

# Women in microbial physiology and metabolism 2022

**Edited by**

Sabine Kleinsteuber and Graciela L. Lorca

**Coordinated by**

Johanna Syrjanen

**Published in**

Frontiers in Microbiology



## FRONTIERS EBOOK COPYRIGHT STATEMENT

The copyright in the text of individual articles in this ebook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this ebook is the property of Frontiers.

Each article within this ebook, and the ebook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this ebook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or ebook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714  
ISBN 978-2-8325-3330-7  
DOI 10.3389/978-2-8325-3330-7

## About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

## Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

## Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: [frontiersin.org/about/contact](https://frontiersin.org/about/contact)



# Women in microbial physiology and metabolism: 2022

## Topic editors

Sabine Kleinsteuber — Helmholtz Centre for Environmental Research,  
Helmholtz Association of German Research Centres (HZ), Germany  
Graciela L. Lorca — University of Florida, United States

## Topic coordinator

Johanna Syrjanen — Cold Spring Harbor Laboratory, United States

## Citation

Kleinsteuber, S., Lorca, G. L., Syrjanen, J., eds. (2023). *Women in microbial physiology and metabolism: 2022*. Lausanne: Frontiers Media SA.  
doi: 10.3389/978-2-8325-3330-7

# Table of contents

- 05 Editorial: Women in microbial physiology and metabolism: 2022  
Graciela L. Lorca and Sabine Kleinsteuber
- 08 A serralyisin-like protein of *Candidatus Liberibacter asiaticus* modulates components of the bacterial extracellular matrix  
Lucila Garcia, Maria Celeste Molina, Kaylie Allyson Padgett-Pagliai, Pablo S. Torres, Roberto E. Bruna, Eleonora García Véscovi, Claudio F. González, Jose Gadea and María Rosa Marano
- 27 The phycoerythrobilin isomerization activity of MpeV in *Synechococcus* sp. WH8020 is prevented by the presence of a histidine at position 141 within its phycoerythrin-I  $\beta$ -subunit substrate  
Lyndsay A. Carrigee, Jacob P. Frick, Xindi Liu, Jonathan A. Karty, Jonathan C. Trinidad, Irin P. Tom, Xiaojing Yang, Louison Dufour, Frédéric Partensky and Wendy M. Schluchter
- 40 The effects of elemene emulsion injection on rat fecal microbiota and metabolites: Evidence from metagenomic exploration and liquid chromatography-mass spectrometry  
Lei Gu, Hao Wu, Yang Zhang, Yousheng Wu, Yuan Jin, Tian Li, Litian Ma and Jin Zheng
- 53 Localization of *Mycobacterium tuberculosis* topoisomerase I C-terminal sequence motif required for inhibition by endogenous toxin MazF4  
Pamela K. Garcia, Rosemarie Martinez Borrero, Thirunavukkarasu Annamalai, Esnel Diaz, Steve Balarezo, Purushottam B. Tiwari and Yuk-Ching Tse-Dinh
- 63 Iron restriction induces the small-colony variant phenotype in *Staphylococcus aureus*  
Shariful Islam, Anna C. Callender, Quynh N. Ho and Catherine A. Wakeman
- 73 Insights into the biotechnology potential of *Methanosarcina*  
Sean Carr and Nicole R. Buan
- 87 Temperature-specific adaptations and genetic requirements in a biofilm formed by *Pseudomonas aeruginosa*  
Karishma Bisht, Alex R. Luecke and Catherine A. Wakeman
- 100 The fate of sulfonamide resistance genes and anthropogenic pollution marker *int1* after discharge of wastewater into a pristine river stream  
Sarah Haenelt, Gangan Wang, Jonas Coelho Kasmanas, Florin Musat, Hans Hermann Richnow, Ulisses Nunes da Rocha, Jochen A. Müller and Niculina Musat

- 112 **Corrigendum: The fate of sulfonamide resistance genes and anthropogenic pollution marker *intl1* after discharge of wastewater into a pristine river stream**  
Sarah Haenelt, Gangan Wang, Jonas Coelho Kasmanas, Florin Musat, Hans Hermann Richnow, Ulisses Nunes da Rocha, Jochen A. Müller and Niculina Musat
- 114 **Metagenome-assembled genomes indicate that antimicrobial resistance genes are highly prevalent among urban bacteria and multidrug and glycopeptide resistances are ubiquitous in most taxa**  
Stefanía Magnúsdóttir, Joao Pedro Saraiva, Alexander Bartholomäus, Majid Soheili, Rodolfo Brizola Toscan, Junya Zhang, Ulisses Nunes da Rocha and CLUE-TERRA consortium
- 125 **An O<sub>2</sub>-sensing diguanylate cyclase broadly affects the aerobic transcriptome in the phytopathogen *Pectobacterium carotovorum***  
Florian J. Fekete, Nick J. Marotta, Xuanyu Liu and Emily E. Weinert



## OPEN ACCESS

EDITED AND REVIEWED BY  
Biswarup Mukhopadhyay,  
Virginia Tech, United States

\*CORRESPONDENCE  
Graciela L. Lorca  
✉ glorca@ufl.edu  
Sabine Kleinstaub  
✉ sabine.kleinstaub@ufz.de

RECEIVED 28 July 2023  
ACCEPTED 02 August 2023  
PUBLISHED 14 August 2023

CITATION  
Lorca GL and Kleinstaub S (2023) Editorial:  
Women in microbial physiology and  
metabolism: 2022.  
*Front. Microbiol.* 14:1268568.  
doi: 10.3389/fmicb.2023.1268568

COPYRIGHT  
© 2023 Lorca and Kleinstaub. This is an  
open-access article distributed under the terms  
of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/)  
(CC BY). The use, distribution or reproduction  
in other forums is permitted, provided the  
original author(s) and the copyright owner(s)  
are credited and that the original publication  
in this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted which  
does not comply with these terms.

# Editorial: Women in microbial physiology and metabolism: 2022

Graciela L. Lorca<sup>1\*</sup> and Sabine Kleinstaub<sup>2\*</sup>

<sup>1</sup>Department of Microbiology and Cell Science, Genetics Institute, Institute of Food and Agricultural Sciences, University of Florida, Gainesville, FL, United States, <sup>2</sup>Department of Environmental Microbiology, Helmholtz Centre for Environmental Research – UFZ, Leipzig, Germany

## KEYWORDS

biofilm, mycobacteria, host-pathogen interface, phyto-pathogen, antimicrobial resistance, methanogens, cyanobacteria, gut microbiota

## Editorial on the Research Topic

### Women in microbial physiology and metabolism: 2022

There is still a gender imbalance in the field of STEM research. According to the UNESCO Institute for Statistics, only 31% of the world's researchers are women. They continue to be underrepresented in the highest academic positions, with only 32% of full professor positions in the U.S. held by women.<sup>1</sup> As highlighted by UNESCO, science and gender equality are essential to ensure sustainable development. The present Research Topic highlights recent work performed by female researchers across the entire breadth of Microbial Physiology and Metabolism. It covers recent insights into the physiology of bacteria that are pathogenic to humans or plants, the prevalence of antimicrobial resistance genes in different environments, the molecular mechanisms of microbial light harvesting systems, and novel application potential of microbial metabolism in biotechnology.

## 1. Human-pathogenic bacteria

*Pseudomonas aeruginosa* is an opportunistic pathogen often associated with nosocomial infections. Its ability to form robust biofilms is related with enhanced virulence and antibiotic resistance. Bisht et al. studied temperature-specific adaptations relevant to both industrial/environmental and medical settings, as well as genetic requirements that affect biofilm formation in *P. aeruginosa*. It was found that temperature can affect production of biofilm through the expression of two previously uncharacterized genes, *PA14\_50070* and *PA14\_67750*, thereby affecting the matrix composition of the biofilm.

The availability of iron is involved in the regulation of the expression of bacterial pathogenicity determinants within the host. Islam et al. investigated the role of iron restriction on *Staphylococcus aureus*. It was found that low iron concentrations result in inhibition of aerobic respiration and the transient display of the small colony variant (SCV) phenotype. Prolonged iron starvation led to the emergence of amino-glycoside resistant SCVs, which potentially serve as a source of antibiotic resistance in iron-restricted host niches.

There is an urgent need to identify new antimicrobial compounds for antibiotic resistant bacteria. A well-defined target for the design of antimicrobials in *Mycobacterium tuberculosis* is the topoisomerase I (TopA). Garcia P. et al. investigated the binding site of the toxin MazF4 in *M. tuberculosis*. Using temperature-sensitive complementation experiments, carboxyterminal truncations and molecular simulations, the TopA C-terminal sequence motif was identified as the mechanistic target of the endogenous toxin MazF4.

<sup>1</sup> <https://www.aauw.org/resources/research/barrier-bias/>



## 2. Phyto-pathogenic bacteria

*Pectobacterium carotovorum* causes bacterial soft rot and is known to be more virulent under low oxygen concentrations (Babujee et al., 2012). Fekete et al. applied transcriptomics, proteomics, phenotypic assays, and inductively coupled plasma-mass spectrometry to elucidate the molecular regulation of oxygen sensing in this important plant pathogen. They found that the soluble globin-coupled sensor protein DgcO interacts with chemotaxis proteins, affects the expression of flagella genes, alters the cellular metal homeostasis, and plays a key role in transition between aerobic and anaerobic growth.

*Candidatus Liberibacter asiaticus* is one of causal agents of huanglongbing disease, which has caused devastation in the citrus industry worldwide. The lack of effective antimicrobials for the management of this disease is a direct consequence of the absence of methods to culture this microorganism in laboratory conditions. Using a variety of bacterial surrogate systems, Garcia L. et al. evaluated the role of a serralyisin-like metalloprotease of *Can. Liberibacter asiaticus* on the modulation of components of the bacterial extracellular matrix. It was found that overexpression of Las1345 negatively impacted cell motility, exopolysaccharide production, and biofilm formation in *Xanthomonas campestris* pv. *campestris* (Xcc).

## 3. Antimicrobial resistance genes

Understanding the environmental fate of antimicrobial resistance genes (ARGs) is crucial within the One Health approach (McEwen and Collignon, 2018) and a prerequisite for more effective regulations on wastewater treatment and discharge. Haenelt et al. monitored the spread of sulfonamide resistance genes and class 1 integron ARG cassettes in a river ecosystem receiving effluents from wastewater treatment. They observed a decrease in ARG abundance with increasing distance from the wastewater discharge but an increasing relative abundance of class 1 integrons. Magnúsdóttir et al. surveyed public metagenome sequences from urban environments for the occurrence and phylogenetic distribution of different ARG classes and found that multidrug and glycopeptide ARGs are ubiquitous among urban bacteria.

## 4. Other topics

In contrast to bacteria and yeasts, archaea are uncommon in biotechnological applications, apart from the central role of methanogenic archaea in producing biogas from organic waste streams. However, this anaerobic digestion process relies on complex microbial consortia dominated by bacteria (De Vrieze and Verstraete, 2016), whereas pure culture applications of methanogens in industrial scale are restricted to the biomethanation of hydrogen in Power-to-Gas approaches (Logroño et al., 2023). Carr and Buan review the opportunities for

archaeal bioproducts beyond methane and describe how the unique and highly efficient metabolism of methanogens can be harnessed for the bioproduction of terpenes and future applications making use of synthetic biology.

Elemene is a plant sesquiterpene with antiproliferative effects on cancer cells, but its use as an antitumor agent has been approved only in China. Although elemene seems to be promising as a chemotherapeutic agent (Jiang et al., 2017), there are still knowledge gaps that hamper its general application. In a preliminary study with rats, Gu et al. observed that intraperitoneal elemene injection modulates the gut microbiome and fecal metabolites, but further studies are needed to understand the physiological effects of elemene on gut microbiota.

Cyanobacteria such as *Synechococcus* contribute to oxygen release in the atmosphere by photosynthesis using phycobilisomes. These complex structures are composed of an allophycocyanin core surrounded by phycocyanin and two phycoerythrin (PE) types: PEI and PEII. Carrigee et al. found that MpeV in *Synechococcus* sp. WH8020 showed differential enzymatic capabilities by having lyase-isomerase activity on the PEII  $\beta$ -subunit but only lyase activity on the PEI  $\beta$ -subunit. Though structural modeling and site directed mutagenesis, the authors show that the phycoerythrobilin isomerization activity of MpeV is modulated depending on the amino acid that is located at position 141 within its phycoerythrin-I  $\beta$ -subunit substrate. These results in combination with structural models led the authors to propose that identity of the residue at this position in the  $\beta$ -subunits may be used to predict which phycobilin is bound on both PEI and PEII.

## Author contributions

SK: Writing—original draft, Writing—review and editing. GL: Writing—original draft, Writing—review and editing.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Babujee, L., Apodaca, J., Balakrishnan, V., Liss, P., Kiley, P. J., Charkowski, A. O., et al. (2012). Evolution of the metabolic and regulatory networks associated with oxygen availability in two phytopathogenic enterobacteria. *BMC Genomics* 13, 110. doi: 10.1186/1471-2164-13-110
- De Vrieze, J., and Verstraete, W. (2016). Perspectives for microbial community composition in anaerobic digestion: from abundance and activity to connectivity. *Environ. Microbiol.* 18, 2797. doi: 10.1111/1462-2920.13437
- Jiang, Z., Jacob, J. A., Loganathachetti, D. S., Nainangu, P., and Chen, B. (2017).  $\beta$ -elemene: Mechanistic studies on cancer cell interaction and its chemosensitization effect. *Front. Pharmacol.* 8, 105. doi: 10.3389/fphar.2017.00105
- Logroño, W., Kleinsteuber, S., Kretzschmar, J., Harnisch, F., De Vrieze, J., and Nikolausz, M. (2023). The microbiology of Power-to-X applications. *FEMS Microbiol. Rev.* 47, fuad013. doi: 10.1093/femsre/fuad013
- McEwen, S. A., and Collignon, P. J. (2018). Antimicrobial resistance: a one health perspective. *Microbiol. Spectr.* 6, ARBA-0009-2017. doi: 10.1128/microbiolspec.ARBA-0009-2017



## OPEN ACCESS

## EDITED BY

Sabine Kleinstüber,  
Helmholtz Association of German Research  
Centres (HZ), Germany

## REVIEWED BY

Juan Camilo Cifuentes-Arenas,  
Fundo de Defesa da Citricultura, Brazil  
Zhanjun Lu,  
Gannan Normal University,  
China  
Beatriz Eugenia Baca,  
Meritorious Autonomous University of  
Puebla, Mexico

## \*CORRESPONDENCE

María Rosa Marano  
marano@ibr-conicet.gov.ar  
Jose Gadea  
jgadea@ibmcp.upv.es

<sup>†</sup>These authors have contributed equally to  
this work and share first authorship

<sup>‡</sup>These authors have contributed equally to  
this work and share senior authorship

<sup>§</sup>PRESENT ADDRESS

Pablo S. Torres,  
Fundación Pablo Cassará,  
Centro de Ciencia y Tecnología "Dr. Cesar  
Milstein", Ciudad de Buenos Aires,  
Argentina

Roberto E. Bruna,  
Department of Pathology and Laboratory  
Medicine, Pennsylvania State College of  
Medicine, Hershey, PA, United States

## SPECIALTY SECTION

This article was submitted to  
Microbial Physiology and Metabolism,  
a section of the journal  
Frontiers in Microbiology

RECEIVED 29 July 2022

ACCEPTED 26 September 2022

PUBLISHED 19 October 2022

## CITATION

García L, Molina MC, Padgett-Pagliai KA,  
Torres PS, Bruna RE, García Vescovi E,  
González CF, Gadea J and  
Marano MR (2022) A serralyisin-like protein  
of *Candidatus Liberibacter asiaticus*  
modulates components of the bacterial  
extracellular matrix.  
*Front. Microbiol.* 13:1006962.  
doi: 10.3389/fmicb.2022.1006962

# A serralyisin-like protein of *Candidatus Liberibacter asiaticus* modulates components of the bacterial extracellular matrix

Lucila García<sup>1,2†</sup>, María Celeste Molina<sup>1,2†</sup>,  
Kaylie Allyson Padgett-Pagliai<sup>3</sup>, Pablo S. Torres<sup>1,4§</sup>,  
Roberto E. Bruna<sup>1†§</sup>, Eleonora García Vescovi<sup>1</sup>,  
Claudio F. González<sup>3</sup>, Jose Gadea<sup>4\*‡</sup> and  
María Rosa Marano<sup>1,2\*‡</sup>

<sup>1</sup>Instituto de Biología Molecular y Celular de Rosario (IBR), Consejo Nacional de Investigaciones Científicas y Tecnológicas (CONICET), Rosario, Argentina, <sup>2</sup>Área Virología, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario (UNR), Rosario, Argentina,

<sup>3</sup>Department of Microbiology and Cell Science, Genetics Institute, Institute of Food and Agricultural Sciences, University of Florida, Gainesville, FL, United States, <sup>4</sup>Instituto de Biología Molecular y Celular de Plantas (IBMCP), Universidad Politécnica de Valencia-C.S.I.C., Ingeniero Fausto Elio, Valencia, Spain

Huanglongbing (HLB), the current major threat for *Citrus* species, is caused by intracellular alphaproteobacteria of the genus *Candidatus Liberibacter* (CaL), with CaL asiaticus (CLAs) being the most prevalent species. This bacterium inhabits phloem cells and is transmitted by the psyllid *Diaphorina citri*. A gene encoding a putative serralyisin-like metalloprotease (CLIBASIA\_01345) was identified in the CLAs genome. The expression levels of this gene were found to be higher in citrus leaves than in psyllids, suggesting a function for this protease in adaptation to the plant environment. Here, we study the putative role of CLAs-serralyisin (Las1345) as virulence factor. We first assayed whether Las1345 could be secreted by two different surrogate bacteria, *Rhizobium leguminosarum* bv. *viciae* A34 (A34) and *Serratia marcescens*. The protein was detected only in the cellular fraction of A34 and *S. marcescens* expressing Las1345, and increased protease activity of those bacteria by 2.55 and 4.25-fold, respectively. In contrast, Las1345 expressed in *Nicotiana benthamiana* leaves did not show protease activity nor alterations in the cell membrane, suggesting that Las1345 do not function as a protease in the plant cell. Las1345 expression negatively regulated cell motility, exopolysaccharide production, and biofilm formation in *Xanthomonas campestris* pv. *campestris* (Xcc). This bacterial phenotype was correlated with reduced growth and survival on leaf surfaces as well as reduced disease symptoms in *N. benthamiana* and *Arabidopsis*. These results support a model where Las1345 could modify extracellular components to adapt bacterial shape and appendages to the phloem environment, thus contributing to virulence.

## KEYWORDS

biofilm, Huanglongbing, protease, surrogate bacteria, virulence factor

## Introduction

Huanglongbing (HLB), also known as citrus greening disease, is a global threat to citrus production (Bové, 2006; Gottwald, 2010; Da Graça et al., 2016). The disease is caused by phloem-limited intracellular Gram-negative alphaproteobacteria belonging to the order of Rhizobiales, family Rhizobiaceae, and genus *Candidatus Liberibacter* (*CaL*; Duan et al., 2009; Kuykendall et al., 2012a; Wang et al., 2017). Three *CaL* species have been found to associated with HLB, *CaL asiaticus* (CLas), *CaL africanus* (CLaf) and *CaL americanus* (CLam; Bové, 2006). Currently, CLas is the most widespread and virulent HLB-associated pathogen worldwide (Bassanezi et al., 2020; Li et al., 2020). CLas is transmitted in a circulative-propagative manner by its insect vector, the Asian citrus psyllid (ACP) *Diaphorina citri* Kuwayama (Hemiptera: Liviidae) during feeding on the phloem of the new shoots (Hall et al., 2016; Lopes and Cifuentes-Arenas, 2021). Upon phloem injection, CLas is confined to sieve elements until the young flush turns into mature leaves (source tissues), then it moves intercellularly through sieve pores following phloem sap from source to sink, including roots and new shoots (Andrade and Wang, 2019; Alves et al., 2021; Pandey et al., 2021; Raiol-Junior et al., 2021). Symptoms of CLas-infected citrus leaves are associated with an inefficient defense response that affect the structure and function of the phloem sieve elements and would normally confine the invading pathogen (Koh et al., 2012; Boava et al., 2017; Granato et al., 2019). However, callus deposition and starch accumulation were not observed in roots despite high bacterial accumulation in this tissue (Etcheberria et al., 2009; Johnson et al., 2014; Pulici et al., 2022). Moreover, the expression levels of both *CALLOSE SYNTHASE 7* (*CaLS7*) and *PLHOEM LECTIN* (*PP2*) were significantly downregulated in HLB-infected roots compared to healthy roots (Achor et al., 2020). Those results support recent investigations which postulate the root system as the primary replication place of CLas at the early infection stages (Alves et al., 2021; Pulici et al., 2022). This differential response between leaves and root may reflect a fine control of plant responses by the pathogen, keeping the balance between defense and nutrient-acquisition (Wang et al., 2016; Martinelli and Dandekar, 2017; Chen et al., 2022).

Bacteria are well known for using flagella to swim through liquids, to swarm across solid surfaces, or to promote host colonization through adherence and biofilm formation (Malamud et al., 2011; Chaban et al., 2015; Rossi et al., 2018). Little is known about systemic movement of *CaL* spp. inside the phloem. Ultra-microscopic analysis of CLas-infected plants showed that unflagellated spherical bacterium cells float freely in the phloem sap by changing morphology from spherical to elongated shape in order to cross the sieve pores without attaching to the sieve tube cell walls or to each other, all without forming biofilms (Achor et al., 2020). However, the presence of *CaL*-associated flagella-like surface structures and aggregates of long rod-shaped cells has been noted inside its psyllid vectors (Cicero et al., 2016; Andrade et al., 2020). Flagella are a highly

conserved microbe-associated molecular patterns (MAMPs/PAMPs), so unflagellated CLas could potentially avoid the elicitation of a strong PAMP-triggered immunity (PTI; Thomma et al., 2011; Zou et al., 2012; Achor et al., 2020; Andrade et al., 2020). In this way, the morphological plasticity of *CaL* spp. may represent an advantage, allowing adaptation in either the citrus phloem sap or in the ACP. Bacterial or host factors that modulate these dynamic morphological changes are yet unknown.

CLas, like other plant-pathogenic bacteria, is believed to secrete effectors (virulence factors) into the cytoplasm of the hosts. The action of these effectors can create favorable environments for colonization and proliferation (Toruño et al., 2016). The expression of effectors needs to be coordinated spatio-temporally to allow the bacteria to thrive and shift between different lifestyles (plant and psyllid). The small CLas genome (~1.2 Mb) only harbors genes for the general protein secretory (Sec) pathway and the Type I secretion system (T1SS) for translocating effectors out of the cell (Duan et al., 2009; Li et al., 2012). Several Sec secreted-dependent effectors (SDEs) have been identified by bioinformatic analysis (Pitino et al., 2016; Clark et al., 2018; Shi et al., 2019; Ying et al., 2019). Although their transgenic expression in model and citrus plants have suggested roles in plant defense modulation, whether they are actually secreted to the extracellular environment is still a matter for discussion (Prasad et al., 2016; Clark et al., 2020; Pang et al., 2020; Du et al., 2021).

The CLas T1SS machinery, similar to other bacterial T1SSs, is composed of three proteins: an inner membrane ATP-binding cassette (ABC) transporter (PrtD), a transmembrane protein (HlyD), and an outer membrane export protein (TolC; Duan et al., 2009; Li et al., 2012; Thapa et al., 2020). The ABC transporters of bacterial T1SSs often show high specificity in binding their unfolded substrates that are translocated to the extracellular side of outer membrane, mostly in a single-step secretion strategy, without periplasmic intermediate (Linhartová et al., 2010; Baumann, 2019; Spitz et al., 2019; Hui et al., 2021). However, the amenability of the T1SS to secrete substrates of different nature and structure has hampered the identification of a canonical -terminal secretion signal.

One of the putative T1SS substrate identified in the CLas genome is CLIBASIA\_01345, a protein with similarity to members of serralyisin-type metalloproteases (Cong et al., 2012; Li et al., 2012). Among the best-characterized secreted metalloproteases is PrtA – also named serralyisin or PrtS – from *S. marcescens*. PrtA plays a key role as bacterial virulence factor, it is involved in cytotoxicity and modulates the immune responses of its host (Lee et al., 2017). *In vitro* studies show that PrtA inhibits the attachment of insect hemocytes and mammal macrophages to tissue surfaces by degrading adhesive molecules (Ishii et al., 2014). Moreover, PrtA contributes to the ability of *S. marcescens* to develop a mature biofilm which could facilitate colonization and host invasion processes (Bruna et al., 2018). In another example, AprA, an extracellular alkaline metalloprotease from *Pseudomonas syringae*, mediates the degradation of flagellin monomers, leading to the



TABLE 1 Bacterial strains and plasmids.

Strain or plasmid	Description	Source or reference
<b>Strain</b>		
<i>R. leguminosarum</i> A34	<i>R. leguminosarum</i> 8401 derivatives carrying pSym plasmid Prl1ji, Str <sup>r</sup>	Russo et al. (2006)
<i>prtA</i>	<i>S. marcescens</i> RM66262; <i>prtA</i> ::pKNOCK, Cm <sup>r</sup>	Bruna et al. (2018)
<i>prtA/PrtA</i>	<i>S. marcescens</i> <i>prtA</i> ; pBBR2:: <i>prtA</i> , Kan <sup>r</sup>	Bruna et al. (2018)
<i>Xcc</i>	<i>X. campestris</i> pv. <i>campestris</i> 8,004, Rif <sup>r</sup>	Daniels et al. (1984)
<i>gumB</i>	<i>X. campestris</i> pv. <i>campestris</i> strain 8,397, [ <i>X. campestris</i> pv. <i>campestris</i> 8,004 <i>gumB</i> ::Tn5, Kan <sup>r</sup> ]	Vojnov et al. (1998)
<i>E. coli</i> S17	<i>E. coli</i> [thi] <i>thr leu tonA lacY 61lic recA</i> ::RP4-2-Tc::Mu lpir] Km <sup>r</sup>	Russo et al. (2006)
<i>E. coli</i> ArticExpress (DE3) RIL	<i>E. coli</i> B F <sup>-</sup> <i>ompT hsdS</i> (rB- mB-) dcm <sup>+</sup> Tet <sup>r</sup> gal λ(DE3) <i>endA</i> Hte [ <i>cpn10 cpn60</i> Gent <sup>r</sup> ] [argU ileY leuW Str <sup>r</sup> ]	Agilent technologies
<i>A. tumefaciens</i>	<i>A. tumefaciens</i> GV3101, Rif <sup>r</sup> , Gm <sup>r</sup>	
<b>Plasmid</b>		
pBBR2	pBBR1-MCS2, Km <sup>r</sup> , broad range	Kovach et al. (1995)
pBBR2::Las1345	pBBR1-MCS2::CLIBASIA01345, Km <sup>r</sup>	This work
pBBR2::HisLas1345	pBBR1-MCS2::His <sub>35</sub> CLIBASIA01345, Km <sup>r</sup>	This work
pMP2444	pBBR1-MCS5::GFP, Gm <sup>r</sup>	Rigano et al. (2007)
p15TV-L	Expression vector, His <sub>35</sub> , Amp <sup>r</sup>	Pagliai et al. (2010)
p15::HisLas1345	p15TV-L::CLIBASIA01345, Amp <sup>r</sup>	This work
pMDC83	Binary vector, 2x35S promoter, C-terminal GFP, Kan <sup>r</sup>	Curtis and Grossniklaus (2003)
pMDC83::Las1345-GFP	35S::CLIBASIA01345-GFP, Kan <sup>r</sup>	This work
pMDC83-GFP	pMDC83 <i>ccd</i> <sup>-</sup> , Kan <sup>r</sup> , allow 35S::GFP expression	This work
pEarlyGate100	Binary vector, 2x35S promoter, Kan <sup>r</sup>	Earley et al. (2006)
pEarlyGate100::Las1345	35S::CLIBASIA01345, Kan <sup>r</sup>	This work

evasion of the first layer of immune responses from *Arabidopsis thaliana* and tomato (Pel et al., 2014).

The function of CLIBASIA\_01345 (hereafter Las1345) in CLas and its contribution to host-CLas interaction is unknown, although its higher expression in the citrus phloem (5.5-fold) as compared to the psyllid vector (Yan et al., 2013) suggests that CLas uses this protein to modulate its lifestyle in the two hosts. CLas as well other pathogenic *CaL* species remain unculturable in artificial media, making the use of surrogate bacterial models the only available approach to analyze putative gene function associated with CLas pathogenesis (Vahling-Armstrong et al., 2012; Kuykendall et al., 2012b; Pagliai et al., 2014; Barnett et al., 2019; Jain et al., 2019). Here we study the role of Las1345 as a potential virulence factor associated with bacterial adaptation to different hosts. *Rhizobium leguminosarum* and *S. marcescens*, as well as the phytopathogen *Xanthomonas campestris* pv. *campestris* (*Xcc*) were used as surrogate models to study Las1345 function in bacteria, whereas the role of Las1345 in planta was investigated by transient expression in model plants.

## Materials and methods

### Sequence data acquisition and multiple sequence alignment

The genome sequences of the three citrus-infecting *CaL* species [Las (CP001677.5), CLaf (CP004021.1) and CLam

(CP006604)] are available at the National Center for Biotechnology Information (NCBI) database. Protein sequences, based on the homology to Las1345, were downloaded from NCBI under the following accession numbers: serralsin-like proteins from CLas strain psy62 (WP\_015452346), from CLaf strain PTSAPSY (AKK19938.1) and from CLso strain ZC1 haplotype B (WP\_013461860); *PrtA* from *S. marcescens* (CAA39139.1), *Ser1* from *S. liquefaciens* FK01 (BAK39731) and *PrtC* from *Dichya chrysanthemii* B374 (WP\_038909783.1). The sequences were aligned using Clustal Omega and MultiAlin (Corpet, 1988; Sievers and Higgins, 2018), following a manual adjustment in the C-terminal sequence, considering the sequence pattern associated with the ABC exporter motif. Las1345 was homology modeled using Swiss Model server (Waterhouse et al., 2018).

### Bacterial strains, grow conditions and cloning

Bacterial strains and plasmids are described in Table 1. *Rhizobium leguminosarum* bv. *viciae* A34 (A34), *S. marcescens* mutant in *prtA* (*prtA*), *Xanthomonas campestris* pv. *campestris* (*Xcc*) wild type and the *gumB* mutant strains were grown in tryptone-yeast (TY), Miller's Luria-Bertani (LB) and peptone-yeast-malt (PYM) medium, respectively (Russo et al., 2006; Torres et al., 2007; Bruna et al., 2018). Plasmids were mobilized into *R. leguminosarum* bv. *viciae* A34 and *S. marcescens* by biparental mating using *Escherichia coli* S17 cultured LB, at 37°C (Vozza

et al., 2016; Bruna et al., 2018). Conjugation was performed at 28°C and 37°C for *R. leguminosarum* and *S. marcescens*, respectively. *Xcc* was transformed by electroporation (Rigano et al., 2007). Bacterial growth was monitored at an optical density of 600 nm (OD<sub>600</sub>).

*CLIBASIA\_01345* was amplified from CLas-infected plant tissue via polymerase chain reaction (PCR) and subsequently cloned into p15TV-L (Pagliai et al., 2010) to obtain p15::HisLas1345. This vector was used as template to obtain pBBR2::HisLas1345 by restriction sites cut with *NcoI* and *HindIII*. The plasmid p15::HisLas1345 was also used as template to amplify tagless *CLIBASIA\_013145* to generate pBBR2::Las1345 using the following primers: 5-GTC GGTACCATGCATAATATAAAACCGG-3/5-GTCAAGCTTT CAGGAAAAATCATGATTTA-3. To express Las1345 in plants, *CLIBASIA\_01345* was amplified from p15::HisLas1345, adding *KpnI*ATG/*HindIII* sites with the following primers 5-GTCGGTACCATGCATAATATAAAACCGG-3 and 5-GTC AAGCTTTCAGGAAAAATCATTTA-3 and cloned into pENTR3c-Las1345 to generate pENTR3c::Las1345. This vector was amplified with AHL primers (5-TAGTTAGTTAC TTAAGCTCGGGC-3/5-CAGAGCTGCAGCTGGATGGC-3) and the PCR product was transferred into binary vectors (pMDC83 and pEarlyGate100) using Gateway LR Clonase II following manufacturer instructions (Thermo Fisher Scientific, Waltham, United States).

## Las1345 expression and purification

The recombinant plasmid p15::HisLas1345 was used to transform *E. coli* ArticExpress competent cells for further expression and purification. Cells were grown at 30°C until OD<sub>600</sub> ~ 0.7. Induction was made using 0.5 mM IPTG and 6-h incubation at 19°C. The cell pellet was suspended in binding buffer (50 mM HEPES pH 7.5, 250 mM NaCl, 5% (v/v) glycerol, 2.5 mM TCEP and 5 mM Imidazol). The thawed cells were passed through a French Press (French pressure cell press 40 K, Thermo Scientific). Three runs at 800 pressure gauge (Phcp) were conducted for each sample. Samples were released at 15 drops/min. The clarified lysed was purified using Ni-NTA resin (Qiagen, Germantown, United States) as described previously (Pagliai et al., 2014). Purified protein was dialyzed against 10 mM HEPES pH 7.5, 50 mM NaCl, 10 mM MgCl<sub>2</sub>, 2 mM CaCl<sub>2</sub>, 5% (v/v) glycerol and 0.25 mM TCEP. Protein was concentrated using Vivaspin centrifugal concentrators (Sartorius, Bohemia, United States) treated with Triton X-100 and the purity was analyzed in gel. The purified protein was refolded by dilution into cold buffer (50 mM Tris, 50 mM NaCl, 0–0.2 mM CaCl<sub>2</sub>, pH 7.8) on ice for 20 min prior to activity. The proteolytic activity of the purified protein (pHisLas1345) was evaluated under different amount of protein and the addition of divalent cations (Zn<sup>2+</sup>, Co<sup>2+</sup>, Cu<sup>2+</sup> and Ca<sup>2+</sup>) in the chloride salt form (Supplementary Figure 1).

## Bacterial cell fractionation

For analysis of secreted and intracellular proteins, bacteria (*R. leguminosarum*, *S. marcescens* and *Xcc*) were grown for 48 h at 28°C in TY, LB or PYM medium, respectively to an OD<sub>600</sub> ~ 0.6. The cell culture (3 ml) was centrifuged at 6,500 rpm and both the extracellular medium (supernatant) and the cellular precipitate were reserved for protein extraction. Supernatant (2 ml) was clarified with a 0.2 µm-syringe filter to obtained a cell free supernatant fraction (SF). Extracellular proteins were concentrated from the SF by precipitation with 10% (v/v) trichloroacetic acid (Merck, Darmstadt, Germany) and incubated on ice for 2 h. Proteins were recovered by centrifugation (12,000 rpm, 30 min) and washed twice with 500 µl acetone (Cicarelli, Santa Fe, Argentina). The precipitate was dried and the proteins were solubilized in the proper buffer for activity or immunodetection. Intracellular proteins were recovered from the soluble and insoluble cell fractions (CFs and CF<sub>i</sub>, respectively) prepared from cellular precipitate. Cells were washed twice with 10 mM Tris, 5 mM EDTA pH 8 and lysed by sonication at a frequency of 20 kHz. The lysate was centrifuged (12,000 rpm, 30 min) at 4°C to recovered the pellet (CF<sub>i</sub>) and the supernatant (CFs) fractions. Proteins in the CFs were precipitated with acetone (Cicarelli, Santa Fe, Argentina) on ice for 2 h and recovered by centrifugation (12,000 rpm, 30 min). CFs and CF<sub>i</sub> pellets were resuspended in the proper buffer for activity or immunodetection.

## Immunodetection

SF, CFs and CF<sub>i</sub> pellets were resuspended in loading buffer [60 mM Tris, 10% (v/v) glycerol, 180 mM β-mercaptoethanol, 0.003% (w/v) bromophenol blue and 2% (w/v) SDS, pH 6.8], separated by 12% sodium dodecylsulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and visualized by staining with Coomassie brilliant blue R-250 (Genbiotech, Buenos Aires, Argentina). Western blot was carried out using standard techniques. Briefly, after SDS-PAGE, proteins were transferred onto pre-wetted polyvinylidene difluoride membranes (PVDF-Immun-BlotV<sup>®</sup>, BioRad, CA, United States). Immunodetection was performed using polyclonal antibodies against His tag (1:6,000 dilution, #SAB1306085 Sigma, Merck, Darmstadt, Germany), *Rhizobium* adhering-protein A1 (RapA1; 1:5,000 dilution; Voza et al., 2016) or *E. coli* Glucose-6-phosphate dehydrogenase (G6PD; 1:500 dilution; Giró et al., 2006) and then revealed using a peroxidase-conjugated goat anti-rabbit IgG (H+L; 1:3,000 dilution, #1706515, BioRad, Des Plaine, IL, United States) and Pierce<sup>™</sup> ECL Western Blotting Substrate (Thermo Scientific, Petaluma, CA, United States) according to the manufacturers. Chemiluminescence was detected using an CCD camera (ChemiDoc<sup>™</sup> XRS<sup>+</sup>, BioRad, Des Plaine, IL, United States). For probe stripping and rehybridization, the filters were incubated with stripping buffer (0.05 M Tris, 1% (w/v) SDS, 0.8% β-mercaptoethanol, pH 6.8) for 45 min at 30°C and washed

five times with TBS (20 mM Tris, 150 mM NaCl, pH 7.6) supplemented with 0.05% (v/v) Tween (Promega, Fitchburg, WI, United States).

## Plant growth and pathogenicity assays

*Arabidopsis thaliana* and *Nicotiana benthamiana* were grown under controlled conditions in a growth chamber with a temperature of 25°C–27°C and a photoperiod of 16 h light/8 h dark and light intensity of 150–200  $\mu\text{E/s m}^2$ . All plant inoculations involved a minimum of 10 plants per strain tested and one leaf per plant was inoculated. Bacterial suspensions of *Xcc* ( $10^7$  CFU/ml in 10 mM  $\text{MgCl}_2$ ) were inoculated by pressure infiltration (Siciliano et al., 2006). Disease progression (water-soaking and necrosis) was monitored phenotypically and by bacterial growth curves. Bacterial population was determined in five samples according to the method described by Rigano et al. (2007). Each sample was obtained from three leaf disks of 1  $\text{cm}^2$  collected randomly from different inoculated leaves. Symptom development were assayed by conductivity measurements using leaf disks submerged in milli-Q water following protocols previously described (Chiesa et al., 2019). Six samples were measured, each one obtained from three leaf disks of 1  $\text{cm}^2$  collected randomly from 10 inoculated leaves. Photosystem II quantum efficiency ( $\phi\text{PSII}$ ) was measurements at 25°C on one dark-adapted leaf from 15 plants using MultispeQ V 2.0 (PhotosynQ INC, East Lansing, MI, United States). These assays were repeated three times.

## Agrobacterium-mediated transient expression and protein localization

To evaluate the function of Las1345 in plants, *N. benthamiana* leaves were agroinfiltrated with *A. tumefaciens* GV3101 transformed with the corresponding plasmid (Table 1), to overexpress Las1345 with and without a GFP tag. Molecular techniques and agroinfiltration were performed as described previously (Sánchez et al., 2010; Enrique et al., 2011). GFP versions were analyzed to study protein localization by confocal laser scanning microscopy (CLSM). Nucleus and membrane were marked with DAPI (Thermo Fisher Scientific, Waltham, United States) and FM 4-64 (Sigma, Merck, Darmstadt, Germany), respectively. Images were taken 2 days after agroinfiltration with a Zeiss LSM 880 confocal laser scanning microscope (Carl Zeiss Microscopy GmbH, Jena, Germany) using the following parameters, GFP (488/527 nm), FM 4-64 (515/604 nm) and DAPI (405/449 nm). For CLSM analysis, each construct was infiltrated into two leaves of 10 plants.

For plant protein extraction, six leaf disks (1  $\text{cm}^2$ ) collected randomly from different agroinfiltrated plants were pulverized in pre-chilled pestle and mortar using liquid nitrogen and the powder was resuspended in 0.5 ml of extraction buffer (50 mM Tris, 100 mM KCl, 10% (v/v) glycerol, pH 7.5). Supernatant was

recovered after centrifugation (12,000 rpm, 20 min) at 4°C and kept at –80 for protease activity measurements.

## Protease assays

Protease activity was assayed using milk or azocasein (#A2765, Sigma, Merck, Darmstadt, Germany) as substrates as described previously (Bruna et al., 2018). For qualitative analysis, the samples were inoculated on LB agar plates supplemented with skim milk at 2% (w/v) and incubated for 16 h at 30°C. Distinct clearing of the milk around the colony was used as a protease activity indicator. For quantitative analysis, protease activity was measured using azocasein as substrate. Samples of SF, CF<sub>s</sub> and CF<sub>i</sub> (50  $\mu\text{l}$ ) were mixed with 10  $\mu\text{l}$  of 1% (w/v) azocasein and 140  $\mu\text{l}$  of phosphate-buffered saline (PBS) and incubated for 1 h at 37°C. The reaction was stopped by addition of 80  $\mu\text{l}$  of 10% (v/v) trichloroacetic acid, and the mix was incubated on ice for 15 min. Protease activity was repeated five times using six replicates of each strain. The same protocol was used to evaluate protease activity with both the purified HisLas1345 protein from *E. coli* and plant protein extracts. For these analyzes, measurements were normalized by protein quantification using the Bradford assay kit (BioRad, Des Plaine, IL, United States) and repeated three times.

## Cell motility and production of the extracellular polysaccharide (EPS) xanthan

Swimming, swarming and sliding motility assays were carried out as previously described (Malamud et al., 2011, 2013). Briefly, overnight cultures were normalized to an  $\text{OD}_{600} \sim 0.8$  and 3  $\mu\text{l}$  were used to inoculate 0.25% (w/v) agar NYGB medium plates (swimming) or 0.5% (w/v) agar NYGB medium (sliding/swarming). Plates were incubated 72 h at 28°C. Motility was assessed quantitatively by measure the circular halo formed by the growing bacterial cells in 10 different culture plates. This assay was repeated three times.

Exopolysaccharide (EPS) xanthan quantification was performed as described previously (Vojnov et al., 1998). To measure EPS xanthan production, strains were grown in PYM medium supplemented with 1% (w/v) D-glucose at 28°C for 24 h with shaking (156 rpm) in 250-mL flask. EPS xanthan was precipitated from culture supernatants using 2 vol of ethanol. The precipitated EPS xanthan was collected, washed with ethanol, dried and weighed. Each assay included eight to 10 replicas and the experiment was repeated three times.

## Biofilm analysis by CLSM

To analyze biofilm structure, GFP-expressing *Xcc* cells were transformed with either pBBR2::Las1345 or pBBR2 as the empty

vector control. Both strains were grown at 28°C on PYM medium supplemented with 50 µg/ml kanamycin. Cultures were diluted in Y minimal medium (YMM) to reach OD<sub>600</sub> ~ 0.02 and aliquots of 500 µl were transferred to 8-wells chambers containing a 1-mm thick borosilicate glass (Nunc, Wiesbaden, Germany), as described by Rigano et al. (2007). Strains were incubated for 4 days at 28°C. Congo red (10 µM) was added to YMM to visualize extracellular components, such as exopolysaccharides, curli amyloid proteins and adhesins which conform the biofilm matrix (Serra and Hengge, 2017). Congo red staining has been used to detect amyloid fibers in *Xanthomonas* spp. (Oh et al., 2007; Kraiselburd et al., 2012). Biofilm formation was monitored using a Zeiss LSM880 confocal laser scanning microscope (Carl Zeiss Microscopy GmbH, Jena, Germany) by excitation at 488 nm with the argon laser line and 20x NA = 0.8 Plan-Apochromat objective (Carl Zeiss Microscopy GmbH, Jena, Germany). To analyze the tridimensional structure of the biofilm, the XY area was screened in 35 Z-intervals (Z-stack) automatically captured up to a 25.5 µm range at green (488/520 nm, pinhole 0.84 A.U.) and red (543/619 nm, pinhole 0.70 A.U.) channels, respectively. Biofilm images were obtained by ZEN BLACK software (Carl Zeiss Microscopy GmbH, Jena, Germany) and analyzed with COMSTAT2 (Heydorn et al., 2000; Vorregaard, 2008).<sup>1</sup> Each assay includes triplicates (3 wells per strain) and it was repeated three times.

## Bacterial RNA extraction, reverse transcription, and quantitative PCR (qPCR)

The expression levels of three (*flaE*, *fliF* and *flgL*) flagellum assembly-related genes of *Xcc* were analyzed by qPCR. *Xcc*/Las1345 and *Xcc*/pBBR2 cells were grown in YMM at 28°C to reach an OD<sub>600</sub> ~ 0.8. The cells (10 ml) were collected and total RNA was extracted using TRIzol Reagent (Thermo Fisher Scientific, Waltham, United States), following treatment with RNase-free DNase (Promega, Wisconsin, United States). Reverse transcription was performed using RevertAid reverse transcriptase (ThermoFisher Scientific, Carlsbad, United States), 2 µg of DNase-treated total RNA and 0.5 µg/µl random hexamers, according with the manufacturer's instructions. cDNAs were used for qPCR using Hot FIREPol® EvaGreen qPCR mix plus (Solis BioDyne, Tartu, Estonia) and 0.25 µM of each primer. Quantitative PCRs (qPCR) were performed for 42 cycles according to the following conditions: denaturation at 95°C for 15 s, annealing at 60°C for 30 s and extension at 72°C for 40 s. After amplification, melting curves were performed to exclude artifactual amplifications. The 16S ribosomal RNA was used as internal control. Primers were designed based on the *Xcc* 8004 (CP029484.1) genome sequence. Flagellar genes were amplified with the following primers, *fliE\_fw* 5'-AGCTTCAGCGAGACCTTGCG-3' and *fliE\_rv*

5'-CAGATCGGCACTGGGGTCAC-3'; *fliF\_fw* 5'-AAGTTCC AAGAGCGCCACCC-3' and *fliF\_rv* 5'-ACGATCTTGC CCTTGGCACC-3'; *flgL\_fw* 5'-ATGGCAATGCGCCTTTCATC and *flgL\_rv* 5'-GCGGATGCGCATGAAGATTT-3'; *16S\_fw* 5'-AGGACCTTCGGGCTT-3' and 5'-TGTCTCAGTTCCA GTG-3' (Yan et al., 2019).

## Statistical analysis

Data were analyzed according to Student's *t*-test through InfoStat Software v2017 (Di Rienzo et al., 2017), excepting protease activity in bacterial cell fractions and EPS production data that were subjected to one-way analysis of variance (ANOVA). Two-way ANOVA was used to analyze the data from protease activity in plant crude extract. In both cases, means were analyzed using Tukey's test *p* < 0.05).

