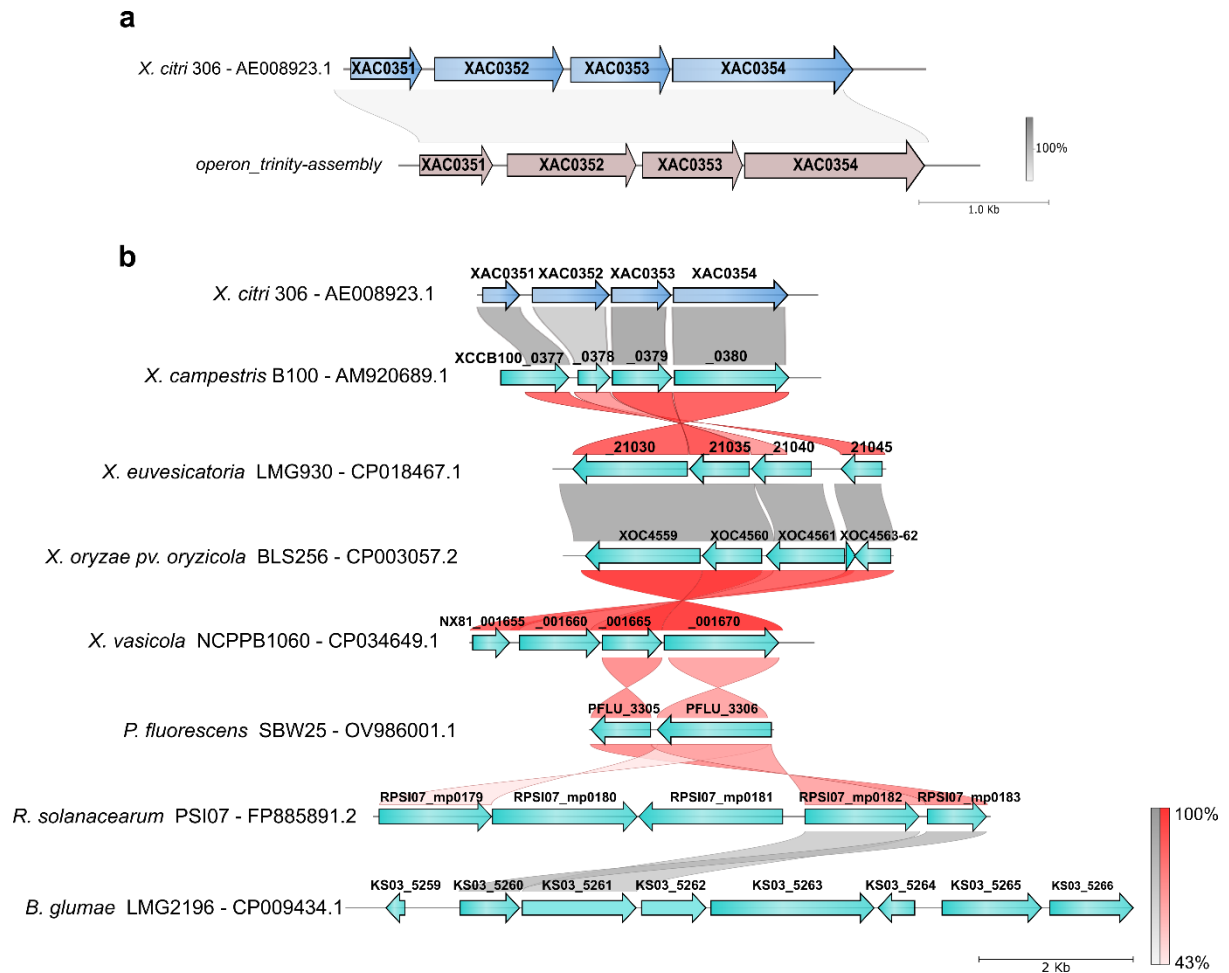
a

XAC0878	MASIIGGIGTS	HVPTIGVAYDKGKQQDPVWKPLFDGYTPVAAWLAAQQADVLVFFYND	HC	60
GalA_P.	MARIIGGLAVS	HTPTIGFAVDHDKQEEAAWAPIFESFEPIRTWLQQRQPDVLFYIFND	HI	60
	**	*****:..*.*.*.*.* *.:**:: . * *.:*: : * : ** : * * *.:*:**		
XAC0878	TTFFFDLYPTFALGVGERFPVADEGAGLRNLPPIRGDVQLQAHIAECLVNDEFDLTVFQD			120
GalA_P.	TSFFFDHYSATLGVDEQYGVADGGNPRDLPPVGGHAALSRHIGQSLMADEFDMSFFRD			120
	*:**** *	:*:***.*:: ***** .. *:***: *.. * . **.:*: : ****:..*:*		
XAC0878	KPIDHGCAAPLPLLWPHVPDWPQTVVPIAINVLQYPLPTARRCYRLGQAVRRAIESYPED			180
GalA_P.	KPLDHGFFSPMSALLPCDESWPVQIVPLQVGVLQLPIPTARRCYKLGQALRRAIESYPED			180
	:* ** :	* * . ** :*: :.* *:*****:*****:*****:*****		
XAC0878	LRVVVVGTTGGLS	HQIHGERTGFNNTEWDMFLDRFQHAPETLTHLTHTDYVRLGGAESV	E	240
GalA_P.	LKVAIVATGGVS	HQVHGERCGFNNPEWDAQFLDLLVNDPQRLTEMTLAEYATLGGMEGA	E	240
	:.:.***:*	*:***** ***** ** :*** : : * : **.:* :*: . *** *..*		
XAC0878	QIMWLMRGALDGPPIRKLHQNYLMTTAMTVVLYEPGEEAAGAPRSAELLART---	AAA		297
GalA_P.	VITWLMRGTLNANVERKHQSYLPSMTGIATLLENRDQALPAPVNERHRQHMQHQLAG			300
	* ** ***:*	.. :. : **.* ** : *.:*: * * :*: * * . . : *		
XAC0878	A-----			298
GalA_P.	AEQLEGTYPYTLERSAKGYRLNKFHRMIEPQWRQFLSEPEALYREAGLSEESDLLRR			360
	*			
XAC0878	-----			298
GalA_P.	RDWRGLIHYGVIFVLEKLGAVLGVSNLDIYAAMRGQSIEDFMKTRNQVRYSVAGKAPN			420

b

XAC0879	-----			0
GalA_P.	MARIIGGLAVSHTPTIGFAVDHDKQEEAAWAPIFESFEPIRTWLQQRQPDVLFYIFNDHI			60
XAC0879	-----			0
GalA_P.	TSFFFDHYSATLGVDEQYGVADGGNPRDLPPVGGHAALSRHIGQSLMADEFDMSFFRD			120
XAC0879	-----			0
GalA_P.	KPLDHGFFSPMSALLPCDESWPVQIVPLQVGVLQLPIPTARRCYKLGQALRRAIESYPED			180
XAC0879	-----			0
GalA_P.	LKVAIVATGGVSHQVHGERCGFNNPEWDAQFLDLLVNDPQRLTEMTLAEYATLGGMEGAE			240
XAC0879	-----			7
GalA_P.	VITWLMRGTLNANVERKHQSYLPSMTGIATLLENRDQALPAPVNERHRQHMQHQLAG		MNPQTHG	300
			*: * *	
XAC0879	MEQIGGTYLFDLRTSNRALRLNRFFWHMIRAPWRDRFLQDAEVLQMQLTEQEKALIRA			67
GalA_P.	AEQLEGTYPYTLERSAKGYRLNKFHRMIEPQWRQFLSEPEALYREAGLSEESDLLRR			360
	:	* : * . * :. *****: :*. **:* **.: * . * :*. *.:* . * :*		
XAC0879	RDWLGLVQYGANFFVIEKFARVVMTNLQVYAIMRGESFEDFMQTRRVPHAR	-----		119
GalA_P.	RDWRGLIHYGVIFVLEKLGAVLGVSNLDIYAAMRGQSIEDFMKTRNQVRYSVAG KAPN			420
	***	***:*. *****:*. * : :*:*:** *****:*****:***		

Supplementary Figure 6. LigAB (XAC0878-79) shares similarity respectively with the N- and C-terminal domains of the enzyme gallate dioxygenase (GalA_P.) from *Pseudomonas putida* KT2440. Amino acid sequence alignment between XAC0878 and GalA_P. (a) and between XAC0879 and GalA_P. (b) using the program Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>). Note that XAC0878 aligns with the N-terminal domain from GalA_P. (1-301) whereas XAC0879 aligns to the GalA_P. C-terminal domain (294-412). In red are highlighted the conserved catalytic residues.



Supplementary Figure 7. De novo reconstruction of the transcript containing the *molRKAB* genes based on RNA-seq data and conservation analysis of the *molRKAB* operon. a) De novo reconstruction of a single mRNA containing the XAC0351-52-53-54 genes (operon *molRKAB*) from RNA-seq data using the Trinity software package⁹⁵. b) Conservation analysis of the *molRKAB* operon in *Xanthomonas* species, two plant pathogens from other bacterial genera, and one plant growth-promoting bacteria (*P. fluorescens*). The bars represent the percentage of amino acid sequence identity. The gray color represents genes in the same direction. The red colors represent genes in the opposite direction.