## Results

### Structural analysis of Las1345 reveals absence of calcium-binding GG repeats in the RTX domains

Genome comparison between CLas and *CaL* solanacearum (CLso, causing zebra chip disease in potato) revealed that the *CaL* serralyisin-like gene (Figure 1A) is largely conserved in terms of size and localization, being next to a gene cluster encoding the T1SS (Li et al., 2012; Ravindran et al., 2018). In this work we have included sequence comparison with the other two HLB-associated *CaL* species, africanus strain PTSAPY (CLaf) and americanus strain São Paulo (CLam; Wulff et al., 2014; Lin et al., 2015). A putative serralyisin gene was found in CLaf genome (Figure 1B). Comparative protein sequence analysis indicates that Las1345 (665 amino acids, aas) has 63.77 and 65.67% amino acid identity with the serralyisin-like protein of CLsoB (577 aas) and CLaf (642 aas), respectively. These two proteins are longer than the characterized serralyisins from *Serratia* spp. (PrtA and Ser1, 504 aas each) and *Dickeya chrysanthemi* (former *Erwinia*; PrtC, 479 aas; Figure 1B). Las1345 share 30.41%, 31.23%, and 31.30% amino acid identity with PrtA, Ser1 and PrtC, respectively. Las1345 conserves the N-terminal metzincin-type metalloprotease motif HExxHxxGxxHP, including the last proline, characteristic for serralyisins. The three histidines in this motif are zinc ligands and the glutamic acid serves as catalytic base (Baumann, 1994, 2019). Twenty-three residues downstream from this motif, the conserved met-turn motif (SxMSYF/W) is found. Three-dimensional prediction shows that the N-terminal part of Las1345 has a folding topology similar to PrtA (Figure 1C). PrtA, Ser1 and PrtC have four conserved motifs (GGxGxDxUx) in the RTX (repeats-in-toxins) domain (Figure 1B). T1SS substrates that have these specific repeats belong to the RTX- protease family (Spitz et al., 2019; Hui et al., 2021). Three-dimensional structure of the RTX

<sup>1</sup> www.comstat.dk



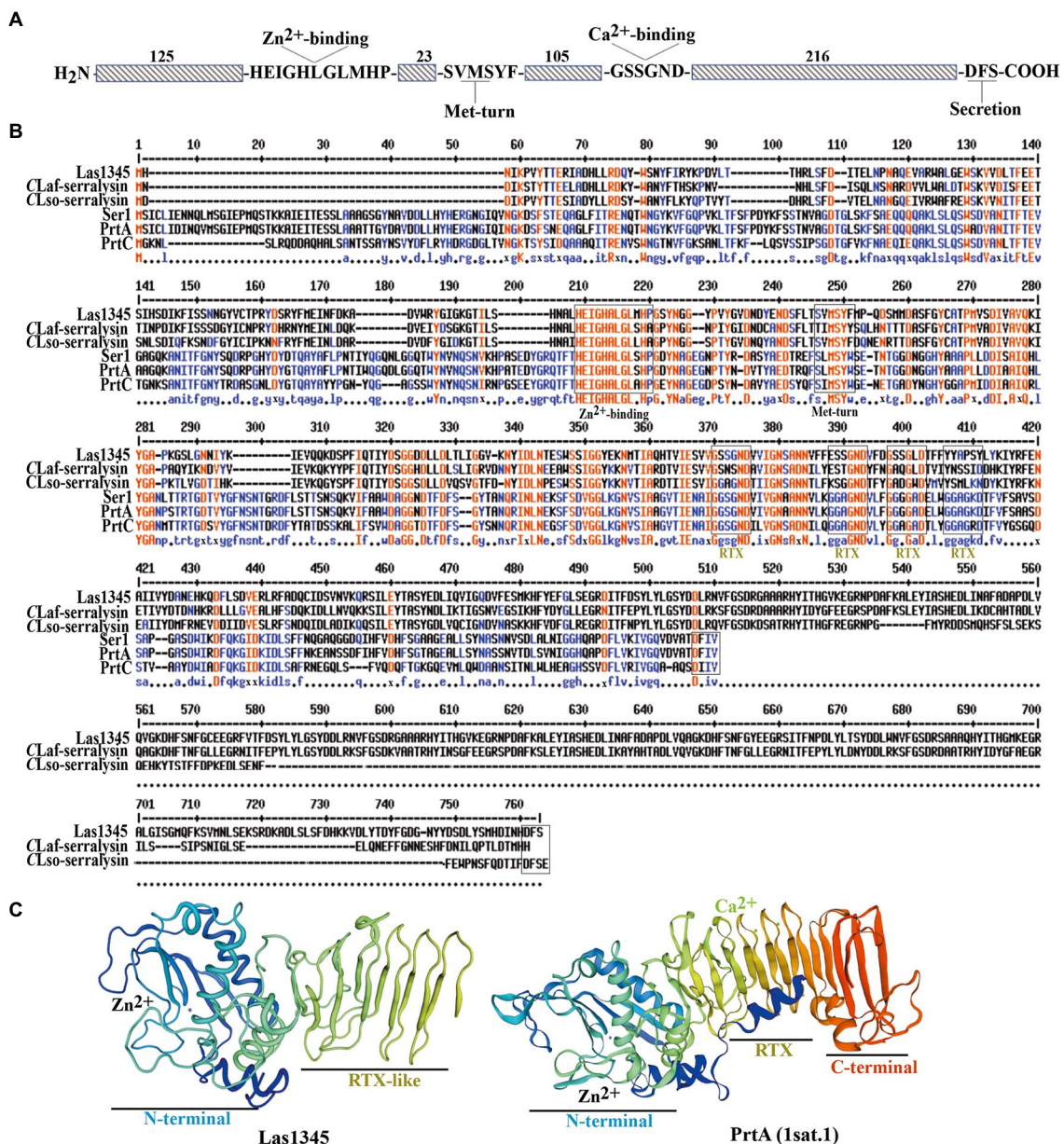


FIGURE 1

Multiple sequence alignment and representative tertiary structure of Las1345. (A) Schematic representation of Las1345 (CLIBASIA\_01345). Conserved motifs associated with serralyisin family protein are shown. HExxHxUGUxH, zinc-binding motif; SxMxY, a Met-turn motif; GGxGxDxUx, calcium-binding motif (glycine-rich repeats) where U represents an aliphatic amino acid and X can be any residue; Dxxx, T1SS secretion signal. (B) Sequence alignment of the serralyisin-like proteins encoded by *Candidatus* Liberibacter (CaL) asiaticus (CLAs) strain psy62 (Las1345, WP\_015452346); CaL africanum strain PTSAPSY (CLaf\_Serralyisin, AKK19938.1); CaL solanacearum strain ZC1 (CLsoB\_Serralyisin, WP\_013461860); *Serratia marcescens* (PrtA, CAA39139.1); *S. liquefaciens* strain FK01 (Ser1, BAK39731) and *Dickeya chrysanthemi* strain B374 (PrtC; PDB 1K71\_A). Alignment was obtained by ClustalW and Multaline. Low consensus alignment (50%) is shown in blue letters while high consensus alignment (90%) is shown in red. Protein domains were identified by Pfam. Protein domains, showing the metalloprotease sequence (HExxHxGxxHP) motif, the met-turn (SxMSYF/W) motif, the four RTX repeats (the classic, GGxGxD and the divergent, GSxGxD) and the ABC exporter motif (Dxxx) are indicated in boxes. (C) Tertiary structure of Las1345 domains based on a homology model derived from PrtA structure (PDB ID 1sat.1; Baumann, 1994; QMEANG 0.65±0.5). The N terminal Zn-binding domain and the RTX β-rolls are colored in rainbow. Zinc and calcium ions are represented by gray and green spheres, respectively.

domain of PrtA shows that this domain consists of four parallel β-rolls, where two GG repeats coordinate one  $\text{Ca}^{2+}$  ion by interaction of the side chain of the aspartate residues and the carbonyl oxygens of the amino acids forming the repeat. Calcium

ion binding is implicated in secretion and extracellular folding of the protein (Baumann, 2019; Spitz et al., 2019; Hui et al., 2021). Unlike PrtA, Las1345 has only one non-canonical RTX repeat, GSSGND, located between G<sub>370</sub> and D<sub>375</sub>. It has been hypothesized

that this consensus motif, instead of the typical nonapeptide repeat, may also be considered as a potential binding site for  $\text{Ca}^{2+}$  (Linhartová et al., 2010). Interestingly, the GSSGND motif in the RTX domain of Las1345 also folds in a right-handed  $\beta$ -roll (GSS) and connects a short strand as in PrtA. However, this RTX repeat is not recognized as a  $\text{Ca}^{2+}$  binding site by the modeling server (Figure 1C). Moreover, prior to the predicted non-cleaved C-terminal T1SS secretion signal (DFS) for serralyins, Las1345 exhibits a duplicated 84-residue sequence (Figure 1B).

These differences between Las1345 and characterized serralyins such as PrtA may influence Las1345 secretion via the T1SS and extracellular folding. However, the metzincin-type metalloprotease domain and the non-classical RTX repeats still supports the possibility of protease activity in Las1345.

## Las1345 has protease activity in the cytoplasm of *rhizobium leguminosarum* and *Serratia marcescens*

CaL is a member of the order *Rhizobiales* and it is phylogenetically related to the *Rhizobium*-type genus (Young et al., 2001; Duan et al., 2009; Kuykendall et al., 2012a). In order to study the putative secretion of Las1345, the symbiont *R. leguminosarum* was used as surrogate model, considering the T1SS similarity between *CaL* spp. and *Rhizobium* (Duan et al., 2009; Kuykendall et al., 2012a). We first evaluated the putative secretion of Las1345 using cell fractionation in *R. leguminosarum* bv. *viciae* strain A34 (A34; Figure 2A). RapA1, an extracellular T1SS-secreted and surfaced associated protein from *R. leguminosarum* (Vozza et al., 2016) and the bacterial intracellular G6PD protein (Giró et al., 2006) were used as subcellular fractionation controls. An N-terminal His-tagged version of Las1345 (HisLas1345) was only detected in the cellular fraction of *R. leguminosarum*, suggesting either that the C-terminal domain of this protein is not recognized by the inner membrane ABC transporter of the T1SS of *R. leguminosarum* or that the N-terminal tag prevents Las1345 secretion to the extracellular medium (Figure 2A). Interestingly, accumulation of HisLas1345 in cellular fraction correlated with an increased protease activity, measured using azocasein as substrate (Figure 2B). To discard a tag-effect during the secretion process, an untagged version of Las1345 was expressed in *R. leguminosarum* and protease activity was measured in the supernatant fraction. In both tagged and untagged versions, Las1345 did not increase proteolytic activity in the supernatant fraction of A34, suggesting that Las1345 might not be secreted, at least in this surrogate system (Figure 2C). Moreover, in both versions of Las1345, an increase in proteolytic activity was detected both in the soluble and insoluble cellular fractions ( $\text{CF}_s$  and  $\text{CF}_i$ ). Interestingly, proteolysis was increased by more than 50% in  $\text{CF}_s$  of both A34/Las1345 ( $100 \pm 3.15$ ) and A34/HisLas1345 ( $78.45 \pm 3.42$ ) compared to the A34/pBBR2 ( $9.67 \pm 2.33$ ) and only 10%–20% in  $\text{CF}_i$  (Figure 2C).

As shown in Figure 1, PrtA from *S. marcescens* shows homology with Las1345, mainly in the catalytic domain. Therefore, a *S. marcescens* *prtA* mutant strain (*prtA*) was transformed with pBBR2::HisLas1345 and pBBR2::Las1345 to evaluate Las1345 secretion and its proteolytic activity in this surrogate model. *prtA* complemented by *trans* expression of *prtA* from the pBBR2::*prtA* plasmid or transformed with pBBR2 were used as positive and negative controls, respectively. Similarly to the findings observed in *Rhizobium*, HisLas1345 was only detected in the cellular fraction of *prtA* (Figure 3A) in contrast with the positive control (PrtA) which was detected in the supernatant fraction (Figure 3A). As expected, based on the cellular localization, protease activity using azocasein as substrate was only detected in the  $\text{CF}_s$  fractions of *prtA*/HisLas1345- and *prtA*/Las1345, increasing the proteolytic activity as compared to the control (pBBR2; Figure 3B). The absence of protease activity in the SF of *prtA*/Las1345 and *prtA*/HisLas1345 (Figure 3C) reinforces the results obtained for the serralyisin-like protein of CLso in *S. liquefaciens* (Ravindran et al., 2018). Proteolytic activity of HisLas1345 was also confirmed using the purified protein (pHisLas1345) from *E. coli* (Supplementary Figure 1).

Taken together, these results indicate that Las1345 is not secreted by bacteria phylogenetically related to CLas, such as *R. leguminosarum*, nor by another surrogate bacterium such as *S. marcescens*. However, this protein shows proteolytic activity in the soluble cell fraction when expressed in these surrogates. We hypothesize that the long C-terminal region and the absence of multiple calcium-binding domains in the RTX motif of Las1345 allows this protein to be in its active conformation within the cell but hinders secretion to the extracellular space.

## Las1345 does not compromise plant cell integrity

Las1345 secretion to the extracellular medium could not be detected in surrogate bacterial models. However, this does not rule out Las1345 secretion in CLas via other non-conserved secretion pathways, or the putative role of Las1345 inside the plant cell. Phenotype changes have been observed via ectopic expression of multiple Sec secreted-dependent effectors *in planta*, even though the secretion of these proteins have not been demonstrated (Pitino et al., 2016; Clark et al., 2020; Pang et al., 2020; Du et al., 2021). Additionally, several extracellular metalloproteases from phytopathogenic bacteria, including AprA and Prt2/3, take part in modulating the plant immune defense, particularly through processes associated with plant cell wall degradation (Dow et al., 1990, 1993; Pel et al., 2014). Thus, observing whether Las1345 affects or not the plant cell in the hypothetical case that it could be secreted, could help to gain insight into the real scenario where Las1345 is functional.

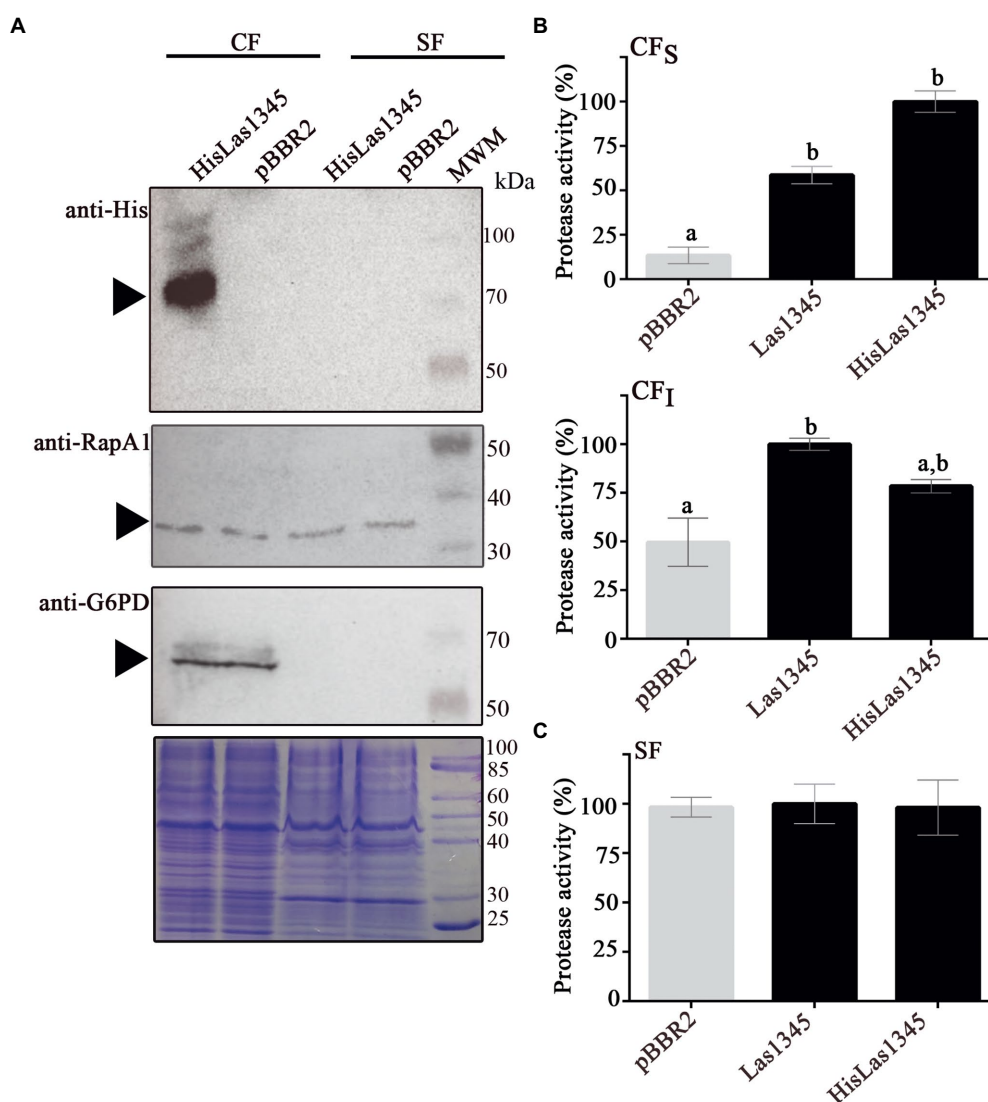


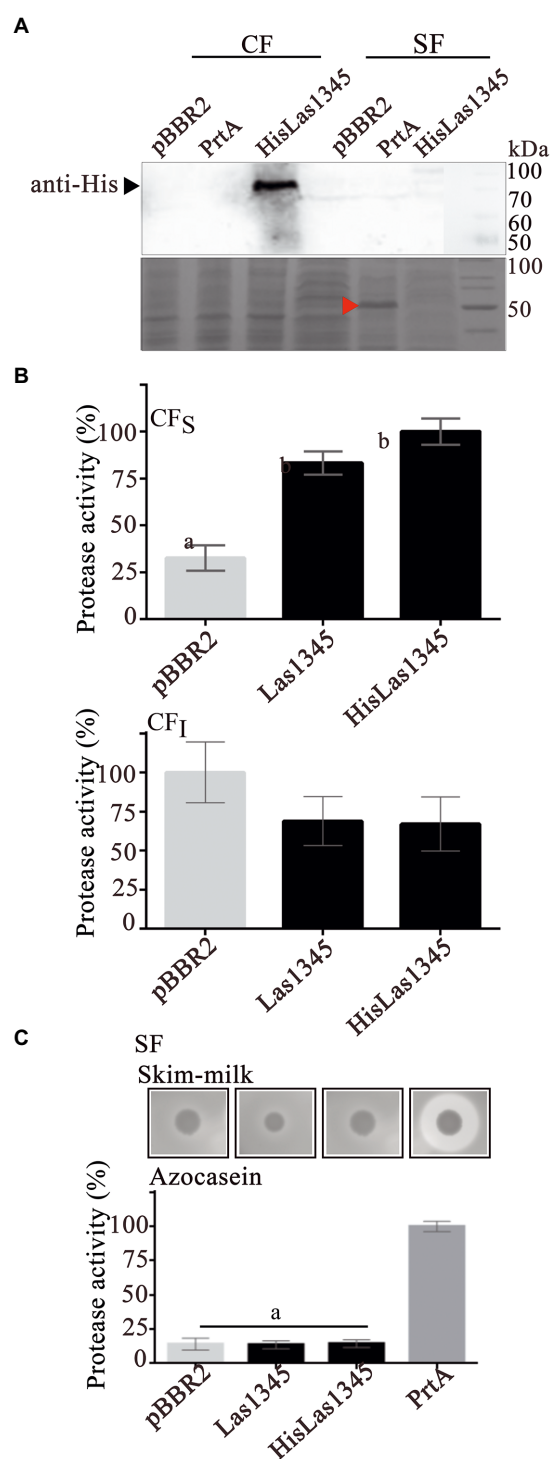
FIGURE 2

Las1345 has intracellular protease activity in *Rhizobium leguminosarum*. (A) Immunodetection of Las1345 (~75kDa) in cellular (CF) and supernatant cell free fraction (SF) from *Rhizobium leguminosarum* A34 (A34) cells expressing Las1345, HisLas1345 or transformed with the empty vector (pBBR2), grown for 16h at 30°C. *Rhizobium* adhering protein A1 (RapA1, ~36kDa) and Glucose-6-phosphate dehydrogenase (G6PD, ~55kDa) were used as extracellular and intracellular controls, respectively. (B) Protease activity in SF from A34 cultures (C), Protease activity in soluble and insoluble CF (CF<sub>S</sub> and CF<sub>I</sub>) from A34 cultures. The relative activity was expressed as the percentage of activity detected with respect to the maximum protease activity in the assay. Values are expressed as means ± standard deviations from six independent biological replicates. Different letters indicate significant differences at  $p < 0.05$  (one-way analysis of variance, Tukey's test).

To explore the effect that Las1345 expression could eventually have in plant tissues, *N. benthamiana* leaves were infiltrated with an *Agrobacterium* culture harboring 35S-directed untagged and C-terminal tagged versions of Las1345. No phenotypic change between *N. benthamiana* plants expressing Las1345 constructs (35S::Las1345GFP or 35S::Las1345) and the control (35S::GFP) were observed. Moreover, plant cell integrity and conductivity were not affected by Las1345 expression in planta (Figure 4A), suggesting that cellular membrane damage is not taking place in *N. benthamiana* leaves when Las1345 is

expressed. Fluorescence microscopy analyzes showed that Las1345GFP co-localizes with the plasma membrane dye FM 4-64, indicating that Las1345 is localized in these subcellular structures of the plant cell, particularly correlated with discrete cytosolic structures as a result of plasma membrane internalization (Figure 4B). Interestingly, no increase in protease activity was detected in Las1345-expressing tissues (Figure 4C). These results suggest that the intracellular environment in the leaf tissue is not favorable for the intrinsic proteolytic activity of Las1345.





**FIGURE 3**  
Las1345 protease activity in *Serratia marcescens*.  
(A) Immunodetection of Las1345 (~75kDa) in cellular (CF) and supernatant cell free fraction (SF) *prtA* cells from *S. marcescens* expressing Las1345, HisLas1345 or transformed with an empty vector (pBBR2), grown for 16h at 30°C. PrtA (~50kDa) was visualized by Coomassie staining (red arrow) and used as control of extracellular secretion (SF). (B) Protease activity in soluble and insoluble CF (CF<sub>s</sub> and CF<sub>i</sub>) from *prtA*/pBBR2, *prtA*/Las1345, *prtA*/

(Continued)

#### FIGURE 3 (CONTINUED)

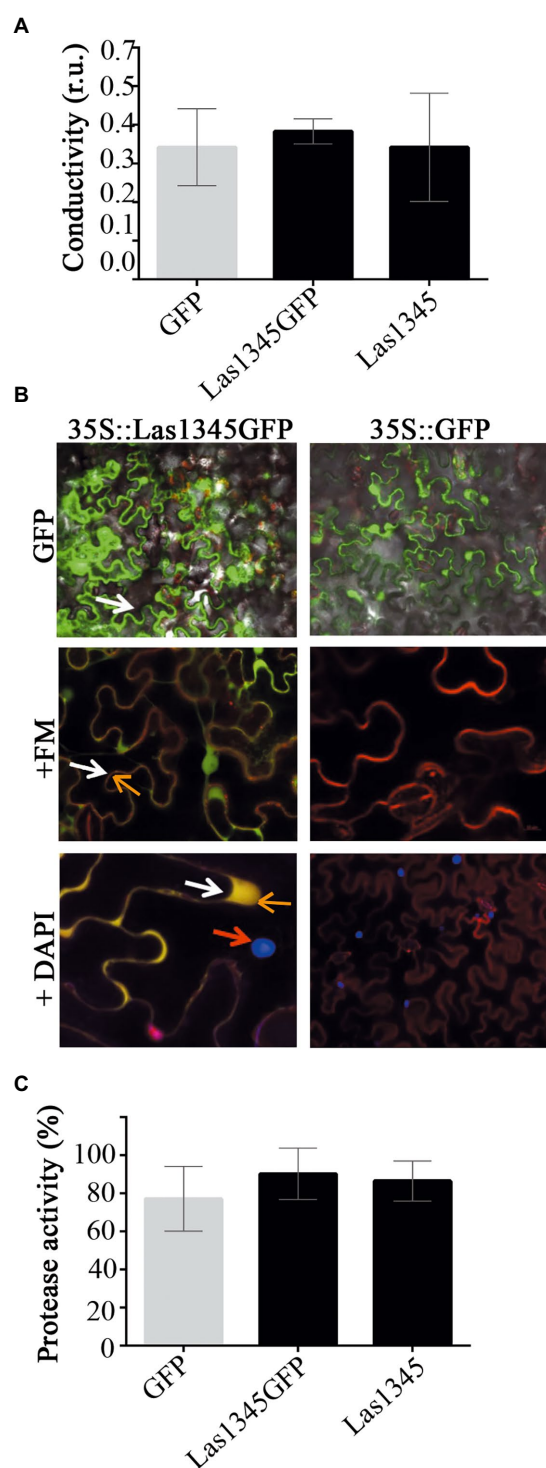
HisLas1345, and *prtA*/PrtA cells. The relative activity was expressed as the percentage of activity detected with respect to the maximum protease activity in the assay (C) Protease activity in SF from *prtA*/pBBR2, *prtA*/Las1345, *prtA*/HisLas1345, and *prtA*/PrtA cells was measured by detecting degradation of milk proteins (LB agar-skim milk plate), seen as a halo around the colony and by the azocaseinase assay. Values are expressed as means  $\pm$  standard deviations from six independent biological replicates. Different letters indicate significant differences at  $p < 0.05$  (one-way analysis of variance, Tukey's test).

### Las1345 affects cell motility and components of the extracellular matrix in the phytopathogen *Xanthomonas campestris* pv. *campestris*

Although growth was similar between Las1345-expressing *R. leguminosarum* and *S. marcescens* and the same bacterial strains transformed with empty pBBR2 plasmids, morphological characteristics of colonies on the surface of agar-solidified media were different (Supplementary Figure 2). Differences in colony phenotypes were also observed when Las1345 was expressed into the phytopathogenic bacteria *Xcc*, whereas again both control and Las1345-expressing bacteria exhibit similar growth kinetics. Las1345-expressing *Xcc* developed dull and dry colonies with slightly thin edges, characteristic of defective cell motility, as compared with the smooth and thick-edges colonies found in the control (Figure 5A). This phenotype resembles that of *Xanthomonas* mutants deficient in the production of xanthan, the main EPS produced by xanthomonads, which affect the extracellular matrix and biofilm development (Yun et al., 2006; Rigano et al., 2007; Torres et al., 2007; Malamud et al., 2011). The *Xcc*/Las1345 colony phenotype, together with the fact that the proteolytic activity was also detected in the CF of *Xcc*/Las1345 (Supplementary Figure 3), lead to the hypothesis that the proteolytic activity may influence the assembly or the structure of the extracellular matrix. Accordingly, significant differences in EPS production were observed between *Xcc*/pBBR2 and *Xcc*/Las1345 ( $8.66 \pm 3.04$  g/g wet cell and  $3.26 \pm 1.81$  g/g wet cell,  $p < 0.05$ ) when grown in liquid medium (Figure 5B). EPS xanthan contribute to cell motility, cell-to-cell aggregation, bacterial fitness, and virulence in many different *Xanthomonas* species (Yun et al., 2006; Rigano et al., 2007; Torres et al., 2007; Malamud et al., 2011, 2013; Yaryura et al., 2015; An et al., 2020). This prompted us to study the effect(s) of Las1345 expression in the extracellular matrix of *Xcc*.

Bacterial motility is mainly controlled by the flagella which is responsible for swimming and swarming (Nakamura and Minamino, 2019). Another type of cell motility, sliding, is a flagella-independent type motility in which xanthan acts as a surfactant or surface-wetting agent to facilitate the movement of bacterial cells (Murray and Kazmierczak, 2008; Malamud et al., 2011). Therefore, we hypothesized that the differential





**FIGURE 4**  
Las1345 expression in plant cells does not affect cell integrity. (A) Quantification of cell death in *N. benthamiana* leaves overexpressing GFP, Las1345-GFP, or Las1345 by measurement of conductivity at 3days post agroinfiltration (dpi). Values are expressed as mean  $\pm$  standard deviation of six samples. Each sample was obtained from three leaf disks of 1cm<sup>2</sup> collected randomly from different agroinfiltrated leaves. Data was analyzed by two-way analysis of variance and Tukey's test at  $p < 0.05$ . Each

(Continued)

#### FIGURE 4 (CONTINUED)

assay was repeated three times. Scale bar, 10 $\mu$ M. (B) *N. benthamiana* leaves expressing Las1345-GFP (green, white arrow) or GFP were imaged by confocal microscopy at 2days post infiltration (dpi). FM4-64 (FM) and DAPI were used to distinguish membrane (red, orange arrow) and nucleus (blue, red arrow), respectively. Co-localization of GFP and FM is shown in yellow. (C) Protease activity using azocasein as substrate in total plant protein extracts (50 $\mu$ g) from *N. benthamiana* leaves overexpressing GFP, Las1345GFP or Las1345 at 2dpi. Activity is expressed as percentage relative to the highest activity value obtained among samples. Values are expressed as mean  $\pm$  standard deviation of three samples. Each sample contains six disks from different agroinfiltrated leaves. Data was analyzed by two-way analysis of variance and Tukey's test at  $p < 0.05$ .

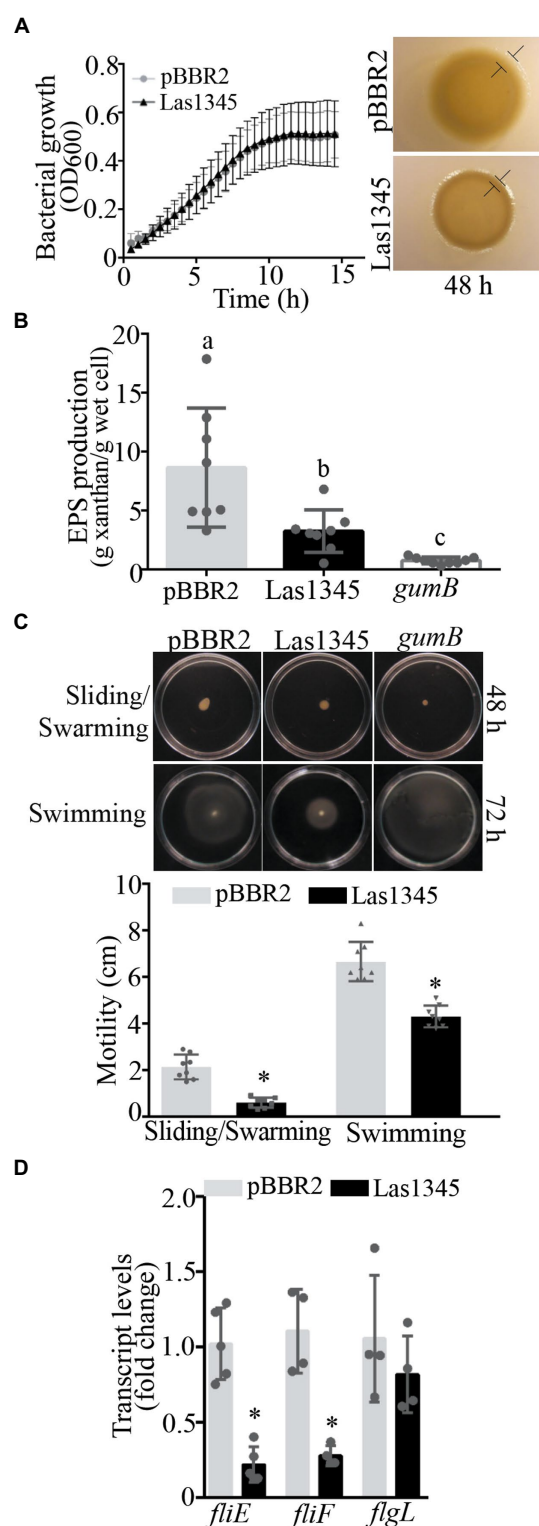
macrocolony phenotype found between *Xcc*/Las1345 and *Xcc*/pBRR2 cells may be associated with differences in motility either associated with the flagella or with the formation of the extracellular matrix. Motility of *Xcc*/Las1345 was assayed in 0.25/0.5% (w/v) agar NYGB over a 2/3-day period. Sliding and swarming motility of *Xcc* was reduced in Las1345-expressing cells at 48 h (pBRR2 2.14  $\pm$  0.53 cm vs. Las1345 0.61  $\pm$  0.21 cm,  $p < 0.0001$ , Student's *t*-test), in agreement with the reduced amount of EPS xanthan production compared with the *Xcc*/pBRR2 control (Figure 5C). A similar phenotype was observed in a *gumB* mutant of *Xanthomonas citri* subsp. *citri* (*X. citri*), as was reported by Malamud et al. (2011). Interestingly, *Xcc*/Las1345 also displayed very little swimming motility at 72 h (pBRR2 6.65  $\pm$  0.084 cm vs. Las1345 4.31  $\pm$  0.46 cm,  $p < 0.0001$ , Student's *t*-test). However, *Xcc gumB* mutant showed a swimming motility similar to the wild-type (*Xcc*/pBRR2; Figure 5C).

In accordance with these results, the relative expression of flagellum assembly-related genes, such as *fliE* and *fliF* (Kan et al., 2018) was significant diminished in *Xcc*/Las1345 compared with control (*Xcc*/pBRR2). In contrast, the expression levels of *flgL* -associated with the hook structure- were similar between Las1345-expressing cells and the control (Figure 5D).

Taken together, all these results suggest that Las1345 expression influences the characteristics of the bacterial extracellular matrix by altering xanthan production, flagellum assembly and biofilm development.

## Las1345 interfere in cell adhesion, biofilm formation and the pathogenesis of *Xanthomonas campestris* pv. *campestris*

We have previously shown that GFP-labeled *Xcc* grown in chambered cover slides, develop after 2 days microcolonies showing intimate lateral interactions, and after 4 days a typical biofilm structure (Torres et al., 2007). Under identical culture conditions, we analyzed the *in vitro* biofilm tridimensional structure of a GFP-labeled strain of *Xcc* harboring either the pBRR2::Las1345 or the pBRR2 empty vector. After 1 day of culture,



**FIGURE 5**  
Las1345 expression reduce xanthan production and cell motility. (A) *Xanthomonas campestris* pv. *campestris* (Xcc) growth in NYGB medium. Values are expressed as means  $\pm$  standard deviation of five biological replicates. This assay was repeated five times. Macrocolony phenotype of Las1345-expressing Xcc and

(Continued)

#### FIGURE 5 (CONTINUED)

the control (Xcc/pBBR2) grown on 1% (w/v) agar NYGB medium at 28°C during 48h. Differences at boundary zone are indicated with T-markers. (B) Quantification of xanthan secreted in Las1345-expressing Xcc cultures grown for 20h at 28°C compared with control Xcc/pBBR2 and a xanthan deficient mutant (*gumB*). Values are expressed as means  $\pm$  standard deviation of eight biological replicates. Different letters indicate significant differences at  $p < 0.05$  (one-way analysis of variance, Tukey's test). This assay was repeated three times. (C) Xcc/Las1345, Xcc/pBBR2 and *gumB* cultures were normalized by OD<sub>600</sub> and inoculated on 0.5% (w/v) agar NYGB medium (sliding/swarming) and on 0.25% (w/v) agar NYGB medium (swimming). Pictures were taken after 48 or 72h of incubation at 28°C. Motility was measured as colony diameter (cm). Values are expressed as means  $\pm$  standard deviation of 10 biological replicates. (D) Quantitative reverse transcription (qPCR) analysis of flagellum assembly-related genes (*fliE*, *fliF*, and *flgL*). Fold change of RNA levels, normalized to 16s and relative to Xcc/pBBR2 is shown. Values are expressed as means  $\pm$  standard deviation (SD) of five samples (replicates). Asterisks indicates significant differences between Xcc/Las1345 and Xcc/pBBR2 at  $p < 0.05$  (Students *t*-test).

Xcc-GFP expressing Las1345 cells were flattened over the bottom and reached the highest z-stacks positions (Las1345; 11828.21 nm vs. pBBR2; 5522.50 nm) without any observable attachment as opposed to control cells (Xcc-GFP/pBBR2), which were grouped at the bottom (lower z-stack) and only contacted the glass surface via one cell pole (Figure 6). By day two, Xcc-GFP/Las1345 cells shaped a homogenous layer, where few cells are interconnected, forming small cellular aggregates separated by extensive water spaces. In contrast, Xcc-GFP/pBBR2 cells developed more complex structures with heterogeneous sizes and shapes, in which bacteria were interconnected side-by-side (Figure 6). Biofilm structure parameters, such as biomass (volume per unit area) and roughness coefficient (biofilm heterogeneity) were analyzed by COMSTAT2. Higher biomass levels were observed in Xcc-GFP/Las1345 expressing Las1345, indicating an increased cell number per slide when compared with Xcc-GFP/pBBR2. The changing in the spatial biomass distribution can be associated with alterations in the hydrodynamic of Las1345-induced biofilm matrix that may result in enhanced availability of nutrient to the cells. In agree with this, the roughness coefficient was increased in the control (Xcc-GFP/pBBR2), showing an intricate biofilm matrix forming channel-like structures that may promote a hydrodynamic biofilm (Figure 6). Similar biofilm structures were also observed in the *gumB* mutant, as reported previously (Torres et al., 2007). Moreover, by day four of growth, structural changes on Las1345-expressing Xcc biofilms were associated with differences on the extracellular matrix disposition, surrounding cell-to-cell contacts, as shown by congo red (red) fluorescence pattern (Figure 6). Red-fluorescence signals were barely detected over the highest green network developed by Xcc-GFP/Las1345, suggesting that other extracellular matrix components, such as amyloid fibers or adhesins, were modified by Las1345 expression. In contrast, cultures of Xcc-GFP/pBBR2 cells showed extensive red areas at the cell boundaries (Figure 6). These results suggest that Las1345

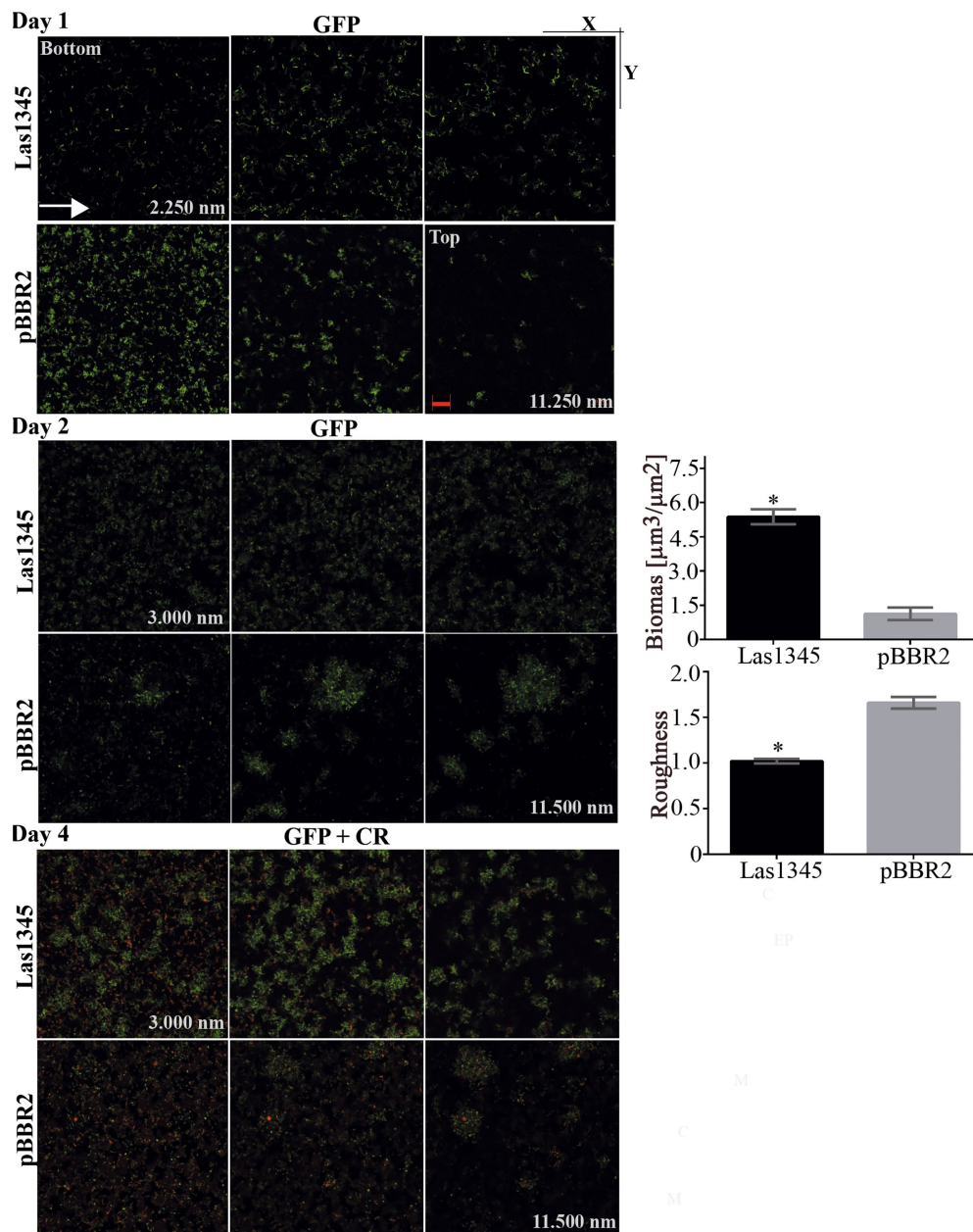


FIGURE 6

Biofilm architecture of *Xanthomonas campestris* pv. *campestris* (Xcc) is disrupted when Las1345 is expressed. GFP-labeled Xcc cells (Xcc-GFP) expressing Las1345 (Xcc-GFP/Las1345) were grown in 8-well chambers with a 1-mm thick borosilicate glass containing Y minimal medium and visualized at different stages of biofilm formation under confocal laser scanning microscopy for 4 days after inoculation at 28°C. Biofilm structure is shown as 2D-images of a single layer in the XY plane at different distances from the bottom well along the Z-axis (0–11898.21nm) at day 1, 2 and 4. Green channel shows Xcc-GFP/Las1345 (Las1345) or Xcc-GFP/pBBR2 (pBBR2) cells. At day 1, a detail (right corner) of the biofilm structure in Xcc-GFP/Las1345 and control cells is shown. At Day 2, Biomass and Roughness coefficient using COMSTAT 2.0 were calculated. Data is shown as means  $\pm$  standard deviation of three replicates. Asterisks indicate significant differences at  $p < 0.05$  (Student's *t*-test). At day 4, biofilm structure is showed by GFP (green channel) and Congo red (CR) emission (red channel) to indicate cells and exopolysaccharides production in the extracellular matrix, respectively. A detail of the biofilm structures is shown for both strains. C, GFP cells; EPS, exopolysaccharide; M, extracellular matrix. Scale bars, 20  $\mu\text{m}$ .

expression in Xcc increases cell aggregates by changing extracellular matrix components to remodel biofilm structures.

We hypothesize that Las1345 could degrade intracellular precursors of the extracellular matrix, or alter the secretion of proteins or EPS to modulate biofilm development, that eventually

would translate into a switch to a planktonic status, that subsequently would favor dispersal of bacteria from the biofilm structure.

The presence of plant-associated biofilms is correlated with pathogenicity in *Xanthomonas*. As previously reported



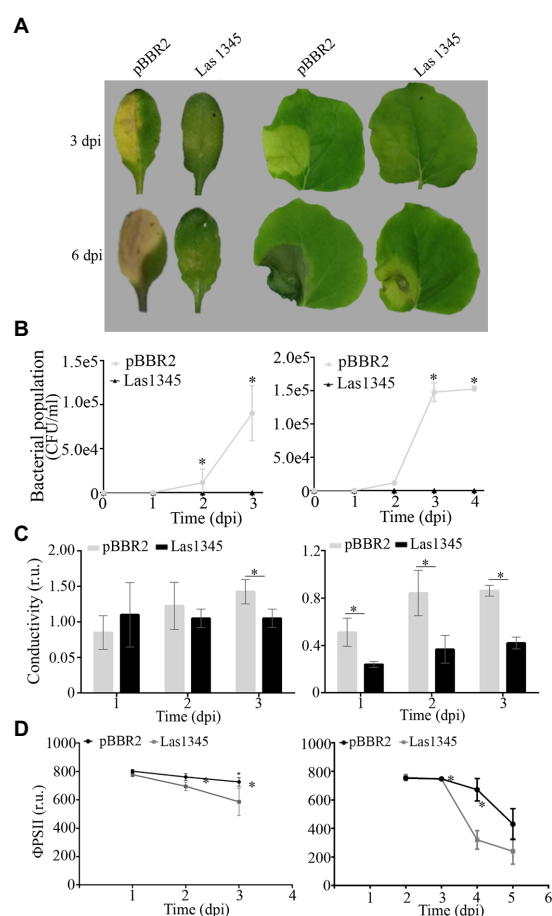


FIGURE 7

Biofilm modifications induced by Las1345 expression alters *Xanthomonas campestris* pv. *campestris* (Xcc) virulence. (A) Characterization of symptom development induced by Las1345-expressing Xcc in host plants. Infection of *Arabidopsis thaliana* and *Nicotiana benthamiana* with Xcc/Las1345 or Xcc/pBBR2 ( $10^7$  CFU/ml) using pressure infiltration. Photos of disease symptoms were taken at different days post inoculation (dpi). (B) Bacterial population of Xcc transformants in *N. benthamiana* and *A. thaliana* leaves. Values are expressed as means  $\pm$  standard deviation from five samples. Each sample was obtained from three leaf disks of 1 cm<sup>2</sup> from different inoculated leaves. (C) Quantification of cell death in leaves treated as described in (A) by measuring electrolyte leakage (conductivity) at different dpi. Values are expressed as means  $\pm$  standard deviation from six samples. Each sample contains three leaf disks of 1 cm<sup>2</sup> from 10 inoculated leaves. (D) Photosystem II quantum efficiency ( $\phi$ PSII) were measured on inoculated leaf. Values are expressed as means  $\pm$  standard deviation from 15 leaves. Asterisk indicates significant differences  $p < 0.05$  (Student's *t*-test).

(Yun et al., 2006; Rigano et al., 2007; Torres et al., 2007; Conforte et al., 2019), changes in both structure and quantity of EPS reduce virulence in host plants. To further understand whether the changes in Xcc biofilm structure as a consequence of Las1345 expression interfere with disease progression, leaves of *A. thaliana* and *N. benthamiana* were inoculated with bacterial suspensions of Xcc/Las1345 and Xcc/pBBR2. Disease symptoms (water-soaking and necrosis) were observed in leaves inoculated with Xcc/pBBR2 at 3 dpi. In contrast, the disease area was reduced in

leaves infected with Xcc/Las1345 (Figure 7A). Quantification of the bacterial population *in planta* revealed no significant differences between the Xcc/pBBR2 and Xcc/Las1345 strain growth until 36 or 48 h post inoculation (hpi), in *A. thaliana* or *N. benthamiana*, respectively (Figure 7B). However, Xcc/pBBR2 population gradually increased more than 10 times over the following monitoring days, while no Xcc/Las1345 could be recovered at soon as 3 dpi. Suppression of the Xcc-associated disease development was shown in *N. benthamiana* leaves infected with Xcc *gumB* mutant (Yun et al., 2006).

As compared with the areas inoculated with Xcc/pBBR2, Xcc/Las1345-infected areas had less plant cellular membrane damage, measured by electrolyte leakage, whereas photosystem II quantum efficiency ( $\phi$ PSII) was greater (Figures 7C,D), indicating that Las1345 expression compromised Xcc-associated disease development.

## Discussion

The identification of virulence factors that contribute to CLAs survival both in the plant phloem as well as in the psyllid vector is necessary to understand HLB disease. In this work, we approach the function of the CLAs serralyisin-like protein Las1345 using the close relatives *R. leguminosarum* and *S. marcescens* as surrogate models. Similar to CLso-serralyisins (Ravindran et al., 2018), Las1345 did not show extracellular proteolytic activity in any of these two systems. Transient expression of Las1345 in *N. benthamiana* leaves did not induce phenotypic changes or alterations in the cell membrane either, which would otherwise be an indicative of protease activity. These facts, together with the absence of proteolytic activity of the protein extract from Las1345-expressing *N. benthamiana* leaves, suggest that Las1345 may not act as a protease in the plant cell.

Las1345 and its ortholog in CLso were classified as serralyisin-like proteins mainly based on their N-terminal zinc metalloprotease domain. However, they have only one non-classical RTX motif (GxxGND) as compared to the four well-established consensus motifs (GGxGxD/N) found in PrtA from *S. marcescens* or Ser from *S. liquefaciens*. Multiple RTXs are needed to maintain an unfolded or loosely-folded state in RTX proteins that enables recognition of its secretion signal by T1SS (Linhartová et al., 2010). Upon extracellular translocation, RTX proteins normally bind Ca<sup>2+</sup>, in a process that promotes their folding and biological activity (Linhartová et al., 2010; Zhang et al., 2012, 2015; Baumann, 2019). Studies in *Rickettsia* suggests that the requirement for secretion is the presence of tandem repeats more than the calcium binding, as RTX-ankyrin proteins which do not possess calcium binding domains can still be T1SS substrates (Jernigan and Bordenstein, 2015). These observations indicate that Las1345 could still be a substrate of T1SS, despite not containing the canonical RTX-Ca<sup>2+</sup> binding motif (Linhartová et al., 2010). However, we could not detect secretion in two surrogate systems. Given the similarities between T1SSs of

Gram-negative including *CaL* spp., our current hypothesis is that Las1345 would not be secreted and would play a role in the bacterial cytoplasm. Additionally, Las1345 has an unusually large C-terminal tail (~200 aas) after the unique RTX motif that could also hinder secretion by T1SSs. In any case, secretion of Las1345 by CLAs cannot be totally discarded as secretion assays using this bacterium cannot be performed yet. Interestingly, intracellular localization and protease activity in *R. leguminosarum* and *S. marcescens* suggested a new role for Las1345 in pathogenesis.

Intracellular localization of Las1345 raises a question regarding a possible function for this protease in the bacterial cell. The higher expression of Las1345 in CLAs present in the citrus phloem as compared to that in the psyllid vector may be an adaptive response of this bacteria to life in every host. Las1345 might be a pathogenicity factor favoring plant colonization through the regulation of bacterial intracellular pathways. Interestingly, in bacterial surrogate models, expression of Las1345 form duller and drier microcolonies, compared with the bright and dome-shape control colonies. These changes were associated with alterations in the extracellular matrix. Biofilm is mainly composed of water (97%) and EPS including polysaccharides, lipopolysaccharides, nucleic acids, and proteins. Also, extracellular proteinaceous bacterial structures such as pili and flagella serve as structural elements that contribute to stabilize and strengthen the biofilm matrix (Limoli et al., 2015; Sena-Vélez et al., 2015). Las1345-expressing *Xcc* produces only 50% of the EPS xanthan found normally in *Xcc*. The flagellum-independent (sliding) motility is modulated by the EPS production as was shown in the xanthan deficient *gumB* mutant (Malamud et al., 2011). Here, flagella-dependent motility (swimming and swarming) is also reduced in Las1345-expressing cells. However, the *gumB* mutant showed a swimming motility similar to that of the wild type, suggesting xanthan dependent and independent modulation of biofilm for Las1345. Swarming motility is associated with the expression of flagella-associated genes in *X. citri* (Malamud et al., 2011). CLAs genome contains most of the known flagellar genes (Duan et al., 2009; Andrade and Wang, 2019). Some of them are expressed in CLA-infected leaves (*fliF*, *flgI*, *flgD* and *motB*) and others, particularly those involved in the hook structure (*flgK* and *flgL*) are preferentially expressed in the psyllid (Yan et al., 2013). A similar pattern was observed with genes involved in adherence. The Tad (tight adherence) pilin *flp3* gene was highly induced in psyllids and its expression is under the control of two regulators belonging to the LuxR transcriptional factor family. CLAs LuxR regulators complement biofilm and motility deficiencies in a surrogate model (Andrade and Wang, 2019). Despite this pattern of expression, CLAs does not produce flagella in the phloem (Yan et al., 2013; Andrade et al., 2020). The lower expression of flagellar-related genes in *Xcc*/Las1345 suggest a function for Las1345 to support the flagellum inhibition of CLAs in the phloem, to avoid triggering of plant defense response. Moreover, the passive movement of CLAs following the phloem sap way is the dominant bacterial movement to colonize the citrus tree (Pandey et al., 2021; Raiol-Junior et al., 2021).

Inefficient flagella assembly provokes a reduction of colony size in different mesophyll cell-infecting pathogens (Malamud et al., 2013; Berleman et al., 2016). On the other hand, flagellated phytopathogens suppress flagella expression in plants, to avoid plant defense responses (Chatnaparat et al., 2016). *Pseudomonas syringae* utilizes AprA, an extracellular metalloprotease with five RTX motifs, to degrade flagellin monomers in the extracellular milieu to avoid PTI induction (Bardoel et al., 2011; Pel et al., 2014). Plant cell death, including callose deposition and induction of defense genes, was associated with transient expression of CLAs flagellin in *N. benthamiana* plants, suggesting that CLAs contains in its genome factors that can trigger a PTI response (Zou et al., 2012; Shi et al., 2018). Therefore, the modulation of the expression of these factors to avoid PTI is a plausible hypothesis.

CLAs cells were observed floating in the phloem-sap without attaching by biofilm structures (Kim et al., 2009; Hartung et al., 2010; Shokrollah et al., 2010; Hilf et al., 2013; Achor et al., 2020). In contrast, an EPS-like matrix surrounded individuals and clusters of CLAs cells was observed in its psyllid-vector (Cicero et al., 2016). Genome analysis indicates that CLAs has the ability to synthesize capsular polysaccharides and surface lipopolysaccharides, suggesting that *CaL* spp. are able to form biofilms (Cong et al., 2012; Wulff et al., 2014; Wang et al., 2017). *In vitro* biofilm formation was recently demonstrated in *L. crescens* (Lcr), the only cultured member of the genus (Fagen et al., 2014; Padgett-Pagliai et al., 2022). Lcr cell-aggregates were supported on a narrow polysaccharide matrix mainly composed by  $\beta$ -glucans stained with calcofluor white (Naranjo et al., 2019). The flagellum and EPS are also key elements that shape and provide structural support for bacterial biofilms in *Xanthomonas* spp. (Rigano et al., 2007; Torres et al., 2007; Malamud et al., 2011, 2013). Las1345-expressing *Xcc* established cell-to-cell contact without formation of a complex matrix which packs cells in channels-like structures. Similar biofilm modifications were observed in a xanthan-deficient mutant of *Xcc* (Torres et al., 2007). Additionally, modifications of biofilm conformation have been observed in mutants that synthesized structural variants of xanthan (Bianco et al., 2016). All these observations in a biofilm-forming bacteria like *Xanthomonas* suggest that CLAs could somehow develop or avoid biofilm structures according to its two biological niches, as an strategy to colonize and replicate in each host. In this regard, CLAs was grown in a biofilm and in the planktonic phase of the membrane biofilm reactor but co-cultured with another bacterial species, showing the need of CLAs to grow in a mutualistic relationship with other bacteria according to nutrient requirements (Ha et al., 2019).

Our results suggest a new hypothesis for CLAs pathogenesis where biofilm formation should be avoided *in planta* to favor CLAs distribution throughout the citrus phloem. Plant intracellular pathogens may inhibit biofilm formation also as a survival mechanism to avoid the host defense response. Future studies, focused on both structure and composition of extracellular matrix using Las1345-expressing *L. crescens* should contribute to understand the lifestyle of CLAs in its two hosts.



## Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding authors.

## Author contributions

Funding was acquired by MRM, JG, CG, and LG. The work was conceptualized by MRM, JG, and LG. LG, MCM, KP-P, and PT conducted the research. Data curation and formal analysis were performed by LG, MRM, and JG. RB and EGV provided mentorship and advice for research. Data curation of the published work was done by LG, MRM, and JG. All authors contributed to the article and approved the submitted version.

## Funding

This work was mainly supported by the Agencia Nacional de Promoción Científica y Tecnológica (PICT-2018-03051) to MRM and (PICT-2016-3108) to LG, by COST (European Cooperation in Science and Technology, COST Action) to JG and by the USDA National Institute of Food and Agriculture (Plant Biotic Interactions Program; award number 2017-03060) to CG.

## Acknowledgments

We thank Juan Pedro Agostini (Instituto Nacional de Tecnología Agropecuaria, Estación Experimental Agropecuaria

Montecarlo) for provide ethanolic samples of CLas-infected tissue and María Victoria Molino for her contribution with *S. marcescens*-activity assays. We also thank Diego Aguirre for technical assistance with plant growth and Rodrigo Venna for his assistance to perform confocal microscopy and with the analysis of biofilm images. We thank Angeles Zorreguieta and Adriana Krap for providing RapA1 and G6PD antibodies, respectively. We also thank Max Dow and Diego Serra for stimulating discussions and critical review of the manuscript.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.1006962/full#supplementary-material>

## References

- Achor, D., Welkerm, S., Ben-Mahmoud, S., Wangm, C., Folimonovam, S. Y., Dutt, M., et al. (2020). Dynamics of 'Candidatus Liberibacter asiaticus' movement and sieve-pore plugging in citrus sink cells. *Plant Physiol.* 182, 882–891. doi: 10.1104/pp.19.01391
- Alves, M. N., Cifuentes-Arenas, J. C., Raiol-Junior, L. L., Ferro, J. A., and Peña, L. (2021). Early population dynamics of 'Candidatus Liberibacter asiaticus' in susceptible and resistant genotypes after inoculation with infected *Diaphorina citri* feeding on young shoots. *Front. Microbiol.* 12:683923. doi: 10.3389/fmicb.2021.683923
- An, A.-Q., Potnis, N., Dow, J. M., Vorhölter, F.-J., He, Y.-Q., Becker, A., et al. (2020). Mechanistic insights into host adaptation, virulence and epidemiology of the phytopathogen *Xanthomonas*. *FEMS Microbiol. Rev.* 44, 1–32. doi: 10.1093/femsre/fuz024
- Andrade, M. O., Pang, Z., Achor, D. S., Wang, H., Yao, T., Singer, B. H., et al. (2020). The flagella of 'Candidatus Liberibacter asiaticus' and its movement in planta. *Mol. Plant Pathol.* 21, 109–123. doi: 10.1111/mpp.12884
- Andrade, M. O., and Wang, N. (2019). The tad pilus apparatus of 'Candidatus Liberibacter asiaticus' and its regulation by Vis NR. *Mol. Plant-Microbe Interact.* 32, 1175–1187. doi: 10.1094/MPMI-02-19-0052-R
- Bardoel, B. W., van der Ent, S., Pel, M. J., Tommassen, J., Pieterse, C. M., and van Kessel, et al., (2011). *Pseudomonas* evades immune recognition of flagellin in both mammals and plants. *PLoS Pathog.* 7:e1002206. doi: 10.1371/journal.ppat.1002206
- Barnett, M. J., Solow-Cordero, D. E., and Long, S. R. (2019). A high-throughput system to identify inhibitors of 'Candidatus Liberibacter asiaticus' transcription regulators. *Proc. Natl. Acad. Sci. U. S. A.* 116, 18009–18014. doi: 10.1073/pnas.1905149116
- Bassanezi, R. B., Lopes, S. A., de Miranda, M. P., Wulff, N. A., Volpe, H. X. L., and Ayres, A. J. (2020). Overview of citrus Huanglongbing spread and management strategies in Brazil. *Trop. Plant Pathol.* 45, 251–264.
- Baumann, U. (1994). Crystal structure of the 50 kDa metallo protease from *Serratia marcescens*. *J. Mol. Biol.* 242, 244–251. doi: 10.1006/jmbi.1994.1576
- Baumann, U. (2019). Structure–function relationships of the repeat domains of RTX toxins. *Toxins* 11:657. doi: 10.3390/toxins11110657
- Berleman, J. E., Zemla, M., Remis, J. P., Liu, H., Davis, A. E., Worth, A. N., et al. (2016). Exopolysaccharide microchannels direct bacterial motility and organize multicellular behavior. *ISME J.* 10, 2620–2632. doi: 10.1038/ismej.2016.60
- Bianco, M. I., Toum, L., Yaryura, P. M., Mielnichuk, N., Gudesblat, G. E., Roeschlin, R., et al. (2016). Xanthan pyruvilation is essential for the virulence of *Xanthomonas campestris* pv. *Campestris*. *Mol. Plant Microbe Interact* 29, 688–699. doi: 10.1094/MPMI-06-16-0106-R
- Boava, L. P., Cristofani-Yaly, M., and Machado, M. (2017). Physiologic, anatomic, and gene expression changes in *citrus sunki*, *Poncirus trifoliata* and their hybrids after *Liberibacter asiaticus* infection. *Phytopathology* 107, 590–599. doi: 10.1094/PHYTO-02-16-0077-R
- Bové, J. M. (2006). Huanglongbing: a destructive, newly-emerging, century-old disease of citrus. *J. Plant Pathol.* 88, 7–37.

- Bruna, R. E., Molino, M. V., Lazzaro, M., Mariscotti, J. F., and Garcia Vescovi, E. (2018). Cpx R-dependent thermoregulation of *Serratia marcescens* Prt a metalloprotease expression and its contribution to bacterial biofilm formation. *J. Bacteriol.* 200, e00006–e00018. doi: 10.1128/JB.00006-18
- Chaban, B., Hughes, H. V., and Beeby, M. (2015). The flagellum in bacterial pathogens: for motility and a whole lot more. *Semin. Cell Dev. Biol.* 46, 91–103. doi: 10.1016/j.semcdb.2015.10.032
- Chatnapat, T., Prathuangwong, S., and Lindow, S. E. (2016). Global pattern of gene expression of *Xanthomonas axonopodis* pv. *Glycines* within soybean leaves. *Mol. Plant Microbe Interact.* 29, 508–522. doi: 10.1094/MPMI-01-16-0007-R
- Chen, Q., Min, A., Luo, S., He, J., Wu, R., Lin, X., et al. (2022). Metabolomic analysis revealed distinct physiological responses of leaves and roots to Huanglongbing in a citrus rootstock. *Int. J. Mol. Sci.* 23:9242. doi: 10.3390/ijms23169242
- Chiesa, M. A., Roeschlin, R. A., Favaro, M. A., Uviedo, F., Campos-Beneyto, L., and D'Andrea, R. y col. (2019). Plant responses underlying nonhost resistance of Citrus limon against *Xanthomonas campestris* pv. *campestris*. *Mol. Plant Pathol.* 20, 254–269. doi: 10.1111/mpp.12752
- Cicero, J. M., Fisher, T. W., and Brown, J. K. (2016). Localization of 'Candidatus Liberibacter solanacearum' and evidence for surface appendages in the potato psyllid vector. *Phytopathology* 106, 142–154. doi: 10.1094/PHYTO-04-15-0088-R
- Clark, K., Franco, J. Y., Schwizer, S., Pang, Z., Hawara, E., Liebrand, T. W. H., et al. (2018). An effector from the Huanglongbing-associated pathogen targets citrus proteases. *Nat. Commun.* 9:1718. doi: 10.1038/s41467-018-04140-9
- Clark, K. J., Pang, Z., Trinh, J., Wang, N., and Ma, W. (2020). Sec-delivered effector 1 (SDE1) of 'Candidatus Liberibacter asiaticus' promotes citrus Huanglongbing. *Mol. Plant-Microbe Interact.* 33, 1394–1404. doi: 10.1094/MPMI-05-20-0123-R
- Conforte, V. P., Yaryura, P. M., Bianco, M. I., Rodríguez, M. C., Daglio, Y., Prieto, E., et al. (2019). Changes in the physico-chemical properties of the xanthan produced by *Xanthomonas citri* subsp. *citri* in grapefruit leaf extract. *Glycobiology* 29, 269–278. doi: 10.1093/glycob/cwy114
- Cong, Q., Kinch, L. N., Kim, B. H., and Grishin, N. V. (2012). Predictive sequence analysis of the 'Candidatus Liberibacter asiaticus' proteome. *PLoS One* 7:e41071. doi: 10.1371/journal.pone.0041071
- Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* 16, 10881–10890. doi: 10.1093/nar/16.22.10881
- Curtis, M. D., and Grossniklaus, U. (2003). A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol.* 133, 462–469. doi: 10.1104/pp.103.027979
- Da Graça, J. V., Douhan, G. W., Halbert, S. E., Keremane, M. L., Lee, R. F., Vidalakis, G., et al. (2016). Huanglongbing: An overview of a complex patosystem ravaging the world's citrus. *J. Integr. Plant Biol.* 58, 373–387. doi: 10.1111/jipb.12437
- Daniels, M., Barber, C., Turner, P., Sawczyk, M., Byrde, R., and Fielding, A. (1984). Cloning of genes involved in pathogenicity of *Xanthomonas campestris* pv. *Campestris* using the broad host range cosmid pLAFRI. *EMBO J.* 3, 3323–3328. doi: 10.1002/j.1460-2075.1984.tb02298.x
- Di Rienzo, J. A., Casanoves, F., Balzarini, M. G., González, L., Tablada, M., and Robledo, C. W. (2017). Info stat version 2017. Grupo info stat, FCA, Universidad Nacional de Córdoba, Argentina. Available at: <http://www.infostat.com.ar> (Accessed September 29, 2020).
- Dow, J. M., Clarke, B. R., Milligan, D. E., Tang, J. L., and Daniels, M. J. (1990). Extracellular proteases from *Xanthomonas campestris* pv. *Campestris*, the black rot pathogen. *Appl. Environ. Microbiol.* 56, 2994–2998. doi: 10.1128/aem.56.10.2994-2998.1990
- Dow, J. M., Fan, M. J., Newman, M. A., and Daniels, M. J. (1993). Differential expression of conserved protease genes in crucifer-attacking pathovars of *Xanthomonas campestris*. *Appl. Environ. Microbiol.* 59, 3996–4003.
- Du, P., Zhang, C., Zou, X., Zhu, Z., Yan, H., Wuriyangan, H., et al. (2021). 'Candidatus Liberibacter asiaticus' secretes nonclassically secreted proteins that suppress host hypersensitive cell death and induce expression of plant pathogenesis-related proteins. *Appl. Environ. Microbiol.* 87, e00019–e00021. doi: 10.1128/AEM.00019-21
- Duan, Y., Zhou, L., Hall, D. G., Li, W., Doddapaneni, H., Lin, H., et al. (2009). Complete genome sequence of citrus huanglongbing bacterium, 'Candidatus Liberibacter asiaticus' obtained through metagenomics. *Mol. Plant-Microbe Interact.* 22, 1011–1020. doi: 10.1094/MPMI-22-8-1011
- Earley, K. W., Haag, J. R., Pontes, O., Oppen, K., Juehne, T., Song, K., et al. (2006). Gateway-compatible vectors for plant functional genomics and proteomics. *Plant J.* 45, 616–629. doi: 10.1111/j.1365-3113X.2005.02617.x
- Enrique, R., Siciliano, F., Favaro, M. A., Gerhardt, N., Roeschlin, R., Rigano, L., et al. (2011). Novel demonstration of RNAi in citrus reveals importance of citr
- callose synthase in defence against *Xanthomonas citri* subsp. *citri*. *Plant Biotechnol. J.* 9, 394–407. doi: 10.1111/j.1467-7652.2010.00555.x
- Etzeberria, E., Gonzalez, P., Achor, D., and Albrigo, G. (2009). Anatomical distribution of abnormally high levels of starch in HLB-affected Valencia orange trees. *Physiol. Mol. Plant Pathol.* 74, 76–83. doi: 10.1016/j.pmpp.2009.09.004
- Fagen, J. R., Leonard, M. T., Coyle, J. F., McCullough, C. M., Davis-Richardson, A. G., Davis, M. J., et al. (2014). *Liberibacter crescens* gen. Nov., sp. nov., the first cultured member of the genus *Liberibacter*. *Int. J. Syst. Evol. Microbiol.* 64, 2461–2466. doi: 10.1099/ij.s.0.063255-0
- Giró, M., Carrillo, N., and Krapp, A. R. (2006). Glucose-6-phosphate dehydrogenase and ferredoxin-NADP (H) reductase contribute to damage repair during the *sox RS* response of *Escherichia coli*. *Microbiology* 152, 1119–1128. doi: 10.1099/mic.0.28612-0
- Gottwald, T. R. (2010). Current epidemiological understanding of citrus Huanglongbing. *Annu. Rev. Phytopathol.* 48, 119–139. doi: 10.1146/annurev-phyto-073009-114418
- Granato, L. M., Galdeano, D. M., Nathália Da Roz, D. A., Breton, M. C., and Machado, M. A. (2019). Callose synthase family genes play an important role in the citrus defense response to 'Candidatus Liberibacter asiaticus'. *Eur. J. Plant Pathol.* 155, 25–38. doi: 10.1007/s10658-019-01747-6
- Ha, P. T., He, R., Killiny, N., Brown, J. K., Omsland, A., Gang, D. R., et al. (2019). Host-free biofilm culture of 'Candidatus Liberibacter asiaticus', the bacterium associated with Huanglongbing. *Biofilms* 1:100005. doi: 10.1016/j.biofilm.2019.100005
- Hall, D. G., Albrecht, U., and Bowman, K. D. (2016). Transmission rates of 'Ca. Liberibacter asiaticus' by Asian citrus psyllid are enhanced by the presence and developmental stage of citrus flush. *J. Econ. Entomol.* 109, 558–563. doi: 10.1093/jee/tow009
- Hartung, J. S., Paul, C., Achor, D., and Bransky, R. H. (2010). Colonization of dodder, *Cuscuta indecora*, by 'Candidatus Liberibacter asiaticus' s and *ca. L. americanus*. *Phytopathology* 100, 756–762. doi: 10.1094/PHYTO-100-8-0756
- Heydorn, A., Nielsen, A. T., Hentzer, M., Sternberg, C., Givskov, M., Ersbøll, B. K., et al. (2000). Quantification of biofilm structures by the novel computer program COMSTAT. *Microbiol. Read. Engl.* 146, 2395–2407.
- Hilf, M. E., Sims, K. R., Folimonova, S. Y., and Achor, D. S. (2013). Visualization of 'Candidatus Liberibacter asiaticus' cells in the vascular bundle of citrus seed coats with fluorescence in situ hybridization and transmission electron microscopy. *Phytopathology* 103, 545–554. doi: 10.1094/PHYTO-09-12-0226-R
- Hui, X., Chen, Z., Zhang, J., Lu, M., Cai, X., Deng, Y., et al. (2021). Computational prediction of secreted proteins in gram-negative bacteria. *Comput. Struct. Biotechnol. J.* 19, 1806–1828. doi: 10.1016/j.csbj.2021.03.019
- Ishii, K., Adachi, T., Hamamoto, H., and Sekimizu, K. (2014). *Serratia marcescens* suppresses host cellular immunity via the production of an adhesion-inhibitory factor against immunosurveillance cells. *J. Biol. Chem.* 289, 5876–5888. doi: 10.1074/jbc.M113.544536
- Jain, M., Cai, L., Fleites, L., Munoz-Bodnar, A., Davis, M., and Gabriel, D. (2019). *Liberibacter crescens* is a cultured surrogate for functional genomics of uncultured pathogenic *Candidatus Liberibacter* spp. and is naturally competent for transformation. *Phytopathology* 109, 1811–1819. doi: 10.1094/PHYTO-04-19-0129-R
- Jernigan, K. K., and Bordenstein, S. R. (2015). Tandem-repeat protein domains across the tree of life. *Peer J.* 3:e732. doi: 10.7717/peerj.732
- Johnson, E. G., Wu, J., Bright, D. B., and Graham, J. H. (2014). Association of 'Candidatus Liberibacter asiaticus' root infection, but not phloem plugging with root loss on Huanglongbing-affected trees prior to appearance of foliar symptoms. *Plant Pathol.* 63, 290–298. doi: 10.1111/ppa.12109
- Kan, J., An, L., Wu, Y., Long, J., Song, L., Fang, R., et al. (2018). A dual role for proline iminopeptidase in the regulation of bacterial motility and host immunity. *Mol. Plant Pathol.* 19, 2011–2024. doi: 10.1111/mpp.12677
- Kim, J. S., Sagaram, U. S., Burns, J. K., Li, J. L., and Wang, N. (2009). Response of sweet orange (*Citrus sinensis*) to 'Candidatus Liberibacter asiaticus' infection: microscopy and microarray analyses. *Phytopathology* 99, 50–57. doi: 10.1094/PHYTO-99-1-0050
- Koh, E. J., Zhou, L., Williams, D. S., Park, J., Ding, N., Duan, Y. P., et al. (2012). Callose deposition in the phloem plasmodesmata and inhibition of phloem transport in citrus leaves infected with 'Candidatus Liberibacter asiaticus'. *Protoplasma* 249, 687–697. doi: 10.1007/s00709-011-0312-3
- Kovach, M. E., Elzer, P. H., Hill, D. S., Robertson, G. T., Farris, M. A., Roop, R. M. 2nd, et al. (1995). Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* 166, 175–176. doi: 10.1016/0378-1119(95)00584-1
- Kraiselburd, I., Alet, A. I., Tondo, M. L., Petrocelli, S., Daurelio, L. D., Monzón, J., Ruiz, O. A., Losi, A., and Orellano, E. G. (2012). A LOV protein modulates the physiological attributes of *Xanthomonas axonopodis* pv. *Citri* relevant for host plant colonization. *PLoS One* 7:e38226. doi: 10.1371/journal.pone.0038226, 7.

- Kuykendall, L. D., Shao, J. Y., and Hartung, J. S. (2012a). Conservation of gene order and content in the circular chromosomes of '*Candidatus Liberibacter asiaticus*' and other Rhizobiales. *PLoS One* 7:e34673. doi: 10.1371/journal.pone.0034673
- Kuykendall, L. D., Shao, J. Y., and Hartung, J. S. (2012b). '*Candidatus Liberibacter asiaticus*' proteins orthologous with pSymA-encoded proteins of *Sinorhizobium meliloti*: hypothetical roles in plant host interaction. *PLoS One* 7:e38725. doi: 10.1371/journal.pone.0038725
- Lee, D. J., Lee, J. B., Jang, H. A., Ferrandon, D., and Lee, B. L. (2017). An antimicrobial protein of the *Riptortus pedestris* salivary gland was cleaved by a virulence factor of *Serratia marcescens*. *Dev. Comp. Immunol.* 67, 427–433. doi: 10.1016/j.dci.2016.08.009
- Li, W., Cong, Q., Pei, J., Kinch, L. N., and Grishin, N. V. (2012). The ABC transporters in '*Candidatus Liberibacter asiaticus*'. *Proteins* 80, 2614–2628. doi: 10.1002/prot.24147
- Li, S., Wu, F., Duan, Y., Singerman, A., and Guan, Z. (2020). Citrus greening: management strategies and their economic impact. *Hort. Sci. Hort.* 55, 604–612. doi: 10.3389/fpls.2021.833328
- Limoli, D. H., Jones, C. J., Wozniak, D. J., and Cruz, S. (2015). Bacterial extracellular polysaccharides in biofilm formation and function. *Microbiol. Spectr.* 3, 1–30. doi: 10.1128/microbiolspec.MB-0011-2014
- Lin, H., Pietersen, G., Han, C., Read, D. A., Lou, B., Gupta, G., et al. (2015). Complete genome sequence of '*Candidatus Liberibacter africanus*', a bacterium associated with citrus Huanglongbing. *Genome Announc.* 22, 1011–1020. doi: 10.1094/mpmi-22-8-1011
- Linhardtová, I., Bumba, L., Mašín, J., Basler, M., Osička, R., and Kamanová, et al., (2010). RTX proteins: a highly diverse family secreted by a common mechanism. *FEMS Microbiol. Rev.* 34, 1076–1112. doi: 10.1111/j.1574-6976.2010.00231.x
- Lopes, S. A., and Cifuentes-Arenas, J. C. (2021). Protocol for successful transmission of '*Candidatus Liberibacter asiaticus*' from citrus to citrus using *Diaphorina citri*. *Phytopathology* 111, 2367–2374. doi: 10.1094/PHYTO-02-21-0076-R
- Malamud, F., Homem, R. A., Conforte, V. P., Yaryura, P. M., Castagnaro, A. P., Marano, M. R., et al. (2013). Identification and characterization of biofilm formation-defective mutants of *Xanthomonas citri* subsp. *citri*. *Microbiology* 159, 1911–1919. doi: 10.1099/mic.0.064709-0
- Malamud, F., Torres, P. S., Roeschlin, R., Rigano, L. A., Enrique, R., Bonomi, H. R., et al. (2011). The *Xanthomonas axonopodis* pv. *Citri* flagellum is required for mature biofilm and canker development. *Microbiology* 157, 819–829. doi: 10.1099/mic.0.044255-0
- Martinelli, F., and Dandekar, A. M. (2017). Genetic mechanisms of the devious intruder *Candidatus Liberibacter* in citrus. *Front. Plant Sci.* 8:904. doi: 10.3389/fpls.2017.00904
- Murray, T. S., and Kazmierczak, B. I. (2008). *Pseudomonas aeruginosa* exhibits sliding motility in the absence of type IV pili and flagella. *J. Bacteriol.* 190, 2700–2708. doi: 10.1128/JB.01620-07
- Nakamura, S., and Minamino, T. (2019). Flagella-driven motility of bacteria. *Biomol. Ther.* 9:279. doi: 10.3390/biom9070279
- Naranjo, E., Merfa, M. V., Ferreira, V., Jain, M., Davis, M. J., Bahar, O., et al. (2019). *Liberibacter* crescens biofilm formation in vitro: establishment of a model system for pathogenic *Candidatus Liberibacter* spp. *Sci. Rep.* 9:5150. doi: 10.1038/s41598-019-41495-5
- Oh, J., Kim, J. G., Jeon, E., Yoo, C. H., Moon, J. S., Rhee, S., et al. (2007). Amyloidogenesis of type III-dependent harpins from plant pathogenic bacteria. *J. Biol. Chem.* 282, 13601–13609. doi: 10.1074/jbc.M602576200
- Padgett-Pagliai, K. A., Pagliai, F. A., da Silva, D. R., Gardner, C. L., Lorca, G. L., and Gonzalez, C. F. (2022). Osmotic stress induces long-term biofilm survival in *Liberibacter crescens*. *BMC Microbiol.* 22:52. doi: 10.1186/s12866-022-02453-w
- Pagliai, F. A., Gardner, C. L., Bojilova, L., Sarnegrim, A., Tamayo, C., Potts, A. H., et al. (2014). The transcriptional activator Ldt R from '*Candidatus Liberibacter asiaticus*' mediates osmotic stress tolerance. *PLoS Pathog.* 10:e1004101. doi: 10.1371/journal.ppat.1004101
- Pagliai, F. A., Gardner, C. L., Pande, S. G., and Lorca, G. L. (2010). LVIS553 transcriptional regulator specifically recognizes novobiocin as an effector molecule. *J. Biol. Chem.* 285, 16921–16930. doi: 10.1074/jbc.M110.111138
- Pandey, S. S., Nogales Da Costa Vasconcelos, F., and Wang, N. (2021). Spatiotemporal dynamics of '*Candidatus Liberibacter asiaticus*' colonization inside citrus plant and Huanglongbing disease development. *Phytopathology* 111, 921–928. doi: 10.1094/PHYTO-09-20-0407-R
- Pang, Z., Zhang, L., Coaker, G., Ma, W., He, S. Y., and Wang, N. (2020). Citrus CsACD2 is a target of '*Candidatus Liberibacter asiaticus*' in Huanglongbing disease. *Plant Physiol.* 184, 792–805. doi: 10.1104/pp.20.00348
- Pel, M. J., van Dijken, A. J., Bardeol, B. W., Seidl, M. F., van der Ent, S., van Strijp, J. A., et al. (2014). *Pseudomonas syringae* evades host immunity by degrading flagellin monomers with alkaline protease Apr a. *Mol. Plant-Microbe Interact.* 27, 603–610. doi: 10.1094/MPMI-02-14-0032-R
- Pitino, M., Armstrong, C., Cano, L., and Duan, Y. (2016). Transient expression of '*Candidatus Liberibacter asiaticus*' effector induces cell death in *Nicotiana benthamiana*. *Front. Plant Sci.* 7:982. doi: 10.3389/fpls.2016.00982
- Prasad, S., Xu, J., Zhang, Y., and Wang, N. (2016). SEC-translocon dependent extracytoplasmic proteins of '*Candidatus Liberibacter asiaticus*'. *Front. Microbiol.* 7:1989. doi: 10.3389/fmicb.2016.01989
- Pulici, J. D. V. S., Murata, M. M., and Johnson, E. G. (2022). Early physiological plant response and systemic effects of Huanglongbing infection in split-root plants. *Phytopathology* 112, 1833–1843. doi: 10.1094/PHYTO-07-21-0293-FI
- Raiol-Junior, L. L., Cifuentes-Arenas, J. C., de Carvalho, E. V., Girardi, E. A., and Lopes, S. A. (2021). Evidence that '*Candidatus Liberibacter asiaticus*' moves predominantly toward new tissue growth in citrus plants. *Plant Dis.* 105, 34–42. doi: 10.1094/PDIS-01-20-0158-RE
- Ravindran, A., Saenkham, P., Levy, J., Tamborindeguy, C., Lin, H., Gross, D. C., et al. (2018). Characterization of the Serralsin-like gene of '*Candidatus Liberibacter solanacearum*' associated with potato zebra chip disease. *Phytopathology* 108, 327–335. doi: 10.1094/PHYTO-02-17-0064-R
- Rigano, L. A., Siciliano, F., Enrique, R., Sendin, L., Filippone, P., Torres, P. S., et al. (2007). Biofilm formation, epiphytic fitness, and canker development in *Xanthomonas axonopodis* pv. *Citri*. *Mol. Plant-Microbe Interact.* 20, 1222–1230. doi: 10.1094/MPMI-20-10-1222
- Rossi, E., Paroni, M., and Landini, P. (2018). Biofilm and motility in response to environmental and host-related signals in gram negative opportunistic pathogens. *J. Appl. Microbiol.* 125, 1587–1602. doi: 10.1111/jam.14089
- Russo, D. M., Williams, A., Edwards, A., Posadas, D. M., Finnie, C., Dankert, M., et al. (2006). Proteins exported via the Prs D-Prs E type I secretion system and the acidic exopolysaccharide are involved in biofilm formation by *rhizobium leguminosarum*. *J. Bacteriol.* 188, 4474–4486. doi: 10.1128/JB.00246-06
- Sánchez, G., Gerhardt, N., Siciliano, F., Vojnov, A., Malcuit, I., and Marano, M. R. (2010). Salicylic acid is involved in the Nb-mediated defense responses to potato virus X in *Solanum tuberosum*. *Mol. Plant-Microbe Interact.* 23, 394–405. doi: 10.1094/MPMI-23-4-0394
- Sena-Vélez, M., Redondo, C., Gell, I., Ferragud, E., Johnson, E., Graham, J. H., et al. (2015). Biofilm formation and motility of *Xanthomonas* strains with different citrus host range. *Plant Pathol.* 64, 767–775. doi: 10.1111/ppa.12311
- Serra, O. D., and Hengge, R. (2017) Experimental detection and visualization of the extracellular matrix in macrocolony biofilms. *Methods Mol. Biol.* 1657, 133–145. doi: 10.1007/978-1-4939-7240-1\_11
- Shi, Q., Febres, V. J., Zhang, S., Yu, F., McCollum, G., Hall, D. G., et al. (2018). Identification of gene candidates associated with Huanglongbing tolerance, using '*Candidatus Liberibacter asiaticus*' flagellin 22 as a proxy to challenge citrus. *Mol. Plant-Microbe Interact.* 31, 200–211. doi: 10.1094/MPMI-04-17-0084-R
- Shi, Q., Pitino, M., Zhang, S., Krystel, J., Cano, L. M., Shatters, R. G., et al. (2019). Temporal and spatial detection of '*Candidatus Liberibacter asiaticus*' putative effector transcripts during interaction with Huanglongbing-susceptible, -tolerant, and -resistant citrus hosts. *BMC Plant Biol.* 19:122. doi: 10.1186/s12870-019-1703-4
- Shokrollah, H., Abdullah, T. L., Sijam, K., and Abdullah, S. N. A. (2010). Ultrastructures of '*Candidatus Liberibacter asiaticus*' and its damage in Huanglongbing (HLB) infected citrus. *Afr. J. Biotechnol.* 9, 5897–5901.
- Siciliano, F., Torres, P., Sendin, L., Bermejo, C., Filippone, P., Vellice, G., et al. (2006). Analysis of the molecular basis of *Xanthomonas axonopodis* pv. *Citri* pathogenesis in *Citrus Limon*. *Electron. J. Biotechnol.* 9, 199–204. doi: 10.2225/vol9-issue3-fulltext-20
- Sievers, F., and Higgins, D. G. (2018). Clustal omega for making accurate alignments of many protein sequences. *Protein Sci.* 27, 135–145. doi: 10.1002/pro.3290
- Spitz, O., Erenburg, I. N., Beer, T., Kanonenberg, K., Holland, I. B., and Schmitt, L. (2019). Type I secretion systems-one mechanism for all? *Microbiol. Spectr.* 7:2. doi: 10.1128/microbiolspec
- Thapa, S. P., De Francesco, A., Trinh, J., Gurung, F. B., Pang, Z., Vidalakis, G., et al. (2020). Genome-wide analyses of *Liberibacter* species provides insights into evolution, phylogenetic relationships, and virulence factors. *Mol. Plant Pathol.* 21, 716–731. doi: 10.1111/mpp.12925
- Thomma, B. P., Nurnberger, T., and Joosten, M. H. (2011). Of PAMPs and effectors: the blurred PTI-ETI dichotomy. *Plant Cell* 23, 4–15. doi: 10.1105/tpc.110.082602
- Torres, P. S., Malamud, F., Rigano, L. A., Russo, D. M., Marano, M. R., Castagnaro, A. P., et al. (2007). Controlled synthesis of the DSF cell-cell signal is

required for biofilm formation and virulence in *Xanthomonas campestris*. *Environ. Microbiol.* 9, 2101–2109. doi: 10.1111/j.1462-2920.2007.01332.x

Toruño, M. T. Y., Stergiopoulos, I., and Coaker, G. (2016). Plant-pathogen effectors: cellular probes interfering with plant defenses in spatial and temporal. *Annu. Rev. Phytopathol.* 54, 419–441. doi: 10.1146/annurev-phyto-080615-100204

Vahling-Armstrong, C. M., Zhou, H., Benyon, L., Morgan, J. K., and Duan, Y. (2012). Two plant bacteria, *S. meliloti* and *ca. Liberibacter asiaticus*, share functional znu ABC homologues that encode for a high affinity zinc uptake system. *PLoS One* 7:e37340. doi: 10.1371/journal.pone.0037340

Vojnov, A. A., Zorreguieta, A., Dow, J. M., Daniels, M. J., and Dankert, M. A. (1998). Evidence for a role for the *gum B* and *gum C* gene products in the formation of xanthan from its pentasaccharide repeating unit by *Xanthomonas campestris*. *Microbiology* 144, 1487–1493. doi: 10.1099/00221287-144-6-1487

Vorregaard, M. (2008). Comstat2-A modern 3D image analysis environment for biofilms. Technical University of Denmark, DTU, DK-2800 Kgs. Lyngby, Denmark.

Vozza, N. F., Abdian, P. L., Russo, D. M., Mongiardini, E. J., Lodeiro, A. R., Molin, S., et al. (2016). A *rhizobium leguminosarum* CHDL- (cadherin-like-) lectin participates in assembly and remodeling of the biofilm matrix. *Front. Microbiol.* 7:1608. doi: 10.3389/fmicb.2016.01608

Wang, N., Pierson, E. A., Setubal, J. C., Xu, J., Levy, J. G., and Zhang, Y., J. et al. (2017). The *Candidatus Liberibacter*-host interface: insights into pathogenesis mechanisms and disease control. *Annu. Rev. Phytopathol.* 55: 451–482, doi:10.1146/annurev-phyto-080516-035513.

Wang, Y., Zhou, L., Yu, X., Stover, E., Luo, F., and Duan, Y. (2016). Transcriptome profiling of Huanglongbing (HLB) tolerant and susceptible citrus plants reveals the role of basal resistance in HLB tolerance. *Front. Plant Sci.* 7:933. doi: 10.3389/fpls.2016.00933

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., et al. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46, 296–303. doi: 10.1093/nar/gky427

Wulff, N. A., Zhang, S., Setubal, J. C., Almeida, N. F., Martins, E. C., Harakava, R., et al. (2014). The complete genome sequence of '*Candidatus Liberibacter americanus*', associated with citrus Huanglongbing. *Mol. Plant-Microbe Interact.* 27, 163–176. doi: 10.1094/MPMI-09-13-0292-R

Yan, Q., Sreedharan, A., Wei, S., Wang, J., Pelz-Stelinski, K., Folimonova, S., et al. (2013). Global gene expression changes in '*Candidatus Liberibacter asiaticus*' during the transmission in distinct hosts between plant and insect. *Mol. Plant Pathol.* 14, 391–404. doi: 10.1111/mppp.12015

Yan, X., Zhang, Q., Zou, J., He, C., and Tao, J. (2019). Selection of optimized reference genes for qRT-PCR normalization in *Xanthomonas campestris* pv. *campestris* cultured in different media. *Curr. Microbiol.* 76, 613–619. doi: 10.1007/s00284-019-01667-y

Yaryura, P. M., Conforte, V. P., Malamud, F., Roeschlin, R., de Pino, V., Castagnaro, A. P., et al. (2015). Xbm R, a new transcription factor involved in the regulation of chemotaxis, biofilm formation and virulence in *Xanthomonas citri* subsp. *citri*. *Environ. Microbiol.* 17, 4164–4176. doi: 10.1111/1462-2920.12684

Ying, X., Wan, M., Hu, L., Zhang, J., Li, H., and Lv, D. (2019). Identification of the virulence factors of '*Candidatus Liberibacter asiaticus*' via heterologous expression in *Nicotiana benthamiana* using tobacco mosaic virus. *Int. J. Mol. Sci.* 20:5575. doi: 10.3390/ijms20225575

Young, J. M., Kuykendall, L. D., Martínez-Romero, E., Kerr, A., and Sawada, H. (2001). A revision of *rhizobium* frank 1889, with an emended description of the genus, and the inclusion of all species of *agrobacterium* conn 1942 and *Allorhizobium undicola* de Lajudie et al., 1998 as new combinations: *rhizobium radiobacter*, *R. rhizogenes*, *R. rubi*, *R. undicola* and *R. vitis*. *Int. J. Syst. Evol. Microbiol.* 51, 89–103. doi: 10.1099/00207713-51-1-89

Yun, M. H., Torres, P. S., El Oirdi, M., Rigano, L. A., Gonzalez-Lamothé, R., Marano, M. R., et al. (2006). Xanthan induces plant susceptibility by suppressing callose deposition. *Plant Physiol.* 141, 178–187. doi: 10.1104/pp.105.074542

Zhang, L., Conway, J. F., and Thibodeau, P. H. (2012). Calcium-induced folding and stabilization of the *Pseudomonas aeruginosa* alkaline protease. *J. Biol. Chem.* 287, 4311–4322. doi: 10.1074/jbc.M111.310300

Zhang, L., Morrison, A. J., and Thibodeau, P. H. (2015). Interdomain contacts and the stability of serralsin protease from *Serratia marcescens*. *PLoS One* 10:e0138419. doi: 10.1371/journal.pone.0138419

Zou, H., Gowda, S., Zhou, L., Hajeri, S., Chen, G., and Duan, Y. (2012). The destructive citrus pathogen, '*Candidatus Liberibacter asiaticus*' encodes a functional flagellin characteristic of a pathogen-associated molecular pattern. *PLoS One* 7:e46447. doi: 10.1371/journal.pone.0046447

#### COPYRIGHT

© 2022 Garcia, Molina, Padgett-Pagliai, Torres, Bruna, García Vescovi, González, Gadea and Marano. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





## OPEN ACCESS

## EDITED BY

Graciela L. Lorca,  
University of Florida, United States

## REVIEWED BY

Ines Abatedaga,  
CONICET Institute of Bionanotechnology  
of NOA (INBIONATEC), Argentina  
Vinod K. Kannaujia,  
Banaras Hindu University, India

## \*CORRESPONDENCE

Wendy M. Schluchter  
wschluch@uno.edu

## SPECIALTY SECTION

This article was submitted to Microbial  
Physiology and Metabolism, a section of  
the journal Frontiers in Microbiology

RECEIVED 03 August 2022

ACCEPTED 12 October 2022

PUBLISHED 15 November 2022

## CITATION

Carrigee LA, Frick JP, Liu X, Karty JA,  
Trinidad JC, Tom IP, Yang X, Dufour L,  
Partensky F and Schluchter WM (2022) The  
phycoerythrobilin isomerization activity of  
MpeV in *Synechococcus* sp. WH8020 is  
prevented by the presence of a histidine at  
position 141 within its phycoerythrin-I  
 $\beta$ -subunit substrate.  
*Front. Microbiol.* 13:1011189.  
doi: 10.3389/fmicb.2022.1011189

## COPYRIGHT

© 2022 Carrigee, Frick, Liu, Karty, Trinidad,  
Tom, Yang, Dufour, Partensky and  
Schluchter. This is an open-access article  
distributed under the terms of the [Creative  
Commons Attribution License \(CC BY\)](#). The  
use, distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in  
this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# The phycoerythrobilin isomerization activity of MpeV in *Synechococcus* sp. WH8020 is prevented by the presence of a histidine at position 141 within its phycoerythrin-I $\beta$ -subunit substrate

Lyndsay A. Carrigee<sup>1,2</sup>, Jacob P. Frick<sup>1</sup>, Xindi Liu<sup>1</sup>, Jonathan A. Karty<sup>3</sup>, Jonathan C. Trinidad<sup>3</sup>, Irin P. Tom<sup>4</sup>, Xiaojing Yang<sup>4</sup>, Louison Dufour<sup>5</sup>, Frédéric Partensky<sup>5</sup> and Wendy M. Schluchter<sup>1\*</sup>

<sup>1</sup>Department of Biological Sciences, University of New Orleans, New Orleans, LA, United States,

<sup>2</sup>Environmental Laboratory, Engineering and Research Development Center, US Army Corps of Engineers, Vicksburg, MS, United States, <sup>3</sup>Department of Chemistry, Indiana University, Bloomington, IN, United States, <sup>4</sup>Department of Chemistry, University of Illinois Chicago, Chicago, IL, United States, <sup>5</sup>Ecology of Marine Plankton Team, UMR 7144 Adaptation and Diversity in the Marine Environment, Station Biologique, Sorbonne Université, CNRS, Roscoff, France

Marine *Synechococcus* efficiently harvest available light for photosynthesis using complex antenna systems, called phycobilisomes, composed of an allophycocyanin core surrounded by rods, which in the open ocean are always constituted of phycocyanin and two phycoerythrin (PE) types: PEI and PEII. These cyanobacteria display a wide pigment diversity primarily resulting from differences in the ratio of the two chromophores bound to PEs, the green-light absorbing phycoerythrobilin and the blue-light absorbing phycourobilin. Prior to phycobiliprotein assembly, bilin lyases post-translationally catalyze the ligation of phycoerythrobilin to conserved cysteine residues on  $\alpha$ - or  $\beta$ -subunits, whereas the closely related lyase-isomerases isomerize phycoerythrobilin to phycourobilin during the attachment reaction. MpeV was recently shown in *Synechococcus* sp. RS9916 to be a lyase-isomerase which doubly links phycourobilin to two cysteine residues (C50 and C61; hereafter C50, 61) on the  $\beta$ -subunit of both PEI and PEII. Here we show that *Synechococcus* sp. WH8020, which belongs to the same pigment type as RS9916, contains MpeV that demonstrates lyase-isomerase activity on the PEII  $\beta$ -subunit but only lyase activity on the PEI  $\beta$ -subunit. We also demonstrate that occurrence of a histidine at position 141 of the PEI  $\beta$ -subunit from WH8020, instead of a leucine in its counterpart from RS9916, prevents the isomerization activity by WH8020 MpeV, showing for the first time that both the substrate and the enzyme play a role in the isomerization reaction. We propose a structural-based mechanism for the role of H141 in blocking isomerization. More generally, the knowledge of the amino acid present at position 141 of the  $\beta$ -subunits may be used to



predict which phycobilin is bound at C50, 61 of both PEI and PEII from marine *Synechococcus* strains.