Supplementary Table 1. GC-MS analysis of compounds identified in LM lignin and Light Bio-oil samples after and before cultivation with *X. citri* 306. ^a post inoculation

Conditions	Normalized peak area	Normalized peak area
LM Lignin	0 h ^a	30 h ^a
lactic acid	6.5	6.71 ± 0.31
malonic acid	0.28	0
4-hydroxybenzaldehyde	1.52	0
vanillin	0.55	0
4-hydroxybenzoic acid	0.29	0.58 ± 0.01
vanillic acid	0.31	0.36 ± 0.01
syringaldehyde	0.62	0
<i>p</i> -coumaric acid	0.88	0.91 ± 0.02
acetosyringone	0.7	0.7 ± 0.02
syringic acid	0.4	0.58 ± 0.02
<i>p</i> -coumaric acid	51.38	56.7 ± 2.12
ferulic acid	8.27	7.19 ± 0.38
4-hydroxybenzyl alcohol	0	2 ± 0.32
vanillyl alcohol	0	0.31 ± 0.05
syringyl alcohol	0	0.12 ± 0.02
Light Bio-oil	0 h ^a	30 h ^a
phenol	0.09 ± 0.01	0.1 ± 0.01
4-ethylphenol	0	0.01 ± 0
catechol	5.32 ± 0.66	6.07 ± 0.48
4-methylcatechol	0.62 ± 0.05	0.3 ± 0.02
hydroquinone	0.1 ± 0.01	0.05 ± 0
4-hydroxybenzaldehyde	0.03 ± 0	0
4-ethyl-catechol	0.47 ± 0.02	0.33 ± 0.02
2,6-dimethoxyphenol	0.36 ± 0.01	0.54 ± 0.1
3-methoxy-catechol	1.18 ± 0.14	0.51 ± 0.06
acetophenone	0.1 ± 0.01	0.12 ± 0.01
4-methyl-2,6-dimethoxyphenol	0.14 ± 0	0.19 ± 0.03
4-hydroxybenzoic Acid	0	0.18 ± 0
4'-hydroxy-3'-methoxyacetophenone	0.03 ± 0	0.03 ± 0
homovanillyl Alcohol	0.04 ± 0	0.02 ± 0

Supplementary Table 2. HPLC analysis of lignin-derived aromatic compounds after and before incubation with *X. citri* 306 in XVM2m(G) medium (except for 4-hydroxybenzoate, which was performed in XVM2m). ^a post inoculation. – not analyzed.

Conditions	Peak area at 274 nm (0 h ^a)	Peak area at 274 nm (15-20 h ^a)	Peak area at 274 nm (40 h ^a)
2 mmol L⁻¹ Aryl alcohols			
<i>p</i> -coumaryl alcohol	10176	9657	8987
coniferyl alcohol	162	117	98
sinapyl alcohol	-	-	-
4-hydroxybenzyl alcohol	280	268	266
vanillyl alcohol	1730	1642	1633
syringyl alcohol	336	329	332
benzyl alcohol	77	76	78
50 μmol L⁻¹ Aryl alcohols			
<i>p</i> -coumaryl alcohol	785	27	-
coniferyl alcohol	417	0	-
sinapyl alcohol	20	0	-
4-hydroxybenzyl alcohol	64	61	-
vanillyl alcohol	125	121	-
syringyl alcohol	43	42	-
benzyl alcohol	9	9	-
1 mmol L⁻¹ Aryl aldehydes			
<i>p</i> -coumaraldehyde	2766	116	471
coniferaldehyde	3125	526	56
sinapaldehyde	13237	338	445
4-hydroxybenzaldehyde	5523	4086	1
vanillin	2588	1	1
syringaldehyde	350	98	2
benzaldehyde	1692	0	0
5 mmol L⁻¹ Aryl acids			
<i>p</i> -coumate	200	195	198
ferulate	168	164	161
sinapate	27	26	27
4-hydroxybenzoate	7479	1423	-
vanillate	3733	3692	3640
syringate	198	194	193
benzoate	512	505	505
50 μmol L⁻¹ Aryl acids			
<i>p</i> -coumate	403	353	391
ferulate	279	273	235
sinapate	68	58	62
4-hydroxybenzoate	246	0	0
vanillate	161	116	0
syringate	539	524	517
benzoate	46	35	44

Supplementary Table 3. Conditions used for RNA-seq assays in XVM2m medium. ^a post inoculation

code	description	Aromatics concentration	glucose concentration	sampling time (h) ^a
1. GLUC	glucose (reference)	-	5 mmol L ⁻¹	5
2. COALC	coniferyl alcohol	2 mmol L ⁻¹	5 mmol L ⁻¹	15
3. VAN	vanillin	3 mmol L ⁻¹	5 mmol L ⁻¹	15
4. 4HBZ	4-hydroxybenzaldehyde	1 mmol L ⁻¹	5 mmol L ⁻¹	15
5. BZD	benzaldehyde	3 mmol L ⁻¹	5 mmol L ⁻¹	8
6. 4HBA	4-hydroxybenzoate	5 mmol L ⁻¹	-	8
7. SYALD	syringaldehyde	3 mmol L ⁻¹	5 mmol L ⁻¹	15
8. MIX	mix of aldehydes (BZD, VAN, 4HBZ, and coniferaldehyde)	0.5 mmol L ⁻¹ (each aldehyde)	28 mmol L ⁻¹	15
9. B.Oil	light bio-oil	0.3 g L ⁻¹	5 mmol L ⁻¹	9
10. LMLig	Low-molecular-mass lignin	1 g L ⁻¹	-	5

Supplementary Table 4. Activity screening of putative alcohol dehydrogenases monitored by absorbance at 340 nm, using NAD(P)⁺ or NAD(P)H as co-substrate. Enzyme activity was defined as $A=\alpha/([E])$, where α (arbitrary unit) is the slope of the linear portion of the time-course curves (initial absorbance rate) and $[E]$ is the enzyme concentration in the reaction ($\mu\text{mol L}^{-1}$). The co-substrate used is shown between parentheses.

[illegible]

Supplementary Table 5. Activity screening of putative aldehyde dehydrogenases monitored by absorbance at 340 nm, using NAD(P)⁺ as co-substrate. Enzyme activity was defined as $A=\alpha/([E])$, where α (arbitrary unit) is the slope of the linear portion of the time-course curves (initial absorbance rate) and $[E]$ is the enzyme concentration in the reaction ($\mu\text{mol L}^{-1}$).

Aldehyde oxidation	XAC0354 (0.3 $\mu\text{mol L}^{-1}$)	
	α/E (NAD ⁺)	α/E (NADP ⁺)
<i>p</i> -coumaraldehyde	0.02	0.03
coniferaldehyde	0.04	0
sinapaldehyde	0.02	0
4-hydroxybenzaldehyde	0	0
vanillin	0	0
silyngaldehyde	0	0
cinnamaldehyde	0	0
benzaldehyde	0	0
3,4-dihydroxybenzaldehyde	0	0

Supplementary Table 6. HPLC analysis of whole-cell activity assays for the putative aldehyde dehydrogenases XAC0129 and XAC0882. ^a post inoculation. Positive hits are highlighted in blue. Data are shown as mean \pm SD of duplicates.

Substrates (5 mmol L ⁻¹)	Peak area at 274 nm (0 h ^a)	Peak area at 274 nm (20 h ^a)		Products	Peak area at 274 nm (0 h ^a)	Peak area at 274 nm (20 h ^a)	
		Control (pET28a)	XAC0882			Control (pET28a)	XAC0882
<i>p</i> -coumaraldehyde	2723 \pm 18	2697 \pm 75	2649 \pm 7.3	<i>p</i> -coumarate	0	4.6 \pm 0.02	4.2 \pm 0.9
coniferaldehyde	3516 \pm 3.1	2474 \pm 33	2766 \pm 42	Ferulate	1.3 \pm 0.03	6.5 \pm 0.1	4.3 \pm 0.04
sinapaldehyde	1885 \pm 53	1697 \pm 1.5	1648 \pm 57	Sinapate	0	3.2 \pm 0	2.6 \pm 0.04
4-hydroxybenzaldehyde	8633 \pm 56	8105 \pm 7	5191 \pm 6	4-hydroxybenzoate	0.4 \pm 0.01	8 \pm 0.1	1466 \pm 9
vanillin	3910 \pm 196	3089 \pm 9	0.4 \pm 0.4	vanillate	0	3 \pm 0	2373 \pm 11
syringaldehyde	386 \pm 1	237 \pm 1.3	0	syringate	1.7 \pm 0.06	3.8 \pm 0.03	2156 \pm 69
3,4-dihydroxybenzaldehyde	4898 \pm 11	4349 \pm 33	3750 \pm 314	protocatechuate	0	1.9 \pm 0.02	118 \pm 6.1
cinnamaldehyde	1868 \pm 8	1312 \pm 19	1108 \pm 35	cinnamate	632 \pm 2	188 \pm 8	173 \pm 1
benzaldehyde	3208	9	0	Benzoate	103	101 \pm 3	102 \pm 8.6
Substrates (5 mmol L ⁻¹)	Peak area at 274 nm (0 h ^a)	Control (pET21b)	XAC0129	Products	Peak area at 274 nm (0 h ^a)	Control (pET21b)	XAC0129
<i>p</i> -coumaraldehyde	4804	4395 \pm 62	4054 \pm 124	<i>p</i> -coumarate	0	0	0
coniferaldehyde	6536	5374 \pm 5	5266 \pm 15	Ferulate	5.8	5 \pm 0.2	15 \pm 0.7
sinapaldehyde	2077	2008 \pm 60	2039 \pm 30	Sinapate	0	0.8 \pm 0	0
4-hydroxybenzaldehyde	5317	4966 \pm 103	2887 \pm 402	4-hydroxybenzoate	1.3	12 \pm 0.2	3172 \pm 273
vanillin	4917 \pm 286	2953 \pm 7.6	2174 \pm 19	vanillate	0.6 \pm 0.6	3.3 \pm 0.04	943 \pm 200
syringaldehyde	739	540 \pm 3.3	353 \pm 15	syringate	2.7	6 \pm 0.15	67 \pm 7
3,4-dihydroxybenzaldehyde	4898 \pm 11	4442 \pm 15	4380 \pm 22	protocatechuate	0	0	22 \pm 2
cinnamaldehyde	1868 \pm 8	1296 \pm 8	1127 \pm 3	cinnamate	632 \pm 2	186 \pm 0.6	644 \pm 13
benzaldehyde	3208 \pm 19	0	0	benzoate	103 \pm 0.8	95 \pm 0.9	313 \pm 6.7

Supplementary Table 7. Genome mining analysis using proteins listed in the eLignin Database¹⁹ or found in the BRENDA database²⁹ as baits. ID(%) = amino acid sequence identity of the reference protein and the homolog found in *X. citri* 306. Predicted enzymes are listed in black, regulators in blue, and transporters in green. Gene names include a prefix consisting of the abbreviated source organism name.