#### KEYWORDS

bilin lyase, cyanobacteria, chromatic acclimation, phycobilisome, phycoerythrin, phycoerythrobilin, phycourobilin, post-translational modification

## Introduction

Marine cyanobacteria are responsible for as much as half of the world's oxygen production and photosynthesis and play a key role in carbon and nutrient cycling (Bengston, 1994; Kasting and Siefert, 2003; Kehoe, 2010). Marine isolates of *Synechococcus* cyanobacteria possess huge light-harvesting complexes (or phycobilisome; hereafter PBS), comprised of up to four types of highly pigmented phycobiliproteins (PBPs). Allophycocyanin constitutes the core of the PBS, surrounded by 6–8 rods made of phycocyanin and up to two types of phycoerythrin (PEI and PEII; Figure 1; Ong and Glazer, 1991; Glazer, 1994; Everroad et al., 2006; Flombaum et al., 2013; Sanfilippo et al., 2019a). This variable PBP content extends the spectral range of the PBS light harvesting capabilities. PEI and PEII are homologous PBP, each composed of an  $\alpha$ - and a  $\beta$ -subunit arranged in a torus-like hetero-hexamers ( $\alpha\beta$ )<sub>6</sub> and stacked together with the help of linker polypeptides to form the distal portion of the PBS rods (Figure 1; Glazer, 1988; Betz, 1997; Schluchter et al., 2010).

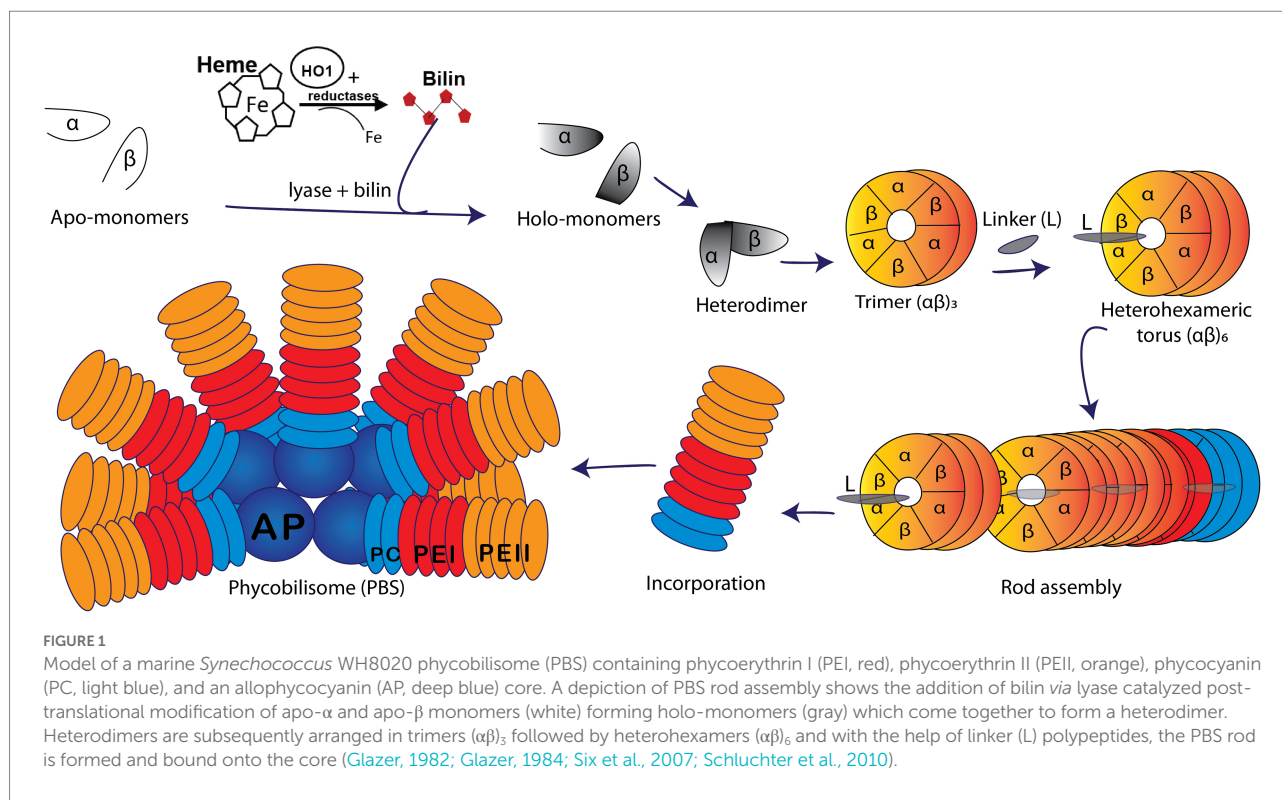
The large pigment diversity in marine strains of *Synechococcus* PBS is not only due to its variable PBP content but also to the variable composition of covalently bound linear tetrapyrrole bilins. The latter are post-translationally added to highly conserved cysteine (C) residues of PBP precursors by a variety of bilin lyases. Three major groups or clans of bilin lyases have been characterized to date: the S/U type, the T type, and the E/F type (Fairchild et al., 1992; Fairchild and Glazer, 1994; Shen et al., 2006; Zhao et al., 2007; Saunée et al., 2008). Bilin chromophore and attachment site specificity as well as primary amino acid sequence similarities are trademarks differentiating members of each clan. The crystal structures of the S/U [Protein Data Bank, (PDB): 3BDR; Zhao et al., 2006; Shen et al., 2008; Schluchter et al., 2010; Kronfel et al., 2013; Overkamp et al., 2014] and T (PDB: 4O4O; Zhou et al.,

2014; Gasper et al., 2017) lyases show that they adopt a similar antiparallel beta-barrel structure akin to the lipocalin protein family (Schluchter et al., 2010; Kronfel et al., 2013; Overkamp et al., 2014). The E/F lyases adopt all helical structures. While the crystal structure of the heterodimeric CpeE/F lyase was reported to adopt an  $\alpha$ -helical structure of a solenoid shape (Zhao et al., 2017), the recent crystal structure of a novel lyase-isomerase, MpeQ (Kumarapperuma et al., 2022), suggested that a question mark-like architecture represents a common protein framework for both single chain and heterodimeric E/F lyases (Andrade et al., 2001; Marcotrigiano et al., 2001; Kozo et al., 2002; Takano and Gusella, 2002; Scheer and Zhao, 2008; Schluchter et al., 2010; Zhao et al., 2017; Kumarapperuma et al., 2022). The S/U lyases exhibit a broad variation in chromophore and PBP substrate specificity but are highly specific with regard to C binding sites (Shen et al., 2006; Saunée et al., 2008; Scheer and Zhao, 2008; Biswas et al., 2010; Schluchter et al., 2010; Kronfel et al., 2013; Zhou et al., 2014; Gasper et al., 2017). In contrast, the E/F lyases display high specificities for the bilin chromophore and bilin binding site on a particular PBP (PDB 5N3U; Fairchild et al., 1992; Swanson et al., 1992; Zhou et al., 1992; Fairchild and Glazer, 1994).

Marine strains of *Synechococcus* containing both PEI and PEII possess a diverse set of E/F lyases (Wilbanks and Glazer, 1993; Six et al., 2007; Shukla et al., 2012), particularly in those strains that undergo Type IV chromatic acclimation (CA4). In the CA4 phenomenon, *Synechococcus* are able to alter the ratio between the blue light-absorbing chromophore phycourobilin (PUB;  $\lambda_{\max} \sim 495$  nm) and the green light-absorbing chromophore phycoerythrobilin (PEB;  $\lambda_{\max} \sim 545$  nm) in both PEI and PEII, thereby extending the light harvesting capabilities of PBS (Figures 1, 2; Palenik, 2001; Everroad et al., 2006; Kehoe, 2010; Humily et al., 2013; Sanfilippo et al., 2016, 2019a,b). This unique CA4 phenomenon is conferred by a small genomic island that occurs in two possible configurations, CA4-A or CA4-B, (Supplementary Figure 1; Humily et al., 2013). Each island encodes a transcriptional activator, a repressor, a protein of unknown function, and an E/F-type bilin lyase (PEB lyase MpeW in the CA4-B island) or lyase-isomerase (MpeZ in the CA4-A island; Supplementary Figure 1). A bilin lyase-isomerase is an enzyme that has an additional activity to isomerize PEB to PUB during the ligation reaction (Everroad et al., 2006; Shukla et al., 2012; Humily et al., 2013; Sanfilippo et al., 2016).

Intriguingly, all marine *Synechococcus* containing a CA4-A island also possess a specific member of the E/F clan, MpeV, which

Abbreviations: BL, Blue light; C, Cysteine residue; CA4, Type IV chromatic acclimation; CpeA/CpeB,  $\alpha$ -/ $\beta$ -subunit of phycoerythrin type I; EICs, Extracted ion chromatograms; GL, Green light; HT, Hexahistidine-tagged; H, Histidine; L, Leucine; MpeA/MpeB,  $\alpha$ -/ $\beta$ -subunit of phycoerythrin type II; MS, Mass spectrometry; MW, Molecular weight; PAGE, Polyacrylamide gel electrophoresis; PBP, Phycobiliprotein(s); PBS, Phycobilisome(s); PDB, Protein Data Bank; PEI, Phycoerythrin I; PEII, Phycoerythrin II; PEB, Phycoerythrobilin; PUB, Phycourobilin; RS, *Synechococcus* sp. RS9916; SDS, Sodium dodecyl sulfate; WH, *Synechococcus* sp. WH8020.



is not involved in the CA4 process (Carrigee et al., 2020b). The *mpeV* gene is located in a large genomic region containing a cluster of genes involved on the biosynthesis of PBS rods, so-called the “PBS genomic region” (Supplementary Figure 1; Wilbanks and Glazer, 1993; Grébert et al., 2021). We recently characterized MpeV from the model *Synechococcus* CA4-A strain RS9916 (hereafter RS9916). In RS9916, we showed that MpeV is integral in the ligation reaction of a doubly-linked PUB at the equivalent position (C50, 61) on the recombinant  $\beta$ -subunits of both PEI (CpeB) and PEII (MpeB), demonstrating that MpeV is a lyase-isomerase capable of attaching the doubly-linked PUB on both subunits (see Figure 2; Carrigee et al., 2020b). This study also showed that RS9916 MpeV activity requires ligation of PEB at C82 position on both CpeB and MpeB by CpeS (a S/U type lyase) and is enhanced by activity of the chaperone-like enzyme CpeZ (a member of the E/F clan; Kronfel et al., 2019a,b; Carrigee et al., 2020b). When all three enzymes (CpeS, CpeZ, and MpeV) from RS9916 were expressed, both CpeB and MpeB possessed a doubly-linked PUB at C50, 61 (Carrigee et al., 2020b).

It has been reported that in *Synechococcus* sp. WH8020 (hereafter WH8020) PEI and PEII, CpeB incorporates PEB while MpeB has PUB at the C50, 61 positions (Ong and Glazer, 1991). Like RS9916, WH8020 is also a CA4-A strain that possesses *mpeV* (Wilbanks and Glazer, 1993; Carrigee et al., 2020b; Grébert et al., 2022). To determine whether MpeV is responsible for ligating PEB on CpeB and PUB on MpeB in WH8020, we employed recombinant protein expression, absorbance and fluorescence spectroscopy, and tandem mass spectrometry. We found that a single amino acid substitution on CpeB, specifically a change from leucine (L) to histidine (H) at position 141, was sufficient to block

the isomerase activity of WH8020 MpeV. We further predict that a few other CA4-A strains containing a CpeB with the same substitution might have a similar bilin pattern, and, more generally, that the knowledge of the amino acid present at position 141 of CpeB or MpeB can be used to predict which phycobilin is bound at C50, 61 of both  $\beta$ -subunits from marine *Synechococcus* strains at large. We also propose a structural-based mechanism for the role of H141 in blocking isomerization.

## Materials and methods

### Plasmids for the characterization of WH8020 MpeV and CpeB

The putative lyase genes *cpeS*, *cpeZ*, and *mpeV* from the RS9916 genome (Supplementary Figure 1) were amplified via polymerase chain reaction (PCR) using a standard *Pfu* DNA polymerase system (ThermoFisher Scientific, Waltham, MA) and synthetic forward and reverse oligonucleotide primers with engineered restriction endonuclease sites (Eurofins MWG Operon, Huntsville, AL). Primers used to amplify genes by PCR for the construction of these expression vectors are previously published (Carrigee, 2020; Carrigee et al., 2020a,b). Amplified fragments were separately cloned into compatible Novagen Duet vectors using corresponding restriction enzymes as listed in Supplementary Tables 1, 2 or were previously published (Carrigee, 2020; Carrigee et al., 2020a,b). Expression vectors used in this study (Supplementary Table 2) include two previously described (Shukla et al., 2012; Kronfel et al., 2019b).

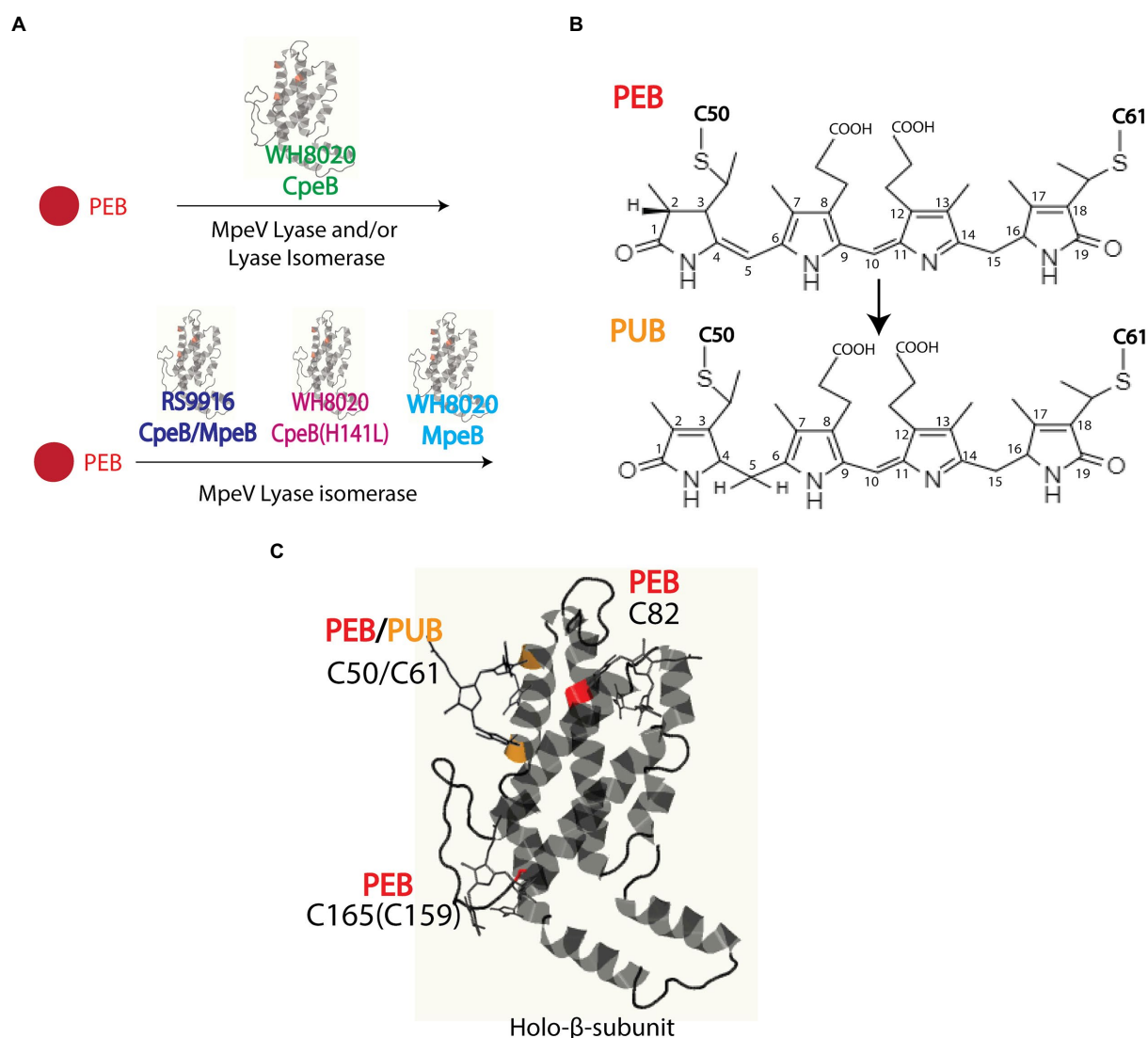


FIGURE 2

(A) Model of *Synechococcus* sp. WH8020 MpeV activity as both a lyase and lyase-isomerase. (B) Chemical structures of phycoerythrobilin (PEB) and doubly-linked phycourobilin (PUB). Post-translational pigment attachment is catalyzed by bilin lyases or lyase isomerases with a thioether linkage at the 3<sup>rd</sup> carbon of the bilin A ring during single attachment or additionally through the 18<sup>th</sup> carbon of the bilin D ring when doubly attached. (C) Bilin structure and ligation sites, representative of RS9916 and WH8020 CpeB. Ribbon diagram generated using Phyre<sup>2</sup> (Kelley et al., 2015), depicting β-subunit of PEI (β-PEI) chromophorylation pattern and lyases involved (bilins in ball and stick). PEB ligated residues are highlighted in red, and PUB ligated residues are highlighted in orange.

The *mpeV* gene from WH8020 was amplified *via* PCR using Platinum SuperFi II DNA Polymerase as a master mix in lieu of the *Pfu* system as previously described (Carrigee et al., 2020b). PCR reactions were performed using the standard High Fidelity PCR protocol from ThermoFisher Scientific (Waltham, MA). WH8020 substrate genes were cloned as an operon for PEI (*cpeBA*) with the beta subunit in frame with a hexahistidine tag (HT) and inserted into pET-DUET independently. WH8020 *mpeV* was cloned in-frame with the coding region of the HT and inserted into pCDF-DUET1. For co-expression of multiple genes required to characterize MpeV, some genes were subcloned into both multiple cloning sites

(MCSs) for compatibility as follows. RS9916 *cpeZ* was cloned using Platinum SuperFi PCR protocol and inserted in frame after the sequence encoding a HT into MCSI of pCOLA-Duet containing non-tagged (NT) *cpeS* in MCSII (Carrigee, 2020; Carrigee et al., 2020b). The RS9916 *cpeA* gene sequence was inserted into multiple cloning site I (MCSI) pET-Duet (Novagen, Madison, WI) in frame with the sequence encoding a hexahistidine tag (HT). *cpeB* was subsequently subcloned into MCSII to achieve (MCSI/MCSII/vector) RS9916 HT*cpeA*/HT*cpeB*/pET-DUET and RS9916 HT*mpeA*/HT*mpeB*/pET-DUET as previously described (Carrigee, 2020; Carrigee et al., 2020b).

A single-site variant of WH8020 CpeB was created by mutating the H141 residue to L (H141L) using combined overlap extension PCR method adapted from (Hussain and Chong, 2016) using the Platinum SuperFi protocol (Supplementary Table 1). A PCR fragment containing the entire mutant WH8020 *cpeBA* operon was cloned into MCS1 of pET-DUET vector containing resulting plasmids listed in Supplementary Table 2. All plasmids were sequenced for verification of clones and mutations (Eurofins Genomics LLC, Louisville, KY).

## *Escherichia coli* growth conditions and recombinant expression

Initial experiments for heterologous protein expression were performed using *E. coli* grown in Luria Bertani (LB) medium. However, we used modified, auto-induced medium for maximal protein yield (Studier, 2005). This involved a 100-ml LB starter culture of *E. coli* cells grown at 37°C overnight; then this was added to a liter of auto-induced medium composed of LB containing 1 mM MgSO<sub>4</sub>, 25 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 50 mM KH<sub>2</sub>PO<sub>4</sub>, 50 mM Na<sub>2</sub>HPO<sub>4</sub>, 0.5% glycerol, and 0.05% glucose at 18°C with appropriate combinations of antibiotics ampicillin (Ap: 100 µg·ml<sup>-1</sup>), chloramphenicol (Cm: 34 µg·ml<sup>-1</sup>), kanamycin (Km: 50 µg·ml<sup>-1</sup>), and spectinomycin (Sp: 100 µg·ml<sup>-1</sup>). With either medium, once the OD<sub>600nm</sub> reached 0.6, cultures were induced with 1 mM isopropyl 1-β-D-thiogalactopyranoside, after which the cells were allowed to grow at 18°C for an additional 24 h before being harvested by centrifugation at 11,000 ×g for 8 min in a Sorvall RC 5C Plus centrifuge (Kendro Laboratory Products, Newtown, CT). Cell pellets were stored at -20°C until ready for purification and analysis. The wet weight of all cell pellets was measured and recorded prior to storage at -20°C.

## Protein purification

The histidine-tagged (HT) proteins were purified as previously described (Carrigee et al., 2020b). Briefly, cell pellets were resuspended at 3.0 ml·g<sup>-1</sup> complete with mini protease cocktail (Thermo Scientific, Waltham, MA), 0.01 mg·ml<sup>-1</sup> lysozyme (Fisher Scientific, Hampton, NH), and passed through a French Pressure Cell Press at 18,000 psi three times. All samples purified by cobalt affinity chromatography, dialyzed to remove imidazole, and then concentrated by ultrafiltration through an Amicon Ultra centrifugal filter unit (10 kDa cutoff; Novagen/EMD Millipore Corp., Darmstadt, Germany) and stored at -20°C.

## Analysis of recombinant protein and bound bilin

Purified protein was quantified using Bradford colorimetric assay (BioRad, Hercules, CA) and diluted to obtain equal

concentrations across co-expressions for direct comparison when possible. Absorbance spectroscopy was performed using Perkin Elmer Lambda 35 UV/VIS or Shimadzu UV-2600 UV-Vis spectrophotometers followed by fluorescence spectroscopy using a Perkin Elmer LS55 (Waltham, MA) with excitation at 490 nm (PEB) or 440 nm (PUB; slit widths were set at 10 nm). Proteins were subsequently resolved by 15% (w/v) polyacrylamide gel electrophoresis (PAGE) in the presence of sodium dodecyl sulfate (SDS) and ultimately visualized by Coomassie blue staining (Sauté et al., 2008). To visualize proteins with bound bilin, gels were subjected to zinc-enhanced fluorescence using ChemiDoc MP imaging system (Bio-Rad, Hercules, CA) with excitation at 460–490 nm (PUB) and 520–545 nm (PEB).

## Growth of cyanobacterial strains

WH8020 cells were obtained from the Roscoff Culture collection.<sup>1</sup> Cultures of WH8020 were grown at 22°C in PCR-S11 media and acclimated for at least 7 days in either blue light (BL) or green light (GL) and PBS were collected as previously described (Sanfilippo et al., 2016; Mahmoud et al., 2017; Sanfilippo et al., 2019b). Fluorescence excitation spectra were recorded at an emission of 575 nm, using a Perkin Elmer LS-50B spectrofluorometer. The fluorescence excitation 495–545 nm ratio was calculated and used as a proxy for the molar PUB to PEB ratio, as described (Humily et al., 2013).

## HPLC separation of PBS, trypsin digestion, and liquid chromatography tandem mass spectrometry

PBS were purified using methods previously described (Shukla et al., 2012; Sanfilippo et al., 2016, 2019a). Samples were dialyzed overnight against 5 mM Na phosphate buffer (pH 7.0) and subsequently purified *via* high performance liquid chromatography (HPLC) using methods outlined in (Shukla et al., 2012; Sanfilippo et al., 2016). Phycobiliproteins were monitored from 210 to 700 nm with specific channels monitoring for total protein (280 nm), PUB (490 nm), and PEB (550 nm). Relevant fractions were vacuum-dried and kept at -20°C prior to digestion for mass spectrometric (MS) analysis as previously described (Biswas et al., 2011; Shukla et al., 2012; Sanfilippo et al., 2016). Purified proteins were dialyzed against 2 mM sodium phosphate buffer (pH 7.0) and 1 mM β-mercaptoethanol. One aliquot of trypsin (dimethylated trypsin from porcine pancreas; Sigma, St. Louis, MO) was added to 2% (w/w) from a 20 µg ml<sup>-1</sup> stock to the denatured protein mixtures and incubated at 30°C for 3 h in the dark (Shukla et al., 2012). The reaction was quenched by adding 30% (v/v) glacial acetic acid. Digested peptides were passed

<sup>1</sup> <https://www.roscoff-culture-collection.org/>



through a pre-equilibrated C8 Sep-Pak cartridge (Waters Corporation, Milford, MA), and thereafter the eluted sample was vacuum dried and stored at  $-80^{\circ}\text{C}$  before LC-MS<sup>2</sup> on a Thermo Orbitrap Fusion Lumos instrument. MS1 scans were obtained with a resolution of 120,000 and mass range of 300–2,000  $m/z$ . Data dependent HCD were acquired with a 3 s cycle time, quadrupole precursor isolation window of two 2  $m/z$  and resolution of 30,000 with 30% relative collision energy. Samples were separated by an Easy NanoLC1200 HPLC (ThermoFisher) equipped with a 75  $\mu\text{m} \times 15\text{ cm}$  Acclaim PepMap100 separating column (Thermo Scientific) downstream of a 2 cm guard column (Thermo Scientific). Buffer A was 0.1% formic acid in water. Buffer B was 0.1% formic acid in 80% acetonitrile. Peptides were separated on a 30 min gradient from 4% B to 33% B. All data processing was performed with Thermo XCalibur 4.0, Proteome Discover 2.1 (Thermo Scientific), and a local copy of ProteinProector 5.22.1 (prospector.ucsf.edu).

## Bioinformatics and structural modeling of proteins

Gene sequences from *Synechococcus* strains RS9916 and WH8020 were retrieved from Cyanorak (Garczarek et al., 2021). Amino acid sequences were analyzed using the ClustalW alignment tool from MacVector software V. 12.7.5 (MacVector Inc., Apex, NC) and Phyre<sup>2</sup> prediction system (Kelley et al., 2015). The model of the MpeV-CpeB complex was obtained using the AlphaFold2 multimer implemented in the ColabFold server based on the protein sequences of MpeV and CpeB from RS9916 (Jumper et al., 2021; Mirdita et al., 2022). The PEB chromophore was then manually docked into the MpeV/MpeB complex structure using Coot (Emsley and Cowtan, 2004).

## Results

### HPLC analyses of WH8020 PBS

Initially, we wanted to confirm that the bilin composition of PEI and PEII from wild type (WT) WH8020 PBS was as previously described (Ong and Glazer, 1991). Cells used in this previous study were isolated from cultures grown in white light, a light condition known to elicit a pigment phenotype equivalent to GL in CA4 strains (Ong and Glazer, 1991; Everroad and Wood, 2006; Humily et al., 2013). Here we isolated PBS from native WH8020 cells grown in two separate light conditions: GL for maximum PEB production and BL for maximum PUB production (Supplementary Figure 2). PBPs were then separated by HPLC with relative absorbance and chromatograms obtained using 550 nm (PEB) and 495 nm (PUB) for bilin content and 280 nm for total protein (Figures 3A–C; Shukla et al., 2012; Carrigee et al., 2020b). The numbered peaks in Figures 3D–G were collected, digested with trypsin, and analyzed by LC/MS/MS to identify the

proteins present in each peak. As shown in Table 1, peaks labeled as MpeA, CpeA, CpeB, and MpeB were verified by MS coverage (Table 1; Figure 3). As expected for this CA4-A strain, variable bilin content was detected among the  $\alpha$ -subunits depending on light quality (GL or BL), and the changes in PUB:PEB ratio observed in the spectra for CpeA and MpeA matched those documented for RS9916 during the CA4 process (Figures 3H,I; Kehoe, 2010; Shukla et al., 2012; Nguyen, 2018; Carrigee et al., 2020b). Unfortunately, we were not able to identify all bilins attached at each light condition for each protein by LC tandem MS, but we can infer the likely bilin content from our spectra and using the previous characterizations of WH8020 in white light (equivalent to GL; Ong and Glazer, 1991) and RS9916 in BL and GL (Shukla et al., 2012) as shown in Figure 3I.

As expected from previous work showing that the CA4 process affects only MpeA and CpeA (Shukla et al., 2012), no changes were observed in the MpeB and CpeB proteins in GL vs. BL (Figures 3E,G). As previously reported by Ong and Glazer for WH8020 grown in white light (Ong and Glazer, 1991) and for RS9916 in both GL and BL (Shukla et al., 2012), WH8020 MpeB binds a doubly-linked PUB at C50, 61 and two PEB (C159 and C82) in a 1:2 ratio, or a PUB:PEB of  $\sim 0.58$  independent of light condition (Figures 3E,G; Ong and Glazer, 1991). However, while RS9916 CpeB maintains bilin composition of one PUB (doubly-linked to C50, 61) and 2 PEB (C82 and C165) for CpeB regardless of growth conditions, HPLC analysis showed that WH8020 CpeB has PEB bound to all sites, including the doubly-linked C50, 61 position, consistent with what Ong and Glazer reported in 1991 (Ong and Glazer, 1991). The fraction collected for CpeB (Figure 3E, peak 2) only contains PEB (no PUB is detected) and no bilin content changes were detected in GL or BL.

### Comparative genomics analysis

MpeV was first suggested as a putative lyase by Wilbanks and Glazer, after sequencing of a large fraction of the PBS rod genomic region from WH8020 (Supplementary Figure 1; Wilbanks and Glazer, 1993). ClustalW analysis of protein sequences from RS9916 and WH8020 revealed 97.3% similarity between the CpeB substrates, 97.8% similarity between MpeB substrates, and 72.2% similarity between the MpeV homologs (Supplementary Figure 3). This led us to hypothesize that MpeV could be a lyase acting at C50, 61 on CpeB and perhaps also the lyase-isomerase acting on C50, 61 on MpeB in WH8020.

A comparison of *Synechococcus* CpeB sequences sorted by pigment type (PT) is shown in Supplementary Figure 4A [for review on PTs, see (Six et al., 2007; Humily et al., 2013; Grébert et al., 2022)]. All strains of *Synechococcus* belonging to PT2 (green light specialists with PEI and PEB only), PT3a (green light specialists with PEI, PEII, and a constitutively low PUB:PEB), about half of PT3dA (CA4-A) including WH8020, and the only representative of PT3eA (RCC307 that is genetically undistinguishable from typical PT3dA but exhibits



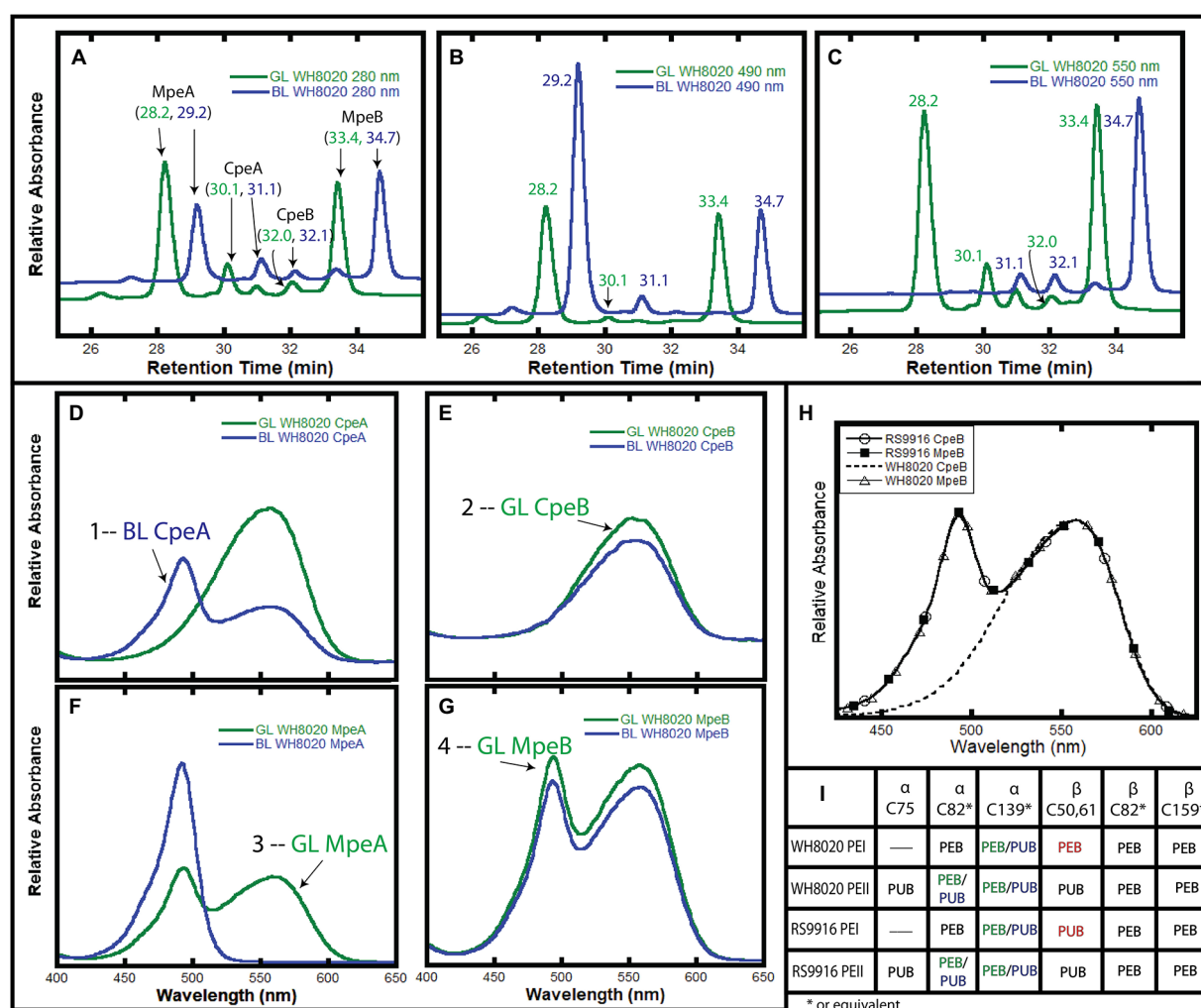


FIGURE 3

HPLC analysis of WH8020 PBS. (A–C). Chromatograms depicting relative absorbance of total protein (280nm), presence of PUB (490nm), and presence of PEB (560nm). (D–G) Relative absorbance spectra of phycoerythrin I (PEI) and II (PEII) phycobiliprotein subunits separated by HPLC and denoted as follows: α-subunit of PEI (CpeA), β-subunit of PEI (CpeB), α-subunit of PEII (MpeA), β-subunit of PEII (MpeB). PBS extracted from WH8020 cells grown in an abundance of green light (GL, green lines) or blue light (BL, blue lines). (H) Comparison of the absorbance spectra of RS9916 and WH8020 β-subunits CpeB and MpeB. Note that WH8020 CpeB contains only PEB. (I) Table depicting the bilin content of these proteins at specific Cys residues (as inferred from Shukla et al., 2012 and Ong and Glazer, 1991). Sites that change for CA4 are shown as PEB (in green light) and PUB (in blue light) separated by a “/”. The CpeB subunit differs at C50,61 between WH8020 and RS9916. Spectra numbered in (D–G) correspond with data in Table 1. These results are representative of two independent replicates.

only faint variations in the PUB:PEB ratio), all possess a H at position 141 in CpeB. In contrast, all strains that are either PT3c (typical BL specialists), PT3f (a rarer type of BL specialists), PT3dB (CA4-B), the other half of PT3dA including RS9916, and a natural mutant strain (BIOS-E4-1 that also display a BL specialist phenotype; Humily et al., 2013; Grébert et al., 2018) possess a L at position 141 in CpeB (Supplementary Figure 4A, blue highlights). By comparison for MpeB, all PT3a (or 3eA) strains exhibited a H whereas most other strains have a L, consistent with the fact that the former have a PEB and all others a PUB at C50, 61 (Supplementary Figure 4B). The only exception to this rule is PT3f strains which have a M at this position.

## Recombinant protein analysis

For simplicity, all recombinant proteins from RS9916 are hereafter prefixed with an “RS” and all WH8020 proteins are hereafter prefixed with a “WH” while generic referral to protein(s) from both strains will remain unprefixed (e.g., CpeA vs. RSCpeA or WHCpeA). We sought to determine the activity of WHMpeV compared to that of RSMpeV using our heterologous *E. coli* expression system with various substrates. All protein co-expressions analyzing β-subunits as substrates were designed to also express α-subunits in an effort to increase solubility of β-subunits as previously shown (Anderson and Toole, 1998; Kronfel et al., 2019b; Carrigee et al., 2020b). Our earlier work with

**TABLE 1** Mass spectrometry coverage of phycobilisome peptide fragments from WH8020. Trypsin digested samples of PEI  $\alpha$ - (CpeA) and  $\beta$ - (CpeB) subunits and PEI  $\alpha$ - (MpeA) and  $\beta$ - (MpeB) subunits showing their MS percent coverage, relative abundance (Rel Abund) and number (Num) of unique peptides found for each protein within each sample. Bold numbers for each sample correspond to numbered absorbance spectra in [Figure 3](#).

1—CpeA sample			2—CpeB sample			3—MpeA sample			4—MpeB sample			Protein name
Num unique	% Cov	Rel Abund	Num Unique	% Cov	Rel Abund	Num Unique	% Cov	Rel Abund	Num Unique	% Cov	Rel Abund	
24	80.5	85.9%	17	66.5	4.4%	15	72	0.2%	8	55.5	0.1%	WH8020_CpeA
8	38.6	6.7%	9	46.7	54.9%	7	42.4	0.3%	10	36.4	2.7%	WH8020_CpeB
8	36.4	7.1%	6	36.4	3.7%	38	98.8	98.7%	11	78.2	0.5%	WH8020_MpeA
14	66.9	0.3%	11	51.7	37.0%	10	51.7	0.8%	28	96.6	96.7%	WH8020_MpeB

RS9916 showed that the enzymes RSCpeS and RSCpeZ were required to obtain enough chromophorylated (attaching PEB at C82 of CpeB and MpeB), soluble RSCpeB and RSMpeB substrate to allow RSMpeV to function ([Carrigee et al., 2020b](#)). Therefore, these genes from RS9916 were expressed concomitantly in trials testing RSMpeV and WHMpeV activity ([Figure 4](#)). Even though detectable amounts of  $\alpha$ -subunits (CpeA) co-purified with their respective  $\beta$ -subunits (CpeB; [Figures 4E–G](#)), CpeA was expectedly not chromophorylated by the available lyases (see section LC–MS–MS analyses of recombinant co-expressions below). CpeB was present in all co-expressions as determined by total protein staining using Coomassie blue ([Figure 4G](#)). Control co-expressions to assess bilin addition and site specificity included CpeB/CpeA expressed without a lyase, or expressed in the presence of RSCpeZ, RSCpeS, and/or MpeV as outlined in [Supplementary Table 3](#). Absorbance spectra from recombinant co-expressions of RSMpeV or WHMpeV enzymes with RSCpeB, WHCpeB, or WHMpeB demonstrate that WHMpeV acts as a PEB lyase-isomerase, doubly ligating PUB on RSCpeB ([Figure 4A](#), co-expression 1, red line; absorbance peak 491.5 nm) and WHMpeB ([Supplementary Figure 5](#); absorbance peak at 493 nm), but it does not isomerize PEB when attaching it to WHCpeB ([Figures 4B,C](#), co-expression 4, black dashed line). The addition of a doubly-ligated PEB on WHCpeB subunit is detected as a prominent shoulder at 534.0 nm in addition to a peak at 556.5 nm from the C82 PEB ligated by RSCpeS ([Figures 4B,C](#), co-expression 4, black dashed line). This is consistent with observations by Kronfel et al. for a paralogous E/F type lyase called CpeF which doubly ligates a PEB on CpeB at an equivalent position in the freshwater cyanobacterium *Fremyella diplosiphon* ([Kronfel et al., 2019b](#)). We conclude that WHMpeV is capable of isomerizing activity on both WHMpeB and RSCpeB but that the WHCpeB substrate prevents the isomerization activity of WHMpeV.

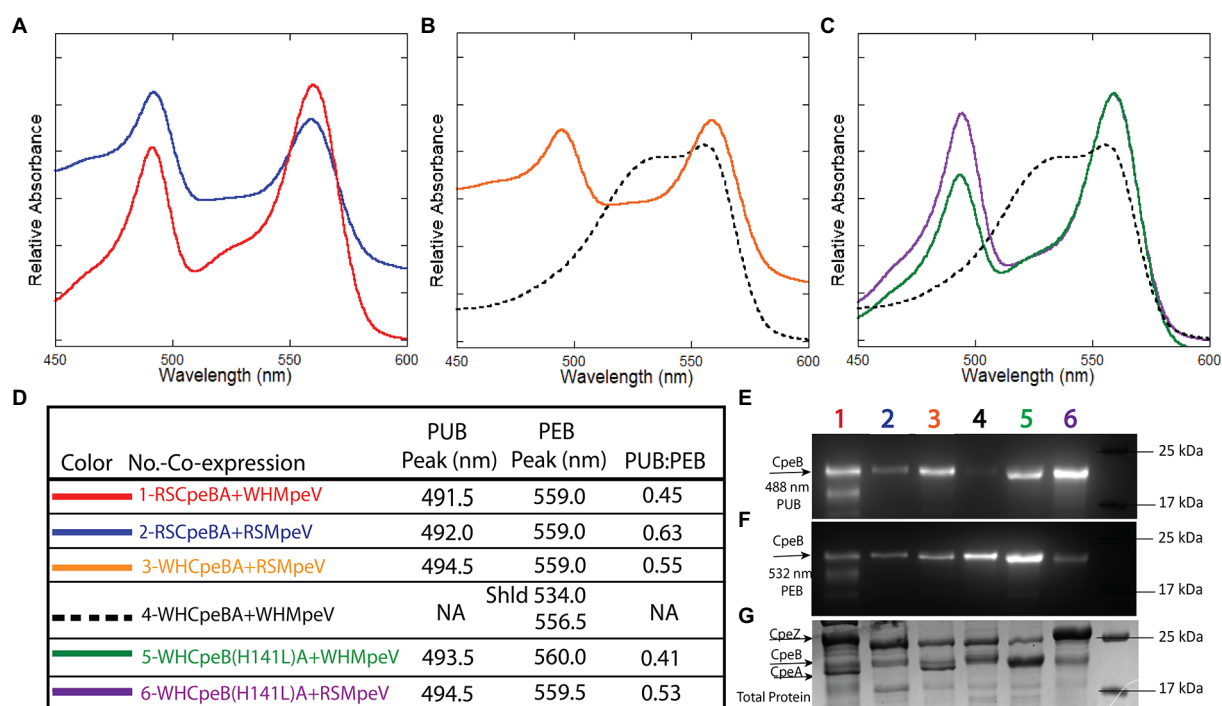
## Analysis of WHMpeV enzymatic activity using structural modeling and site-directed mutagenesis

As mentioned above, WHCpeB contains H at position 141 whereas RSCpeB contains L at this position ([Supplementary Figures 3, 4](#)). Molecular modeling of the structure

of WHCpeB suggested that H141 should be positioned close to the C50, 61 bilin position and at  $\sim 7.4$  Å from C50 ([Figure 5](#)). When PEB is docked with its ring A close to C50 and ring D close to C61, aspartic acid (D)54 of WHCpeB is poised toward the bilin, potentially interacting with the pyrrole nitrogen atoms of rings B and C ([Figure 5](#)). The arginine (R)57 residue on WHMpeV points toward the propionate chain of the C ring of PEB whereas MpeV-D116/R89 line the binding pocket near D ring ([Figure 5B](#)). Linkage at the A ring of PEB to C50 is a critical step in chromophore attachment, isomerization and stability for RSCpeB ([Carrigee et al., 2020b](#)). A site-directed mutation was introduced into WHCpeB converting H141 to leucine (H141L; [Figure 5](#); [Supplementary Table 1](#)) to test whether this residue affected isomerization activity by WHMpeV. Absorbance data from the co-expressions including RSCpeZ, RSCpeS, and WHMpeV as outlined in [Supplementary Table 3](#) revealed that this single substitution provided sufficient change(s) within the binding pocket to allow isomerization to occur ([Figures 4, 5](#)). Indeed, when native WHCpeB is expressed in the presence of RSCpeZ, RSCpeS, and WHMpeV, we see evidence of a doubly-ligated PEB at the C50, 61 position ([Figures 4B–D](#), co-expression 4, black dashed line, and [Table 2](#)). However, when we perform this same co-expression using pHTCpeB(H141L)A ([Supplementary Table 2](#)), we see a doubly-linked PUB at the C50, 61 position ([Figures 4C,D](#), co-expression 6 purple line, co-expression 5 green line, and [Table 2](#)). These findings indicate that the microenvironment of the WHCpeB binding pocket surrounding C50 plays a major role in WHMpeV isomerization activity at C50, 61.

## LC–MS–MS analyses of recombinant co-expressions

[Table 2](#) shows peak areas obtained for extracted ion chromatograms ( $M+4H$ )<sup>4+</sup> ions of the tryptic peptide for CpeB containing residues 38–78. The peptide sequence is the same in both the RSCpeB and WHCpeB and is shown in [Supplementary Tables 4–7](#). The negative control sample (WHCpeBA, no MpeV) shows no bilin attachment and interestingly, MS–MS data suggest C61 and C73 are linked by a disulfide bridge. Complete ligation of bilin to C50 and C61 was observed for the positive control sample (WHCpeB + WHMpeV). RSMpeV appears to modify WHCpeB and



**FIGURE 4**  
Absorbance spectra and SDS-PAGE for *Synechococcus* RS9916 and WH8020 CpeB co-expressions with MpeV. (A–C) Relative absorbance spectra depicting bilin addition to RSCpeB and WHCpeB when expressed in the presence of either WHMpeV or RSMpeV (as indicated in the legend in panel D). Purified protein substrates were resolved via SDS-PAGE and imaged with zinc-enhanced fluorescence at 460–490nm (E) which excites PUB and at 520–545nm (F) which excites PEB. Lanes are labeled with co-expression numbers and colors as listed in panel D. The same gel was then stained with Coomassie blue (G) to visualize proteins. All expressions contained RS9916 *cpeS*, *cpeZ*, and *ho1/pebS*. This study is representative of two independent replicates. NA means not applicable.