Funneling pathway	Gene	Protein	Source Organism	ID (%)	<i>X. citri</i> 306 homolog
Coniferyl alcohol	<i>Pshr199_calA</i>	coniferyl alcohol dehydrogenase	<i>Pseudomonas</i> sp. HR199	60.94%	XAC0353
	<i>Pshr199_calB</i>	coniferaldehyde dehydrogenase	<i>Pseudomonas</i> sp. HR199	37.97%	XAC4238
Ferulic acid	<i>Sph_ferA</i>	feruloyl-coa synthetase	<i>Sphingobium</i> sp. SYK-6	40% 29%	XAC0881 XAC4077
	<i>Sph_ferB</i>	trans-feruloyl-coa hydratase	<i>Sphingobium</i> sp. SYK-6	53.96%	XAC0883
	<i>Sph_ferB2</i>	trans-feruloyl-coa hydratase	<i>Sphingobium</i> sp. SYK-6	-	-
	<i>Str_ech</i>	enoyl-coa hydratase/aldolase	<i>Streptomyces</i> sp. strain V-1	31% 26%	XAC1853 XAC1584
	<i>Pp_fca</i>	ferulic acid hydratase	<i>P. putida</i>	28.12%	XAC1314
	<i>Ab_hcaC</i>	caffeate coenzyme a ligase	<i>Acinetobacter baylyi</i> ADP1	-	-
	<i>Ab_hcaA</i>	enoyl-coa hydratase/lyase	<i>Acinetobacter baylyi</i> ADP1	53.38%	XAC0883
p-coumaric acid	<i>Ab_hcaB</i>	hydroxybenzaldehyde dehydrogenase	<i>Acinetobacter baylyi</i> ADP1	36%	XAC0354
	<i>Pp_pobA</i>	p-hydroxybenzoate 3-monooxygenase	<i>Pseudomonas fluorescens</i>	64%	XAC0356
	<i>Pp_pobR</i>	poba regulatory protein	<i>P. putida</i> WCS358	41%	XAC0355
	<i>Pp_PcaK</i>	4-hydroxybenzoate transporter	<i>P. putida</i>	28%	XAC0349
Benzoic acid	<i>Pp_benA</i>	benzoate 1,2-dioxygenase large subunit	<i>P. putida</i> KT2440	25.13%	XAC1960
	<i>Pp_benB</i>	benzoate 1,2-dioxygenase small subunit	<i>P. putida</i> KT2440	28%	XAC3960
	<i>Pp_benC</i>	benzoate dioxygenase reductase subunit	<i>P. putida</i> PRS2000	33%	XAC4186
	<i>Pp_benD</i>	cis-diol dehydrogenase	<i>P. putida</i> KT2440	34%	XAC1178
	<i>Pp_benK</i>	benzoate transport protein	<i>P. putida</i>	-	-
	<i>Ab_benM</i>	transcriptional activator	<i>Acinetobacter</i> sp. ADP1	37%	XAC0316
	<i>Am_GcoA</i>	p450 aryl-o-demethylase	<i>Amycolatopsis</i> sp. 75iv2 (ATCC 39116)	30%	XAC4066
Guaiacol	<i>Am_GcoB</i>	p450 aryl-o-demethylase	<i>Amycolatopsis</i> sp. 75iv2 (ATCC 39116)	-	-
	<i>Str_vdcBCD</i>	non-oxidative vanillate decarboxylase	<i>Streptomyces</i> sp. D7	-	-
Caffeic acid	<i>Ab_hcaD</i>	acyl coenzyme a (acyl-coa) dehydrogenase	<i>Acinetobacter baylyi</i> ADP1	33%	XAC0265
m-cresol	<i>Cg_nahG</i>	3-hydroxybenzoate 6-hydroxylase	<i>Corynebacterium glutamicum</i>	23%	XAC1667
p-cresol	<i>Pp_pchC</i>	4-methylphenol methylhydroxylase	<i>P. putida</i> NCIMB 9866	-	-
	<i>Pp_pchF</i>	p-cresol methylhydroxylase (pcmh)	<i>P. putida</i> NCIMB 9866	-	-
Vanillin	<i>Sph_ligV</i>	vanillin dehydrogenase	<i>Sphingobium</i> sp. SYK-6	33%	XAC2469
	<i>Sph_ligM</i>	vanillate monooxygenase	<i>Sphingobium</i> sp. SYK-6	-	-
	<i>Pshr199_VanA</i>	vanillate o-demethylase oxygenase	<i>Pseudomonas</i> sp. HR199	44% 26.16%	XAC0363 XAC0311
	<i>Pshr199_VanB</i>	vanillate o-demethylase oxidoreduct.	<i>Pseudomonas</i> sp. HR199	47% 30%	XAC0310 XAC0362

Supplementary Table 7. Continued.

Funneling pathway	Gene	Protein	Source Organism	ID (%)	<i>X. citri</i> 306 homolog
Syringic acid	<i>Sph_desA</i>	syringate o-demethylase	<i>Sphingobium</i> sp. SYK-6	-	-
	<i>Sph_desB</i>	gallate dioxygenase	<i>Sphingobium</i> sp. SYK-6	53%	XAC0878
	<i>Sph_ligJ</i>	4-oxalomesaconate hydratase	<i>Sphingobium</i> sp. SYK-6	72%	XAC4157
	<i>Sph_ligK</i>	hmg aldolase	<i>Sphingobium</i> sp. SYK-6	71%	XAC4155
	<i>Sph_ligU</i>	4-oxalomesaconate tautomerase	<i>Sphingobium</i> sp. SYK-6	67%	XAC4156
	<i>Sph_desZ</i>	3-o-methylgallate 3,4-dioxygenase	<i>S. paucimobilis</i>	-	-
	<i>Sph_ligI</i>	2-pyrone-4,6-dicarboxylate hydrolase	<i>Sphingobium</i> sp. SYK-6	-	-
Resorcinol	<i>Cgl1159</i>	maleylacetate reductase i-ii	<i>C. glutamicum</i> ATCC13032	-	-
	<i>Cgl1160</i>	hydroxyquinol 1,2-dioxygenase i-ii	<i>C. glutamicum</i> ATCC13032	-	-
	<i>Cgl1158</i>	resorcinol 4-hydroxylase	<i>C. glutamicum</i> ATCC13032	-	-
Salicylic acid	<i>Str_sdgA</i>	salicyl-coa synthase / adenylyltransferase	<i>Streptomyces</i> sp. WA46	30%	XAC0881
	<i>Str_sdgA</i>	salicylate	<i>Streptomyces</i> sp. WA46	-	-
	<i>Str_sdgB</i>	adenylyltransferase	<i>Streptomyces</i> sp. WA46	31%	XAC4077
Syringalde-hyde	<i>Psm1_salA</i>	salicyl-coa synthase	<i>Pseudomonas</i> sp. MT1	-	-
	<i>desV</i>	salicylate 1-hydroxylase	<i>Sphingobium</i> sp. SYK-6	66%	XAC0882
		aldehyde dehydrogenase			
Fission pathway	Gene	Protein	Source Organism	ID (%)	<i>X. citri</i> 306 homolog
Gentisate-cleavage	<i>Cg_nagI</i>	gentisate 1,2-dioxygenase	<i>Ralstonia</i> sp. U2	-	-
	<i>Cg_nagK</i>	fumarylpyruvate hydrolase	<i>Ralstonia</i> sp. U2	43%	XAC3307
	<i>Cg_nagL</i>	maleylpyruvate isomerase	<i>Ralstonia</i> sp. U2	47%	XAC3608
β-ketoadipate Catechol branch	<i>Pp_catA</i>	catechol 1,2-dioxygenase	<i>P. putida</i> KT2440	-	-
	<i>Pp_catB</i>	muconate cycloisomerase	<i>P. putida</i> KT2440	30%	XAC3862
	<i>Pp_catC</i>	muconolactone isomerase	<i>P. putida</i> KT2440	-	-
	<i>Pp_catR</i>	transcriptional activator (lysr family)	<i>P. putida</i> KT2440	32%	XAC0905
	<i>Pp_pcaG</i>	protocatechuate 3,4-dioxygenase	<i>P. putida</i> KT2440	41%	XAC0368
β-ketoadipate: Protocatechuate branch	<i>Pp_pcaH</i>	protocatechuate 3,4-dioxygenase	<i>P. putida</i> KT2440	47%	XAC0367
	<i>Pp_pcaB</i>	3-carboxy-cis,cis-muconate cycloisomerase	<i>P. putida</i> KT2440	38%	XAC0369
	<i>Pp_pcaC</i>	4-carboxymuconolactone decarboxylase	<i>Acinetobacter</i> sp. ADP1	46%	XAC0371
	<i>Pp_PcaT</i>	β-ketoadipate transporter	<i>P. putida</i>	41%	XAC4308
	<i>Pae_praA</i>	protocatechuate 2,3-dioxygenase	<i>Paenibacillus</i> sp. JJ-1b	25%	XAC3706
β-ketoadipate pathway	<i>Pp_PcaF</i>	β-ketoadipyl coa thiolase	<i>P. putida</i>	67%	XAC0366
	<i>Pp_PcaI</i>	beta-ketoadipate succinyl-coa transferase	<i>P. putida</i>	49.30%	XAC3578
	<i>Pp_PcaJ</i>	β-ketoadipate succinyl-coa transferase	<i>P. putida</i>	52.71%	XAC3577
	<i>Ab_PcaD</i>	β-ketoadipate enolactone hydrolase i	<i>Acinetobacter</i> sp. ADP1	35.08%	XAC0370
	<i>Cn_xylE</i>	catechol 2,3-dioxygenase	<i>Cupriavidus necator</i> N-1	-	-

Supplementary Table 7. Continued.

Fission pathway	Gene	Protein	Source Organism	ID (%)	<i>X. citri</i> 306 homolog
Protocatechuate 2,3-cleavage	<i>Pae_praB</i>	2-hydroxymuconate-6-semialdehyde dehydrogenase	<i>Paenibacillus</i> sp. JJ-1b	41%	XAC0129
	<i>Pae_praC</i>	4-oxalocrotonate tautomerase	<i>Paenibacillus</i> sp. JJ-1b	-	-
	<i>Pae_praD</i>	2-oxo-3-hexenedioate decarboxylase	<i>Paenibacillus</i> sp. JJ-1b	-	-
Protocatechuate 4,5-cleavage	<i>Sph_ligA</i>	protocatechuate 4,5-dioxygenase	<i>Sphingobium</i> sp. SYK-6	25%	XAC0879
	<i>Sph_ligB</i>	protocatechuate 4,5-dioxygenase	<i>Sphingobium</i> sp. SYK-6	44%	XAC0878
	<i>Sph_ligC</i>	2-hydroxy-4-carboxymuconate semialdehyde dehydrogenase	<i>Sphingobium</i> sp. SYK-6	-	-
Catechol meta-cleavage pathway	<i>Ppmt2_xylG</i>	2-hydroxymuconate-6-semialdehyde hydrolase	<i>P. putida</i> mt-2	-	-
	<i>Cn_bpHI</i>	4-hydroxy-2-oxovalerate aldolase	<i>C. necator</i> N-1	-	-

Supplementary Table 8. *X. citri* 306 growth curves conditions. The growth was performed in XVM2m or in XVM2m with 5 mmol L⁻¹ glucose (G).

Conditions	Concentration	Medium
Aryl alcohols		
<i>p</i> -coumaryl alcohol	5 mmol L ⁻¹	XVM2m(G)
coniferyl alcohol	5 mmol L ⁻¹	XVM2m(G)
sinapyl alcohol	5 mmol L ⁻¹	XVM2m(G)
4-hydroxybenzyl alcohol	5 mmol L ⁻¹	XVM2m(G)
vanillyl alcohol	5 mmol L ⁻¹	XVM2m(G)
syringyl alcohol	5 mmol L ⁻¹	XVM2m(G)
benzyl alcohol	5 mmol L ⁻¹	XVM2m(G)
Aryl aldehydes		
<i>p</i> -coumaraldehyde	5 mmol L ⁻¹	XVM2m(G)
coniferaldehyde	5 mmol L ⁻¹	XVM2m(G)
sinapaldehyde	5 mmol L ⁻¹	XVM2m(G)
4-hydroxybenzaldehyde	5 mmol L ⁻¹	XVM2m(G)
vanillin	5 mmol L ⁻¹	XVM2m(G)
syringaldehyde	5 mmol L ⁻¹	XVM2m(G)
benzaldehyde	5 mmol L ⁻¹	XVM2m(G)
Aryl acids		
<i>p</i> -coumarate	5 mmol L ⁻¹	XVM2m(G)
ferulate	5 mmol L ⁻¹	XVM2m(G)
sinapate	5 mmol L ⁻¹	XVM2m(G)
4-hydroxybenzoate	5 mmol L ⁻¹	XVM2m
vanillate	5 mmol L ⁻¹	XVM2m(G)
syringate	5 mmol L ⁻¹	XVM2m(G)
benzoate	5 mmol L ⁻¹	XVM2m(G)
Lignin-derived samples		
LM Lignin	1 g L ⁻¹	XVM2m
Light Bio-oil	0.3 g L ⁻¹	XVM2m

Supplementary Table 9. Growth conditions for the HPLC analysis of aromatics consumption by *X. citri* 306. ^a post inoculation.

Conditions	Concentration	Medium	Sampling time points (h) ^a
Aryl alcohols			
<i>p</i> -coumaryl alcohol	50 $\mu\text{mol L}^{-1}$	XVM2m(G)	20
coniferyl alcohol	50 $\mu\text{mol L}^{-1}$	XVM2m(G)	20
sinapyl alcohol	50 $\mu\text{mol L}^{-1}$	XVM2m(G)	20
4-hydroxybenzyl alcohol	50 $\mu\text{mol L}^{-1}$	XVM2m(G)	20
vanillyl alcohol	50 $\mu\text{mol L}^{-1}$	XVM2m(G)	20
syringyl alcohol	50 $\mu\text{mol L}^{-1}$	XVM2m(G)	20
benzyl alcohol	50 $\mu\text{mol L}^{-1}$	XVM2m(G)	20
Aryl aldehydes			
<i>p</i> -coumaraldehyde	1 mmol L^{-1}	XVM2m(G)	15 or 40
coniferaldehyde	1 mmol L^{-1}	XVM2m(G)	15 or 40
sinapaldehyde	1 mmol L^{-1}	XVM2m(G)	15 or 40
4-hydroxybenzaldehyde	1 mmol L^{-1}	XVM2m(G)	15 or 40
vanillin	1 mmol L^{-1}	XVM2m(G)	15 or 40
syringaldehyde	1 mmol L^{-1}	XVM2m(G)	15 or 40
benzaldehyde	1 mmol L^{-1}	XVM2m(G)	15 or 40
Aryl acids			
<i>p</i> -coumarate	50 $\mu\text{mol L}^{-1}$	XVM2m(G)/NYG	15 or 40
ferulate	50 $\mu\text{mol L}^{-1}$	XVM2m(G)/NYG	15 or 40
sinapate	50 $\mu\text{mol L}^{-1}$	XVM2m(G)/NYG	15 or 40
4-hydroxybenzoate	50 $\mu\text{mol L}^{-1}$	XVM2m	15 or 40
vanillate	50 $\mu\text{mol L}^{-1}$	XVM2m(G)/NYG	15 or 40
syringate	50 $\mu\text{mol L}^{-1}$	XVM2m(G)/NYG	15 or 40
benzoate	50 $\mu\text{mol L}^{-1}$	XVM2m(G)/NYG	15 or 40
Mix of aldehydes			
coniferaldehyde	0.5 mmol L^{-1} (each aldehyde)	XVM2m(G)	15 or 40
4-hydroxybenzaldehyde			
vanillin			
benzaldehyde			

Supplementary Table 10. Primers and expression vectors used in this work to clone and express putative dehydrogenases, predicted Ros-type *O*-demethylases (XAC0362-63 and XAC0310-11) and a hydroxylase (PobA) from *X. citri* 306. Enzyme restriction sites are underlined. *All forward primers contain a site for the restriction enzyme NdeI that provides the first ATG. The XAC0353 and XAC0354 forward primers also have EcoRI restriction site. The reverse primers include an EcoRI restriction site, except for XAC0129, XAC0310, XAC0353, XAC0354, and XAC0362, which feature the XhoI site.

GENE (PROTEINID)	PRIMERS*	EXPRESSIO N VECTOR
XAC0129 (AAM35021.1)	FW 5'-AAGGAGATATAC <u>CATATG</u> AACGCGCTTGCCTCTGCA RV 5'-GGTGGTGGTGC <u>TCGAGG</u> AAAAACCCCAGCGCCTT	pET21b(+)
XAC0310 (AAM35202.1)	FW 5'-AAGGAGATATAC <u>CATATG</u> AGCCTGCATGAAGTGCG RV 5'-GGTGGTGGTGC <u>TCGAGA</u> AAGATCCAGCACCAACCGCG	pET21b(+)
XAC0311 (AAM35203.1)	FW 5'-CGCGCGGCAGCC <u>CATATG</u> GATGTCGCCGTCCCC RV 5'-GACGGAGCTC <u>GAATTC</u> TTAGCGCAGGAAGAAGTCGCT	pET28a(+)
XAC0353 (AAM35245.1)	FW 5'- <u>GAATTC</u> CATATGAATTTGCAGAACAAAGACTATCGTGGTG RV 5'- <u>CTCGAG</u> TTACAGCGTTGTGGAAGCCAGC	pET28a(+)
XAC0354 (AAM35246.1)	FW 5'- <u>GAATTC</u> CATATGAATGCTCTTACCAGCGC RV 5'- <u>CTCGAG</u> GAACGGATACTGCGGCACTTC	pET21b(+)
XAC0356- POBA- (AAM35248.1)	FW 5'-CGCGCGGCAGCC <u>CATATG</u> CGAACGCAGGTTGCCATCATCG RV 5'-GACGGAGCTC <u>GAATTC</u> TCAATCAGACAGCGCAGCCA	pET28a(+)
XAC0362- VANB- (AAM35254.1)	FW 5'-AAGGAGATATAC <u>CATATG</u> CGTAAAGACACCCAGTGGC RV 5'-GGTGGTGGTGC <u>TCGAGG</u> CCCAGCTCGCTGCGATAGC	pET21b(+)
XAC0363— VANA- (AAM35255.1)	FW 5'-CGCGCGGCAGCC <u>CATATG</u> TCGCAGTCCAAGCCG RV 5'-GACGGAGCTC <u>GAATTC</u> TTAGCCACCATGCGCAAC	pET28a(+)
XAC0580 (AAM35469.1)	FW 5'-CGCGCGGCAGCC <u>CATATG</u> ACGATCGAAAACAAGTCTGTGG RV 5'-GACGGAGCTC <u>GAATTC</u> TCTACGCTTGAGATGTAGGGCG	pET28a(+)
XAC0882 (AAM35770.1)	FW 5'-CGCGCGGCAGCC <u>CATATG</u> GAAGCCTGTCCGAAACGC RV 5'-GACGGAGCTC <u>GAATTC</u> TCTAGCTCAGGTAGAGGGTTCTGG	pET28a(+)
XAC1011 (AAM35894.1)	FW 5'-CGCGCGGCAGCC <u>CATATG</u> AAAGCTGTCGCCCTCACCC RV 5'-GACGGAGCTC <u>GAATTC</u> TTACCAGCCGCCAGCACC	pET28a(+)
XAC1481 (AAM36351.1)	FW 5'-CGCGCGGCAGCC <u>CATATG</u> GGCACGGCATGGAATTATCCC RV 5'-GACGGAGCTC <u>GAATTC</u> TCTATGGGGAAATCTCTCGGG	pET28a(+)
XAC1484 (AAM36354.1)	FW 5'-CGCGCGGCAGCC <u>CATATG</u> AGCAACACCAAGATCGCATTGG RV 5'-GACGGAGCTC <u>GAATTC</u> TTACCATGGCAACACCTCTCCC	pET28a(+)
XAC2896 (AAM37741.1)	FW 5'-CGCGCGGCAGCC <u>CATATG</u> CAAGCCATTCGTCTTCGCC RV 5'-GACGGAGCTC <u>GAATTC</u> TCAGAACTCCAGCGCGATC	pET28a(+)
XAC3477 (AAM38320.1)	FW 5'-CGCGCGGCAGCC <u>CATATG</u> CATCAGCCACGCTCACTGGAC RV 5'-GACGGAGCTC <u>GAATTC</u> TCAGGTGCGCAGCTGCAC	pET28a(+)