WHCpeB(H141L) at a slower rate than WHMpeV modifies WHCpeB as we observed unmodified, disulfide linked 38–78 peptide in the mixture. Intriguingly, the co-expressions where WHMpeV acted on RSCpeB or WHCpeB(H141L) showed evidence of bilins singly attached to the 38–78 peptide as well as normal double attachment and some unmodified substrate with no bilins attached at all. It is likely that the disulfide bond formed during purification and processing, as the cytoplasm of *E. coli* is generally a reducing environment (Gąciarz et al., 2017). However, its formation may indicate a lack of sufficient activity by MpeV or insufficient folding by CpeB to achieve the doubly ligated chromophore in some of these combinations (Table 2). Representative mass spectra supporting these observations comprise Supplementary Figures 5–7; Supplementary Tables 4–7, contain lists of tandem mass spectral fragment matches. In all samples containing RSCpeS, a bilin attached at C82 was observed, as previously described (Carrigee et al., 2020b). We did not observe any bilin modifications on CpeA in these co-expressions.

## Modeling of MpeV with substrates suggests the role of His141

To explore the possible role of CpeB H141 in conferring the lyase activity of MpeV, we built a structure model of MpeV in

complex with two substrates: CpeB and PEB, using manual docking aided by AlphaFold2 implemented in the ColabFold server (Jumper et al., 2021; Mirdita et al., 2022; Figure 5). In this model, the bilin pigment adopts an *anti*, *syn*, *anti*-configuration in which the ring A and ring D are in close proximity of their respective C anchors (i.e., C50 and C61 of CpeB) and the pyrrole nitrogen atoms of PEB rings B/C are in hydrogen-bonding distances with CpeB-D54 (Figure 5). Interestingly, we note that the side chain at the 141 position of CpeB would directly interact with the ring B propionate (Figure 5). For example, in WH8020 CpeB, H141 potentially forms a hydrogen bond with the ring B propionate while L141 of RS9916 CpeB cannot. This difference may explain the absence of the isomerase activity of MpeV on the CpeB substrate of WH8020 because the hydrogen bond between H141 and ring B propionate group may confer steric hinderance for the bilin transformation required by the PEB to PUB isomerization during a ligation reaction (Yang et al., 2007).

## Discussion

Among the three clans of lyases, only some members of the E/F clan have been reported to have the capability to isomerize bilins during attachment (Everroad et al., 2006; Shukla et al.,

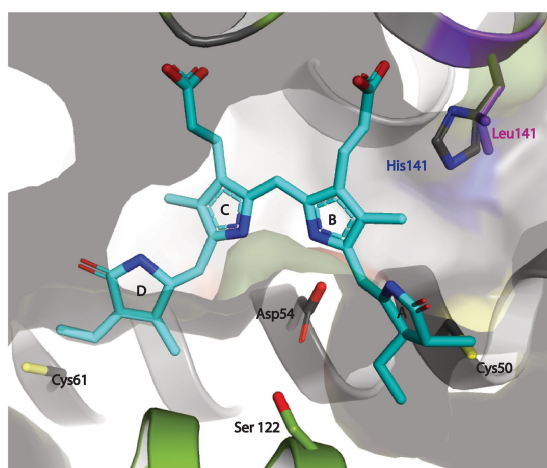


FIGURE 5

Docking of PEB at the putative binding cleft between the MpeV enzyme and WH8020 CpeB substrate. The PEB chromophore (cyan) is modeled in its final *anti, syn, anti*-conformation with ring A and ring D positioned next to their respective covalent anchors, (Cys50 and Cys61 of CpeB in yellow) bound between CpeB (gray ribbon), and MpeV (green ribbon). The pyrrole nitrogen atoms of rings B and C of PEB are stabilized via the side chain of the highly conserved Asp54 of CpeB. This model also places the side chain of His141 from WH8020 CpeB ~3.5 Å away from the ring B propionate group of PEB. We propose that this hydrogen bond between His141 and PEB potentially plays an important role in preventing the isomerization of PEB to PUB during the ligation reaction. The positioning of the conserved Ser122 of MpeV suggests it may be part of a putative active site for lyase activity.

2012; Humily et al., 2013; Sanfilippo et al., 2016); moreover very few studies regarding the mechanism among substrate-bilin-enzyme interactions during the isomerization reaction have been explored (Zhao et al., 2017; Kumarapperuma et al., 2022). The ability of a member of the E/F-type lyase family (in the present case MpeV) to behave as both a lyase and an isomerase to two very similar substrates has never been reported before and suggests the CpeB (and likely MpeB) substrate environment plays a role in the isomerization process. *Synechococcus* sp. WH8020 is a CA4-A capable strain whose  $\beta$ -subunits share a high sequence similarity with RS9916  $\beta$ -subunits (Supplementary Figures 1, 3A); however, the doubly ligated bilin at C50, 61 of WHCpeB is a PEB (Ong and Glazer, 1991) rather than a PUB, as on RSCpeB (Shukla et al., 2012), matching the chromophore pattern that is seen *in vivo* in WH8020 PBS (Figure 4). We observed that PEB isomerization to PUB is possible on recombinant WHCpeB at the doubly ligated C50, 61 position as demonstrated when it was co-expressed with RSMpeV in *E. coli* (Figure 5B, co-expression 4, orange line; Table 2), leading us to conclude that the interaction(s) among the lyase, bilin, and  $\beta$ -PE substrate are all important for the isomerization process. These data suggest the WHCpeB substrate may play a role in blocking the isomerization by the WHMpeV bilin lyase during ligation. In examining the alignment between the RSCpeB and WHCpeB sequences and

**TABLE 2** LC–MS–MS results for MpeV co-expressions. Extracted ion chromatograms were made using  $m/z$  1200.8254 for 38–78 peptides with a bilin attached; 1053.7513 was extracted for the 38–78 peptide lacking a bilin. Sample abbreviations are defined in **Supplementary Table 1**.

Sample	38–78 peptide cysteine state <sup>a</sup>	Retention time <sup>d</sup>	EIC area <sup>b</sup>
WHCpeBA + RSCpeZ/	2B	N/D	N/D
CpeS, no MpeV	1B	N/D	N/D
	0B	22.34	3.34E+08
WHCpeBA + RSCpeZ/	2B	20.94	2.35E+07
CpeS + WHMpeV	1B	N/D	N/D
	0B	N/D	N/D
WHCpeBA + RSCpeZ/	2B	21.19	7.73E+08
CpeS + RSMpeV	1B	N/D	N/D
	0B	22.6	1.08E+08
RSCpeBA + RSCpeZ/	2B	25.71,26.44	4.57E+07
CpeS + WHMpeV <sup>c</sup>	1B	27.25	3.59E+07
	0B	27.23	2.32E+07
WHCpeB(H141L)A +	2B	21.14	4.42E+09
RSCpeZ/CpeS +	1B	23.2	6.79E+07
RSMpeV <sup>c</sup>	0B	22.54	8.07E+07
WHCpeB(H141L)A +	2B	25.83,26.48	2.32E+08
RSCpeZ/CpeS +	1B	27.27	2.41E+07
WHMpeV <sup>c</sup>	0B	21.97	7.82E+07

Representative tandem mass spectra for each form of the peptide can be found in [Supplementary Figures 5–7](#).

<sup>a</sup>The 38–78 peptide has C50, C61 and C73. Peptides with doubly-linked bilin bridging C50 and C61 are labeled 2B. Peptides with a singly-linked bilin at C50 are labeled 1B; peptides with no bilin but a disulfide between C60 and C73 are labeled 0B.

<sup>b</sup>All EICs were made using an 8 ppm mass window and a 3 point boxcar smoothing routine. Areas were integrated with the ICIS algorithm of Thermo QualBrowser 4.3.73.11. Results are representative of two independent replicates.

<sup>c</sup>Sample was run separately from the others using a slightly modified gradient.

<sup>d</sup>N/D means not detected.

our molecular model of the WHCpeB C50, 61 binding pocket, we identified H141 as potentially important for affecting isomerization by preventing WHMpeV (but not RSMpeV) from isomerizing PEB to PUB under native conditions (Figures 3, 5). By altering WHCpeB with the H141L point mutation, we showed that WHMpeV was able to isomerize PEB to PUB at C50, 61 on this WHCpeB (H141L) mutant during the attachment reaction. One explanation for this phenomenon is steric hinderance of H141 within the binding pocket of WHCpeB due to its bulky side chain and its proximity to C50 on WHCpeB, the site of PEB A-ring attachment and subsequent location of the bond isomerization (Figure 5). Alternatively, H141 may be involved in protonation/deprotonation or hydrogen-bonding dependent activity affecting the bilin conformation within the binding pocket (Kumarapperuma et al., 2022). Our modeling suggested that H141 in CpeB is close to the ring B propionate of PEB. This type of interaction has been seen in bacteriophytochromes previously (Yang et al., 2007). H141 may hydrogen bond to this group, blocking the



movement that has to occur during the isomerization reaction. L (or M) at this position would not be so constrained, allowing the isomerization to occur.

Historically, E/F-type lyases including MpeV exhibit broad variation in chromophore and PBP substrate specificity, while demonstrating high binding-site specificity (e.g., which C residue). Kumarapperuma and collaborators recently determined the crystal structure of the lyase-isomerase MpeQ, proposed a mechanism for the reaction, and compared it to the mechanism for the related PEB lyase MpeW (Kumarapperuma et al., 2022). For MpeQ, Tyrosine 318 was proposed to activate PEB, polarizing the C3=C3<sup>1</sup> double bond. The side chain of V319 creates a steric conflict with the A-ring of PEB. This conflict is resolved by the PEB/PUB isomerization reaction and then ligation can happen, resulting in the MpeA-PUB product. In the PEB lyase MpeW, ligation happens directly after Tyrosine 318 activates that bond because there is no steric hindrance by the smaller residue of G319 on MpeW. In the case of WHMpeV, we hypothesize that the H141 residue in WHCpeB hydrogen bonds to the ring B propionate, constraining the movement that may occur during the isomerization reaction, favoring ligation before isomerization can take place (Figures 2, 5). RSMpeV is capable of isomerizing WHCpeB (Figure 4B, orange line), so there must be subtle differences within the MpeV enzymes that also control this isomerization reaction. In order to better understand this, we would need a structure of the substrate CpeB with the enzyme MpeV and the substrate PEB, a goal we are currently working toward.

This H at 141 is conserved in all green light specialists, i.e., *Synechococcus* strains that have either no PUB (PT2) or a constitutively low PUB/PEB ratio (PT3a; Supplementary Figure 3; Six et al., 2007; Grébert et al., 2022). Instead of MpeV, these strains possess the PEB lyase CpeF, the most ancient member of the CpeF-MpeV-MpeU E/F lyase family, which ligates doubly-linked PEB at C50/61 (Kronfel et al., 2019b; Carrigee et al., 2020b). In contrast, the blue light specialists, i.e., PT3c and PT3f which constitutively have a high PUB/PEB ratio, and the PT3dB strains, i.e., chromatic acclimators possessing a CA4-B island (Humily et al., 2013; Grébert et al., 2021), all have a L at CpeB-141, and we posit that these strains must have a PUB at C50, 61 of CpeB. All of them lack CpeF and MpeV, so we hypothesize that the lyase-isomerase necessary to bind a PUB at both CpeB and MpeB C50, 61 in these strains is MpeU, a partially characterized member of this E/F lyase family (Mahmoud et al., 2017; Carrigee et al., 2020b; Grébert et al., 2022). Interestingly, while all PT3dA strains, i.e., chromatic acclimators possessing a CA4-A island, possess MpeV, only half of them, including RS9916, MITS9220, BL107, PROS-9-1 and the natural mutant strain BIOS-E4-1 which has become a BL specialist by selective loss of *fciA*, *fciB*, and *mpeY* genes (Humily et al., 2013; Grébert et al., 2018), have L at CpeB-141, so we hypothesize that they all have PUB at C50, 61 of CpeB, as previously demonstrated for RS9916 (Shukla et al., 2012). The other half of the PT3dA strains, including WH8020, CC9311, BIOS-U3-1 and RCC307, which shows only faint variations of its

PUB/PEB ratio (Humily et al., 2013) all have H at this 141 position and consequently must have PEB at CpeB 50–61, like was first shown in WH8020 by Ong and Glazer (1991) and checked by us in the present study. The identity of the amino-acid at position 141 is seemingly also important for MpeB, with all GL specialists that have a PEB at C50, 61 having a H, and all BL specialists and chromatic acclimators that have a PUB at C50, 61 having a L. The only exception to this rule is PT3f strains which have a M at the MpeB 141 position. Even though the phycobiliprotein chromophorylation of strains representative of this pigment type has never been formally determined, it is most likely that they bind a PUB at C50, 61 on both CpeB and MpeB. So, altogether it appears that the chromophore bound at C50, 61 of CpeB and/or MpeB in any PE-containing marine *Synechococcus* can be reliably predicted from the amino acid present at position 141 of these PE  $\beta$ -subunits.

Combined with the structural and mechanistic analyses previously performed on MpeQ (Kumarapperuma et al., 2022), the present study demonstrated that there are residues within the substrate that influence the isomerization reaction of lyase/isomerases, a novel finding. A complete understanding of the mechanisms of these isomerase-capable lyases within the E/F clan is a main challenge for future studies.

## Data availability statement

The original contributions presented in the study are publicly available. This data can be found at: MassIVE (<https://massive.ucsd.edu/ProteoSAFe/static/massive.jsp>), dataset ID number MSV000090019.

## Author contributions

LC and WS conceived of the study through discussions with the coauthors. WS supervised the work and together with LC wrote the draft manuscript. LC, JF, and XL conducted the recombinant protein analyses. LD and FP performed the growth and analyses of *Synechococcus* sp. WH8020 cells for PBS purification and performed the bioinformatic analyses of CpeB and MpeB. JK and JT performed the tandem mass spectrometry and analyzed the data. IT and XY modeled the structure of MpeV with CpeB and PEB. All authors contributed to the article and approved the submitted version.

## Funding

The Orbitrap Fusion Lumos was purchased with funds from the Precision Health Initiative of the Indiana University Bicentennial Grand Challenges Program. This research project has been supported by awards from the National Science Foundation to WS (MCB 2017171) and XY (MCB 2017274).



and from the Agence Nationale de la Recherche (ANR) program EFFICACY (ANR-19-CE02-0019) to FP.

## Acknowledgments

We are grateful to Kes Lynn Joseph for helpful discussions and for assistance with culture prep and maintenance. We are also very grateful to David Kehoe for helpful discussions.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

- Anderson, L. K., and Toole, C. M. (1998). A model for early events in the assembly pathway of cyanobacterial phycobilisomes. *Mol. Microbiol.* 30, 467–474. doi: 10.1046/j.1365-2958.1998.01081.x
- Andrade, M. A., Petosa, C., O'Donoghue, S. I., Müller, C. W., and Bork, P. (2001). Comparison of ARM and HEAT protein repeats. *J. Mol. Biol.* 309, 1–18. doi: 10.1006/jmbi.2001.4624
- Bengston, S. (1994). *Early Life On Earth. Nobel Symposium No. 84*. Columbia Univ. press: New York.
- Betz, M. (1997). One century of protein crystallography: the phycobiliproteins. *Biol. Chem.* 378, 167–176.
- Biswas, A., Boutaghou, M. N., Alvey, R. M., Kronfel, C. M., Cole, R. B., Bryant, D. A., et al. (2011). Characterization of the activities of the CpeY, CpeZ, and CpeS Bilin lyases in phycoerythrin biosynthesis in *Fremyella diplosiphon* strain UTEX 481. *J. Biol. Chem.* 286, 35509–35521. doi: 10.1074/jbc.M111.284281
- Biswas, A., Vasquez, Y. M., Dragomani, T. M., Kronfel, M. L., Williams, S. R., Alvey, R. M., et al. (2010). Biosynthesis of cyanobacterial phycobiliproteins in *Escherichia coli*: chromophorylation efficiency and specificity of all Bilin lyases from *Synechococcus* sp. strain PCC 7002. *Appl. Environ. Microbiol.* 76, 2729–2739. doi: 10.1128/AEM.03100-09
- Carrigee, L. (2020). Characterizing the function of lyases involved in the biosynthesis of phycoerythrin I and II from marine *Synechococcus* cyanobacteria: A PhD dissertation. University of New Orleans, New Orleans, LA.
- Carrigee, L. A., Frick, J. P., Karty, J. A., Garczarek, L., Partensky, F., and Schluchter, W. M. (2020b). MpeV is a lyase isomerase that ligates a doubly linked phycocoubrilin to the  $\beta$ -subunit of phycoerythrin I and II in marine *Synechococcus*. *J. Biol. Chem.* 296, 1–13. doi: 10.1074/jbc.RA120.015289
- Carrigee, L., Mahmoud, R. M., Sanfilippo, J. E., Frick, J. P., Strnat, J. A., Karty, J. A., et al. (2020a). CpeY is a phycoerythrobilin lyase for cysteine 82 of the phycoerythrin I  $\alpha$ -subunit in marine *Synechococcus*. *BBA-Bioenergetics* 1861:148215. doi: 10.1016/j.bbabin.2020.148215
- Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D* 60, 2126–2132. doi: 10.1107/S0907444904019158
- Everroad, C., Six, C., Partensky, F., Thomas, J.-C., Holtzendorff, J., and Wood, A. M. (2006). Biochemical bases of type IV chromatic adaptation in marine *Synechococcus* spp. *J. Bacteriol.* 188, 3345–3356. doi: 10.1128/jb.188.9.3345-3356.2006
- Everroad, R. C., and Wood, A. M. (2006). Comparative molecular evolution of newly discovered picocyanobacterial strains reveals a phylogenetically informative variable region of beta-phycoerythrin. *J. Phycol.* 42, 1300–1311. doi: 10.1111/j.1529-8817.2006.00282.x
- Fairchild, C. D., and Glazer, A. N. (1994). Oligomeric structure, enzyme kinetics, and substrate specificity of the phycocyanin  $\alpha$  subunit phycocyanobilin lyase. *J. Biol. Chem.* 269, 8686–8694. doi: 10.1016/S0021-9258(17)37022-9
- Fairchild, C. D., Zhao, J., Zhou, J., Colson, S. E., Bryant, D. A., and Glazer, A. N. (1992). Phycocyanin  $\alpha$  subunit phycocyanobilin lyase. *Proc. Natl. Acad. Sci., USA*, 89, 7017–7021. doi: 10.1073/pnas.89.15.7017
- Flombaum, P., Gallegos, J. L., Gordillo, R. A., Rincon, J., Zabala, L. L., Jiao, N., et al. (2013). Present and future global distributions of the marine cyanobacteria *Prochlorococcus* and *Synechococcus*. *Proc. Natl. Acad. Sci. USA*, 110, 9824–9829. doi: 10.1073/pnas.1307701110
- Gąciarz, A., Khatri, N. K., Velez-Suberbie, M. L., Saaranen, M. J., Uchida, Y., Keshavarz-Moore, E., et al. (2017). Efficient soluble expression of disulfide bonded proteins in the cytoplasm of *Escherichia coli* in fed-batch fermentations on chemically defined minimal media. *Microb. Cell Fact.* 16:108. doi: 10.1186/s12934-017-0721-x
- Garczarek, L., Guyet, U., Doré, H., Farrant, G. K., Hoebeke, M., Brillet-Guéguen, L., et al. (2021). Cyanorak v2.1: a scalable information system dedicated to the visualization and expert curation of marine and brackish picocyanobacteria genomes. *Nucleic Acids Res.* 49, D667–D676. doi: 10.1093/nar/gkaa958
- Gasper, R., Schwach, J., Hartmann, J., Holtkamp, A., Wiethaus, J., Reidel, N., et al. (2017). Distinct features of Cyanophage-encoded T-type Phycobiliprotein Lyase  $\Phi$ CpeT: the role of auxiliary metabolic genes. *J. Biol. Chem.* 292, 3089–3098. doi: 10.1074/jbc.M116.769703
- Glazer, A. N. (1982). Phycobilisomes: structure and dynamics. *Annu. Rev. Microbiol.* 36, 173–198. doi: 10.1146/annurev.mi.36.100182.001133
- Glazer, A. N. (1984). Phycobilisome: a macromolecular complex optimized for light energy transfer. *Biochim. Biophys. Acta.* 768, 29–51.
- Glazer, A. N. (1988). Phycobilisomes. *Methods Enzymol.* 167, 304–312. doi: 10.1016/0076-6879(88)67035-2
- Glazer, A. N. (1994). Phycobiliproteins - a family of valuable, widely used fluorophores. *J. Appl. Phycol.* 6, 105–112. doi: 10.1007/BF02186064
- Grébert, T., Doré, H., Partensky, F., Farrant, G. K., Boss, E. S., Picheral, M., et al. (2018). Light color acclimation is a key process in the global ocean distribution of *Synechococcus* cyanobacteria. *Proc. Natl. Acad. Sci. U. S. A.*, 115, E2010–E2019. doi: 10.1073/pnas.1717069115
- Grébert, T., Garczarek, L., Daubin, V., Humily, F., Marie, D., Ratin, M., et al. (2022). Diversity and evolution of pigment types in marine *Synechococcus* cyanobacteria. *Genome Biol. Evol.* 14:evac035. doi: 10.1093/gbe/evac035
- Grébert, T., Nguyen, A. A., Pokhrel, S., Joseph, K. L., Ratin, M., Dufour, L., et al. (2021). Molecular bases of an alternative dual-enzyme system for light color acclimation of marine *Synechococcus* cyanobacteria. *Proc. Natl. Acad. Sci. U. S. A.* 118. doi: 10.1073/pnas.2019715118
- Humily, F., Partensky, F., Six, C., Farrant, G. K., Ratin, M., Marie, D., et al. (2013). A gene island with two possible configurations is involved in chromatic acclimation in marine *Synechococcus*. *PLoS One* 8:e84459. doi: 10.1371/journal.pone.0084459
- Hussain, H., and Chong, N. F.-M. (2016). Combined overlap extension PCR method for improved site directed mutagenesis. *Hindawi Publish. Corporat. BioMed. Res. Int.* 2016, 1–7. doi: 10.1155/2016/8041532
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589. doi: 10.1038/s41586-021-03819-2
- Kasting, J. F., and Siefert, J. L. (2003). Life and the evolution of Earth's atmosphere. *Science* 299:1015. doi: 10.1126/science.1071184
- Kehoe, D. M. (2010). Chromatic adaptation and the evolution of light color sensing in cyanobacteria. *Proc. Natl. Acad. Sci. U. S. A.* 107, 9029–9030. doi: 10.1073/pnas.1004510107
- Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. (2015). The Phyre2 web portal for protein modeling prediction and analysis. *Nat. Protoc.* 10, 845–858. doi: 10.1038/nprot.2015.053

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.1011189/full#supplementary-material>

- Kozo, M., Kazuaki, N., and Masato, N. (2002). Identification of a novel prokaryotic HEAT-repeats-containing protein which interacts with a cyanobacterial IscA homolog. *FEBS Lett.* 519, 123–127. doi: 10.1016/s0014-5793(02)02736-9
- Kronfel, C. M., Biswas, A., Frick, J. P., Gutu, A., Blensdorf, T., Karty, J. A., et al. (2019a). The roles of the chaperone-like protein CpeZ and the phycoerythrobilin lyase CpeY in phycoerythrin biogenesis. *Biochim. Biophys. Acta-Bioenerg.* 1860, 249–261. doi: 10.1016/j.bbabo.2019.06.001
- Kronfel, C. M., Hernandez, C. V., Frick, J. P., Hernandez, L. S., Gutu, A., Karty, J. A., et al. (2019b). CpeF is the Bilin lyase that ligates the doubly linked phycoerythrobilin on  $\beta$ -phycoerythrin in the cyanobacterium *Fremyella diplosiphon*. *J. Biol. Chem.* 294, 3987–3999. doi: 10.1074/jbc.RA118.007221
- Kronfel, C. M., Kuzin, A. P., Forouhar, F., Biswas, A., Su, M., Lew, S., et al. (2013). Structural and biochemical characterization of the Bilin lyase CpcS from *Thermosynechococcus elongatus*. *Biochemistry* 52, 8663–8676. doi: 10.1021/bi401192z
- Kumarapperuma, I., Joseph, K. L., Wang, C., Biju, L. M., Tom, I. P., Weaver, K. D., et al. (2022). Crystal structure and molecular mechanism of an E/F type Bilin lyase-isomerase. *Structure* 30, 564–574.e3. doi: 10.1016/j.str.2022.01.007
- Mahmoud, R. M., Sanfilippo, J. E., Nguyen, A. A., Strnat, J. A., Partensky, F., Garczarek, L., et al. (2017). Adaptation to blue light in marine *Synechococcus* requires MpeU, an enzyme with similarity to phycoerythrobilin lyase isomerases. *Front. Microbiol.* 8. doi: 10.3389/fmicb.2017.00243
- Marcotrigiano, J., Lomakin, I. B., Sonenberg, N., Pestova, T. V., Hellen, C. U., and Burley, S. K. (2001). A conserved HEAT domain within eIF4G directs assembly of the translation initiation machinery. *Mol. Cell* 7, 193–203. doi: 10.1016/S1097-2765(01)00167-8
- Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M. (2022). ColabFold: making protein folding accessible to all. *Nat. Methods* 19, 679–682. doi: 10.1038/s41592-022-01488-1
- Nguyen, A. (2018). *Characterization of genes involved in the biosynthesis of Phycoerythrin I and II in cyanobacteria: A PhD dissertation*. New Orleans, LA: University of New Orleans.
- Ong, L. J., and Glazer, A. N. (1991). Phycoerythrins of marine unicellular cyanobacteria. I. Bilin types and locations and energy transfer pathways in *Synechococcus* spp. phycoerythrins. *J. Biol. Chem.* 266, 9515–9527. doi: 10.1016/S0021-9258(18)92851-6
- Overkamp, K. E., Gasper, R., Kock, K., Herrmann, C., Hofmann, E., and Frankenburg-Dinkel, N. (2014). Insights into the biosynthesis and assembly of Cryptophycean Phycobiliproteins. *J. Biol. Chem.* 289, 26691–26707. doi: 10.1074/jbc.M114.591131
- Palenik, B. (2001). Chromatic adaptation in marine *Synechococcus* strains. *Appl. Environ. Microbiol.* 67, 991–994. doi: 10.1128/AEM.67.2.991-994.2001
- Sanfilippo, J. E., Garczarek, L., Partensky, F., and Kehoe, D. M. (2019a). Chromatic acclimation in cyanobacteria: a diverse and widespread process for optimizing photosynthesis. *Annu. Rev. Microbiol.* 73, 407–433. doi: 10.1146/annurev-micro-020518-115738
- Sanfilippo, J. E., Nguyen, A. A., Garczarek, L., Karty, J. A., Pokhrel, S., Strnat, J. A., et al. (2019b). Interplay between differentially expressed enzymes contributes to light color acclimation in marine *Synechococcus*. *Proc. Natl. Acad. Sci. U. S. A.* 1–6. doi: 10.1073/pnas.1810491116
- Sanfilippo, J. E., Nguyen, A. A., Karty, J. A., Shukla, A., Schluchter, W. M., Garczarek, L., et al. (2016). Self-regulating genomic island encoding tandem regulators confers chromatic acclimation to marine *Synechococcus*. *Proc. Natl. Acad. Sci. U. S. A.* 113, 6077–6082. doi: 10.1073/pnas.201600625
- Saunee, N. A., Williams, S. R., Bryant, D. A., and Schluchter, W. M. (2008). Biogenesis of phycobiliproteins. II. CpcS-I and CpcU comprise the heterodimeric Bilin lyase that attaches phycocyanobilin to Cys-82 of  $\beta$ -phycocyanin and Cys-81 of allophycocyanin subunits in *Synechococcus* sp. PCC 7002. *J. Biol. Chem.* 283, 7513–7522. doi: 10.1074/jbc.M708165200
- Scheer, H., and Zhao, K. (2008). Biliprotein maturation: the chromophore attachment. *Mol. Microbiol.* 68, 263–276. doi: 10.1111/j.1365-2958.2008.06160.x
- Schluchter, W. M., Shen, G., Alvey, R. M., Biswas, A., Saunee, N. A., Williams, S. R., et al. (2010). Phycobiliprotein biosynthesis in cyanobacteria: structure and function of enzymes involved in post-translational modification. *Adv. Exp. Med. Biol.* 675, 211–228. doi: 10.1007/978-1-4419-1528-3\_12
- Shen, G., Saunee, N. A., Williams, S. R., Gallo, E. F., Schluchter, W. M., and Bryant, D. A. (2006). Identification and characterization of a new class of Bilin lyase: the *cpcT* gene encodes a Bilin lyase responsible for attachment of phycocyanobilin to Cys-153 on the  $\beta$  subunit of phycocyanin in *Synechococcus* sp. PCC 7002. *J. Biol. Chem.* 281, 17768–17778. doi: 10.1074/jbc.M602563200
- Shen, G., Schluchter, W. M., and Bryant, D. A. (2008). Biogenesis of phycobiliproteins. I. *cpcS*-I and *cpcU* mutants of the cyanobacterium *Synechococcus* sp. PCC 7002 define a heterodimeric phycocyanobilin lyase specific for  $\beta$ -phycocyanin and allophycocyanin subunits. *J. Biol. Chem.* 283, 7503–7512. doi: 10.1074/jbc.M708164200
- Shukla, A., Biswas, A., Blot, N., Partensky, F., Karty, J. A., Hammad, L. A., et al. (2012). Phycoerythrin-specific bilin lyase-isomerase controls blue-green chromatic acclimation in marine *Synechococcus*. *Proc. Natl. Acad. Sci. U. S. A.* 109, 20136–20141. doi: 10.1073/pnas.1211771109
- Six, C., Thomas, J.-C., Garczarek, L., Ostrowski, M., Dufresne, A., Blot, N., et al. (2007). Diversity and evolution of phycobilisomes in marine *Synechococcus* spp.: a comparative genomics study. *Genome Biol.* 8:R259. doi: 10.1186/gb-2007-8-12-r259
- Studier, F. W. (2005). Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* 41, 207–234. doi: 10.1016/j.pep.2005.01.016
- Swanson, R. V., Zhou, J., Leary, J. A., Williams, T., de Lorimier, R., Bryant, D. A., et al. (1992). Characterization of phycocyanin produced by *cpeE* and *cpcF* mutants encoding an intergenic suppressor of the defect in Bilin attachment. *J. Biol. Chem.* 267, 16146–16154. doi: 10.1016/S0021-9258(18)41979-5
- Takano, H., and Gusella, J. (2002). The predominantly HEAT-like motif structure of huntingtin and its association and coincident nuclear entry with dorsal, an NF-kB/Rel/dorsal family transcription factor. *BMC Neurosci.* 3:15. doi: 10.1186/1471-2202-3-15
- Wilbanks, S. M., and Glazer, A. N. (1993). Rod structure of a phycoerythrin II-containing phycobilisome I: organization and sequence of the gene cluster encoding the major phycobiliprotein rod components in the genome of marine *Synechococcus* sp. WH8020. *J. Biol. Chem.* 268, 1226–1235. doi: 10.1016/S0021-9258(18)54064-3
- Yang, X., Stojković, E. A., Kuk, J., and Moffat, K. (2007). Crystal structure of the chromophore binding domain of an unusual bacteriophytochrome, RbBphP3, reveals residues that modulate photoconversion. *Proc. Natl. Acad. Sci. U. S. A.* 104, 12571–12576. doi: 10.1073/pnas.0701737104
- Zhao, C., Hoppner, A., Xu, Q.-Z., Gartner, W., Scheer, H., Zhou, M., et al. (2017). Structures and enzymatic mechanisms of phycobiliprotein lyases CpcE/F and PecE/F. *Proc. Natl. Acad. Sci. U. S. A.* 114, 13170–13175. doi: 10.1073/pnas.1715495114
- Zhao, K. H., Su, P., Li, J., Tu, J. M., Zhou, M., Bubenzner, C., et al. (2006). Chromophore attachment to phycobiliprotein  $\beta$ -subunits: phycocyanobilin:cystein- $\beta$ 84 phycobiliprotein lyase activity of CpeS-like protein from *anabaena* sp. PCC7120. *J. Biol. Chem.* 281, 8573–8581. doi: 10.1074/jbc.M513796200
- Zhao, K.-H., Su, P., Tu, J.-M., Wang, X., Liu, H., Ploscher, M., et al. (2007). Phycobilin:cysteine-84 biliprotein lyase, a near-universal lyase for cysteine-84-binding sites in cyanobacterial phycobiliproteins. *Proc. Natl. Acad. Sci. U. S. A.* 104, 14300–14305. doi: 10.1073/pnas.0706209104
- Zhou, W., Ding, W.-L., Zheng, X.-L., Dong, L.-L., Zhao, B., Zhou, M., et al. (2014). Structure and mechanism of the phycobiliprotein Lyase CpcT. *J. Biol. Chem.* 289, 26677–26689. doi: 10.1074/jbc.M114.586743
- Zhou, J., Gasparich, G. E., Stirewalt, V. L., de Lorimier, R., and Bryant, D. A. (1992). The *cpcE* and *cpcF* genes of *Synechococcus* sp. PCC 7002: construction and phenotypic characterization of interposon mutants. *J. Biol. Chem.* 267, 16138–16145. doi: 10.1016/S0021-9258(18)41978-3



## OPEN ACCESS

## EDITED BY

Graciela L. Lorca,  
University of Florida, United States

## REVIEWED BY

Dengke Yin,  
Anhui University of Chinese  
Medicine, China  
Jiajun Wu,  
City University of Hong Kong, Hong  
Kong SAR, China

## \*CORRESPONDENCE

Jin Zheng  
zjddln@163.com  
Litian Ma  
malitian1234@163.com  
Tian Li  
fmmult@foxmail.com

<sup>†</sup>These authors have contributed  
equally to this work and share first  
authorship

## SPECIALTY SECTION

This article was submitted to  
Microbial Physiology and Metabolism,  
a section of the journal  
Frontiers in Microbiology

RECEIVED 05 April 2022

ACCEPTED 31 October 2022

PUBLISHED 24 November 2022

## CITATION

Gu L, Wu H, Zhang Y, Wu Y, Jin Y, Li T,  
Ma L and Zheng J (2022) The effects of  
elemene emulsion injection on rat  
fecal microbiota and metabolites:  
Evidence from metagenomic  
exploration and liquid  
chromatography-mass spectrometry.  
*Front. Microbiol.* 13:913461.  
doi: 10.3389/fmicb.2022.913461

## COPYRIGHT

© 2022 Gu, Wu, Zhang, Wu, Jin, Li, Ma  
and Zheng. This is an open-access  
article distributed under the terms of  
the [Creative Commons Attribution  
License \(CC BY\)](#). The use, distribution  
or reproduction in other forums is  
permitted, provided the original  
author(s) and the copyright owner(s)  
are credited and that the original  
publication in this journal is cited, in  
accordance with accepted academic  
practice. No use, distribution or  
reproduction is permitted which does  
not comply with these terms.

# The effects of elemene emulsion injection on rat fecal microbiota and metabolites: Evidence from metagenomic exploration and liquid chromatography-mass spectrometry

Lei Gu<sup>1†</sup>, Hao Wu<sup>2†</sup>, Yang Zhang<sup>3†</sup>, Yousheng Wu<sup>4</sup>, Yuan Jin<sup>5</sup>,  
Tian Li<sup>6\*</sup>, Litian Ma<sup>2,7\*</sup> and Jin Zheng<sup>2\*</sup>

<sup>1</sup>Department of Cardiology, Xi'an International Medical Center Hospital Affiliated to Northwest University, Xi'an, China, <sup>2</sup>Department of Traditional Chinese Medicine, Tangdu Hospital, Air Force Medical University, Xi'an, China, <sup>3</sup>Health Center of 95816 of the People's Liberation Army, Wuhan, China, <sup>4</sup>National Demonstration Center for Experimental Preclinical Medicine Education, Air Force Medical University, Xi'an, China, <sup>5</sup>Department of Internal Medicine, The Third Affiliated Hospital of Xinxiang Medical College, Xinxiang, China, <sup>6</sup>School of Basic Medicine, Fourth Military Medical University, Xi'an, China, <sup>7</sup>Department of Gastroenterology, Tangdu Hospital, Air Force Medical University, Xi'an, China

**Objective:** Elemene emulsion injection (EEI) has been approved for interventional and intracavitary chemotherapy in treating malignant ascites in China, but few studies have focused on the effects of EEI on gut microbiota and metabolites. In this study, we investigated the effects of EEI on the fecal microbiota and metabolites in healthy Sprague-Dawley (SD) rats.

**Methods:** We randomly assigned 18 male SD rats to three groups ( $n = 6$  in each group): the sham group (group S), the low-concentration EEI group (L-EEI), and the high-concentration EEI group (H-EEI). The L-EEI and H-EEI rats were administered 14 days of consecutive EEI, 20 mg/kg, and 40 mg/kg intraperitoneally (IP). Group S rats were administered the same volume of normal saline. On day 14, each animal's feces were collected for metagenomic sequencing and metabolomic analysis, and the colonic contents were collected for 16S rRNA sequencing.

**Results:** EEI could alter the  $\beta$ -diversity but not the  $\alpha$ -diversity of the fecal microbiota and induce structural changes in the fecal microbiota. Different concentrations of EEI affect the fecal microbiota differently. The effects of different EEI concentrations on the top 20 bacteria with significant differences at the species level among the three groups were roughly divided into three categories: (1) A positive or negative correlation with the different EEI concentrations. The abundance of *Ileibacterium Valens* increased as the EEI concentration increased, while the abundance of *Firmicutes bacteria* and *Clostridium sp.* CAC:273 decreased. (2) The microbiota showed a tendency to increase first, then decrease or decrease first, and then increase as EEI concentration increased—the abundance of *Prevotella sp.* PCHR, *Escherichia coli*, and *Candidatus Amuliumraptor caecigallinaris* tended to decrease with L-EEI but significantly increased with H-EEI. In

contrast, L-EEI significantly increased *Ruminococcus bromii* and *Dorea sp.* 5–2 abundance, and *Oscillibacter sp.* 1–3 abundance tended to increase, while H-EEI significantly decreased them. (3) L-EEI and H-EEI decreased the abundance of bacteria (*Ruminococcaceae bacterium*, *Romboutsia ilealis*, and *Staphylococcus xylosus*). Fecal metabolites, like microbiota, were sensitive to different EEI concentrations and correlated with fecal microbiota and potential biomarkers.

**Conclusion:** This study shows that intraperitoneal EEI modulates the composition of rat fecal microbiota and metabolites, particularly the gut microbiota's sensitivity to different concentrations of EEI. The impact of changes in the microbiota on human health remains unknown, particularly EEI's efficacy in treating tumors.

#### KEYWORDS

elemene emulsion injection, fecal microbiota, metabolomics, 16S rRNA, metagenome

## Introduction

Elemenes are sesquiterpene compounds derived from the traditional Chinese medicinal plant, *Curcuma wenyujin*. Elemene emulsion injection (EEI) contains the active ingredients  $\beta$ -,  $\gamma$ -, and  $\delta$ -elemene and excipients such as soybean lecithin, cholesterol, ethanol, disodium hydrogen phosphate, and sodium dihydrogen phosphate. The National Medical Products Administration of China has approved it for interventional and intracavitary chemotherapy and for treating cancerous pleural ascites. Furthermore, combining EEI with conventional radiotherapy or chemotherapy could improve their therapeutic effects (Jiang et al., 2017; Tong et al., 2020) against lung cancer (Chen et al., 2020), nasopharyngeal cancer, brain tumors (Liu S. et al., 2020), bone metastases, and others (Cai et al., 2021) while decreasing their adverse effects (Chen et al., 2020; Liu S. et al., 2020; Hashem et al., 2021).

Elemene-containing hyperthermic intraperitoneal chemotherapy has been used to treat peritoneal metastatic advanced gastric cancer with minimal myelosuppression (Zheng et al., 2014). Recent studies have shown that gut microbiota is critical to human health (Markowiak and Slizewska, 2017; Singh et al., 2017; Sommer et al., 2017; Gentile and Weir, 2018). Natural selection and adaptation maintain the microbiota-host-environment system in dynamic equilibrium, creating mutual constraints (Smits et al., 2013). Studies have linked gut microbiota to gastric and breast cancers (Chen et al., 2019). In addition, long-term antibiotic use can change gut microbes, increasing the risk of colorectal and prostate cancers (Boursi et al., 2015a,b; Dik et al., 2016; Ianaro et al., 2016; Chen et al., 2019). Elemene is an anticancer drug with unknown effects on the gut microbiota. The only study showing that  $\beta$ -elemene improves brain metabolites

in obese C57BL/6 male mice fed a high-fat diet (HFD) and reversed HFD-induced changes in gut bacterial composition and content in mice (Zhou et al., 2021). Based on the above evidence, it is not yet possible to link the neuroprotective effect of EEI to the intestinal microbiota, nor can it be proved that there is a causal relationship between gut microbiota and EEI on neuroprotective effects. Likewise, it is not yet known whether the regulation of the gut microbiota would affect the therapeutic effect of EEI on tumors and the protective effect on nerves. Therefore, this study used meta-genomic sequencing and untargeted metabolomics techniques to explore the effects of different EEI concentrations on the rat fecal microbiota and metabolites and provide evidence that EEI administration could regulate the microbiota and metabolites in the rat.

## Materials and methods

### Animals

Eighteen male Sprague-Dawley (SD) rats (weighing 220–250 g, aged 9–11 weeks) were acquired from the Air Force Medical University under certificate number SCXK (Shaan) 2019-001. The rats were kept in separate cages, each with three rats. A 12-h light/dark cycle was implemented in the rearing environment, with a temperature of 20–25 °C and a relative humidity of 50–65%. The Animal Welfare and Ethics Committee, Laboratory Animal Center, and Air Force Military Medical University (IACUC-20220522) approved this study. All animal procedures were performed following the US National Institutes of Health's (NIH) Guide for the Care and Use of Laboratory Animals (NIH publication no. 85–23, revised).



## Materials

EEI was bought from Huali Jingang Pharmaceutical Co., Ltd. (Liaoning, China). According to literature reports, the detection data of 24 batches of samples showed that the content of  $\beta$ -syringene was 6.0 to 8.4%, with an average value of 7.2%, the content of  $\beta$ -elemene chiral isomer (RRT about 0.96) was 3.2–4.6%, with an average value of 3.9%, and the total of the two was 9.2–12.7%, with an average of 11.1% (Zhihua, 2018). The NEXTFLEX Rapid DNA-Seq Kit (Bioo Scientific, USA) and NovaSeq Reagent Kits/HiSeq X Reagent Kits (Illumina, America) were used as the primary kits for metagenomic sequencing. The main reagents and instruments for liquid chromatography-mass spectrometry (LC-MS) untargeted metabolomics include an ultra-high pressure liquid chromatography (UHPLC) liquid chromatography system (Vanquish Horizon system, Thermo), a mass spectrometer (Q-Exactive, Thermo), methanol and acetonitrile (Fisher Chemical), 2-propanol (Merck), and 2-Chloro-L-Phenylalanine (Adamas-beta).

## Grouping and administration

Eighteen male SD rats were randomly divided into three groups of six rats each: the sham group (group S), the low-concentration EEI group (L-EEI), and the high-concentration EEI group (H-EEI). Before the experiment started, rats were fed adaptively for 1 week and had free access to water and food. Different EEI concentrations were selected based on previously published studies (Ma et al., 2021; Sun et al., 2022; Wu et al., 2022). For 14 days, L-EEI rats received elemene emulsion [20 mg/(kg·d)] intraperitoneally (IP), H-EEI rats received elemene emulsion [40 mg/(kg·d)] IP, and group S rats received the same volume of normal saline IP as H-EEI. To equalize volume, L-EEI rats were supplemented with normal saline to equal the volume in H-EEI before injection.

## Amplification of the 16S rRNA gene and sequencing of colon contents

Fresh colon contents samples from 3 groups ( $n = 18$ ) were collected in cryogenic vials and stored immediately in liquid nitrogen. Amplification and sequencing of 16S rRNA genes of colon contents were performed on the Illumina platform (Illumina, San Diego, USA) according to the standard protocols of Majorbio Bio-Pharm Technology Co. Ltd. (Shanghai, China). The V3–V4 region of the bacterial 16S rRNA gene was amplified with primer pairs 338F (5'-ACTCCTACGGGAGGCAGCAG-3') and 806R (5'-GGACTACHVGGGTWTCTAAT-3') (Liu et al., 2016). Raw FASTQ files were de-multiplexed using an in-house Perl script, then quality-filtered by Fast version 0.19.6

(Chen et al., 2018) and merged by FLASH version 1.2.7 (<https://ccb.jhu.edu/software/FLASH/index.shtml>) (Mago and Salzberg, 2011) with the following criteria: (1) The 300 bp reads were truncated at any site receiving an average quality score of <20 over a 50 bp sliding window, and the truncated reads shorter than 50 bp were discarded; reads containing ambiguous characters were also discarded. (2) Only overlapping sequences longer than 10 bp were assembled according to their overlapped sequence. The maximum mismatch ratio of the overlap region is 0.2. Readings that could not be assembled were discarded. After that, operational taxonomic units (OTUs) clustering analysis and taxonomic analysis were performed (UPARSE, version 7.0.1090, <http://drive5.com/uparse/>) (Edgar, 2013). OTU clustering was performed on non-repetitive sequences (excluding single sequences) according to 97% similarity, and the chimeras were removed in the clustering process to obtain the representative sequences of OTUs. The sequences with more than 97% similarity to the representative sequences were selected, and the OTU table was generated. To obtain the corresponding species classification information of each OTU, compare the following databases: the bacterial and archaeal 16S rRNA databases [Silva (Release 138 <http://www.arb-silva.de>) and Greengene (Release 13.5 <http://greengenes.secondgenome.com/>)]. First, the  $\alpha$ -diversity analysis (<http://www.mothur.org/wiki/Calculators>) of colonic contents was assessed by community richness (ACE, Sobs, and Chao), community diversity (Shannon and Simpson), and community coverage (coverage) (Mothur, version 1.30.2, <https://mothur.org/wiki/calculators/>). Based on the results of OTU clustering analysis, a Venn diagram was used to display the common and unique microbiota among the three groups (R, version 3.3.1) at the genus level. Second, the composition and relative abundance were analyzed at the genus level in each group (R, version 3.3.1). Partial least squares discriminant analysis (PLS-DA) (Gromski et al., 2015) was used to analyze the similarity among the three groups at the genus level (<http://fiehnlab.ucdavis.edu/staff/kind/Statistics/Concepts/OPLS-PLSDA>, R, version 3.3.1). Permutational multivariate analysis of variance (PERMANOVA) (Kelly et al., 2015) was used to analyze the degree of explanation of different grouping factors on differences between samples (Bray-Curtis). Third, the non-metric multidimensional scaling (NMDS) was used to evaluate the  $\beta$ -diversity of colon contents among three groups at the genus level.

## Metagenomic sequencing and assembly

Total genome DNA was extracted from 200–300 mg of stool samples. Before analyzing the raw data generated using the Illumina sequencing platform (Illumina, San Diego, USA), the software Fast (<https://github.com/OpenGene/fastp>) was used to perform statistical and quality control

on the raw data to ensure subsequent analysis accuracy. About 300 bp of fragments were sequenced in metagenomics. [Supplementary Table 1](#) lists the original sequence and clean data statistics. Then, the software BWA v0.7.17 was used to decontaminate the sample's host genome. The host name was *Vertebrates: rattus\_norvegicus*. [Supplementary Table 1](#) lists the data after removing the host genome. Sequences of different sequencing depths were assembled using tEGAHIT v1.1.2 (<https://github.com/voutcn/megahit>) (Li et al., 2015). After splicing and assembly, the shortest contig length retained was 300 bp. [Supplementary Table 2](#) lists the data after assembly. Afterward, an open reading frame (ORF) prediction was performed on the contigs in the splicing results using Prodigal (<https://github.com/hyatt/Prodigal>). Genes with a nucleic acid length  $\geq$  of 100 bp were selected and translated to amino acid sequences to generate a statistical table of gene prediction results for each sample ([Supplementary Table 3](#)). CD-HIT software (<http://www.bioinformatics.org/cd-hit/>) (Fu et al., 2012) was used to cluster the predicted gene sequences of all samples (the default parameters are 90% identity and 90% coverage). The representative sequence for constructing a gene set with no duplicate genes was the longest gene in each class. The SOAPaligner software (<http://soap.genomics.org.cn/>) (Li et al., 2009) was used to align each sample's high-quality reads with the non-redundant gene set (default parameter: 95% identity) and count the genes in the corresponding sample's abundance information. Gene abundance was calculated using reads per kilobase million (Lawson et al., 2017). The linear discriminant analysis effect size (LEfSe) differential discriminant analysis ([http://huttenhower.sph.harvard.edu/galaxy/root?tool\\_id=lefse\\_upload](http://huttenhower.sph.harvard.edu/galaxy/root?tool_id=lefse_upload)) (Segata et al., 2011; Zhang et al., 2013) was used to identify the species that best explains the differences between groups in multiple samples.