Supplementary Table 11. Protein purification conditions. IMAC = ion metal affinity chromatography. SEC = size exclusion chromatography.

<i>Protein</i>	<i>Lysis buffer</i>	<i>IMAC buffer</i>	<i>Dialysis buffer</i>	<i>SEC buffer</i>
<i>XAC0353</i>	20 mmol L ⁻¹ sodium phosphate, pH 7.5, 500 mmol L ⁻¹ NaCl, 5 mmol L ⁻¹ imidazole, 1 mmol L ⁻¹ PMSF, 5 mmol L ⁻¹ benzamidine, and 30 µg MI ⁻¹ DNase	20 mmol L ⁻¹ sodium phosphate, pH 7.5, 500 mmol L ⁻¹ NaCl, 5-500 mmol L ⁻¹ imidazole	20 mmol L ⁻¹ HEPES, pH 7.5, 150 mmol L ⁻¹ KCl	20 mmol L ⁻¹ HEPES, pH 7.5, 150 mmol L ⁻¹ KCl
<i>XAC1011</i>	20 mmol L ⁻¹ HEPES pH 7.4, 300 mmol L ⁻¹ NaCl, 1 mmol L ⁻¹ PMSF, and 30 µg. MI ⁻¹ DNase	20 mmol L ⁻¹ HEPES pH 7.4, 300 mmol L ⁻¹ NaCl, 10-500 mmol L ⁻¹ imidazole	20 mmol L ⁻¹ HEPES pH 7.4, 300 mmol L ⁻¹ NaCl	-
<i>XAC1481</i>	20 mmol L ⁻¹ HEPES pH 7.4, 300 mmol L ⁻¹ NaCl, 1 mmol L ⁻¹ PMSF, and 30 µg. MI ⁻¹ DNase	20 mmol L ⁻¹ Bicine pH 7.8, 300 mmol L ⁻¹ NaCl, 10-500 mmol L ⁻¹ imidazole	20 mmol L ⁻¹ Bicine pH 9, 300 mmol L ⁻¹ NaCl	-
<i>XAC1484</i>	20 mmol L ⁻¹ HEPES pH 7.4, 300 mmol L ⁻¹ NaCl, 1 mmol L ⁻¹ PMSF, and 30 µg. MI ⁻¹ DNase	20 mmol L ⁻¹ HEPES pH 7.4, 300 mmol L ⁻¹ NaCl, 10-500 mmol L ⁻¹ imidazole	20 mmol L ⁻¹ HEPES pH 7.4, 150 mmol L ⁻¹ KCl	-
<i>XAC3477</i>	20 mmol L ⁻¹ HEPES pH 7.4, 300 mmol L ⁻¹ NaCl, 1 mmol L ⁻¹ PMSF, and 30 µg. MI ⁻¹ DNase	20 mmol L ⁻¹ Bicine pH 7.8, 300 mmol L ⁻¹ NaCl, 10-500 mmol L ⁻¹ imidazole	20 mmol L ⁻¹ Bicine pH 7.8, 300 mmol L ⁻¹ NaCl	-
<i>XAC2896</i>	20 mmol L ⁻¹ HEPES pH 7.4, 300 mmol L ⁻¹ NaCl, 1 mmol L ⁻¹ PMSF, and 30 µg. MI ⁻¹ DNase	20 mmol L ⁻¹ HEPES pH 7.4, 300 mmol L ⁻¹ NaCl, 10-500 mmol L ⁻¹ imidazole	20 mmol L ⁻¹ HEPES pH 7.4, 300 mmol L ⁻¹ NaCl	-
<i>XAC0354</i>	20 mmol L ⁻¹ sodium phosphate, pH 7.5, 500 mmol L ⁻¹ NaCl, 5 mmol L ⁻¹ imidazole, 1 mmol L ⁻¹ PMSF, 5 mmol L ⁻¹ benzamidine, and 30 µg MI ⁻¹ DNase	20 mmol L ⁻¹ sodium phosphate, pH 7.5, 500 mmol L ⁻¹ NaCl, 5-500 mmol L ⁻¹ imidazole	20 mmol L ⁻¹ HEPES, pH 7.5, 150 mmol L ⁻¹ KCl	20 mmol L ⁻¹ HEPES, pH 7.5, 150 mmol L ⁻¹ KCl
<i>XAC0129</i>	20 mmol L ⁻¹ sodium phosphate, pH 7.0, 300 mmol L ⁻¹ NaCl, 1 mmol L ⁻¹ PMSF, 1 mmol β-mercaptoethanol, and 30 µg MI ⁻¹ DNase	20 mmol L ⁻¹ sodium phosphate, pH 7.0, 300 mmol L ⁻¹ NaCl, 5-500 mmol L ⁻¹ imidazole	20 mmol L ⁻¹ sodium phosphate, pH 7.0, 50 mmol L ⁻¹ NaCl	-

Supplementary Table 12. Enzyme concentration used in the activity screening.

<i>Protein ID</i>	<i>Protein classification</i>	<i>Concentration</i>
<i>Putative alcohol dehydrogenases</i>		
<i>XAC0353</i>	short chain dehydrogenase	0.3 $\mu\text{mol L}^{-1}$
<i>XAC1011</i>	Oxidoreductase	1 $\mu\text{mol L}^{-1}$
<i>XAC1481</i>	Dehydrogenase	1 $\mu\text{mol L}^{-1}$
<i>XAC1484</i>	short chain dehydrogenase	0.4 $\mu\text{mol L}^{-1}$
<i>XAC3477</i>	aldo-keto reductase	1 $\mu\text{mol L}^{-1}$
<i>XAC2896</i>	alcohol dehydrogenase	1 $\mu\text{mol L}^{-1}$
<i>Putative aldehyde dehydrogenases</i>		
<i>XAC0354</i>	aldehyde dehydrogenase	0.4 $\mu\text{mol L}^{-1}$
<i>XAC0129</i>	aldehyde dehydrogenase	2.5 $\mu\text{mol L}^{-1}$

Supplementary Table 13. List of substrates used for enzyme activity screening and preparation method.

Substrate	Supplier (catalog number)	Stock concentration (mmol L ⁻¹)	Preparation
Aryl alcohols			
<i>p</i> -coumaryl alcohol	TRC (C755420)	10	water at 50 °C
coniferyl alcohol	Sigma-Aldrich (223735)	5	water at 50 °C
sinapyl alcohol	Sigma-Aldrich (404586)	3	4% acetonitrile solution
4-hydroxybenzyl alcohol	Sigma-Aldrich (H20806)	10	water at room temperature
vanillyl alcohol	Sigma-Aldrich (W373702)	10	water at room temperature
syringyl alcohol	Biosynth (FD71202)	2	water at 50 °C
cinnamyl alcohol	Sigma-Aldrich (108197)	10	water at 50 °C
benzyl alcohol	Sigma-Aldrich (W213713)	10	water at room temperature
Aryl aldehydes			
<i>p</i> -coumaraldehyde	Combi-Blocks (QH-0771)	5	alkaline water, pH 10
coniferaldehyde	Sigma-Aldrich (382051)	5	alkaline water, pH 10
sinapaldehyde	ChemScene (CS0016716)	4	water at 50 °C
4-hydroxybenzaldehyde	Sigma-Aldrich (144088)	15	water at room temperature
vanillin	Sigma-Aldrich (V1104)	15	water at room temperature
syringaldehyde	Sigma-Aldrich (S7602)	15	water at 50 °C
3,4-dihydroxybenzaldehyde	Sigma-Aldrich (D108405)	15	water at room temperature
cinnamaldehyde	Sigma-Aldrich (C80687)	10	water at 50 °C
benzaldehyde	Sigma-Aldrich (B1334)	15	water at room temperature

Supplementary Table 14. Primers used for the construction of *pobA* (XAC0356), XAC0310, XAC0311, *VanAB* (XAC0362-63) and *pcaHG* (XAC00367-68) gene deletion cassettes and internal primers for mutant confirmation. The Up_Fw primers contain overlaps regions to the pNTPS138 vector (underline) and the Up_Rv primers contain a downstream homology arm (bold). The Down_Fw and Down_Rv primers contain overlaps to the upstream homology arm (bold) and the pNTPS138 vector (underline), respectively.

Primer	Sequence 5' → 3'
Up-pobA-Fw	<u>GTGGATCCAGATATCCTGCA</u> ACCGACTCGAAAGCGCGTC
Up-pobA-Rv	ACAGCGCAGCTCGCATGCCCTGCTCCTG
Down-pobA-Fw	GGGCATGCGAGCTGCGCTGTCTGATTGATG
Down-pobA-Rv	<u>TGGCGCCAAGCTTCTCTGC</u> ACGATCCGGAACATGTCTGAAC
Internal-pobA-Fw	CAAGCCGAATGCAGGGACTG
Internal-pobA-Rv	ACTCAGGTGGTCTGTGAGGA
Up-310-Fw	<u>GATATCTGGATCCACGAATT</u> CACTGCACCTTCGACGCTG
Up-310-Rv	TTCAAAGATCCAGGCTCATGCGGGTGTC
Down-310-Fw	CATGAGCCTGGATCTTTGAACGCTGTTTCCGC
Down-310-Rv	<u>ACGGCCGAAGCTAGCGAATT</u> CATTTCACCAGCACGAACACC
Internal-310-Fw	GTTTGGTGATCATGAATGCCG
Internal-310-Rv	GTTGACGGCAGGACAAAGGA
Up-311-Fw	<u>GATATCTGGATCCACGAATT</u> CCCACTGGCGGGAGTAAGG
Up-311-Rv	GCAGGAAGAACATGGGGCGGTCTCTGGC
Down-311-Fw	CCGCCCCATGTTCTTCCTGCGCTGACGG
Down-311-Rv	<u>ACGGCCGAAGCTAGCGAATT</u> CTGAAAGGCAAGCGAGGCG
Internal-311-Fw	TCTTCCTCGGCCTTGAGGAT
Internal-311-Rv	GGTGCAGCGGGAACAGATT
Up-362-Fw	<u>GATATCTGGATCCACGAATT</u> CGCCTGGGACCATGAGATCAAG
Up-362-Rv	CCAGCTCGCTCATGAAGGGCTCCAGGCAATG
Down-362-Fw	GCCCTTCATGAGCGAGCTGGGCTGAGTC
Down-362-Rv	<u>GATATCTGGATCCACGAATT</u> CATCGAAGCGTGACGGTC
Internal-362-Fw	CACCCGGATCACACCTTCTAC
Internal-362-Rv	GCCAATCACCGCATGGCAAG
Up-363-Fw	<u>GATATCTGGATCCACGAATT</u> CATCGAAGCGTGACGGTC
Up-363-Rv	GCTCCAGGCACTGCGACATGACCGTGACTC
Down-363-Fw	CATGTCGCAGTGCCCTGGAGCCCTTCATG
Down-363-Rv	<u>ACGGCCGAAGCTAGCGAATT</u> CATAGCCGGTGTGCGATGCTG
Internal-363-Fw	ATCCGCGAGAGCGACGATTT
Internal-363-Rv	GGTTCGCCCACCAACGAATAC
Up-367-68-Fw	<u>GTGGATCCAGATATCCTGCA</u> AACTGGATCTTGCCATTG
Up-367-68-Rv	CGGATCAATAATCGCGCATCTTGAAATC
Down-367-68-Fw	GATGCGCGATTATTGATCCGCATGCCGAC
Down-367-68-Rv	<u>TGGCGCCAAGCTTCTCTGC</u> ACAACGGCAACTCCAGAGTG
Internal-367-68-Fw	ATTCGGGCTGGCCAGCATGT
Internal-367-68-Rv	CGGAGACACCCTGGCGGTTT

Supplementary Table 15. Primers used for the construction of *MolA* (XAC0353) and *MolB* (XAC0354) gene deletion cassettes and internal primers for mutant confirmation. Restriction sites are underlined: GGATCC (BamHI), CATATG (NdeI), AAGCTT (HindIII) e GAATTC (EcoRI). F = forward. R = reverse. Internal primers were used for DNA sequencing.

Primer	Sequence 5'→3'
XAC0353BamHI F1	<u>GGATCC</u> GCCTGACGTTTGTGATGTGGCACG
XAC0353 R1	<u>CATATG</u> CATCCTCCAAAGGAGCCAGCCGT
XAC0353 F2	<u>CATATG</u> CTGGCTTCCACAACGCTGTAATCC
XAC0353HindIII R2	<u>AAGCTT</u> GCTGCGACGCTATCGTTGACGATG
XAC0354BamHI F1	<u>GGATCC</u> ACTACATCTACAACCTTCGCAGCGAG
XAC0354 R1	<u>CATATG</u> CTAGCCTCCTGCTAAAAACGTGTG
XAC0354 F2	<u>CATATG</u> CCGCAGTATCCGTTCTGATCGC
XAC0354EcoRI R2	<u>GAATTC</u> ACGCTGCTGAATTGCTGCATTGGG
Internal-XAC0353 F	ATGCCTATTACCTGACCCAGTTC
Internal-XAC0353 R	AGAACACCACCGATTTCGTCCA
Internal-XAC0354 F	ACGACACCTGCTACCAATACTTCA
Internal-XAC0354 R	AAATGCGTCCCTATAAGCTCCTTG

Supplementary Table 16. *X. citri* 306 mutant strains growth conditions. The growth was performed in XVM2m or in XVM2m with 5 mmol L⁻¹ glucose (G). ^a post inoculation

Gene deleted	Aromatic compound	Concentration	Medium	Sampling time points (h) ^a
<i>ΔpobA</i>	4-hydroxybenzoate	5 mmol L ⁻¹	XVM2m	15
	benzaldehyde	5 mmol L ⁻¹	XVM2m(G)	15
<i>ΔmolA</i>	<i>p</i> -coumaryl alcohol	50 μmol L ⁻¹	XVM2m(G)	15 or 40
	sinapyl alcohol	50 μmol L ⁻¹	XVM2m(G)	15 or 40
	coniferyl alcohol	50 μmol L ⁻¹	XVM2m(G)	15 or 40
<i>ΔmolB</i>	<i>p</i> -coumaraldehyde	50 μmol L ⁻¹	XVM2m(G)	15 or 40
	coniferaldehyde	50 μmol L ⁻¹	XVM2m(G)	15 or 40
	sinapaldehyde	50 μmol L ⁻¹	XVM2m(G)	15 or 40
<i>ΔXAC0362-63</i>	vanillin + syringaldehyde	1.5 mmol L ⁻¹ (each one)	XVM2m(G)	40
<i>ΔXAC0310-11</i>	vanillin + syringaldehyde	1.5 mmol L ⁻¹ (each one)	XVM2m(G)	40
<i>ΔpcaHG</i>	4-hydroxybenzoate	0.5 mmol L ⁻¹	XVM2m	15
	vanillin	0.5 mmol L ⁻¹	XVM2m(G)	15
	benzaldehyde	0.5 mmol L ⁻¹	XVM2m(G)	15

Reference

1. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* **29**, 644–652 (2011).
2. Brink, D. P., Ravi, K., Lidén, G. & Gorwa-Grauslund, M. F. Mapping the diversity of microbial lignin catabolism: experiences from the eLignin database. *Appl Microbiol Biotechnol* **103**, 3979–4002 (2019).
3. Schomburg, I. *et al.* The BRENDA enzyme information system—From a database to an expert system. *J Biotechnol* **261**, 194–206 (2017).

CONSIDERAÇÕES FINAIS

Atualmente a maior parte dos compostos aromáticos utilizados nos mais diversos setores industriais são artificialmente produzidos por síntese química a partir de derivados do petróleo, empregando processos que requerem alta demanda energética e tendem a produzir substâncias intermediárias tóxicas e perigosas. Em vista disso, a lignina, o segundo biopolímero mais abundante na Terra e um dos principais resíduos agroindustriais ricos em carbono, surge como uma matéria-prima renovável atraente na busca por alternativas ao petróleo para a produção de compostos químicos de valor agregado. A conversão de misturas complexas de compostos aromáticos, gerados na etapa de despolimerização da lignina, em produtos alvo específicos por meio da biocatálise microbiana tem sido amplamente estudada nas últimas décadas. Essa abordagem mostra-se promissora para a obtenção de produtos com melhores rendimentos e em maiores concentrações, o que facilita o processo de purificação geralmente necessário para se chegar no produto final.

Neste contexto, a exploração de métodos de fragmentação da lignina, buscando melhorias em rendimento, juntamente com a avaliação da compatibilidade dessas ligninas fragmentadas com a conversão microbiana e a identificação de rotas envolvidas na bioconversão desses compostos em novos organismos, são fundamentais para o desenvolvimento de estratégias biológicas mais eficientes para a aplicação prática e sustentável na produção de produtos químicos de valor agregado a partir de compostos aromáticos derivados da lignina.

No capítulo I, demonstramos a viabilidade e potencialidades da despolimerização hidrotérmica para a produção de bio-óleos enriquecido em aromáticos. O bio-óleo leve de melhor rendimento (31% em massa) foi gerado nas condições de 350 °C, 90 min e 1:50 m V⁻¹, contendo uma mistura complexa de aromáticos, tais como guaiacol, siringol, fenol, creosol, catecol, 3-metoxicatecol e 2-etil-fenol. Esse bio-óleo demonstrou capacidade de promover o crescimento de nove bactérias, incluindo *X. citri* 306. Notavelmente, as cepas *Pseudomonas sp.* e *Burkholderia sp.* foram identificadas como mais promissoras para serem usadas no desenvolvimento de chassis microbianos customizados para a valorização dos produtos da despolimerização hidrotérmica de lignina de bagaço de cana de açúcar.