## Taxonomic annotation and difference analysis

The non-redundant gene set was aligned with the non-redundant protein sequence database (NR database) using DIAMOND (Buchfink et al., 2015, 2021) (<https://github.com/bbuchfink/diamond>) (parameters: BLASTP; E-value  $\leq 1 \times 10^{-5}$ ). The taxonomic information database correlating to the NR database was used to get species annotation results. The species abundance was calculated by adding up the abundance of its genes. Venn plots were used to count shared and unique species across multiple groups, and community column charts were used to visually study the dominant species in a community.

## Fecal metabolomics

Different groups of fecal metabolites were detected using LC-MS, and metabolites with differential expressions were found. First, a 50 mg sample was weighed accurately, and then 400  $\mu$ L of extraction solution [methanol: water = 4:1 (v: v)] containing 0.02 mg/mL of the internal standard (L-2-chlorophenyl alanine) was added. Second, the sample was ground with a frozen tissue grinder for 6 min ( $-10^{\circ}\text{C}$ , 50 Hz), and ultrasonic extraction was used for 30 min ( $5^{\circ}\text{C}$ , 40 kHz) at low temperature. Third, let the sample stand at  $-20^{\circ}\text{C}$  for 30 min, centrifuge for 15 min (13,000 g,  $4^{\circ}\text{C}$ ), and then transfer the supernatant to the inlet with an inner cannula. In addition, 20  $\mu$ L of supernatant was pipetted from each sample and mixed as a quality control sample (Zheng et al., 2021; Zhu et al., 2022).

The raw data were imported into the metabolomics processing software ProgenesisQI (Waters Corporation, Milford, USA) for baseline filtering, peak identification, integration, retention time correction, peak alignment, and other tasks before a data matrix containing retention time, mass-to-charge ratio, and peak intensity was obtained. The software was then used to perform a library search of characteristic peaks, matching the MS and MS/MS mass spectral information to the metabolic database. The software was then used to search for characteristic peaks in a library by matching the information from the MS and MS/MS mass spectra to the metabolic database. The MS mass error was set to less than 10 ppm, and metabolites were identified based on secondary mass spectrometry matching scores. The main databases are <http://www.hmdb.ca/> and <https://metlin.scripps.edu/>. The total ion chromatograms of the quality control samples in positive and negative ion modes show that the peak shape is acceptable, and the distribution is relatively uniform under this detection condition ([Supplementary Figure 3](#)). The PLS-DA model quality parameters and the corresponding permutation test ( $R^2$  intercept, 0.8907;  $Q^2$  intercept,  $-0.7653$ ) demonstrate the statistical validity of the analysis and indicate distinct metabolic profiles among the three groups ([Supplementary Figure 2](#)).

Principal component analysis (Worley and Powers, 2013) was used to identify the "main" elements and structures in the data to reduce the dimensionality of the high-dimensional data space while minimizing the loss of data information. We also used the PLS-DA (Gromski et al., 2015) method to reduce the dimensionality of the data to better obtain the different information between groups. To prevent overfitting in the PLS-DA, we used the permutation test to determine whether the PLS-DA model is overfitting. The Kruskal-Wallis (H) test was used to compare differential metabolites among multiple groups. The metabolites that were common or unique among the differential metabolites between the two groups were displayed using Venn diagrams.

## Association of fecal differential metabolites with fecal microbiota

The *Pearson* correlation analysis found correlation coefficients between fecal microbiota and metabolites. The correlation between significantly differential fecal microbiota within the genus ( $P < 0.01$ ) and differential metabolites ( $P < 0.01$ ) is shown in Figure 6. The clustering algorithm was hierarchical clustering, the distance algorithm was Euclidean, and the hierarchical clustering method was complete (took the distance between the two data points that were the farthest from the two combined data points as the distance between the two combined data points).

## Association of fecal differential metabolites and metabolic pathways

Using *Pearson* correlation analysis, we perform a metabolic pathway study of our metagenomic data and link it to the chemicals identified in the stool samples. The correlation between significantly differential metabolites ( $P < 0.01$ ) and significantly differential metabolic pathways ( $P < 0.01$ ) in level 3 of the Kyoto Encyclopedia of Genes and Genomes (KEGG, <https://www.kegg.jp/>) (Figure 7).

## Statistical analysis

Sequencing was performed on the rats' feces ( $n = 18$ ). The Wilcoxon rank-sum test was used to calculate  $\alpha$ -diversity among three groups, and the results were expressed as mean  $\pm$  standard deviation.  $P < 0.05$  was deemed significant. The Kruskal-Wallis test (H test) was used to compare the differential microbiota among multiple groups, and the Tukey-Kramer procedure was used for *post-hoc* testing.  $P < 0.05$  was deemed significant. The Kruskal-Wallis test and Scheffe's *post-hoc* test were also used to find differences in metabolites between more than two groups.  $P < 0.05$  was considered significant (Edwards and Berry, 1987; Lee and Lee, 2018).

## Results

### H-EEI altered the $\beta$ -diversity of fecal microbiota

Different concentrations of elemene did not significantly affect the  $\alpha$ -diversity of fecal microbiota (Supplementary Table 4). The Venn diagrams showed bacteria that are unique and shared among different groups at the genus level (Figure 1A). There were 45 and 32 genera specific to groups S and L-EEI, respectively, and 60 genera specific to

H-EEI. Deeper taxonomic levels, including species, genera, and families, were investigated to learn more about the fecal microbiota of different groups. The relative abundance of the genus level in each group is shown in Figure 1B. The relative abundance of the species and families in each group is shown in Supplementary Figure 1. The NMDS analysis of different groups showed a distinct formation of clusters at the genus level (Figure 1C) and species level (Figure 1D). The differences between L-EEI and group S were less evident than those between H-EEI and group S, which formed quite different clusters. There is a clear visual distinction between group S and the H-EEI (stress = 0.150 for genera and stress = 0.193 for species).

### Distinctive sensitivity of fecal microbiota to different concentrations of EEI

We further compared the differences among the three groups to determine the effect of different concentrations of EEI on rat fecal microbiota. Figure 2A shows the abundance of the top 20 bacteria with significant differences at the species level among the three groups (Kruskal-Wallis test;  $P < 0.05$  was considered significant). Following the multiple-group test, we compared the groups using pairwise comparisons (Tukey-Kramer,  $P < 0.05$  was considered significant, Figure 2B). The effects of different EEI concentrations on the top 20 bacteria with significant differences at the species level among the three groups could be roughly divided into three categories: (1) A positive or negative correlation with the concentration of the EEI. The abundance of *Ileibacterium Valens* increased as the EEI concentration increased, while the abundance of *Firmicutes bacterium* and *Clostridium sp. CAC: 273* decreased. (2) The microbiota showed a tendency to increase first, then decrease or decrease first, and then increase as EEI concentration increased. Compared to group S, the abundance of *Prevotella sp. PCHR*, *Escherichia coli*, and *Candidatus Amulumpurator caecigallinarius* tended to decrease by L-EEI but significantly increased by H-EEI ( $P < 0.05$ ). Inversely, L-EEI significantly increased the abundance of *Ruminococcus bromii* ( $P < 0.01$ ) and *Dorea sp. 5-2* ( $P < 0.05$ ), and then, *Oscillibacter sp. 1-3* tended to increase, but H-EEI significantly decreased the abundance of these three microorganisms. (3) Both L-EEI and H-EEI tend to decrease the abundance of the bacteria (*Ruminococcaceae bacterium*, *Romboutsia ilealis*, and *Staphylococcus xylosus*).

First, the abundance of *Firmicutes bacterium* and *Clostridium sp. CAC: 273* decreased gradually as EEI concentration increased, and the H-EEI was significantly lower than group S ( $P < 0.05$ ). *Ileibacterium Valens* abundance increased gradually with the increasing EEI concentration, and the H-EEI was significantly higher than group S ( $P < 0.05$ ). Second, for *Prevotella sp. PCHR*, *Escherichia coli*, and

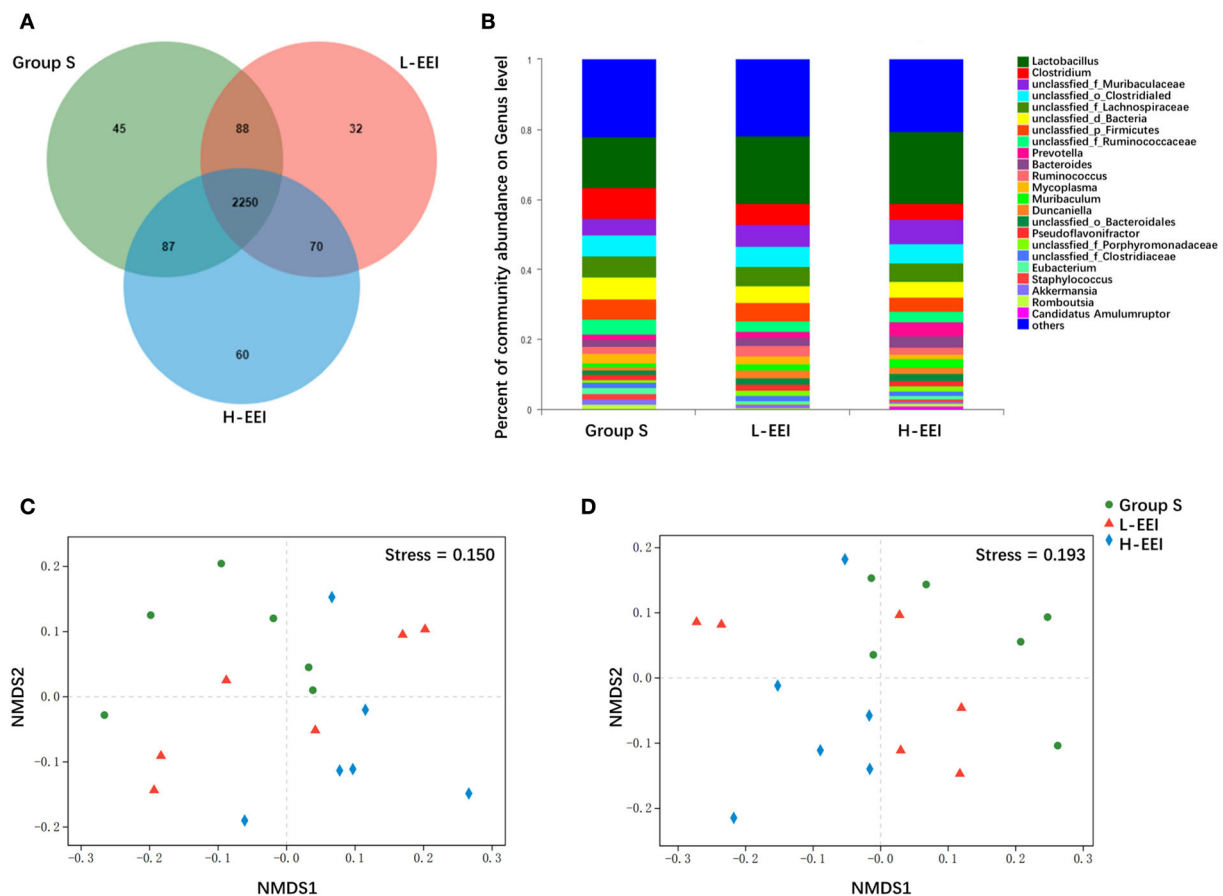


FIGURE 1

H-EEI altered the  $\beta$ -diversity of fecal microbiota. (A) The Venn diagrams show that bacteria that are unique and shared by different groups at the genus level. The overlapping part represents the bacteria shared by different sample groups, the non-overlapping part represents the bacteria unique to the sample group, and the number represents the corresponding species. There were 45 and 32 genera specific to groups S and L-EEI, respectively, and 60 genera specific to H-EEI. (B) The relative abundance of the genus level in each group is shown. The horizontal axis represents the sample name, and the vertical axis represents the proportion of microbiota in the sample. Colored columns represent different microbiota, and column length represents a proportion of the microbiota. (C,D) The NMDS analysis of different groups shows a distinct formation of clusters at the genus level (C), and according to species level (D). Bray–Curtis. NMDS is a data analysis method that simplifies research objects in multidimensional space into low-dimensional space for positioning, analysis, and classification while retaining the original relationship between objects. The pros and cons of NMDS analysis results are measured by stress. It is generally believed that when stress  $< 0.2$ , it can be represented by a two-dimensional point graph of NMDS, and this graph has a certain explanatory significance. The distance between the points indicates the degree of difference, and the horizontal and vertical coordinates indicate the relative distance, which has no practical significance. The data used in the analysis were derived from metagenomics sequencing data of fecal microbiota ( $n = 6$  in each group).

*Candidatus Amuluruptor caecigallinarius*, the L-EEI tended to decrease the abundance compared to group S, whereas the H-EEI significantly increased the abundance compared to the L-EEI ( $P < 0.05$ ). Third, for *Ruminococcus bromii*, *Oscillibacter* sp. 1–3, and *Dorea* sp. 5–2, L-EEI tended to increase the abundance compared to group S. In contrast, H-EEI significantly decreased the abundance compared to L-EEI ( $P < 0.05$ ). Finally, both L-EEI and H-EEI decreased the abundance of the *Ruminococcaceae* bacterium, *Romboutsia ilealis*, and *Staphylococcus xylosus*.

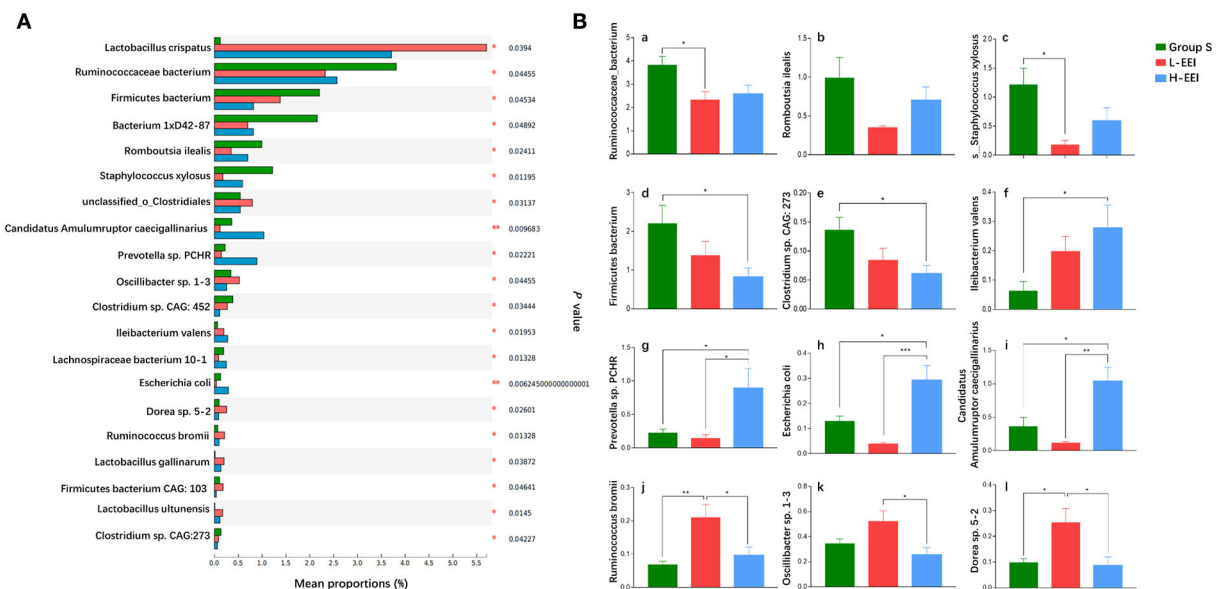
LEfSe analysis was applied to investigate the biomarkers among the three groups. We found 17 differentially abundant

taxa among the three groups, all with an LDA score  $> 3.0$  and  $P < 0.05$  (Figures 3A,B). LEfSe analysis identified ten genera with increased abundance in Group S, one genus with increased abundance in the L-EEI, and six genera with increased abundance in the H-EEI.

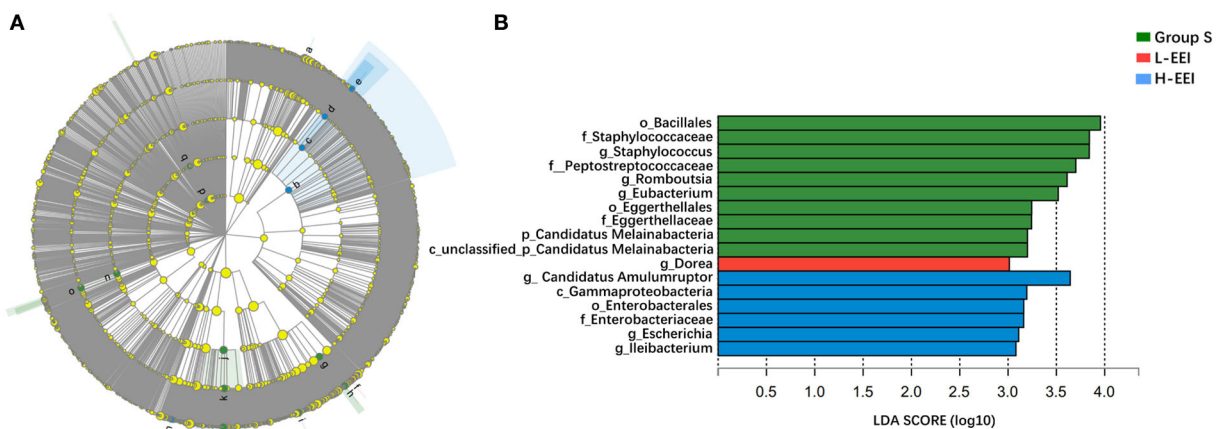
## L-EEI and H-EEI-induced differential fecal metabolites

We used LC-MS to directly test fecal metabolites to find out how different concentrations of EEI affected the metabolites of





**FIGURE 2** Distinctive sensitivity of fecal microbiota to different concentrations of EEI. (A) The abundance of the top 20 fecal bacteria with significant differences at the species level was determined using fecal metagenomics data (Kruskal–Wallis test,  $P < 0.05$  was considered significant). The vertical axis represents the species names, the horizontal axis represents the percentage value of a species' abundance in the sample, and different colors represent different groups (green to group S, red to L-EEI, and blue to H-EEI). (B) Post-hoc analysis of the top 20 fecal bacteria with significant differences in (A) (Tukey–Kramer,  $P < 0.05$  was considered significant).  $n = 6$  in each group, \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ .



**FIGURE 3** Biomarkers of L-EEI and H-EEI: *g\_Dorea* and *g\_Candidatus Amulnruptor*. (A) The hierarchy of linear discriminant analysis effect size (LEfSe) was based on fecal metagenomics data. Nodes in the figure with different colors represent microbial groups that are significantly enriched in the corresponding group and have a significant effect on the differences among the three groups; light yellow nodes represent microbial groups with no significant difference between groups or have no significant effect on the differences among the three groups. (B) The linear discriminant analysis (LDA) score (log10) for various microbial groups. The higher the LDA score, the greater the influence of microbiota abundance on the differential effect.  $P < 0.05$ , LDA > 3.0.

the fecal microbiota. Fecal metabolomics is the consequence of both host and microbiota interaction. The chemical signatures were identified based on internal standards (Majorbio Bio-Pharm Technology Co. Ltd.), the Kyoto Encyclopedia of

Genes and Genomes (KEGG, <http://www.genome.jp/kegg/>), the Human Metabolome Database (HMDB 5.0, [www.hmdb.ca](http://www.hmdb.ca)) and METLIN 2019 (<http://metlin.scripps.edu>). The analysis of fecal metabolites revealed variations in different EEI

concentrations. A total of 1,091 metabolites were identified, including vitamins and cofactors, peptides, nucleic acids, hormones and neurotransmitters, steroids, organic acids, lipids, and carbohydrates. A Venn diagram shows the common or unique metabolites found following a pairwise comparison of the three groups (Figure 4A). Figure 4B shows the findings of the PLS-DA model obtained from multivariate statistical comparisons of groups S, L-EEI, and H-EEI.

Similar to the fecal microbiota, fecal metabolites differ in their sensitivity to different EEI concentrations. Figure 5 shows the effect of different EEI concentrations on fecal metabolites. A total of 73 metabolites (Supplementary Table 5) showed highly significant ( $P < 0.01$ , Kruskal–Wallis test) differences among the three groups. Among the top 20 metabolites in the relative abundance of identified metabolites, two metabolites [Cis-9, 10-epoxystearic acid (HMDB0247617) and (+/-)-enterolactone (HMDB0006101)] show significant differences among the three groups ( $P < 0.05$ ). To be specific, compared with Group S, L-EEI has no significant effect on the relative abundance of Cis-9, 10-epoxystearic acid, while H-EEI could induce a significant increase compared with L-EEI ( $P < 0.01$ ); compared with Group S, both L-EEI and H-EEI could decrease the relative abundance of (+/-)-enterolactone significantly ( $P < 0.05$ ), while there was no significant difference between L-EEI and H-EEI.

## Many differentially expressed microorganisms are negatively correlated with metabolites

Fecal microbiota and fecal metabolite alterations are significantly related to EEI concentration. At the genus level, we linked microbiota to metabolites that were significantly different between the three groups (Figure 6). For example, *g\_Sediminibacillus* had a highly significant positive correlation with (3E, 5Z)-3,5-octadien-1-ol and a substantial negative correlation with 3-3-nor-3-oxopanasinsan-6-ol. In general, most bacteria are negatively correlated with metabolites.

## Many differential metabolites are negatively correlated with metabolic pathways

To conduct more in-depth correlation studies, we performed a metabolic pathway study on our metagenomic data and linked it to the chemicals identified in the stool samples (Figure 7). The results indicate that most differential metabolites (49/73) are negatively related to the following eight signaling pathways significantly: apoptosis (ko04214), carotenoid biosynthesis (ko00906), furfural degradation (ko00365), steroid biosynthesis (ko00100), phospholipase D signaling pathway (ko04072),

axon regeneration (ko04361), choline metabolism in cancer (ko05231), and oxytocin signaling pathway (ko04921). Except for the eight signaling pathways, there are another five signaling pathways worth being concerned about: autophagy yeast, flavone and flavonol biosynthesis, MAPK signaling pathways, measles, and toxoplasmosis.

## Discussion

EEI, a Chinese anti-tumor drug derived from the traditional Chinese medicinal plant *Curcuma wenyujin*, is primarily composed of  $\beta$ -elemene. The main component,  $\beta$ -elemene, has been shown in isotope labeling studies to penetrate the blood–brain barrier (Wu et al., 2009). Its content is comparable to other tissues, with a low incidence of myelosuppression during the medication process (Chen et al., 2022). EEI has been used clinically to treat cancerous pleural effusion by pleural and cancerous ascites by intraperitoneal perfusion (Luo et al., 2019; Qureshi et al., 2019; Zhai et al., 2019), but few studies have examined its impact on fecal microbiota.

EEI alters the fecal microbiota's  $\beta$ -diversity but not the  $\alpha$ -diversity. This indicates that EEI could alter fecal microbiota structure. Following EEI administration, some of the previously dominant microbiota are no longer dominant, and new dominant microbiota emerge. For example, the *Candidatus Amuluruptor* of *Muribaculaceae* was the most prominent bacterium in the H-EEI (LEfSe). The function of *Candidatus Amuluruptor* has not yet been reported. *Muribaculaceae* members use mucin monosaccharide (Pereira et al., 2020) and are abundant in the mice that had been fed an HFD (Liu et al., 2021).

*Ruminococcus bromii* and *Dorea sp.* 5-2 are particularly interesting because L-EEI significantly increased their abundance while H-EEI decreased it. In our previous study, *Ruminococcus* was the dominant bacterium in the EEI group using 16S rRNA technology (<https://www.ncbi.nlm.nih.gov/sra/PRJNA821627>), but the impact of *Ruminococcus* on human health is complex. Its benefits include the following: *Ruminococcus albus* has an inverse relationship with ulcerative colitis (Li et al., 2020) and protects infants from allergies (Wang et al., 2021). The harmful aspects include the association of *Ruminococcus gnavus* with Crohn's disease, which has been identified as causing Crohn's symptoms (Henke et al., 2019). In addition, *Ruminococcus* has been linked to irritable bowel syndrome (Baumgartner et al., 2021). In this study, we used metagenomic technology to confirm that L-EEI (20 mg/kg·d) increased the abundance of *Ruminococcus bromii*. According to current research, *Ruminococcus bromii* is a type of bacteria beneficial to humans; this bacterium can produce short-chain fatty acids and thus alleviate type 2 diabetic symptoms (Lordan et al., 2020; Yao et al., 2020). There are currently no reports on the function of *Dorea*

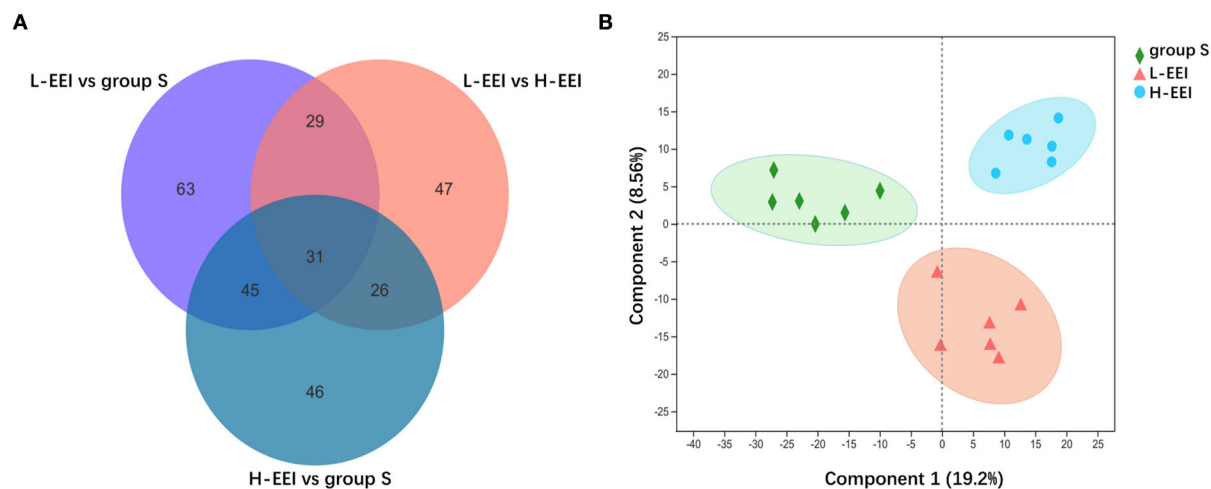


FIGURE 4

L-EEI and H-EEI induced differential fecal metabolites. (A) The common or unique metabolites are found after a pairwise comparison of three groups. In the figure, the overlapping part represents the number of metabolites shared by multiple metabolic sets, the non-overlapping part represents the number of metabolites unique to the metabolic set, and the number represents the number of corresponding metabolites. (B) PLS-DA score plot. The PLS-DA score map is used to visually represent the classification effect of the model. The greater the degree of separation among the three groups of samples in the figure, the more significant the classification effect. Component 1 is the first principal component explainability, and Component 2 is the second principal component explainability. The data used in the analysis were derived from LC-MS data of feces ( $n = 6$  in each group).

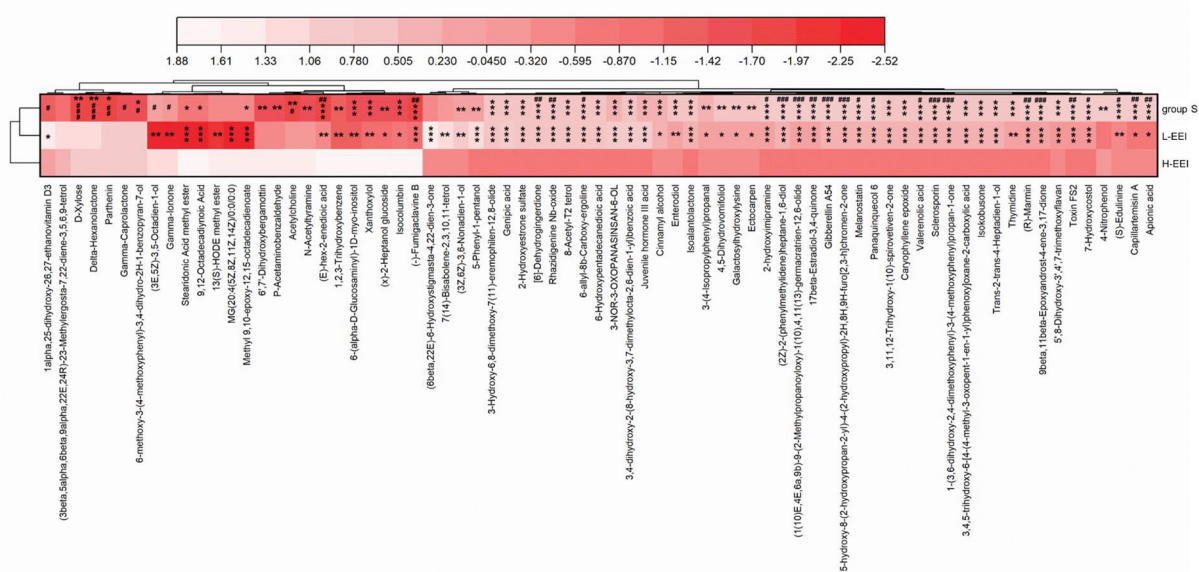
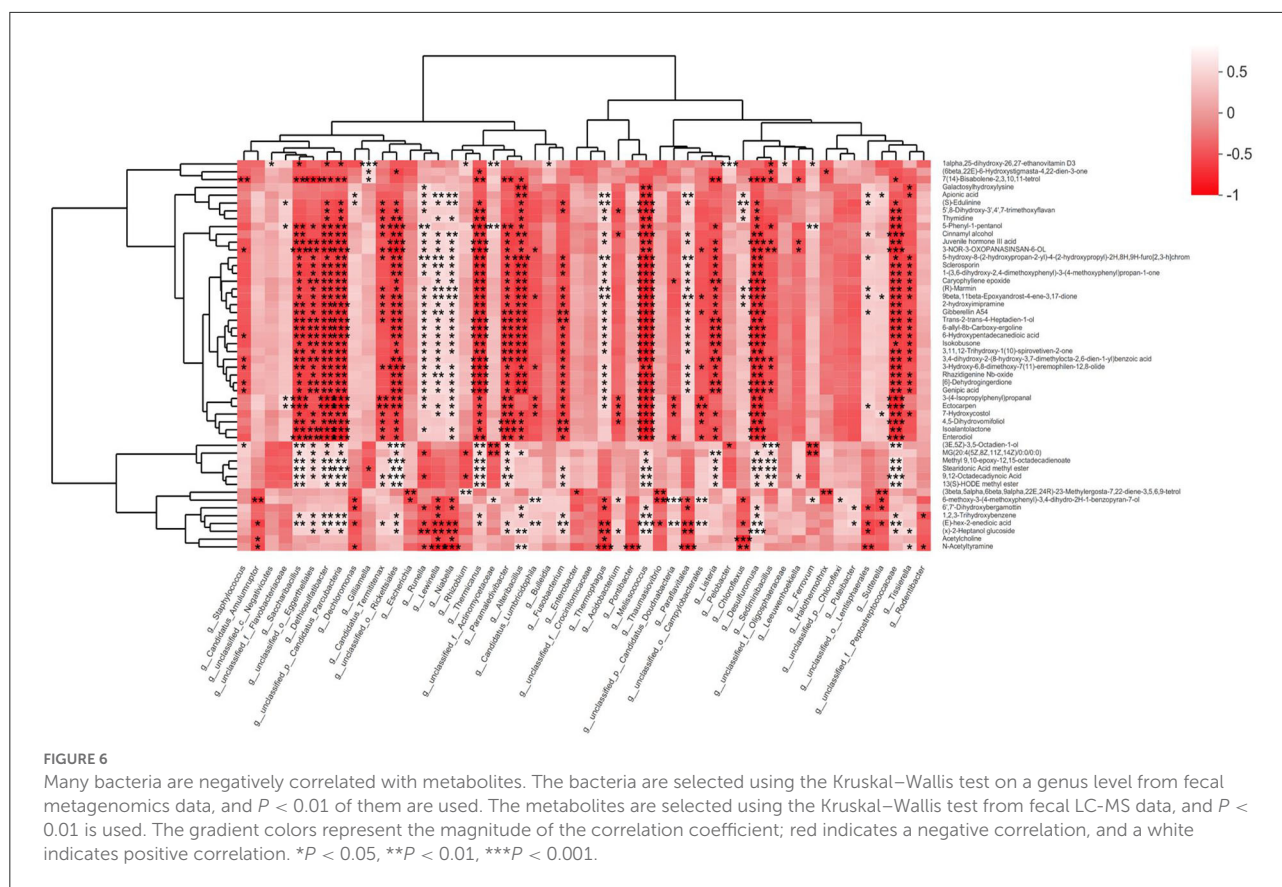


FIGURE 5

Differential metabolites among the three groups based on LC-MS data of feces. Darker colors indicate lower metabolite abundance, while lighter colors indicate higher metabolite abundance. The Kruskal–Wallis H test compares multiple groups, followed by the Scheffe *post-hoc* tests for metabolites with  $P < 0.01$ .  $n = 6$  in each group,  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$ , compared with group S;  $\#P < 0.05$ ,  $\#\#\#P < 0.001$ , compared with the L-EEI.

*sp.* 5-2, and most studies focus on *Dorea*. The impact of *Dorea* on human health is unclear, but studies have shown that *Dorea* is more abundant in patients with Parkinson's

(Petrov et al., 2017) and obesity (Jiao et al., 2018). It is unknown why the abundance of *Ruminococcus bromii* and *Dorea sp.* 5-2 significantly increased at L-EEI but decreased



at H-EEI. This is also one of the future topics we intend to investigate.

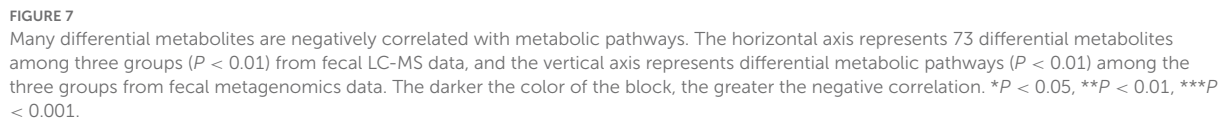
Fecal metabolomics is the consequence of both host and microbiota interaction. Figure 5 displays the effect of different concentrations of EEI on metabolites. A total of 73 metabolites had extremely significant statistical differences among the three groups ( $P < 0.01$ ). In the abundance of the top 20 differential metabolites, two kinds of metabolites [Cis-9,10-Epoxyoctadecanoic acid and (+/-)-Enterolactone] displayed significant statistical differences among the three groups ( $P < 0.05$ ). Compared to L-EEI, H-EEI had a higher Cis-9,10-epoxyoctadecanoic acid content. The effects of Cis-9,10-Epoxyoctadecanoic acid on humans are unclear. Cis-9,10-Epoxyoctadecanoic acid dose- and time-dependently increased the number and size of cellular lipid droplets in the human hepatocarcinoma cell line HepG2, decreasing cell viability and causing cell death (Liu et al., 2018; Liu Y. et al., 2020). Compared to group S, both L-EEI and H-EEI could decrease the abundance of (+/-)-enterolactone significantly ( $P < 0.05$ ). Enterolactone's health implications are debated. Enterolactone, a bioactive phenolic metabolite from dietary lignans, may help protect against different stages of breast, prostate, colon, gastric, and lung cancer (Mali et al., 2019; Senizza et al., 2020). Unfortunately, a case-control study found no association between serum EL levels and breast cancer risk (Kilkinen et al., 2004). A New York prospective study found comparable

outcomes (Zeleniuch-Jacquotte et al., 2004).  $\beta$ -elemene partially corrected HFD-induced alterations in mouse gut bacteria composition and metabolites (Zhou et al., 2021). Because EEI is used clinically for intraperitoneal infusion therapy of cancerous ascites, we examined its effects on rat fecal microbiota and metabolites (Jiang et al., 2012; Zhu et al., 2019).

In addition to fecal microbiota, we used 16S rRNA technology to investigate the effects of different EEI concentrations on colon contents. EEI concentrations, like fecal microbiota, did not affect  $\alpha$ -diversity of colonic microbiota (Supplementary Table 6). Genus-level community composition is shown in Supplementary Figure 4B. The NMDS analysis of different groups revealed distinct genus-level clusters. Similar to fecal microbiota, L-EEI and group S differed less than H-EEI, and group S. Group S and H-EEI stood out visually (Supplementary Figure 4D). Unlike the fecal microbiota, the most prominent biomarkers in the colon contents of low and high elemene groups were *g\_Marvinbryantia* and *g\_norank\_f\_Erysipelotrichaceae*, respectively (Supplementary Figure 5, LEfSe).

IP injection of EEI alters the microbiota and metabolites in rat feces. Since the gut microbiota is linked to several cancers and EEI inhibits the growth of various cancers, we hope to research the effect of EEI on the feces of patients with cancer.





that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be

evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.913461/full#supplementary-material>

## References

- Baumgartner, M., Lang, M., Holley, H., Crepaz, D., Hausmann, B., Pjevac, P., et al. (2021). Mucosal biofilms are an endoscopic feature of irritable bowel syndrome and ulcerative colitis. *Gastroenterology* 161, 1245–56.e20. doi: 10.1053/j.gastro.2021.06.024
- Boursi, B., Haynes, K., Mamtani, R., and Yang, Y. X. (2015a). Impact of antibiotic exposure on the risk of colorectal cancer. *Pharmacoevid. Drug. Saf.* 24, 534–542. doi: 10.1002/pds.3765
- Boursi, B., Mamtani, R., Haynes, K., and Yang, Y. X. (2015b). Recurrent antibiotic exposure may promote cancer formation—Another step in understanding the role of the human microbiota? *Eur. J. Cancer* 51, 2655–2664. doi: 10.1016/j.ejca.2015.08.015
- Buchfink, B., Reuter, K., and Drost, H. G. (2021). Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat. Methods* 18, 366–368. doi: 10.1038/s41592-021-01101-x
- Buchfink, B., Xie, C., and Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59–60. doi: 10.1038/nmeth.3176
- Cai, S. Z., Xiong, Q. W., Zhao, L. N., Ji, Y. T., Luo, Z. X., and Ma, Z. R. (2021).  $\beta$ -Elemene Triggers ROS-dependent apoptosis in glioblastoma cells through suppressing STAT3 Signaling pathway. *Pathol. Oncol. Res.* 27, 594299. doi: 10.3389/pore.2021.594299
- Chen, D., Wu, J., Jin, D., Wang, B., and Cao, H. (2019). Fecal microbiota transplantation in cancer management: Current status and perspectives. *Int. J. Cancer* 145, 2021–2031. doi: 10.1002/ijc.32003
- Chen, P., Li, X., Zhang, R., Liu, S., Xiang, Y., Zhang, M., et al. (2020). Combinative treatment of  $\beta$ -elemene and cetuximab is sensitive to KRAS mutant colorectal cancer cells by inducing ferroptosis and inhibiting epithelial-mesenchymal transformation. *Theranostics* 10, 5107–5119. doi: 10.7150/thno.44705
- Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i90. doi: 10.1093/bioinformatics/bty560
- Chen, Z. X. L., i. J., Liu, W. B., Zhang, S. R., and Sun, H. (2022). Elemene-containing hyperthermic intraperitoneal chemotherapy combined with chemotherapy for elderly patients with peritoneal metastatic advanced gastric cancer. *World J. Clin. Cases* 10, 1498–1507. doi: 10.12998/wjcc.v10.i5.1498
- Dik, V. K., van Oijen, M. G., Smeets, H. M., and Siersema, P. D. (2016). Frequent use of antibiotics is associated with colorectal cancer risk: results of a nested case-control study. *Dig. Dis. Sci.* 61, 255–264. doi: 10.1007/s10620-015-3828-0
- Edgar, R. C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* 10, 996–998. doi: 10.1038/nmeth.2604
- Edwards, D., and Berry, J. J. (1987). The efficiency of simulation-based multiple comparisons. *Biometrics* 43, 913–928. doi: 10.2307/2531545
- Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28, 3150–3152. doi: 10.1093/bioinformatics/bts565
- Gentile, C. L., and Weir, T. L. (2018). The gut microbiota at the intersection of diet and human health. *Science* 362, 776–780. doi: 10.1126/science.aau5812
- Gromski, P. S., Muhamadali, H., Ellis, D. I., Xu, Y., Correa, E., Turner, M. L., et al. (2015). A tutorial review: Metabolomics and partial least squares-discriminant analysis—a marriage of convenience or a shotgun wedding. *Anal. Chim. Acta* 879, 10–23. doi: 10.1016/j.aca.2015.02.012
- Hashem, S., Nisar, S., Sageena, G., Macha, M. A., Yadav, S. K., Krishnankutty, R., et al. (2021). Therapeutic effects of curcumin in several diseases; an overview. *Nutr. Cancer* 73, 181–195. doi: 10.1080/01635581.2020.1749676
- Henke, M. T., Kenny, D. J., Cassilly, C. D., Vlamakis, H., Xavier, R. J., Clardy, J., et al. (2019). *Ruminococcus gnavus*, a member of the human gut microbiome associated with Crohn's disease, produces an inflammatory polysaccharide. *Proc. Natl. Acad. Sci. U S A* 116, 12672–12677. doi: 10.1073/pnas.1904099116
- Ianiro, G., Tilg, H., and Gasbarrini, A. (2016). Antibiotics as deep modulators of gut microbiota: between good and evil. *Gut* 65, 1906–1915. doi: 10.1136/gutjnl-2016-312297
- Jiang, Z., Jacob, J. A., Loganathachetti, D. S., Nainangu, P., and Chen, B. (2017).  $\beta$ -elemene: Mechanistic studies on cancer cell interaction and its chemosensitization effect. *Front. Pharmacol.* 8, 105. doi: 10.3389/fphar.2017.00105
- Jiang, Z. Y., Qin, S. K., Yin, X. J., Chen, Y. L., and Zhu, L. (2012). Synergistic effects of Endostar combined with  $\beta$ -elemene on malignant ascites in a mouse model. *Exp. Ther. Med.* 4, 277–284. doi: 10.3892/etm.2012.583
- Jiao, N., Baker, S. S., Nugent, C. A., Tsompana, M., Cai, L., Wang, Y., et al. (2018). Gut microbiome may contribute to insulin resistance and systemic inflammation in obese rodents: a meta-analysis. *Physiol. Genomics* 50, 244–254. doi: 10.1152/physiolgenomics.00114.2017
- Kelly, B. J., Gross, R., Bittinger, K., Sherrill-Mix, S., Lewis, J. D., Collman, R. G., et al. (2015). Power and sample-size estimation for microbiome studies using pairwise distances and PERMANOVA. *Bioinformatics* 31, 2461–2468. doi: 10.1093/bioinformatics/btv183
- Kilkinen, A., Virtamo, J., Vartiainen, E., Sankila, R., Virtanen, M. J., Adlercreutz, H., et al. (2004). Serum enterolactone concentration is not associated with breast cancer risk in a nested case-control study. *Int. J. Cancer* 108, 277–280. doi: 10.1002/ijc.11519
- Lawson, C. E., Wu, S., Bhattacharjee, A. S., Hamilton, J. J., McMahon, K. D., Goel, R., et al. (2017). Metabolic network analysis reveals microbial community interactions in anammox granules. *Nat. Commun.* 8, 15416. doi: 10.1038/ncomms15416
- Lee, S., and Lee, D. K. (2018). What is the proper way to apply the multiple comparison test? *Korean J. Anesthesiol.* 71, 353–360. doi: 10.4097/kja.d.18.00242
- Li, D., Liu, C. M., Luo, R., Sadakane, K., and Lam, T. W. (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31, 1674–1676. doi: 10.1093/bioinformatics/btv033
- Li, Q., Ding, X., Liu, K., Marcella, C., Liu, X., Zhang, T., et al. (2020). Fecal microbiota transplantation for ulcerative colitis: the optimum timing and gut microbiota as predictors for long-term clinical outcomes. *Clin. Transl. Gastroenterol.* 11, e00224. doi: 10.14309/ctg.0000000000000224
- Li, R., Yu, C., Li, Y., Lam, T. W., Yiu, S. M., Kristiansen, K., et al. (2009). SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* 25, 1966–1967. doi: 10.1093/bioinformatics/btp336
- Liu, C., Zhao, D., Ma, W., Guo, Y., Wang, A., Wang, Q., et al. (2016). Denitrifying sulfide removal process on high-salinity wastewaters in the presence of *Halomonas* sp. *Appl. Microbiol. Biotechnol.* 100, 1421–1426. doi: 10.1007/s00253-015-7039-6
- Liu, S., Li, Q., Li, G., Zhang, Q., Zhuo, L., Han, X., et al. (2020). The mechanism of m(6)A methyltransferase METTL3-mediated autophagy in

reversing gefitinib resistance in NSCLC cells by  $\beta$ -elemene. *Cell Death Dis.* 11, 969. doi: 10.1038/s41419-020-03148-8

Liu, Y., Cheng, Y., Li, J., Wang, Y., and Liu, Y. (2018). Epoxy stearic acid, an oxidative product derived from oleic acid, induces cytotoxicity, oxidative stress, and apoptosis in HepG2 cells. *J. Agric. Food Chem.* 66, 5237–5246. doi: 10.1021/acs.jafc.8b01954

Liu, Y., Li, J., and Liu, Y. (2020). Effects of epoxy stearic acid on lipid metabolism in HepG2 cells. *J. Food Sci.* 85, 3644–3652. doi: 10.1111/1750-3841.15405

Liu, Y., Yang, K., Jia, Y., Shi, J., Tong, Z., Fang, D., et al. (2021). Gut microbiome alterations in high-fat-diet-fed mice are associated with antibiotic tolerance. *Nat. Microbiol.* 6, 874–884. doi: 10.1038/s41564-021-00912-0

Lordan, C., Thapa, D., Ross, R. P., and Cotter, P. D. (2020). Potential for enriching next-generation health-promoting gut bacteria through prebiotics and other dietary components. *Gut. Microbes.* 11, 1–20. doi: 10.1080/19490976.2019.1613124

Luo, H., Vong, C. T., Chen, H., Gao, Y., Lyu, P., Qiu, L., et al. (2019). Naturally occurring anti-cancer compounds: shining from Chinese herbal medicine. *Chin. Med.* 14, 48. doi: 10.1186/s13020-019-0270-9