No capítulo II aprofundamos nossa análise nas vias moleculares empregadas pelo fitopatógeno *X. citri* 306 para metabolizar compostos aromáticos relacionados à lignina. Embora essa bactéria não tenha tido a melhor performance de bioconversão no primeiro estudo, nosso interesse em desvendar suas vias de metabolismo de compostos aromáticos foi motivado pelo diferente nicho ecológico que ela habita em relação às bactérias modelo geralmente

estudadas nessa área de pesquisa. Partimos da premissa de que, ao habitar um nicho ecológico diferente, sendo exposta a células vegetais vivas e responsivas a ataques microbianos, essa bactéria poderia ter evoluído estratégias singulares e eventualmente complementares às existentes em bactérias que habitam o solo, as quais costumam agir como decompositoras de biomassas vegetais. *X. citri* 306 foi capaz de metabolizar 16 diferentes aromáticos relacionados a lignina, além de crescer em duas misturas complexas de fragmentos de lignina, evidenciando a capacidade desta cepa de metabolizar diferentes frações solúveis de lignina. Além disso, identificamos um novo *operon* (*molRKAB*) envolvido nas duas primeiras etapas do catabolismo dos três principais monolignóis (álcoois *p*-cumarílico, coniferílico e sinapílico), e demonstramos todas as etapas enzimáticas necessárias para canalizá-los para o ciclo do ácido tricarboxílico. Vale ressaltar que, até o momento, apenas o catabolismo microbiano do álcool coniferílico havia sido reportado na literatura¹⁹. Nosso estudo evidenciou a presença de vias catabólicas para o álcool *p*-cumarílico e sinapílico, contribuindo significativamente para uma compreensão mais abrangente das vias metabólicas microbianas relacionadas aos compostos aromáticos da lignina. Concomitantemente, duas novas aril aldeído redutases (XAC1484 e XAC3477), as quais compartilham menos de 41% de identidade de sequência de aminoácidos com enzimas funcionalmente caracterizadas foram identificadas, representando novas aril aldeído redutases com potencial aplicação na síntese de álcoois arílicos de interesse comercial.

Além das descobertas demonstradas no capítulo 2, os dados de transcriptoma obtidos oferecem uma perspectiva ampla e promissora para futuras análises. Diferentes campos de investigação ainda podem ser explorados, incluindo a exploração de genes que foram regulados negativamente, genes relacionados ao estresse oxidativo, e ativação de vias metabólicas não diretamente relacionadas aos aromáticos, mas que foram reguladas positivamente na presença deles. Essas análises podem revelar insights sobre a regulação gênica e a expressão de genes envolvidos na resposta de *X. citri* a esses compostos aromáticos, contribuindo para uma compreensão mais profunda dos mecanismos subjacentes à adaptação e interação da bactéria a seu ambiente, além de uma visão geral sobre o estado de estresse das células sob essas condições. Esses dados também podem ser utilizados para uma análise de conservação das vias metabólicas relacionadas aos aromáticos, identificadas neste estudo, dentro da família *Xanthomonadaceae*. Essa análise pode revelar espécies mais bem adaptadas, que demonstrem uma capacidade catabólica superior e/ou maior resistência à toxicidade imposta por esses compostos, possibilitando a identificação de outros genes relacionados ao catabolismo de aromáticos, bem como de espécies que estejam em processo de perda desses genes. Essa abordagem gera discussões interessantes sobre as capacidades catabólicas dessas espécies em

relação ao seu habitat, hospedeiro, local de infecção e à importância da presença de vias para o metabolismo de aromáticos dentro dessa família. Em última análise, essa abordagem poderia fornecer insights valiosos sobre a adaptação e evolução dessas espécies em resposta aos compostos aromáticos relacionados à lignina.

Por fim, os dois capítulos apresentados se destacam por promover um avanço significativo na compreensão sobre a viabilidade e os desafios da conversão biológica de produtos da despolimerização hidrotérmica de lignina, além de revelar novas enzimas e vias metabólicas de compostos relacionados a lignina, fornecendo conhecimento instrumental para futuras pesquisas envolvendo engenharia genética e desenvolvimento de plataformas microbianas para a transformação de aromáticos derivados da lignina em químicos de relevância industrial.

Referências

1. Clauser, N. M. *et al.* Biomass Waste as Sustainable Raw Material for Energy and Fuels. *Sustainability* **13**, 794 (2021).
2. Rajesh Banu, J. *et al.* A review on biopolymer production via lignin valorization. *Bioresour Technol* **290**, 121790 (2019).
3. Barros, J., Serk, H., Granlund, I. & Pesquet, E. The cell biology of lignification in higher plants. *Ann Bot* **115**, 1053–1074 (2015).
4. Wyman, C. E. & Ragauskas, A. J. Lignin Bioproducts to Enable Biofuels. *Biofuels, Bioproducts and Biorefining* **9**, 447–449 (2015).
5. Morya, R., Kumar, M., Singh, S. S. & Thakur, I. S. Genomic analysis of Burkholderia sp. ISTR5 for biofunneling of lignin-derived compounds. *Biotechnol Biofuels* **12**, 277 (2019).
6. del Río, J. C. *et al.* Differences in the chemical structure of the lignins from sugarcane bagasse and straw. *Biomass Bioenergy* **81**, 322–338 (2015).
7. Kärkäs, M. D., Matsuura, B. S., Monos, T. M., Magallanes, G. & Stephenson, C. R. J. Transition-metal catalyzed valorization of lignin: the key to a sustainable carbon-neutral future. *Org Biomol Chem* **14**, 1853–1914 (2016).
8. Schutyser, W. *et al.* Chemicals from lignin: an interplay of lignocellulose fractionation, depolymerisation, and upgrading. *Chem Soc Rev* **47**, 852–908 (2018).
9. Chio, C., Sain, M. & Qin, W. Lignin utilization: A review of lignin depolymerization from various aspects. *Renewable and Sustainable Energy Reviews* **107**, 232–249 (2019).
10. Zhou, N., Thilakarathna, W. P. D. W., He, Q. S. & Rupasinghe, H. P. V. A Review: Depolymerization of Lignin to Generate High-Value Bio-Products: Opportunities, Challenges, and Prospects. *Front Energy Res* **9**, (2022).
11. Roy, R., Rahman, M. S., Amit, T. A. & Jadhav, B. Recent Advances in Lignin Depolymerization Techniques: A Comparative Overview of Traditional and Greener Approaches. *Biomass* **2**, 130–154 (2022).
12. Kawamoto, H. Lignin pyrolysis reactions. *Journal of Wood Science* **63**, 117–132 (2017).
13. Linger, J. G. *et al.* Lignin valorization through integrated biological funneling and chemical catalysis. *Proceedings of the National Academy of Sciences* **111**, 12013–12018 (2014).
14. Beckham, G. T., Johnson, C. W., Karp, E. M., Salvachúa, D. & Vardon, D. R. Opportunities and challenges in biological lignin valorization. *Curr Opin Biotechnol* **42**, 40–53 (2016).

15. Harwood, C. S. & Parales, R. E. THE β -KETOADIPATE PATHWAY AND THE BIOLOGY OF SELF-IDENTITY. *Annu Rev Microbiol* **50**, 553–590 (1996).
16. Fuchs, G., Boll, M. & Heider, J. Microbial degradation of aromatic compounds — from one strategy to four. *Nat Rev Microbiol* **9**, 803–816 (2011).
17. Liu, Z.-H. *et al.* Identifying and creating pathways to improve biological lignin valorization. *Renewable and Sustainable Energy Reviews* **105**, 349–362 (2019).
18. Xu, Z., Lei, P., Zhai, R., Wen, Z. & Jin, M. Recent advances in lignin valorization with bacterial cultures: microorganisms, metabolic pathways, and bio-products. *Biotechnol Biofuels* **12**, 32 (2019).
19. Brink, D. P., Ravi, K., Lidén, G. & Gorwa-Grauslund, M. F. Mapping the diversity of microbial lignin catabolism: experiences from the eLignin database. *Appl Microbiol Biotechnol* **103**, 3979–4002 (2019).
20. Kohlstedt, M. *et al.* From lignin to nylon: Cascaded chemical and biochemical conversion using metabolically engineered *Pseudomonas putida*. *Metab Eng* **47**, 279–293 (2018).
21. Perez, J. M. *et al.* Integrating lignin depolymerization with microbial funneling processes using agronomically relevant feedstocks. *Green Chemistry* **24**, 2795–2811 (2022).
22. Schutyser, W. *et al.* Chemicals from lignin: an interplay of lignocellulose fractionation, depolymerisation, and upgrading. *Chem Soc Rev* **47**, 852–908 (2018).
23. Odier, E. & Monties, B. Biodegradation de la lignine de ble par *Xanthomonas*. *Ann Microbiol (Paris)* **129**, 361–377 (1978).
24. Kern, H. W. Bacterial degradation of dehydropolymers of coniferyl alcohol. *Arch Microbiol* **138**, 18–25 (1984).
25. Kern, H. W. & Kirk, T. K. Influence of Molecular Size and Ligninase Pretreatment on Degradation of Lignins by *Xanthomonas* sp. Strain 99. *Appl Environ Microbiol* **53**, 2242–2246 (1987).
26. Jiménez, D. J. *et al.* Characterization of three plant biomass-degrading microbial consortia by metagenomics- and metasecretomics-based approaches. *Appl Microbiol Biotechnol* **100**, 10463–10477 (2016).
27. Vieira, P. S. *et al.* Xyloglucan processing machinery in *Xanthomonas* pathogens and its role in the transcriptional activation of virulence factors. *Nat Commun* **12**, 4049 (2021).
28. Giuseppe, P. O., Bonfim, I. M. & Murakami, M. T. Enzymatic systems for carbohydrate utilization and biosynthesis in *Xanthomonas* and their role in pathogenesis and tissue specificity. *Essays Biochem* **67**, 455–470 (2023).

29. Schomburg, I. *et al.* The BRENDA enzyme information system—From a database to an expert system. *J Biotechnol* **261**, 194–206 (2017).
30. Wang, J.-Y. *et al.* A functional 4-hydroxybenzoate degradation pathway in the phytopathogen *Xanthomonas campestris* is required for full pathogenicity. *Sci Rep* **5**, 18456 (2015).
31. Chen, B. *et al.* The phytopathogen *Xanthomonas campestris* utilizes the divergently transcribed *pobA* / *pobR* locus for 4-hydroxybenzoic acid recognition and degradation to promote virulence. *Mol Microbiol* **114**, 870–886 (2020).
32. Chen, B. *et al.* The phytopathogen *Xanthomonas campestris* scavenges hydroxycinnamic acids in planta via the hca cluster to increase virulence on its host plant. *Phytopathology Research* **4**, 12 (2022).

ANEXOS

Anexo I - Termo de autorização do comitê de biossegurança

Uso exclusivo da CIBio:

Número de projeto / processo: 2019-07

Formulário de encaminhamento de projetos de pesquisa com OGMs para análise da CIBio - CNPEM

1. Título do projeto: Investigação das rotas metabólicas utilizadas por *Xanthomonas axonopodis* pv *citri* para a utilização de compostos aromáticos derivados da lignina

2. Pesquisador responsável: Priscila Oliveira de Giuseppe

3. Experimentador(es):

Damaris Batião Martim, Anna Julyana Brilhante, Douglas Antonio Alvaredo Paixão

Nível do treinamento do experimentador: ☐ Iniciação científica, ☐ [AJB]-mestrado, ☐ DBM ☐-doutorado, ☐ ☐-doutorado direto, ☐ ☐-pós-doutorado, ☐ ☐-nível técnico, ☐ DAAP ☐-outro, especifique:especialista

4. Unidade operativa: ☐ LNLS ☐ LNNano ☒ CTBE ☐ LNBio

5. Maior Classe de risco de OGM deste projeto: ☒ Risco I ☐ Risco II ☐ Risco III ☐ Risco IV

6. O projeto é confidencial? ☒ não ☐ sim

7. No caso de projeto confidencial, o título do projeto pode constar em lista aberta no CNPEM? ☐ não ☐ sim

8. Qual é o objetivo do projeto? os objetivos desse projeto consistem em (1) avaliar a capacidade da bactéria *Xanthomonas axonopodis* pv *citri* cepa 306, também chamada de *Xanthomonas citri* susp. *citri* cepa 306, de utilizar compostos aromáticos derivados da lignina como fonte de carbono, (2) avaliar quais vias metabólicas são ativadas por esses compostos, (3) identificar as enzimas chaves desse processo, (4) caracterizar a estrutura e função de enzimas de função desconhecida que participem dessas vias.

9. Informe um número e nome para cada OGM, organismo receptor, organismo doador, o transgene e classe de risco do OGM. (1) *X. citri* 306 KO1, risco I, (2) *X. citri* 306 KO2, (3) Ecoli-XAC0353, *E. coli*, *X. citri* 306, XAC0353, risco I (4) Ecoli-XAC0354, *E. coli*, *X. citri* 306, XAC0354, risco I.

10. Descreva brevemente a função dos transgenes de cada OGM: (1) deleção do gene XAC0352, potencial transportador de compostos aromáticos, (2) deleção dos genes XAC0353-XAC0354 para validação da via de degradação de compostos aromáticos. (3 e 4) expressão heteróloga de genes para purificação das enzimas e caracterização bioquímica e estrutural por cristalografia de raios-X.

11. Algum OGM produz proteína tóxica, oncogênica ou pode gerar produtos deletérios para saúde humana, animal ou meio ambiente? Não.

12. Algum OGM é agente patogênico esporulante? ☒ Não ☐ Sim: _____

13. Algum OGM é agente patogênico e pode se propagar pelo ar? ☒ Não ☐ Sim : _____

14. Algum transgene confere infectividade ou patogenicidade para os OGMs? Descreva. Os OGMs 2 e 3 apresentam resistência ao antibiótico canamicina. Nenhum transgene confere infectividade ou patogenicidade ao organismo receptor.

15. Com relação aos cuidados preventivos associados a manipulação dos organismos, será necessária alguma avaliação médica periódica para experimentadores? ☒ Não ☐ Sim. Que tipo de avaliação? (Ex: consulta com médico, exames laboratoriais etc...) Qual periodicidade? Onde será realizada esta avaliação?

16. Com relação aos cuidados preventivos associados a manipulação dos organismos, será necessária alguma vacinação preventiva para experimentadores? ☒ Não ☐ Sim. Qual periodicidade? Onde será realizada esta vacinação?

17. No caso de uma eventual contaminação com organismos patogênicos ou toxinas, descreva medidas emergenciais para tratamento de pessoas envolvidas, descontaminação de equipamentos, instalações e meio ambiente.

Uso exclusivo da CIBio:

Número de projeto / processo: 2019-07

Formulário de encaminhamento de projetos de pesquisa com OGMs para análise da CIBio - CNPEM

18. Projetos que façam uso de organismos ou genes associados ao patrimônio genético brasileiro precisam de cadastro na plataforma SISGEN (www.sisgen.gov.br). É de total responsabilidade do pesquisador responsável esse cadastramento e cumprimento da legislação. O projeto envolve manipulação, transferência, modificação, armazenamento, coleta de Organismos e derivados relativos ao patrimônio genético brasileiro? (X) SIM, () Não. No caso de responder sim, mencionar a seguir quais os códigos de acesso do cadastro no SISGEN: A8702DF

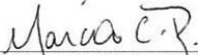
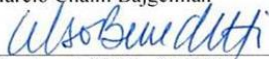
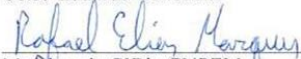
O pesquisador principal tem conhecimento de que conforme a RDC 50 de 21/02/2002 da Anvisa, é responsável por determinar a classificação de riscos de seu projeto, assim como determinar EPIs e medidas de segurança necessárias para prevenir a contaminação de experimentadores, equipamentos, instalações, terceiros e meio ambiente. O pesquisador responsável também precisará providenciar rotina para realização de exames médicos e laboratoriais para sua equipe, bem como vacinações quando aplicável. Todos os experimentadores envolvidos devem ser supervisionados pelo pesquisador principal, que é o responsável pelo treinamento de biossegurança adequado às suas necessidades para a manipulação, armazenamento, descarte e transporte de OGMs, atendendo a legislação e normativas preconizadas pela CTNBio, Anvisa e outros órgãos e agências regulamentadoras e fiscalizadoras.

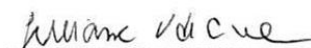
Assinatura eletrônica do pesquisador responsável:

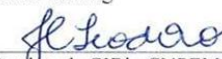
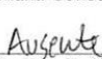
A CIBio analisou este projeto em reunião realizada no dia: 6/5/2019 .

Parecer final: ☒ -projeto aprovado, [] -projeto recusado, [] -projeto com deficiências.

comentários da CIBio:


Presidente da CIBio CNPEM
Marcio Chaim Bajgelman

Membro da CIBio CNPEM
Celso Eduardo Benedetti

Membro da CIBio CNPEM
Rafael Elias Marques Pereira Silva

Membro da CIBio CNPEM
Douglas Galante

Membro da CIBio CNPEM
Juliana Velasco de Castro Oliveira

Membro da CIBio CNPEM
Daniel Kolling

Membro da CIBio CNPEM
Juliana Conceição Teodoro

Membro da CIBio CNPEM
Diego Stefani Teodoro Martinez

Anexo II. Autorização da Editora para inclusão de artigo já publicado



Dear Damaris Martim

We hereby grant you permission to reprint the material below at no charge in your thesis subject to the following conditions:

RE: Exploring the compatibility between hydrothermal depolymerization of alkaline lignin from sugarcane bagasse and metabolization of the aromatics by bacteria, International Journal of Biological Macromolecules, Volume 223, Part A, 2022, Pages 223-230, Menezes et al.

1. If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies.

2. Suitable acknowledgment to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"This article was published in Publication title, Vol number, Author(s), Title of article, Page Nos, Copyright Elsevier (or appropriate Society name) (Year)."

3. Your thesis may be submitted to your institution in either print or electronic form.

4. Reproduction of this material is confined to the purpose for which permission is hereby given. The material may not be reproduced or used in any other way, including use in combination with an artificial intelligence tool (including to train an algorithm, test, process, analyse, generate output and/or develop any form of artificial intelligence tool), or to create any derivative work and/or service (including resulting from the use of artificial intelligence tools).

5. This permission is granted for non-exclusive world English rights only. For other languages please reapply separately for each one required. Permission excludes use in an electronic form other than submission. Should you have a specific electronic project in mind please reapply for permission.

6. As long as the article is embedded in your thesis, you can post/share your thesis in the University repository.

7. Should your thesis be published commercially, please reapply for permission.

8. Posting of the full article/ chapter online is not permitted. You may post an abstract with a link to the Elsevier website www.elsevier.com, or to the article on ScienceDirect if it is available on that platform.

Roopa

Thanks & Regards,
Roopa Lingayath

Sr. Copyrights Specialist, Copyrights Team
ELSEVIER | Health Content Management
+91 44 3378 4167 office
r.lingayath@elsevier.com

Anexo III. Declaração de direitos autorais**Declaração**

As cópias de artigos de minha autoria ou de minha co-autoria, já publicados ou submetidos para publicação em revistas científicas ou anais de congressos sujeitos a arbitragem, que constam da minha Dissertação/Tese de Mestrado/Doutorado, intitulada **Investigação do metabolismo de compostos aromáticos relacionados à lignina em um fitopatógeno Xanthomonas**, não infringem os dispositivos da Lei n.º 9.610/98, nem o direito autoral de qualquer editora.

Campinas, 05 de janeiro de 2024

Assinatura : Damaris B. Martim
Nome do(a) autor(a): **Damaris Batistão Martim**
RG n.º 10.310.817-9

Assinatura : Priscila Ol. de Giuseppe
Nome do(a) orientador(a): **Priscila Oliveira de Giuseppe**
RG n.º 41.092.927-X