Ma, L. T., Bai, Y., Li, J., Qiao, Y., Liu, Y., Zheng, J., et al. (2021). Elemene emulsion injection administration reduces neuropathic pain by inhibiting astrocytic NDRG2 expression within spinal dorsal horn. *Chin. J. Integr. Med.* 27, 912–918. doi: 10.1007/s11655-021-3438-3

Magoč, T., and Salzberg, S. L. (2011). FLASH. fast length adjustment of short reads to improve genome assemblies. *Bioinformatics.* 27, 2957–2963. doi: 10.1093/bioinformatics/btr507

Mali, A. V., Padhye, S. B., Anant, S., Hegde, M. V., and Kadam, S. S. (2019). Anticancer and antimetastatic potential of enterolactone: Clinical, preclinical and mechanistic perspectives. *Eur. J. Pharmacol.* 852, 107–124. doi: 10.1016/j.ejphar.2019.02.022

Markowiak, P., and Slizewska, K. (2017). Effects of probiotics, prebiotics, and synbiotics on human health. *Nutrients* 9, 1021. doi: 10.3390/nu9091021

Pereira, F. C., Wasmund, K., Cobankovic, I., Jehmlich, N., Herbold, C. W., Lee, K. S., et al. (2020). Rational design of a microbial consortium of mucosal sugar utilizers reduces *Clostridiodes difficile* colonization. *Nat. Commun.* 11, 5104. doi: 10.1038/s41467-020-18928-1

Petrov, V. A., Saltykova, I. V., Zhukova, I. A., Alifirova, V. M., Zhukova, N. G., Dorofeeva, Y. B., et al. (2017). Analysis of gut microbiota in patients with Parkinson's disease. *Bull. Exp. Biol. Med.* 162, 734–737. doi: 10.1007/s10517-017-3700-7

Qureshi, M. Z., Attar, R., Romero, M. A., Sabitaliyevich, U. Y., Nurmurazayevich, S. B., Ozturk, O., et al. (2019). Regulation of signaling pathways by  $\beta$ -elemene in cancer progression and metastasis. *J. Cell Biochem.* 120, 12091–12100. doi: 10.1002/jcb.28624

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S., et al. (2011). Metagenomic biomarker discovery and explanation. *Genome Biol.* 12, R60. doi: 10.1186/gb-2011-12-6-r60

Senizza, A., Rocchetti, G., Mosele, J. I., Patrone, V., Callegari, M. L., Morelli, L., et al. (2020). Lignans and gut microbiota: An interplay revealing potential health implications. *Molecules* 25, 5709. doi: 10.3390/molecules2525709

Singh, R. K., Chang, H. W., Yan, D., Lee, K. M., Ucmak, D., Wong, K., et al. (2017). Influence of diet on the gut microbiome and implications for human health. *J. Transl. Med.* 15, 73. doi: 10.1186/s12967-017-1175-y

Smits, L. P., Bouter, K. E., de Vos, W. M., Borody, T. J., and Nieuwdorp, M. (2013). Therapeutic potential of fecal microbiota transplantation. *Gastroenterology.* 145, 946–953. doi: 10.1053/j.gastro.2013.08.058

Sommer, F., Anderson, J. M., Bharti, R., Raes, J., and Rosenstiel, P. (2017). The resilience of the intestinal microbiota influences health

and disease. *Nat. Rev. Microbiol.* 15, 630–638. doi: 10.1038/nrmicro.2017.58

Sun, W., Kim, D. H., Byon, C. H., Choi, H. I., Park, J. S., Bae, E. H., et al. (2022).  $\beta$ -elemene attenuates renal fibrosis in the unilateral ureteral obstruction model by inhibition of STAT3 and Smad3 signaling via suppressing MyD88 expression. *Int. J. Molec. Sci.* 23, 5553. doi: 10.3390/ijms23105553

Tong, H., Liu, Y., Jiang, L., and Wang, J. (2020). Multi-targeting by  $\beta$ -elemene and its anticancer properties: A good choice for oncotherapy and radiochemotherapy sensitization. *Nutr. Cancer.* 72, 554–567. doi: 10.1080/01635581.2019.1648694

Wang, S., Wei, Y., Liu, L., and Li, Z. (2021). Association between breastmilk microbiota and food allergy in infants. *Front. Cell Infect. Microbiol.* 11, 770913. doi: 10.3389/fcimb.2021.770913

Worley, B., and Powers, R. (2013). Multivariate analysis in metabolomics. *Curr. Metabol.* 1, 92–107. doi: 10.2174/2213235X11301010092

Wu, Q., Shi, X., Pan, Y., Liao, X., Xu, J., Gu, X., et al. (2022). The chemopreventive role of  $\beta$ -elemene in cholangiocarcinoma by restoring PCDH9 expression. *Front. Oncol.* 12, 874457. doi: 10.3389/fonc.2022.874457

Wu, X. S., Xie, T., Lin, J., Fan, H. Z., Huang-Fu, H. J., Ni, L. F., et al. (2009). An investigation of the ability of elemene to pass through the blood-brain barrier and its effect on brain carcinomas. *J. Pharm. Pharmacol.* 61, 1653–1656. doi: 10.1211/jpp.61.12.0010

Yao, Y., Yan, L., Chen, H., Wu, N., Wang, W., Wang, D., et al. (2020). Cyclocarya paliurus polysaccharides alleviate type 2 diabetic symptoms by modulating gut microbiota and short-chain fatty acids. *Phytomedicine.* 77, 153268. doi: 10.1016/j.phymed.2020.153268

Zeleniuch-Jacquotte, A., Adlercreutz, H., Shore, R. E., Koenig, K. L., Kato, I., Arslan, A. A., et al. (2004). Circulating enterolactone and risk of breast cancer: a prospective study in New York. *Br. J. Cancer.* 91, 99–105. doi: 10.1038/sj.bjc.6601893

Zhai, B., Zhang, N., Han, X., Li, Q., Zhang, M., Chen, X., et al. (2019). Molecular targets of  $\beta$ -elemene, a herbal extract used in traditional Chinese medicine, and its potential role in cancer therapy: A review. *Biomed. Pharmacother.* 114, 108812. doi: 10.1016/j.biopha.2019.108812

Zhang, C., Li, S., Yang, L., Huang, P., Li, W., Wang, S., et al. (2013). Structural modulation of gut microbiota in life-long calorie-restricted mice. *Nat. Commun.* 4, 2163. doi: 10.1038/ncomms3163

Zheng, C., Cai, X., Wu, S., Liu, Z., Shi, Y., Zhou, W., et al. (2014). Enhancing effect of  $\beta$ -elemene emulsion on chemotherapy with harringtonine, aclinomycin, and Ara-c in treatment of refractory/relapsed acute myeloid leukemia. *Pakistan J. Med. Sci.* 30, 1270–1272. doi: 10.12669/pjms.306.5207

Zheng, J., Ye, C., Hu, B., Yang, H., Yao, Q., Ma, J., et al. (2021). Bile acid profiles in bile and feces of obese mice by a high-performance liquid chromatography-tandem mass spectrometry. *Biotechnol. Appl. Biochem.* 68, 1332–1341. doi: 10.1002/bab.2055

Zhihua, Y. U. E. H. L. (2018). Analysis of improvement of national standards of elemene and its preparations. *Chin. J. Pharmaceut.* 49, 1187–1190. doi: 10.16522/j.cnki.cjph.2018.08.023

Zhou, Y., Qiu, W., Wang, Y., Wang, R., Takano, T., Li, X., et al. (2021).  $\beta$ -Elemene suppresses obesity-induced imbalance in the microbiota-gut-brain axis. *Biomedicine* 9, 704. doi: 10.3390/biomedicine9070704

Zhu, J., Li, B., Ji, Y., Zhu, L., Zhu, Y., Zhao, H., et al. (2019).  $\beta$ -elemene inhibits the generation of peritoneum effusion in pancreatic cancer via suppression of the HIF1A-VEGFA pathway based on network pharmacology. *Oncol. Rep.* 42, 2561–2571. doi: 10.3892/or.2019.7360

Zhu, T., Hu, B., Ye, C., Hu, H., Yin, M., Zhang, Z., et al. (2022). Bletilla striata oligosaccharides improve ulcerative colitis by regulating gut microbiota and intestinal metabolites in dextran sulfate sodium-induced mice. *Front. Pharmacol.* 13, 867525. doi: 10.3389/fphar.2022.867525



## OPEN ACCESS

## EDITED BY

Graciela L. Lorca,  
University of Florida,  
United States

## REVIEWED BY

Christina Bourne,  
University of Oklahoma,  
United States  
Stephanie Petrella,  
Université de Paris,  
France

## \*CORRESPONDENCE

Yuk-Ching Tse-Dinh  
ytsedinh@fiu.edu

## SPECIALTY SECTION

This article was submitted to  
Microbial Physiology and Metabolism,  
a section of the journal  
Frontiers in Microbiology

RECEIVED 30 August 2022

ACCEPTED 17 November 2022

PUBLISHED 05 December 2022

## CITATION

Garcia PK, Martinez Borrero R, Annamalai T,  
Diaz E, Balarezo S, Tiwari PB and  
Tse-Dinh YC (2022) Localization of  
*Mycobacterium tuberculosis*  
topoisomerase I C-terminal sequence motif  
required for inhibition by endogenous toxin  
MazF4.

Front. Microbiol. 13:1032320.

doi: 10.3389/fmicb.2022.1032320

## COPYRIGHT

© 2022 Garcia, Martinez Borrero,  
Annamalai, Diaz, Balarezo, Tiwari and  
Tse-Dinh. This is an open-access article  
distributed under the terms of the [Creative  
Commons Attribution License \(CC BY\)](#). The  
use, distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in  
this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Localization of *Mycobacterium tuberculosis* topoisomerase I C-terminal sequence motif required for inhibition by endogenous toxin MazF4

Pamela K. Garcia<sup>1</sup>, Rosemarie Martinez Borrero<sup>1</sup>,  
Thirunavukkarasu Annamalai<sup>1,2</sup>, Esnel Diaz<sup>1</sup>, Steve Balarezo<sup>1</sup>,  
Purushottam B. Tiwari<sup>3</sup> and Yuk-Ching Tse-Dinh<sup>1,2\*</sup>

<sup>1</sup>Department of Chemistry and Biochemistry, Florida International University, Miami, FL, United States,

<sup>2</sup>Biomolecular Sciences Institute, Florida International University, Miami, FL, United States,

<sup>3</sup>Department of Oncology, Georgetown University, Washington, DC, United States

Only about half the multi-drug resistant tuberculosis (MDR-TB) cases are successfully cured. Thus, there is an urgent need of new TB treatment against a novel target. *Mycobacterium tuberculosis* (*Mtb*) topoisomerase I (TopA) is the only type IA topoisomerase in this organism and has been validated as an essential target for TB drug discovery. Toxin-antitoxin (TA) systems participate as gene regulators within bacteria. The TA systems contribute to the long-term dormancy of *Mtb* within the host-cell environment. *Mtb*'s toxin MazF4 (Rv1495) that is part of the MazEF4 TA system has been shown to have dual activities as endoribonuclease and topoisomerase I inhibitor. We have developed a complementary assay using an *Escherichia coli* strain with temperature-sensitive *topA* mutation to provide new insights into the MazF4 action. The assay showed that *E. coli* is not sensitive to the endoribonuclease activity of *Mtb* MazF4 but became vulnerable to MazF4 growth inhibition when recombinant *Mtb* TopA relaxation activity is required for growth. Results from the complementation by *Mtb* TopA mutants with C-terminal deletions showed that the lysine-rich C-terminal tail is required for interaction with MazF4. Site-directed mutagenesis is utilized to identify two lysine residues within a conserved motif in this C-terminal tail that are critical for MazF4 inhibition. We performed molecular dynamics simulations to predict the *Mtb* TopA-MazF4 complex. Our simulation results show that the complex is stabilized by hydrogen bonds and electrostatic interactions established by residues in the TopA C-terminal tail including the two conserved lysines. The mechanism of *Mtb* TopA inhibition by MazF4 could be useful for the discovery of novel inhibitors against a new antibacterial target in pathogenic mycobacteria for treatment of both TB and diseases caused by the non-tuberculosis mycobacteria (NTM).

## KEYWORDS

topoisomerase, tuberculosis, MazF toxin, TopA, mycobacteria



## Introduction

*Mycobacterium tuberculosis* (*Mtb*) is the microorganism that causes Tuberculosis (TB), a leading infectious disease in the world with 10.0 million individuals experiencing active-TB, while 1.5 million succumbed in 2020 (Dean et al., 2022). The pathogen's ability to remain dormant (also known as latent-TB) by eluding the host's immune system, synergy between TB and AIDS, and active TB's long 6-month therapy requirement have attributed to TB's persistence (Kwan and Ernst, 2011; World Health Organization, 2020; Seki et al., 2021). The growing threat of drug resistance in *Mtb* poses challenge for effective treatments outcome (Kempker et al., 2015; Dean et al., 2022). As antimicrobial resistance represents a global public health threat (Boucher, 2020; eClinicalMedicine, 2021); treatments that can target TB at its latent as well as active stages are needed.

Topoisomerases are enzymes present in all domains of life (Bush et al., 2015; Garnier et al., 2021). The requirement of topoisomerases is due to its role in removing topological barriers that arise during important genomic processes, including transcription, replication, and recombination (Vos et al., 2011; Pommier et al., 2016; Brochu et al., 2020; McKie et al., 2021). Topoisomerases are classified as type I or type II topoisomerases according to the number of strands that they cleave during catalysis. Type I topoisomerases are further subclassified according to their sequence/structure homology and mechanism of actions (Baker et al., 2009). Bacterial topoisomerase I belongs to the type IA subfamily (Garnier et al., 2018; Dasgupta et al., 2020). The genome sequencing of *Mtb* confirmed that this organism encodes topoisomerase I (TopA) as the only type I topoisomerase along with gyrase as the only type II topoisomerase. Gyrase is well known as an important target for fluoroquinolones (Bush et al., 2020) including moxifloxacin used currently for treatment of drug resistant TB (World Health Organization, 2022). Genetic studies using saturation mutagenesis (Sasseti et al., 2003) and conditional knockdown (Ahmed et al., 2014; Ravishankar et al., 2015) have shown that *Mtb* TopA is essential for viability *in vitro* and *in vivo*, validating *Mtb* TopA as a target for discovery of new TB therapy (Nagaraja et al., 2017).

The high number (>80 putative) of toxin-antitoxin (TA) systems identified in *Mtb* suggests that the TA systems may contribute to *Mtb*'s long-term dormancy and tenacity in the host (Ramage et al., 2009; Sala et al., 2014). The toxin Rv1495 (MazF4) belongs to the MazEF TA family, where MazE is the antitoxin and MazF is the toxin with sequence specific endoribonuclease activity (Zhu et al., 2006, 2008). A previous study reported that in addition to exhibiting ribonucleolytic activity, *Mtb* MazF4 interacts directly with the *Mtb* TopA and mutually inhibit each other's enzymatic activity (Huang and He, 2010) through interaction involving the C-terminal domains (amino acid 622-934) of *Mtb* TopA. Bacterial topoisomerase I has highly conserved N-terminal domains (D1–D4) that form the active site for cleaving and rejoining of a single-strand of DNA (Baker et al., 2009; Dasgupta et al., 2020). The more variable C-terminal region (Diaz et al., 2022) also contributes to the relaxation of supercoiled DNA. The *Mtb* TopA C-terminal region has the same organization with other mycobacteria, consisting of four repeated Topo\_C\_Rpt

domains (D5–D8) that can bind ssDNA (Tan et al., 2016; Cao et al., 2020), and a lysine-rich C-terminal tail ~20 amino acids in length that also plays a role in the *Mtb* TopA relaxation activity (Ahmed et al., 2013). The inhibition of *Mtb* TopA by MazF4 could potentially serve as a model for novel small molecule antibacterial compounds selective for pathogenic mycobacteria. Such antibacterial compounds may have advantage over broad-spectrum antibacterial compounds targeting TopA active site in avoiding harmful effects on the human microbiome. We report here results that further characterize the selective inhibition of *Mtb* TopA catalytic activity by MazF4, and determination of the binding site of *Mtb* MazF4 in the TopA C-terminal region.

## Materials and methods

### Cloning of *Mtb* TopA, MazF4 and *Escherichia coli* TopA

Cloning of *Mtb* TopA coding sequence in strain H37Rv into pLIC-HK vector (Corn et al., 2005) to generate plasmid pLIC-MTOP (with T7-promoter, N-terminal His-tag and kanamycin selection) has been described (Annamalai et al., 2009). The cloning of *E. coli* TopA coding sequence into the same position of pLIC-HK vector has also been reported previously (Cheng et al., 2008). The MazF4 (Rv1495) toxin gene amplified from H37Rv genomic DNA (from ATCC) was cloned using the Gibson assembly cloning procedures (Gibson, 2011) in the plasmid pBAD/Thio (from Invitrogen) using the NEB Gibson Assembly Master Mix to generate plasmid pBAD-Rv1495 (with BAD promoter, N-terminal thioredoxin tag and ampicillin resistance). Primers used in the cloning are listed in Supplementary Table S1. The clones were isolated as transformants of *Escherichia coli* NEB® 5-alpha competent cells (from New England BioLabs).

### Random and site-directed mutagenesis of *Mtb* TopA

A library of mutant pLIC-MTOP plasmid with random mutations (Muteeb and Sen, 2010) in the TopA coding region was generated by first passing the plasmid through *E. coli* XL1-Red strain with *mutS*, *mutD*, and *mutT* mutations (from Agilent). The pLIC-MTOP plasmid extracted from XL1-Red was amplified in *E. coli* XL1-Blue. The coding region for TopA was excised with restriction enzymes BamHI and NcoI and ligated to the vector backbone generated from digestion of the original plasmid. The mutant pLIC-MTOP plasmid library was obtained from transformation of XL1-Blue with the ligation product.

pLIC-MTOP plasmids that express truncated version of *Mtb* TopA terminating after residue 840 or 910 were made by assembly cloning of vector and insert fragments generated with primers listed in Supplementary Table S1. Alanine substitutions of specific lysine residues in the ~20 a.a. C-terminal tail of *Mtb* TopA were later introduced into pLIC-MTOP plasmid using the NEB HiFi mutagenesis Master Mix and primers listed in Supplementary Table S1.

## Growth complementation assay

*Escherichia coli* strain AS17 *topA17(am)* and *Tet<sup>r</sup>supD43,74* with temperature sensitive *topA* is not viable at 42°C (Wang et al., 2002). The growth complementation of AS17 at the non-permissive temperature by recombinant *Mtb* TopA (Narula et al., 2010) can be tested for inhibition by MazF4 toxin when co-expressed. For the growth complementation assays, AS17 transformants of pLIC-MTOP and pBAD-Rv1495 or control plasmids were grown from single colonies overnight at 30°C in LB broth (Miller's) in presence of kanamycin (50 µg/ml) and carbenicillin (100 µg/ml). The OD<sub>600</sub> of the overnight cultures was adjusted to 1.0 and 10-fold serial dilutions were made with LB medium. Five microliters of serially diluted cells were spotted on LB plates (with antibiotics kanamycin 50 µg/ml and carbenicillin 100 µg/ml) containing 0.2% arabinose to induce the expression of the MazF4 toxin from the BAD promoter in pBAD-Rv1495. The plates in duplicates were incubated at 30°C for 2–3 days or 42°C for 1–2 days before they were photographed. Each complementation assay was repeated 3 times to confirm that the results are similar.

## Western blot analysis

Western blot analysis using rabbit polyclonal antibodies raised against *Mtb* TopA was conducted to compare the levels of *Mtb* TopA expression in *E. coli* AS17. Overnight cultures of AS17 transformants of pLIC-MTOP grown at 30 or 42°C were resuspended in sterile water to a concentration of OD<sub>600</sub> = 10. Equal volumes of resuspended cell suspension were mixed with 2×-SDS gel loading buffer and boiled for 5 min to yield whole cell lysates. Whole cell lysates corresponding to equal cell numbers for different transformants or conditions were electrophoresed on a 10% SDS gel and transferred on to a nylon membrane for western blot analysis using rabbit anti-*Mtb* TopA (polyclonal) as primary antibody (Sandhaus et al., 2016) and goat anti-rabbit-HRP (Santa Cruz Biotechnology) as secondary antibody. The C-DiGit blot scanner (LI-COR) was used to measure the chemiluminescent western blot signal. Blotted membranes were also stained (Pierce™ reversible stain) to confirm for equal loading and gel transfer of proteins.

## Molecular dynamics simulation

We created a full length *Mtb* TopA (MTOP-FL) structure using RaptorX (Källberg et al., 2012) and Modeller (Šali and Blundell, 1993). MazF4 dimer structure was taken from PDB entry 5XE3 (Ahn et al., 2017). At least 50 ns molecular dynamics (MD) simulations were conducted for both MTOP-FL and MazF4 using NAMD simulation package (Phillips et al., 2005). CHARMM-GUI (Lee et al., 2016) was used to create simulation input files. Protein coordinates at the end of the simulations were used to predict a complex structure between *Mtb* TopA and MazF4 using ZDOCK webserver (Pierce et al., 2011). Finally, the docked complex was simulated for 150 ns using the NAMD simulation package. VMD (Humphrey et al., 1996) was used to visualize and analyze the

simulated structures as well as to generate pictures of the complex structures. GraphPad Prism was used to plot graphs.

## Results

### *Mtb* MazF4 inhibits specifically the growth complementation of *Escherichia coli topA<sup>ts</sup>* mutation by *Mtb* TopA

Incubation of plates spotted with serial dilutions of overnight culture grown at 30°C (Figure 1) showed that *E. coli* strain AS17 with *topA<sup>ts</sup>* mutation is not viable at 42°C (Wang et al., 2002) unless complemented by recombinant *E. coli* TopA (from plasmid pLIC-ETOP) or *Mtb* TopA (from plasmid pLIC-MTOP). When an additional plasmid pBAD-Rv1495 expressing the MazF4 toxin was present in the cell, growth complementation by recombinant *E. coli* TopA was not affected, while growth of AS17 complementation by recombinant *Mtb* TopA was reduced by up to 10<sup>3</sup>-fold. The CFU/ml values were measured in three replicated experiments to confirm the reduction of growth complementation by *Mtb* TopA by >200-fold due to the presence of the *Mtb* MazF4 toxin (Supplementary Table S2). The resistance of *E. coli* TopA can be explained by the absence of protein structure and sequence homology between the C-terminal domains of TopA from *E. coli* and *Mtb* (Tan et al., 2015, 2016; Dasgupta et al., 2020; Diaz et al., 2022). The lack of growth inhibition at 30°C and complementation by pLIC-ETOP provided evidence that the ribonuclease activity of MazF4 has no effect on growth of *E. coli*. *Mtb* MazF4 inhibition of AS17 growth complementation by *Mtb* TopA but not *E. coli* TopA provided support that inhibition of *Mtb* TopA growth complementation results from reduction of *in vivo* activity of *Mtb* TopA but not *E. coli* TopA. There is no T7 RNA polymerase gene in *E. coli* AS17. Western blot analysis showed that expression of recombinant *Mtb* TopA from the unknown promoter in pLIC-HK derived clone is higher at 42°C than 30°C (Supplementary Figure S1). This complementation assay provides a tool to further study the mechanism of inhibition of *Mtb* TopA by MazF4.

### Selection of mutant pLIC-MTOP clones resistant to inhibition of growth complementation by pBAD-Rv1495

We prepared a library of mutated pLIC-MTOP plasmids by passage through the mutator strain XL1-Red. The TopA coding region was excised and ligated to the vector fragment of the original pLIC-MTOP clone. The resulting LIC-pMTOP1 library was then transformed into *E. coli* AS17 cells that have plasmid pBAD-Rv1495. Individual pLIC-MTOP plasmids were isolated from colonies that grew well at 42°C. Transformation and growth complementation assay was repeated to identify isolates of mutant pLIC-MTOP plasmid that could complement 42°C growth of AS17/pBAD-Rv1495 at much greater efficiency than the original pLIC-MTOP plasmid. Mutant-3 shown in Figure 2A

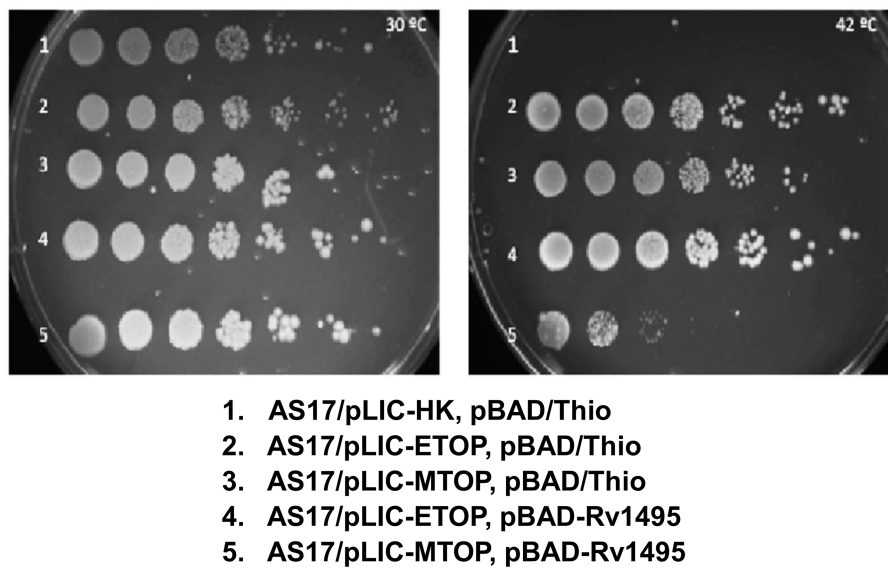


FIGURE 1

Assay for growth complementation of *Escherichia coli* AS17 with temperature sensitive *topA*. Ten-fold serial dilution of *E. coli* AS17 transformed with indicated plasmids expressing recombinant topoisomerase I (pLIC-ETOP and pLIC-MTOP), MazF4 toxin (pBAD-Rv1495) or control vectors (pLIC-HK and pBAD/Thio) were spotted on LB agar plates with antibiotics and 0.2% arabinose. The set of duplicated plates was incubated at 30°C (left) and 42°C (right).

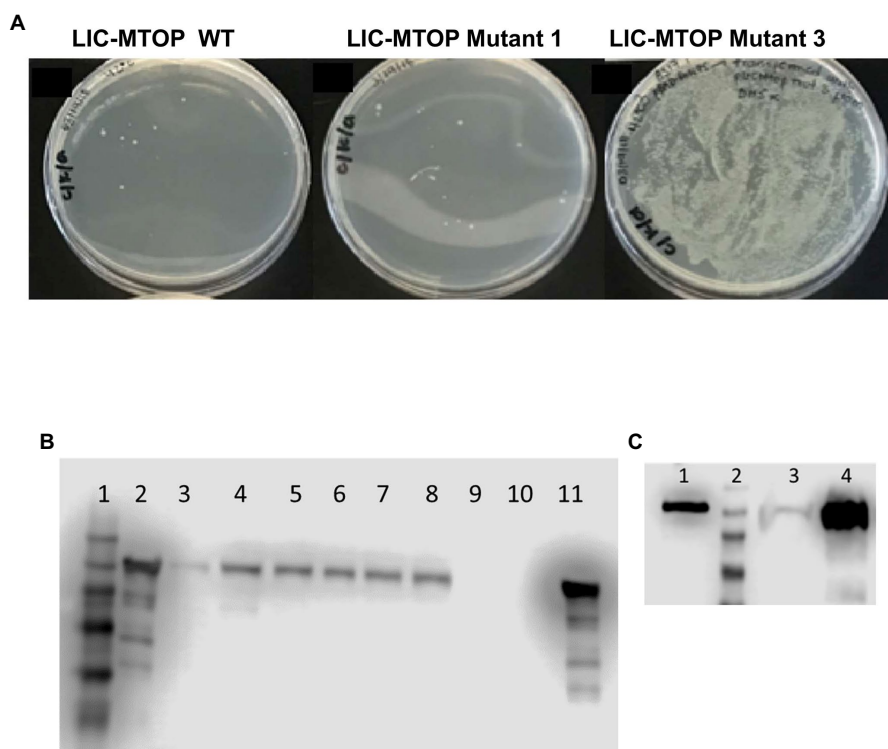


FIGURE 2

Mutant pLIC-MTOP Selected for Resistance to pBAD-Rv1495. (A) pLIC-MTOP plasmids were extracted from colonies of AS17/pBAD-Rv1495 transformed with mutated pLIC-MTOP library that grew at 42°C. Transformation of AS17/pBAD-Rv1495 was repeated with the extracted plasmids. Viable transformants at 42°C for the mutant-3 isolate of plasmid pLIC-MTOP showed an increase of  $\sim 10^4$  fold when compared to wild-type and mutant-1 isolates. (B,C) Western blot analysis of *Mtb* TopA expression with rabbit polyclonal antibodies raised against *Mtb* TopA. (B) Lane 1: prestained molecular weight standards; Lanes 2, 11: 100 ng of purified *Mtb* TopA protein; Lanes 3–8: whole cell lysate of AS17/pBAD-Rv1495 transformed with selected mutant-1 to mutant-6 isolates of pLIC-MTOP. Lanes 9, 10: whole cell lysate of AS17/pBAD-Rv1495 transformed with original pLIC-MTOP clone. (C) Lane 1: 1ng of purified *Mtb* TopA protein; Lane 2: prestained molecular weight standards; Lane 3: whole cell lysate of AS17/pBAD-Rv1495 transformed with original pLIC-MTOP clone; Lane 4: whole cell lysate of AS17/pBAD-Rv1495 transformed with selected mutant-3 isolate of pLIC-MTOP. According to OD<sub>600</sub> of cultures used, lane 3 corresponds to lysate from twice the number of cells compared to lysate in lane 4.



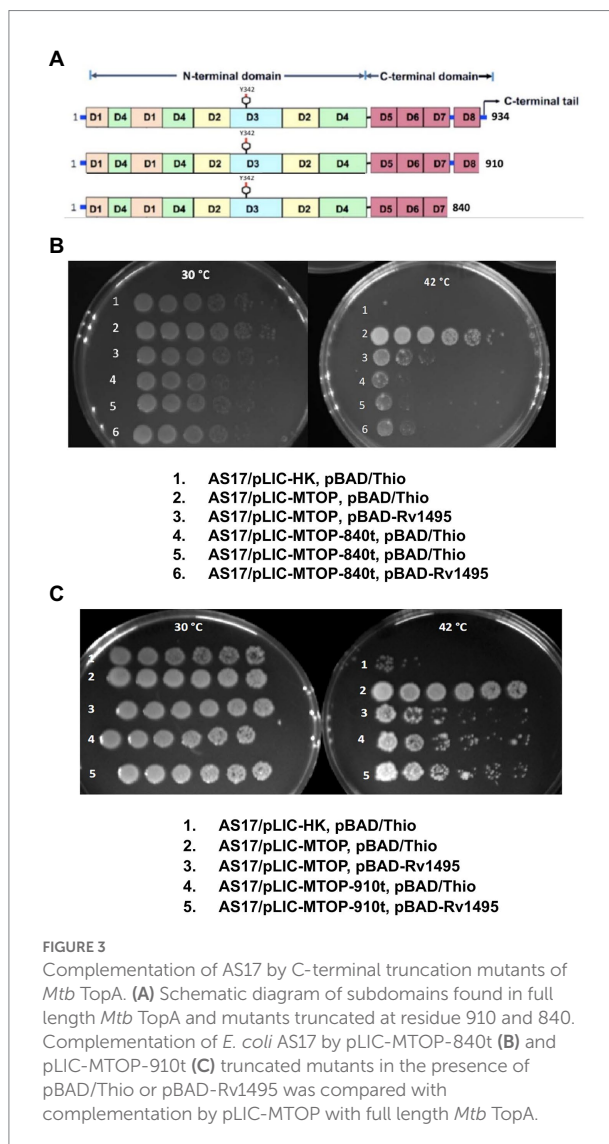
supported growth of AS17/pBAD-Rv1495 better than mutant-1. However, when the TopA coding region of six selected pLIC-MTOP plasmids including mutant-1 and mutant-3 were analyzed by Sanger sequencing, no mutation could be found in the TopA coding region. When protein expression of the AS17/pBAD-Rv1495 transformants of these selected pLIC-MTOP plasmids was examined, they were found to express greatly increased level of the recombinant *Mtb* TopA protein when compared to the original pLIC-MTOP clone (Figures 2B,C). This mechanism of resistance most likely confers *in vivo* resistance to MazF4 inhibition of *Mtb* TopA effectively without cost of decreasing the relaxation activity of the *Mtb* TopA protein. Although the result did not further inform on the region of TopA that may interact directly with MazF4, it is consistent with reduction of *Mtb* TopA catalytic activity as the basis for the AS17 growth complementation inhibition by pBAD-Rv1495.

## Effect of *Mtb* TopA C-terminal truncation

*Mtb* TopA (934 residues in length) is composed of two main domains: N-terminal domain and C-terminal domain (Figure 3A). The active site with Y342 for linking to DNA in the covalent intermediate is formed by the N-terminal subdomains D1–D4. The C-terminal subdomains D5–D8 consisting of repeated Topo\_C\_Rpt motifs (Tan et al., 2016) and the positively charged C-terminal tail that follow also interact with DNA during catalysis (Ahmed et al., 2013; Cao et al., 2020). The truncated TopA protein encoded by plasmid pLIC-MTOP-840t lacks subdomain D8 and the C-terminal tail. Only the C-terminal tail is missing in the truncated TopA encoded in plasmid pLIC-MTOP-910t. The growth complementation of *E. coli* AS17 at 42°C with pLIC-MTOP-840t was weak, consistent with the reduced *in vitro* catalytic activity reported previously for *Mtb* TopA truncated at residue 840 (Cao et al., 2020), but above the background of pLIC-HK (Figure 3B). The *in vitro* catalytic activity for the mutant *Mtb* TopA truncated at residue 910 (Cao et al., 2020) and partial growth complementation by pLIC-MTOP-910t was stronger in comparison to pLIC-MTOP-840t (Figure 3C). The degree of complementation by both pLIC-MTOP-840t and pLIC-MTOP-910t was not affected by the MazF4 toxin expressed by pBAD-Rv1495. These results indicate that the presence of the TopA C-terminal tail is required for the inhibition of *Mtb* TopA activity by MazF4.

## Effect of lysine to alanine substitutions in *Mtb* TopA C-terminal tail

Sequence alignment of TopA from mycobacterial species shows that they share a conserved sequence of PAKKA in the lysine-rich C-terminal tail (Figure 4A). We hypothesized that a negatively charged surface of the *Mtb* MazF4 toxin (Ahn et al., 2017) interacts with the positively charged C-terminal tail of



*Mtb* TopA. Site-directed alanine substitutions were made to test the potential role of lysine residues in the conserved sequence for interaction with MazF4. The plasmid pLIC-MTOP-KK encodes mutant TopA-KK with alanine substitutions at K928 and K929 within the conserved PAKKA sequence. Complementation assay (Figure 4B) showed that the TopA-KK mutant had slightly less activity as wild-type TopA for growth support of AS17 at the non-permissive temperature. The growth complementation by pLIC-MTOP-KK did not appear to be affected by the presence of pBAD-Rv1495. Measurement of CFU/ml numbers (Supplementary Table S2) at 30 and 42°C from three experimental replicates showed that there is no significant reduction in complementation (ratio of CFU at 42°C/CFU at 30°C) by pLIC-MTOP-KK when pBAD-Rv1495 is present versus pBAD/Thio (value of p from two-tailed test > 0.05). Therefore, loss of the side chains at these two lysine residues in the conserved PAKKA motif greatly reduced the inhibition of *Mtb* TopA activity by MazF4 expressed by pBAD-Rv1495.



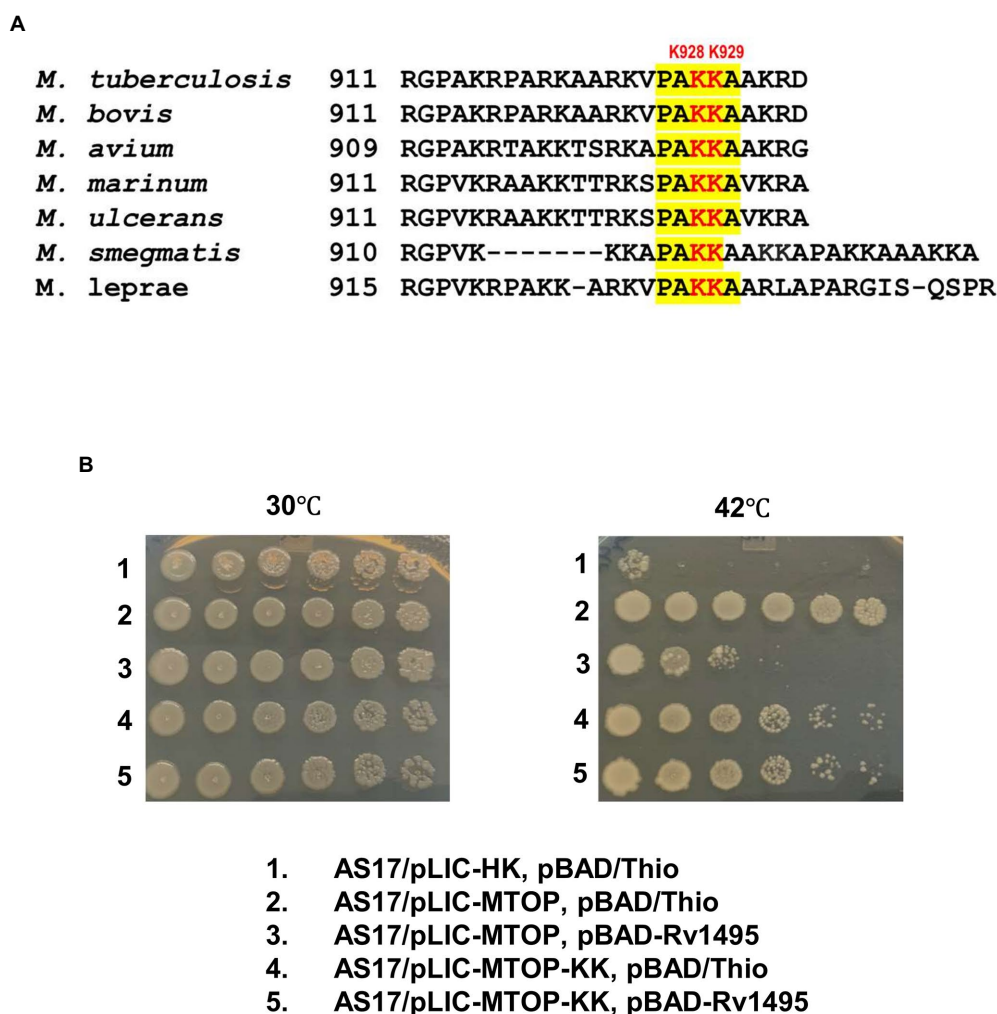


FIGURE 4

Alanine substitutions of lysines in conserved sequence present in C-terminal tail of *Mtb* TopA. (A) Alignment of mycobacterial TopA sequences. The conserved sequence PAKKA is highlighted. (B) Complementation assay of pLIC-MTOP-KK plasmid with K928A and K929A substitutions in *Mtb* TopA.

## Molecular simulations to predict and optimize the *Mtb* TopA-MazF4 complex

Figure 5A shows a *Mtb* TopA-MazF4 complex predicted from molecular docking as described in the Methods section. To reduce the computational cost, we selected structure as enclosed within dotted region, which includes amino acid 786-934 portion of the *Mtb* TopA C-terminal domain (CTD) and complete MazF4 dimer. We performed MD simulations to access stability and structural integrity of the docked complex (Tiwari et al., 2016). Figure 5B shows the complex structure after 150 ns MD simulation. The complex was stable for the 150 ns simulation time as revealed by the RMSD measurements shown in Figure 5C. As shown in Figures 5D,E, several amino acid residues were found to establish the *Mtb* TopA-MazF4 complex formation through hydrogen bond and electrostatic interactions. These include K928, K929 and other amino acids in the *Mtb* TopA C-terminal tail.

## Discussion

A previous report has indicated that the binding site for MazF4 toxin is located within the 312 residue long C-terminal region of *Mtb* TopA (Huang and He, 2010). The C-terminal region of mycobacterial topoisomerase I is required for the relaxation of supercoiled DNA (Ahmed et al., 2013; Cao et al., 2020) and protein-protein interactions (Banda et al., 2017). The goal of this study is to further characterize the binding site and elucidate the mechanism of inhibition of *Mtb* TopA by MazF4.

The specific inhibition of *Mtb* TopA but not *E. coli* TopA by *Mtb* MazF4 overexpressed in the *E. coli* AS17 *topA*<sup>ts</sup> complementation assay provides interesting insights on the action of MazF4. It can be inferred that any specific cleavage of single-stranded RNA in *E. coli* that might occur at previously characterized sequence of U<sup>~</sup>CGCU (~denoting the cleavage site) by *Mtb* MazF4 (Zhu et al., 2008) did not affect the growth of *E. coli* AS17 at either 30°C or

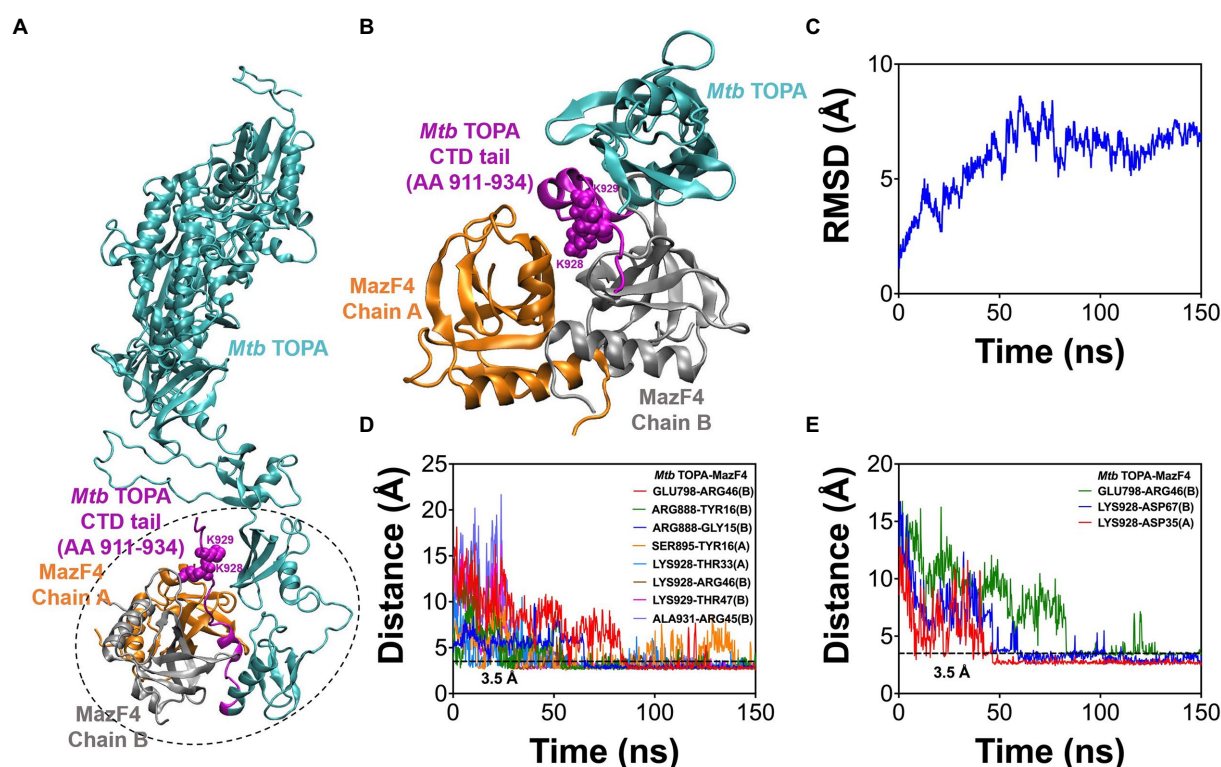


FIGURE 5

Molecular simulations of *Mtb* TopA-MazF4 complex. (A) Molecular docking result showing a complex formed between *Mtb* TopA and MazF4 dimer. Structures inside the dotted enclosed region were selected in molecular dynamics (MD) simulations of the complex. (B) *Mtb* TopA and MazF4 complex structure after 150ns MD simulations. (C) RMSD measurements. (D) Hydrogen bonding and (E) salt bridge distances calculated from the 150ns simulation trajectory.

42°C. The overexpressed *Mtb* MazF4 prevented growth of *E. coli* AS17 when the relaxation activity of recombinant *Mtb* TopA is required to complement the temperature sensitive *topA* mutation at the non-permissive temperature. This demonstrates that inhibition of the in-cell activity of *Mtb* TopA by MazF4 can result in loss of bacterial viability. The growth inhibition is specific for *Mtb* TopA as target as co-expressed *Mtb* MazF4 has no effect on the complementation by recombinant *E. coli* TopA. The finding we made that overexpression of recombinant *Mtb* TopA results in resistance to MazF4 growth inhibition provides further support that reduction of *Mtb* TopA relaxation activity is the molecular basis of the observed growth inhibition by MazF4. This mechanism of resistance to MazF4 inhibition does not affect the efficiency of complementation by *Mtb* TopA. The mutants from our attempt of targeted mutagenesis of the *Mtb* TopA coding region may not be able to produce resistant mutant *Mtb* TopA proteins without affecting the catalytic activity of *Mtb* TopA and efficiency of complementation.

Truncation of the *Mtb* TopA C-terminal domain in the LIC-MTOP-840t (missing subdomain D9 and C-terminal tail) and LIC-MTOP-910t (missing the 24 residue long C-terminal tail only) reduced the degree of growth complementation at 42°C because the C-terminal subdomains and C-terminal tail play important roles in interaction with DNA substrate during relaxation of supercoiled DNA (Ahmed et al., 2013; Cao et al., 2020). Nevertheless, the extent

of complementation by these truncated *Mtb* TopA was not affected by the presence of *Mtb* MazF4, suggesting that the C-terminal tail is required for inhibition of *Mtb* TopA relaxation activity by MazF4.

The crystal structure of the MazEF4 complex shows a negatively charged pocket displayed by the toxin MazF4 when it is binding to the antitoxin MazE4 (Ahn et al., 2017). We hypothesize that the highly basic *Mtb* TopA C-terminal tail can be bound by the negatively charged surface of MazF4 to inhibit the TopA C-terminal domain function during catalysis. Sequence alignment of mycobacterial TopA sequences shows the presence of conserved PAKKA motif in the C-terminal tail. Alanine substitutions of the two positively charged lysine residues in *Mtb* TopA encoded in pLIC-MTOP-KK plasmid resulted in loss of nearly all the MazF4 growth inhibition in the complementation assay. Molecular dynamics simulation of *Mtb* TopA-MazF4 binding provided further support of the conserved lysine residues K928, K929 in the *Mtb* TopA C-terminal tail being involved in interaction with MazF4. In addition to electrostatic interactions, molecular simulations showed that these conserved lysine residues can also form hydrogen bonds with residues in MazF4. Other amino acid residues in the *Mtb* TopA C-terminal domain contribute to the interactions with MazF4 as well.

The model of relaxation of supercoiled DNA by bacterial TopA is shown in Figure 6. Previous biochemical studies (Ahmed

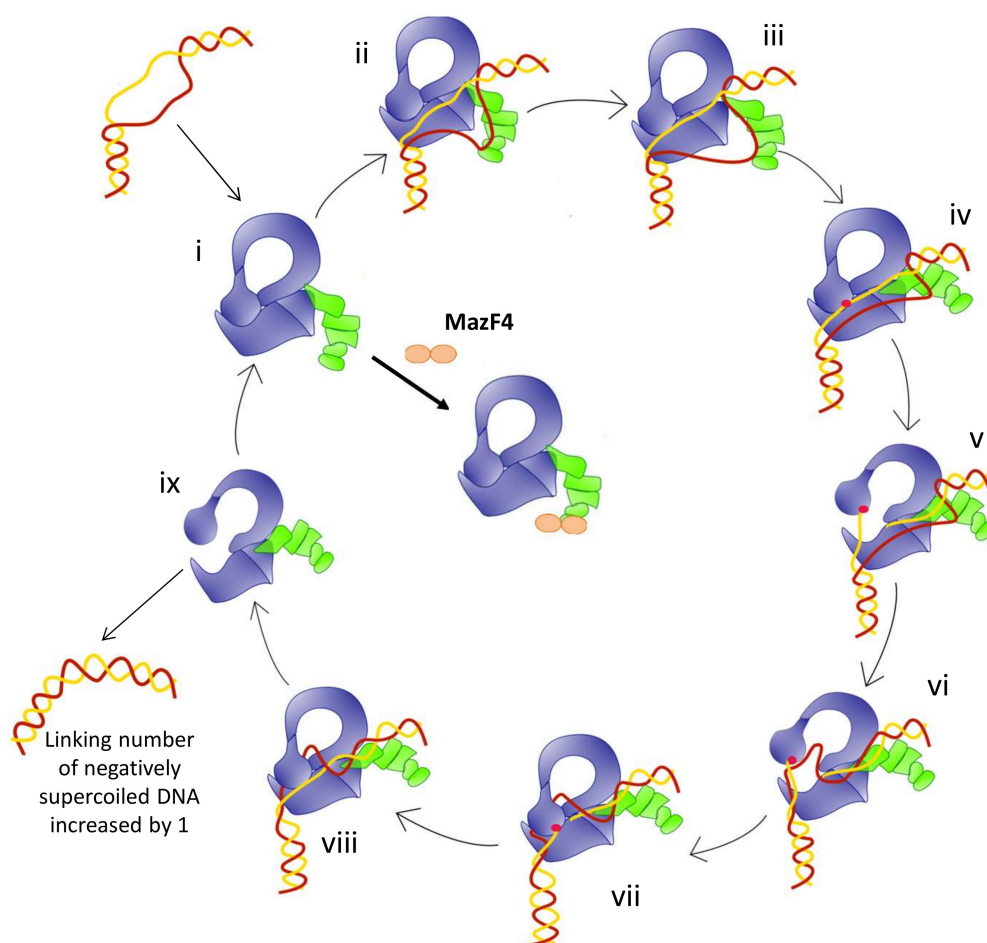


FIGURE 6

Model for the relaxation of supercoiled DNA by bacterial topoisomerase I based on the crystal structures of *E. coli* and Mycobacteria topoisomerase I. (i) Apo enzyme; (ii) C-terminal domains (green) recognizes ssDNA region in unwound DNA duplex as T-strand (red); (iii) ssDNA or G-strand (yellow) binds the N-terminal domains (blue); (iv) Active site tyrosine (red circle) becomes accessible; (v) Cleavage of the G-strand and gate opening with T-strand DNA approaching toroid hole; (vi) Passage of T-strand inside the toroid; (vii) Gate closing and trapping of T-strand; (viii) Religation of the G-strand; (ix) Gate opening and release of dsDNA. Figure reproduced from Dasgupta et al. (2020).

et al., 2013; Cao et al., 2020) on mycobacterial TopA showed that the high affinity of the C-terminal domains and basic tail for ssDNA is important for recognition of the unwound duplex DNA substrate and promote strand passage (step ii and step v, Figure 6). The formation of a stable *Mtb* TopA-MazF4 complex is likely to inhibit the catalytic function of the *Mtb* TopA.

*Mtb* TopA is a validated target for discovery of novel TB drug leads (Godbole et al., 2015; Ravishankar et al., 2015; Nagaraja et al., 2017). Inhibitors that act at the highly conserved N-terminal domains of bacterial TopA might be more likely to be broad-spectrum in antibacterial action and could have undesirable effect on the normal human microbiota. This study has led to further understanding of the action of *Mtb* toxin MazF4. It suggests that the *in vivo* relaxation activity of *Mtb* TopA can be inhibited through specific targeting of its C-terminal tail to achieve select growth inhibition of *Mtb* and possibly related mycobacteria. We have not included in this study characterization of the *in vitro* interaction between *Mtb* TopA and MazF4. Such characterization is needed in

future studies to compare the action of MazF4 with small molecule inhibitors that might also be binding at the *Mtb* TopA C-terminal tail.

## Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

## Author contributions

Y-CT-D and TA conceived the idea and designed the study. PG, RB, ED, and PT executed the experiments. TA, PG, and PBT performed the analyses. Y-CT-D wrote the draft of the manuscript. All authors contributed to the article and approved the submitted version.

## Funding

This research was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number R35GM139817, R01GM054226, and Supplement GM054226-S1 (to Y-CT-D), Colciencias (Convocatoria 586 de becas en el exterior to PG). PBT acknowledges the computational resources acquired with a research restart discretionary fund from the office of the Dean for Research at Georgetown University Medical Center.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

- Ahmed, W., Bhat, A. G., Leelaram, M. N., Menon, S., and Nagaraja, V. (2013). Carboxyl terminal domain basic amino acids of mycobacterial topoisomerase I bind Dna to promote strand passage. *Nucleic Acids Res.* 41, 7462–7471. doi: 10.1093/nar/gkt506
- Ahmed, W., Menon, S., Godbole, A. A., Karthik, P. V., and Nagaraja, V. (2014). Conditional silencing of topoisomerase I gene of mycobacterium tuberculosis validates its essentiality for cell survival. *FEMS Microbiol. Lett.* 353, 116–123. doi: 10.1111/1574-6968.12412
- Ahn, D. H., Lee, K. Y., Lee, S. J., Park, S. J., Yoon, H. J., Kim, S. J., et al. (2017). Structural analyses of the Mazef4 toxin-antitoxin pair in mycobacterium tuberculosis provide evidence for a unique extracellular death factor. *J. Biol. Chem.* 292, 18832–18847. doi: 10.1074/jbc.M117.807974
- Annamalai, T., Dani, N., Cheng, B., and Tse-Dinh, Y. C. (2009). Analysis of DNA relaxation and cleavage activities of recombinant mycobacterium tuberculosis DNA topoisomerase I from a new expression and purification protocol. *BMC Biochem.* 10:18. doi: 10.1186/1471-2091-10-18
- Baker, N. M., Rajan, R., and Mondragón, A. (2009). Structural studies of type I topoisomerases. *Nucleic Acids Res.* 37, 693–701. doi: 10.1093/nar/gkn1009
- Banda, S., Cao, N., and Tse-Dinh, Y. C. (2017). Distinct mechanism evolved for mycobacterial Rna polymerase and topoisomerase I protein-protein interaction. *J. Mol. Biol.* 429, 2931–2942. doi: 10.1016/j.jmb.2017.08.011
- Boucher, H. W. (2020). Bad bugs, no drugs 2002–2020: Progress, challenges, and call to action. *Trans. Am. Clin. Climatol. Assoc.* 131, 65–71.
- Brochu, J., Breton, E. V., and Drolet, M. (2020). Supercoiling, R-loops, replication and the functions of bacterial type 1A topoisomerases. *Genes (Basel)* 11, 249. doi: 10.3390/genes11030249
- Bush, N. G., Diez-Santos, I., Abbott, L. R., and Maxwell, A. (2020). Quinolones: mechanism, lethality and their contributions to antibiotic resistance. *Molecules* 25. doi: 10.3390/molecules25235662
- Bush, N. G., Evans-Roberts, K., and Maxwell, A. (2015). DNA Topoisomerases. *EcoSal Plus* 6. doi: 10.1128/ecosalplus.ESP-0010-2014
- Cao, N., Tan, K., Zuo, X., Annamalai, T., and Tse-Dinh, Y. C. (2020). Mechanistic insights from structure of mycobacterium smegmatis topoisomerase I with ssDNA bound to both N- and C-terminal domains. *Nucleic Acids Res.* 48, 4448–4462. doi: 10.1093/nar/gkaa201
- Cheng, B., Sorokin, E. P., and Tse-Dinh, Y. C. (2008). Mutation adjacent to the active site tyrosine can enhance Dna cleavage and cell killing by the Toprim Gly to Ser mutant of bacterial topoisomerase I. *Nucleic Acids Res.* 36, 1017–1025. doi: 10.1093/nar/gkm1126
- Corn, J. E., Pease, P. J., Hura, G. L., and Berger, J. M. (2005). Crosstalk between primase subunits can act to regulate primer synthesis in trans. *Mol. Cell* 20, 391–401. doi: 10.1016/j.molcel.2005.09.004
- Dasgupta, T., Ferdous, S., and Tse-Dinh, Y. C. (2020). Mechanism of type IA topoisomerases. *Molecules* 25, 4709. doi: 10.3390/molecules25204769
- Dean, A. S., Tosas Auguet, O., Glaziou, P., Zignol, M., Ismail, N., Kasaeva, T., et al. (2022). 25 years of surveillance of drug-resistant tuberculosis: achievements, challenges, and way forward. *Lancet Infect. Dis.* 22, e191–e196. doi: 10.1016/S1473-3099(21)00808-2
- Diaz, B., Mederos, C., Tan, K., and Tse-Dinh, Y. C. (2022). Microbial type IA topoisomerase C-terminal domain sequence motifs, distribution and combination. *Int. J. Mol. Sci.* 23, 8709. doi: 10.3390/ijms23158709
- eClinicalMedicine (2021). Antimicrobial resistance: a top ten global public health threat. *eClinicalMedicine* 41:101221. doi: 10.1016/j.eclinm.2021.101221
- Garnier, F., Couturier, M., Débat, H., and Nadal, M. (2021). Archaea: a gold mine for topoisomerase diversity. *Front. Microbiol.* 12:661411. doi: 10.3389/fmicb.2021.661411
- Garnier, F., Debat, H., and Nadal, M. (2018). Type Ia Dna topoisomerases: a universal Core and multiple activities. *Methods Mol. Biol.* 1703, 1–20. doi: 10.1007/978-1-4939-7459-7\_1
- Gibson, D. G. (2011). Enzymatic assembly of overlapping DNA fragments. *Methods Enzymol.* 498, 349–361. doi: 10.1016/B978-0-12-385120-8.00015-2
- Godbole, A. A., Ahmed, W., Bhat, R. S., Bradley, E. K., Ekins, S., and Nagaraja, V. (2015). Targeting mycobacterium tuberculosis topoisomerase I by small-molecule inhibitors. *Antimicrob. Agents Chemother.* 59, 1549–1557. doi: 10.1128/AAC.04516-14
- Huang, F., and He, Z. G. (2010). Characterization of an interplay between a mycobacterium tuberculosis MazF homolog, Rv1495 and its sole Dna topoisomerase I. *Nucleic Acids Res.* 38, 8219–8230. doi: 10.1093/nar/gkq737
- Humphrey, W., Dalke, A., and Schulten, K. (1996). Vmd: visual molecular dynamics. *J. Mol. Graph.* 14, 33–38. doi: 10.1016/0263-7855(96)00018-5
- Källberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H., et al. (2012). Template-based protein structure modeling using the RaptorX web server. *Nat. Protoc.* 7, 1511–1522. doi: 10.1038/nprot.2012.085
- Kempker, R. R., Kipiani, M., Mirtskhulava, V., Tukvadze, N., Magee, M. J., and Blumberg, H. M. (2015). Acquired drug resistance in mycobacterium tuberculosis and poor outcomes among patients with multidrug-resistant tuberculosis. *Emerg. Infect. Dis.* 21, 992–1001. doi: 10.3201/eid2106.141873
- Kwan, C. K., and Ernst, J. D. (2011). HIV and tuberculosis: a deadly human syndemic. *Clin. Microbiol. Rev.* 24, 351–376. doi: 10.1128/CMR.00042-10
- Lee, J., Cheng, X., Swails, J. M., Yeom, M. S., Eastman, P. K., Lemkul, J. A., et al. (2016). Charmm-Gui input generator for Namd, Gromacs, Amber, Openmm, and Charmm/Openmm simulations using the Charmm36 additive force field. *J. Chem. Theory Comput.* 12, 405–413. doi: 10.1021/acs.jctc.5b00935
- Mckie, S. J., Neuman, K. C., and Maxwell, A. (2021). DNA topoisomerases: advances in understanding of cellular roles and multi-protein complexes via structure-function analysis. *BioEssays* 43:e2000286. doi: 10.1002/bies.202000286
- Muteeb, G., and Sen, R. (2010). Random mutagenesis using a mutator strain. *Methods Mol. Biol.* 634, 411–419. doi: 10.1007/978-1-60761-652-8\_29
- Nagaraja, V., Godbole, A. A., Henderson, S. R., and Maxwell, A. (2017). DNA topoisomerase I and DNA gyrase as targets for Tb therapy. *Drug Discov. Today* 22, 510–518. doi: 10.1016/j.drudis.2016.11.006

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.1032320/full#supplementary-material>



- Narula, G., Becker, J., Cheng, B., Dani, N., Abrenica, M. V., and Tse-Dinh, Y. C. (2010). The DNA relaxation activity and covalent complex accumulation of *Mycobacterium tuberculosis* topoisomerase I can be assayed in *Escherichia coli*: application for identification of potential fret-dye labeling sites. *BMC Biochem.* 11:41. doi: 10.1186/1471-2091-11-41
- Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., et al. (2005). Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26, 1781–1802. doi: 10.1002/jcc.20289
- Pierce, B. G., Hourai, Y., and Weng, Z. (2011). Accelerating protein docking in Zdock using an advanced 3D convolution library. *PLoS One* 6:e24657. doi: 10.1371/journal.pone.0024657
- Pommier, Y., Sun, Y., Huang, S. N., and Nitiss, J. L. (2016). Roles of eukaryotic topoisomerases in transcription, replication and genomic stability. *Nat. Rev. Mol. Cell Biol.* 17, 703–721. doi: 10.1038/nrm.2016.111
- Ramage, H. R., Connolly, L. E., and Cox, J. S. (2009). Comprehensive functional analysis of mycobacterium tuberculosis toxin-antitoxin systems: implications for pathogenesis, stress responses, and evolution. *PLoS Genet.* 5:e1000767. doi: 10.1371/journal.pgen.1000767
- Ravishankar, S., Ambady, A., Awasthy, D., Mudugal, N. V., Menasinakai, S., Jatheendranath, S., et al. (2015). Genetic and chemical validation identifies *Mycobacterium tuberculosis* topoisomerase I as an attractive anti-tubercular target. *Tuberculosis (Edinb.)* 95, 589–598. doi: 10.1016/j.tube.2015.05.004
- Sala, A., Bordes, P., and Genevaux, P. (2014). Multiple toxin-antitoxin systems in *Mycobacterium tuberculosis*. *Toxins (Basel)* 6, 1002–1020. doi: 10.3390/toxins6031002
- Šali, A., and Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234, 779–815. doi: 10.1006/jmbi.1993.1626
- Sandhaus, S., Annamalai, T., Welmaker, G., Houghten, R. A., Paz, C., Garcia, P. K., et al. (2016). Small-molecule inhibitors targeting topoisomerase I as novel Antituberculosis agents. *Antimicrob. Agents Chemother.* 60, 4028–4036. doi: 10.1128/AAC.00288-16
- Sassetti, C. M., Boyd, D. H., and Rubin, E. J. (2003). Genes required for mycobacterial growth defined by high density mutagenesis. *Mol. Microbiol.* 48, 77–84. doi: 10.1046/j.1365-2958.2003.03425.x
- Seki, M., Choi, H., Kim, K., Whang, J., Sung, J., and Mitarai, S. (2021). Tuberculosis: a persistent unpleasant neighbour of humans. *J. Infect. Public Health* 14, 508–513. doi: 10.1016/j.jiph.2021.01.005
- Tan, K., Cao, N., Cheng, B., Joachimiak, A., and Tse-Dinh, Y. C. (2016). Insights from the structure of *Mycobacterium tuberculosis* topoisomerase I with a novel protein fold. *J. Mol. Biol.* 428, 182–193. doi: 10.1016/j.jmb.2015.11.024
- Tan, K., Zhou, Q., Cheng, B., Zhang, Z., Joachimiak, A., and Tse-Dinh, Y. C. (2015). Structural basis for suppression of hypernegative DNA supercoiling by *E. coli* topoisomerase I. *Nucleic Acids Res.* 43, 11031–11046. doi: 10.1093/nar/gkv1073
- Tiwari, P. B., Chapagain, P. P., Banda, S., Darici, Y., Uren, A., and Tse-Dinh, Y. C. (2016). Characterization of molecular interactions between *Escherichia coli* RNA polymerase and topoisomerase I by molecular simulations. *FEBS Lett.* 590, 2844–2851. doi: 10.1002/1873-3468.12321
- Vos, S. M., Tretter, E. M., Schmidt, B. H., and Berger, J. M. (2011). All tangled up: how cells direct, manage and exploit topoisomerase function. *Nat. Rev. Mol. Cell Biol.* 12, 827–841. doi: 10.1038/nrm3228
- Wang, Y., Lynch, A. S., Chen, S. J., and Wang, J. C. (2002). On the molecular basis of the thermal sensitivity of an *Escherichia coli* topA mutant. *J. Biol. Chem.* 277, 1203–1209. doi: 10.1074/jbc.M109436200
- World Health Organization. (2020). *The who global task force on Tb impact measurement, march 2020*. Geneva: World Health Organization.
- World Health Organization. (2022). *TRapid communication: key changes to the treatment of drug-resistant tuberculosis*. Geneva: World Health Organization.
- Zhu, L., Phadtare, S., Nariya, H., Ouyang, M., Husson, R. N., and Inouye, M. (2008). The mRNA interferases, MazF-mt3 and MazF-mt7 from *Mycobacterium tuberculosis* target unique pentad sequences in single-stranded RNA. *Mol. Microbiol.* 69, 559–569. doi: 10.1111/j.1365-2958.2008.06284.x
- Zhu, L., Zhang, Y., Teh, J. S., Zhang, J., Connell, N., Rubin, H., et al. (2006). Characterization of mrna interferases from *Mycobacterium tuberculosis*. *J. Biol. Chem.* 281, 18638–18643. doi: 10.1074/jbc.M512693200



## OPEN ACCESS

## EDITED BY

Graciela L. Lorca,  
University of Florida,  
United States

## REVIEWED BY

Artur J. Sabat,  
University Medical Center Groningen,  
Netherlands  
Kaylie Allyson Padgett-Pagliai,  
University of Florida,  
United States

## \*CORRESPONDENCE

Catherine A. Wakeman  
catherine.wakeman@ttu.edu

## SPECIALTY SECTION

This article was submitted to  
Microbial Physiology and Metabolism,  
a section of the journal  
Frontiers in Microbiology

RECEIVED 26 June 2022

ACCEPTED 21 November 2022

PUBLISHED 08 December 2022

## CITATION

Islam S, Callender AC, Ho QN and  
Wakeman CA (2022) Iron restriction  
induces the small-colony variant  
phenotype in *Staphylococcus aureus*.  
*Front. Microbiol.* 13:978859.  
doi: 10.3389/fmicb.2022.978859

## COPYRIGHT

© 2022 Islam, Callender, Ho and Wakeman.  
This is an open-access article distributed  
under the terms of the [Creative Commons  
Attribution License \(CC BY\)](#). The use,  
distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in  
this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Iron restriction induces the small-colony variant phenotype in *Staphylococcus aureus*

Shariful Islam, Anna C. Callender, Quynh N. Ho and Catherine A. Wakeman\*

Department of Biological Sciences, Texas Tech University, Lubbock, TX, United States

Pathogens such as *Staphylococcus aureus* must overcome host-induced selective pressures, including limited iron availability. To cope with the harsh conditions of the host environment, *S. aureus* can adapt its physiology in multiple ways. One of these adaptations is the fermenting small-colony variant (SCV) phenotype, which is known to be inherently tolerant to certain classes of antibiotics and heme toxicity. We hypothesized that SCVs might also behave uniquely in response to iron starvation since one of the major cellular uses of iron is the respiration machinery. In this study, a respiring strain of *S. aureus* and fermenting SCV strains were treated with different concentrations of the iron chelator, 2,2'-dipyridyl (DIP). Our data demonstrate that a major impact of iron starvation in *S. aureus* is the repression of respiration and the induction of the SCV phenotype. We demonstrate that the SCV phenotype transiently induced by iron starvation mimics the aminoglycoside recalcitrance exhibited by genetic SCVs. Furthermore, prolonged growth in iron starvation promotes increased emergence of stable aminoglycoside-resistant SCVs relative to the naturally occurring subpopulation of SCVs within an *S. aureus* community. These findings may have relevance to physiological and evolutionary processes occurring within bacterial populations infecting iron-limited host environments.

## KEYWORDS

*Staphylococcus aureus*, small-colony variant, host-pathogen interface, nutritional immunity, aerobic respiration

## Introduction

*Staphylococcus aureus* is a non-motile, Gram-positive, and facultative anaerobe often observed as microflora on vertebrate skin and residing in the nostrils of almost one-third of the world's population (Noble et al., 1967; Kuehnert et al., 2006). This pathogen can cause many severe diseases by breaching the hosts' epithelia, including osteomyelitis, bacteremia, endocarditis, pneumonia, and septic shock (Klebens et al., 2007). Studies have emphasized the pathogen as the second most common agent that causes bloodstream infection and is the most significant cause of bloodstream infection-associated fatalities (Laupland et al., 2013). Invading pathogens, including *S. aureus*, encounter various selective pressures such

as the limitation of metal nutrients (e.g., iron, zinc, and manganese) while surviving in host environments (Palmer and Skaar, 2016; Baishya and Wakeman, 2019). Therefore, bacteria adopt several mechanisms to withstand these diverse host environments in response to host-induced selective pressures. The successful outcome of *S. aureus* as a chronic pathogen is contingent on its capability to sense these diverse environments, challenge host immunity, alter its metabolic reactions properly, and acquire or secrete essential nutrients (Krishna and Miller, 2012; Dastgheyb and Otto, 2015; Thammavongsa et al., 2015; Richardson, 2019; Lung et al., 2020). One of the most significant adaptations of *S. aureus* observed in clinical isolates is the development of small-colony variants (SCVs; Proctor et al., 2006, 2014; Bui and Kidd, 2015). Compared to wild-type counterparts, the SCVs of *S. aureus* show some phenotypic differences, such as slower growth, reduced hemolysis, and the lack of carotenoid pigmentation (Proctor et al., 1994). The formation of SCV phenotype provides the *S. aureus* increased tolerance against aminoglycoside antibiotics and heme toxicity and may enable increased intracellular persistence (Wakeman et al., 2012; Tuchscher et al., 2016; Rollin et al., 2017; Herrin Brittan et al., 2021).

Previous studies have assessed how chemicals present in human hosts can influence the SCV phenotype. For example, one study demonstrated that certain *S. aureus* strains transition to biofilm and SCV-associated lifestyles to survive under the influence of host generated reactive aldehydes (Bui et al., 2015). Another study revealed that various staphylococcal species, such as *S. aureus*, *S. lugdunensis*, and *S. epidermidis*, may produce SCV phenotypes in response to exposures to environmental challenges such as pH 5, 10% NaCl, 4°C, and the penicillin G, and vancomycin (Onyango et al., 2013). The staphylococcal response to reactive oxygen species (ROS) is also frequently studied as these molecules are associated with the host innate immune response. One such study indicated that when *S. aureus* is exposed to sublethal levels of the ROS hydrogen peroxide ( $H_2O_2$ ), there is a distinct, dose-dependent rise in the population of SCVs that are resistant to gentamicin (Painter et al., 2015). Due to its encouragement of adaptive and defensive responses that change cell physiology in ways that also increase antibiotic resistance, stress is a specific environmental factor affecting bacteria's susceptibility to antibiotics (Poole, 2012). When *P. aeruginosa* biofilms are treated with stressors, including metal ions, antibiotics, and oxidative agents in laboratories, various kinds of colony morphological variants, such as antibiotic-resistant small-colony variants (SCVs) and rugose small-colony variants (RSCVs), are recovered with increasing frequency. This shows that colony morphological variations may be crucial to a biofilm population's ability to survive when it is subjected to environmental challenges (Boles et al., 2004; Davies et al., 2007; Harrison et al., 2007; Workentine et al., 2010). Metal nutrients, such as Cu and Zn, have a long history of being associated with the emergence of antibiotic resistance in environmental bacteria (Calomiris et al., 1984; Wales and Davies, 2015).

Metal nutrients, such as iron, are essential for most living organisms, including *S. aureus* (Bullen et al., 1999). The pathogenicity of *S. aureus* is significantly reliant upon the iron status of the infected hosts. While defending against pathogens (e.g., *S. aureus*), host immune proteins acquire transition metals such as iron, manganese, and zinc. Therefore, bacteria evolve their strategies to access those metal nutrients to counteract the host-induced defense mechanisms (Lopez and Skaar, 2018). The original hypothesis of this study was that SCVs might also be more resistant to iron restriction since one of the major cellular uses of iron is the respiration machinery that SCVs lack. Therefore, the effect of iron limitation on the respiring wild type and the fermentative SCV *S. aureus* is anticipated to be different in the aerobic condition than the effect in the anaerobic condition. The data herein indicate that one of the major impacts of iron restriction in *S. aureus* is inhibition of respiration and the transient adoption of the SCV phenotype, including the antibiotic recalcitrance characteristic of SCVs. Additionally, prolonged exposure to an iron-starved environment enabled increased emergence of antibiotic-resistant stable SCV subpopulations within an *S. aureus* community.

## Materials and methods

### Chemicals and reagents

2,2' dipyridyl (DIP) was purchased from Sigma Aldrich, United States. All other additional reagents and media were purchased from VWR, USA, and Fisher Scientific, United States.

### Bacterial strains

Newman (Duthie and Lorenz, 1952), a clinical isolate of *S. aureus*, was used for all the experiments and will be referred to as wild type (WT) herein. Additionally, SCV mutant strains were derived from the parental background of the WT Newman isolate. These include the  $\Delta menB$  SCV strain lacking menaquinone production (Wakeman et al., 2012), the  $\Delta hemB$  SCV lacking endogenous heme production (Hammer et al., 2013), and the  $\Delta cydBA\Delta qoxB$  (*cyto*<sup>-</sup>) mutant lacking cytochrome oxidases (Hammer et al., 2013).

### 2,2' dipyridyl treatment in liquid medium

WT and SCV strains were streaked on tryptic soy agar (TSA) plates and incubated at 37°C overnight and for more than 24 h, respectively. Individual colonies from each of the strains were picked in triplicates, then they were grown in tryptic soy broth (TSB) medium in a round bottom 96-well plate and then incubated overnight at 37°C with shaking condition at 150 rpm. Then, the cells were normalized to identical OD<sub>600</sub> (~0.145). DIP

solutions were freshly dissolved in absolute ethanol prior to use for each treatment. Firstly, 1 M DIP stock solution was prepared in absolute ethanol and then diluted to 10 mM DIP by adding TSB. All the appropriate concentrations were prepared in TSB medium in 96-well round bottom plates. Control wells were treated with the exact amounts of ethanol as DIP-treated wells.

Then, 2  $\mu$ l of normalized cells were added to each well. The ethanol control and the DIP-treated cells were incubated at 37°C on a shaker for 24 h. The 24 h treated samples were homogenized by pipetting up and down before taking the OD<sub>600</sub> reading on the BioTek plate reader. The data was then graphed and analyzed using one-way ANOVA in GraphPad Prism.

## DIP treatment in solid medium

WT and SCV mutants were streaked on TSA plates and incubated at 37°C overnight and for more than 24 h, respectively. Individual colonies from each of the strains were picked in triplicates, and then they were incubated overnight in a round bottom 96-well plate containing TSB at 37°C in shaking conditions at 150 rpm. Then, the cells were normalized to identical OD<sub>600</sub> (~0.145). DIP solutions were freshly dissolved in absolute ethanol prior to use for the preparation of all the TSA plates with or without different concentrations of DIP or the exact amounts of ethanol (ethanol control). Firstly, 1 M DIP stock solution was prepared in absolute ethanol and then diluted to 10 mM DIP by adding TSB. Next, warm TSA was poured into 50 ml falcon tubes. The 10 mM DIP and ethanol (with added TSB) were added to the falcon tubes accordingly to make different concentrations of DIP and ethanol control. Then, the medium with different concentrations of DIP and ethanol was poured on the Petri dishes labeled accordingly. The plates were kept at room temperature for one to 2 days to dry, and then each freshly made plate was partitioned (labeled) for all four strains. Cells (2  $\mu$ l) from the normalized overnight cells were poured and streaked out on the fresh plates (1–2 days old at room temperature). Additionally, 10  $\mu$ l of normalized cells were added to 90  $\mu$ l of 1X phosphate buffered saline (1X PBS) and then serially diluted in the 1X PBS. After streaking and pouring of serially diluted cells on DIP-treated and ethanol (no DIP)-treated plates, the plates were incubated at 37°C and observed for the following 3 days. The DIP-treated cells and SCVs could not clearly be observed on 1-day or 2-day post incubation because of their slow growth. However, on the third day, the colonies were clearly visible. Therefore, 3-day post incubation images were chosen for the manuscript.

## Reversion experiment on solid medium

Three-day-old colonies from iron-starved solid medium were streaked on the iron-sufficient TSA plates. The morphology of colonies on iron-sufficient media was observed after 3 days of incubation at 37°C.

## Anaerobic growth condition

The BD GasPak EZ anaerobic system (BD Biosciences, United States) was used to create an anaerobic condition. All the Plates containing the treated and untreated cells were set inside the anaerobic jar containing CO<sub>2</sub> gas packs and an indicator to observe the proper maintenance of the anaerobic condition. The anaerobic system containing the plates was incubated inside a 37°C incubator, and the data was recorded after the same time periods as the aerobic conditions.

## Gentamicin treatment

22-h DIP-treated and untreated cells of WT, and  $\Delta menB$  strains (a representative of SCVs) were normalized to 0.1 OD<sub>600</sub>. The untreated cells were normalized to iron sufficient medium, whereas the DIP-treated cells were normalized to the same DIP-treated TSB medium. After 4 h of gentamicin (5  $\mu$ g/ml) treatment in shaking aerobic condition, the cells were diluted and poured on plain TSA plates and incubated at 37°C for ~40 h to observe the CFUs.

## In vitro experimental evolution

WT bacteria were streaked on TSA plates and incubated at 37°C overnight. Individual colonies from each of the strains were picked in six replicates, and then they were incubated overnight in 15 ml culture tubes containing 5 ml TSB at 37°C in shaking conditions at 150 rpm. Then, the cells were normalized to identical OD<sub>600</sub> (~0.2). DIP (1 mM) was prepared in 5 ml TSB medium in 15 ml culture tubes containing 5 ml TSB. Control tubes were treated with the exact amounts of ethanol as DIP-treated wells.

Then, 5  $\mu$ l of normalized cells were added to the culture tubes containing 1 mM DIP in 5 ml TSB. The ethanol control and the DIP-treated cells were incubated at 37°C on a shaker for 24 h. On every day (until 9 days), 5  $\mu$ l of cells from DIP treated and control was transferred to fresh medium containing the 1 mM DIP or the same amount of ethanol containing TSB medium in new tubes. Every time, DIP-treated cells were transferred to DIP-Treated fresh medium, and similarly, control cells into the fresh medium with controlled environments. The experiment continued for 9 days. On the 3rd day, the cells were normalized to 0.2 OD<sub>600</sub>, and 10  $\mu$ l of normalized cells were serially diluted to observe the number of CFUs/ mL on plain TSA plates for each replicate. Additionally, the normalized (0.2 OD<sub>600</sub>) 1 ml cells were centrifuged and resuspended in 100  $\mu$ l (900  $\mu$ l supernatant discarded). The resuspended 100  $\mu$ l of cells from all the replicated tubes were poured on the 30  $\mu$ g/ml of gentamicin-containing TSA plates. Sterile glass beads were used to spread the cells on the entire plate. The number of gentamicin-resistant SCV colonies for each replicate was observed after 2 days at a 37°C incubator. Then, the



gentamicin-resistant SCV colonies/mL were calculated compared to their corresponding Typical CFUs/mL on plain TSA plates. A similar procedure was followed to measure the gentamicin-resistant SCVs per mL after 9 days. The data was then graphed and analyzed using a one-way ANOVA (Tukey's multiple comparisons) in GraphPad Prism.

## Cross-streaking for growth rescue of SCVs

Gentamicin-resistant individual isolates that emerged from experimental evolution were cocultured (cross-streaked) with either laboratory-derived  $\Delta menB$  or  $\Delta hemB$  on solid TSA media to test for increased growth indicative of heme or menaquinone exchange (Hammer et al., 2014). Twenty-four isolates were individually cocultured (cross-streaked) with the  $\Delta menB$  or  $\Delta hemB$  SCV strains. Growth rescue of the  $\Delta menB$  SCV but not of the  $\Delta hemB$  SCV by cross-streaking with a mutant indicated that the mutant emerging from was capable of menaquinone production but lacked heme production. Growth rescue of the  $\Delta hemB$  SCV but not the  $\Delta menB$  SCV by cross-streaking with a mutant indicated the mutant produced heme but not menaquinone.

## Results and discussion

### Iron starvation represses the growth of aerobically respiring *Staphylococcus aureus* to the level of fermenting *Staphylococcus aureus*

Oxygen and iron are important molecules influencing the virulence factors in pathogenic bacteria like *S. aureus* that serve important enzymatic functions and contribute to electron transport system-related reactions. Therefore, an alteration in oxygen and/or iron availability may change the bacterial metabolome (Ledala et al., 2014). Additionally, many studies indicated that the alteration in the availability of oxygen and iron could affect the growth and synthesis of numerous virulent molecules, as well as the antibiotic tolerance of bacteria (Cohen et al., 1967; Dassy and Fournier, 1996; Trivier and Courcol, 1996; Johnson et al., 2008). To assess the impact of iron starvation on WT versus SCV *S. aureus*, we used 2,2' dipyridyl (DIP) treatment which is a well-characterized iron chelator. It has previously been demonstrated that *S. aureus* SCVs commonly exhibit auxotrophy for hemin or menadione and have faulty respiration (Schaaff et al., 2003; Proctor et al., 2006; Sifri et al., 2006; Dean et al., 2014; Masoud-Landgraf et al., 2016) which results in the *hemB* and *menB* SCV mutants chosen for this work. Additionally, a cytochrome deficient mutant  $\Delta cydBA\Delta qoxB$  (Hammer et al., 2013) was selected as a final control for specifically monitoring the role of aerobic respiration

since hemin and menadione have other cellular roles including anaerobic respiration (Wissenbach et al., 1990; Edwards et al., 2020).

Previous studies have often used DIP treatment in the low millimolar range to induce iron starvation in multiple types of microorganisms, including *S. aureus* (Poole et al., 1993; Friedman et al., 2006; Minandri et al., 2016). Therefore, our initial studies used a similar range (0.25–4 mM). Adding DIP caused a significant reduction in the growth rate of respiring WT *S. aureus* but yielded no effect on any of the fermenting SCV strains even at 4 mM DIP as observed by optical density (Figure 1A). The growth suppression of WT was reduced to exactly match the level of SCVs and no further. This finding suggested that a major impact of DIP treatment was to inhibit aerobic respiration. As a further confirmation of this hypothesis, we grew cells under anaerobic conditions. In anaerobic growth, all strains relied on fermentative metabolism and DIP treatment displayed no impact on *S. aureus* growth (Figure 1B).

### Iron starvation causes transient emergence of an SCV phenotype in WT *Staphylococcus aureus*

Because the major effect of iron starvation on *S. aureus* using standard DIP concentrations appeared to be suppression of aerobic respiration, we hypothesized that WT *S. aureus* would exhibit SCV colony morphology when grown on an iron-limited agar plate. Indeed, we observed that genetically WT but iron-starved *S. aureus* adopted a traditional SCV phenotype characterized by a reduction in both colony size and yellow pigmentation (Figure 2A). SCV conversion can be a phenotype induced by the acquisition of a genetic mutant induced by exposure to a stressor (Wilson and Sanders, 1976; Bjorkman et al., 2000; McNamara and Proctor, 2000; Nagaev et al., 2001). Alternatively, SCV conversion could simply be a transient physiological response that will immediately revert when iron levels are restored. To differentiate between these possibilities, we isolated the colonies from an iron-starved plated and restreaked onto iron-sufficient media. Complete and immediate conversion to the normal colony morphology was observed in WT *S. aureus* (Figure 2B).

As none of the SCVs regardless of the genetic origin of the mutation appeared to be differently impacted by DIP treatment, the  $\Delta menB$  SCV will serve as the SCV control for the remainder of the experiments outlined in this manuscript since menadione auxotrophs are one of the most frequently observed SCV types in the clinic (Proctor et al., 2006). To determine if there were any additional toxicities associated with DIP treatment besides the suppression of respiration, we plated a serial dilution of cells onto DIP-containing plates. At 1.5 mM a complete transition to SCV phenotype can be observed with no reduction in viable cell counts. However, at higher DIP concentrations, other toxicities associated with DIP exposure/iron starvation may be emerging as overall reduction in viable cell counts occur in addition to the SCV phenotypic conversion at 2 mM DIP treatment (Figure 2C).

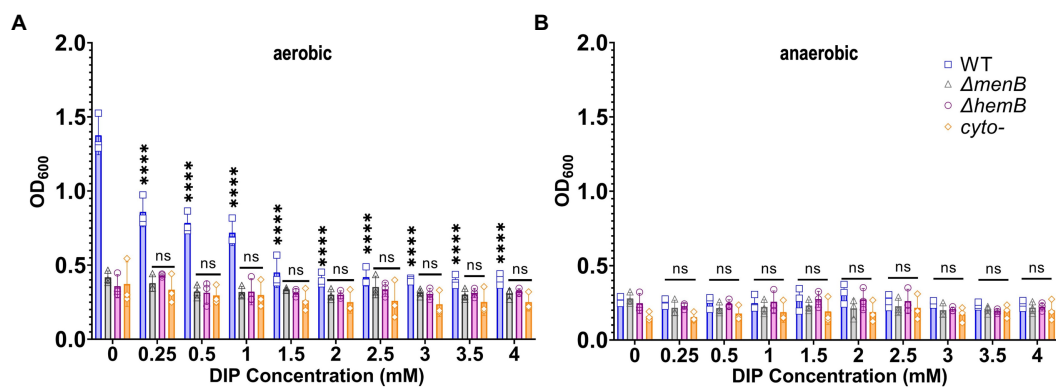


FIGURE 1

Growth of WT *Staphylococcus aureus* was diminished to the level of SCV growth under iron or oxygen limitation. (A) WT and SCV strains were treated with a range of DIP concentrations and cultured aerobically and growth was detected via OD<sub>600</sub> readings. DIP-induced growth suppression is only observed for WT *S. aureus* and only decreases to the level of baseline SCV growth. (B) WT and SCV strains were treated with a range of DIP concentrations and cultured anaerobically and growth was detected via OD<sub>600</sub> readings. No DIP-induced growth suppression could be observed for any strain tested. Each of the data points on the graph represents the average of triplicates on each day. Error bars represent standard deviation. A one-way ANOVA (Tukey's multiple comparisons) was applied to understand the statistical significance of the differences between the control (untreated, 0mM DIP) and the treatments ('\*\*\*\*' is the indication for  $p < 0.0001$ ; 'NS' is for non-significant data).

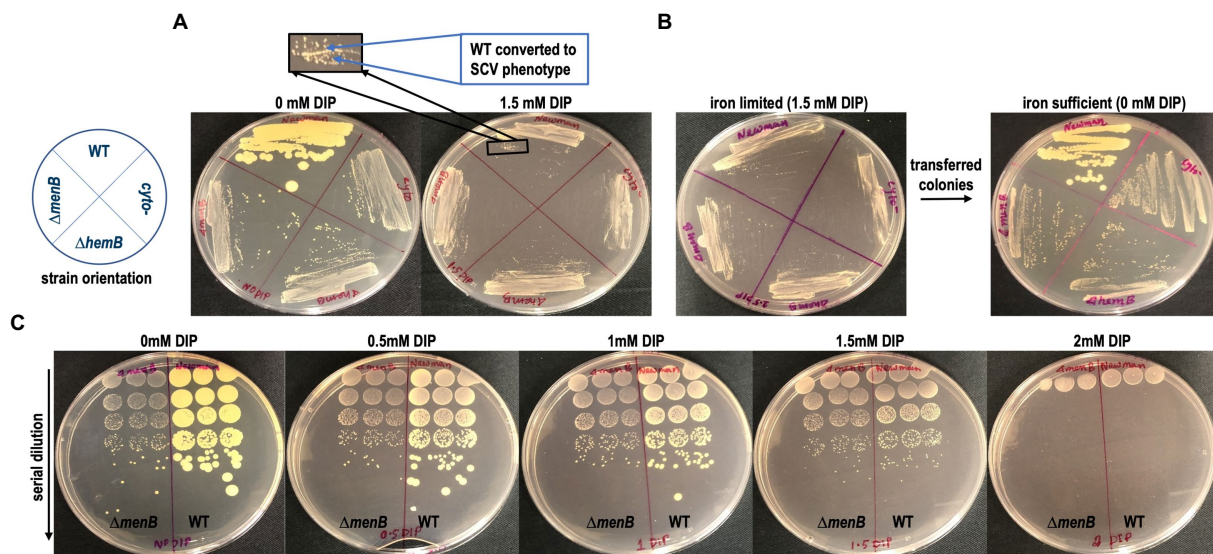


FIGURE 2

Transient SCV phenotypes are induced in genetically WT *Staphylococcus aureus* during iron restriction. (A) WT colonies on TSA are larger and more pigmented than SCV colonies yet are indistinguishable from SCV colonies when grown on TSA supplemented with DIP. (B) The SCV conversion is a physiological adaptation to iron limitation rather than a genetic conversion to SCV form as the cells immediately revert to their large, pigmented colony form when subcultured onto iron replete TSA plates. (C) The SCV conversion occurs over a large concentration range of DIP treatment and is identical to the SCV conversion induced by anaerobic growth. These plates are representative of data observed on multiple plates cultured on at least 3 separate days.

## Transient SCVs induced by iron starvation exhibit similar antibiotic resistance to genetic SCVs

In addition to colony morphology, it was anticipated that the transient SCVs induced by iron starvation might share other fundamental traits with genetic SCVs, including the

characteristic antibiotic tolerances. To evaluate this hypothesis, WT *S. aureus* and a representative SCV strain ( $\Delta menB$ ) were exposed to gentamicin for 4 h in liquid media in the presence or absence of DIP prior to serial dilution and plating onto untreated media for colony forming unit (CFU) determination (Figure 3). In all conditions, the inherently resistant  $\Delta menB$  strain was unaffected by either iron starvation or antibiotic treatment and

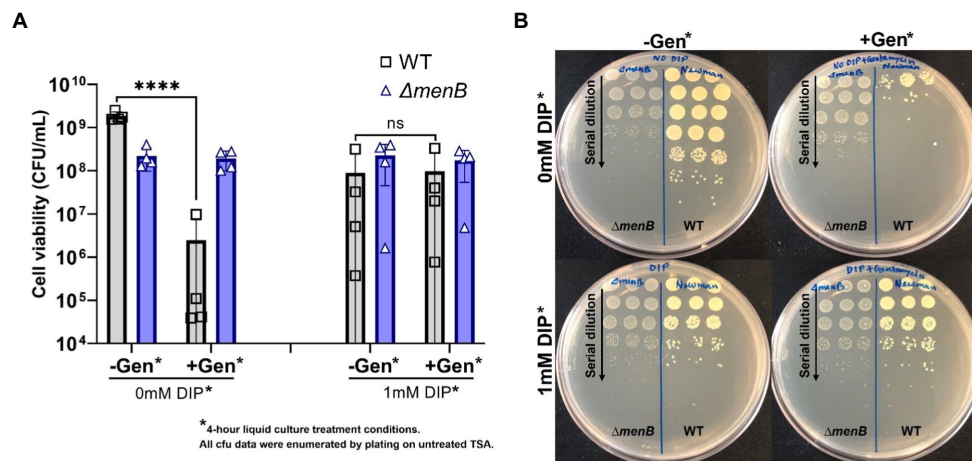


FIGURE 3

Iron restriction enables WT *S. aureus* to exhibit the aminoglycoside resistance characteristic of the SCV phenotype. (A) WT *S. aureus* is highly susceptible to gentamicin treatment but becomes equivalently resistant to this antibiotic as a  $\Delta menB$  strain when iron-restricted with DIP treatment. Each of the data points on the graph represents the average of triplicates on each day and this was repeated for 4 separate days. Error bars represent standard deviation. A one-way ANOVA (Tukey's multiple comparisons) was applied to understand the statistical significance of the differences ("\*\*\*\*" is the indication for  $p < 0.0001$ ; "NS" is for non-significant data). (B) The serial dilution plates depict representative CFU data from a single day and were determined by plating onto untreated TSA plates. Treatment conditions listed in the figure represent the conditions from the 4h incubation period prior to plating on untreated TSA plates for CFU enumeration.

grew to equivalent CFUs. WT cells that were not starved of iron were sensitive to gentamicin treatment. On the other hand, when WT cells were transiently iron starved for the 4-h period during the antibiotic exposure, they achieved CFUs similar to the  $\Delta menB$  strain in the presence or absence of gentamicin treatment. Our finding is supported by a previous study that stated that iron starvation induces gentamicin-resistant *Escherichia coli* regulated by small RNA, RyhB (Chareyre et al., 2019).

## Iron starvation enables the emergence of gentamicin-resistant stable SCVs over time

We further hypothesized that prolonged growth in iron-starved conditions may enable the emergence of stable gentamicin-resistant SCVs, which has potential clinical implications for chronic infections occurring within iron-resisted host niches. The rationale for this hypothesis is that typically naturally occurring SCV mutations would be selected against within an aerobic *S. aureus* population since respiring *S. aureus* grows more rapidly than fermenting *S. aureus*. However, in iron-starved conditions, our data indicate that there would be no such competitive advantage for the maintenance of respiratory function. This hypothesis was tested using experimental evolution of WT cells passaged in 1 mM DIP for 9 days. The number of gentamicin-resistant colonies was quantified after 3 days and 9 days of DIP treatment (Figure 4A). SCVs were displaying a trend toward

increased emergence in iron starvation yet this trend did not achieve statistical significance when analyzed with ANOVA. Of note, the comparison was significant when analyzed by t-test (data not shown). This trend for increased emergence of SCVs under iron starvation appeared more dramatic and did achieve statistical significance by day 9. A previous study indicated that low nutrition combined with a long-time span causes cell heterogeneity in the *S. aureus* population, and stress favors stable SCV cell types (Bui et al., 2015). Therefore, we subcultured several of the gentamicin-resistant SCVs that emerged from the 9-day passaging under iron starvation to assess the phenotypic stability of these populations.

From each of the six replicated 9-day passaged cultures, four gentamicin-resistant SCVs were selected for further characterization for a total of 24 isolates (herein designated isolates 1–24). When streaked out on untreated TSA, all SCVs that emerged from the passaging experiments exhibited a stable SCV phenotype mimicking that of the  $\Delta menB$  SCV (Figure 4B). Of the 24 isolates, 17 were stable following multiple subcultures on untreated media whereas 7 exhibited the ability to revert to a WT growth phenotype following repeated subculturing (Table 1). To define an apparent auxotrophy for each isolate, were individually cross-streaked with the laboratory-derived  $\Delta menB$  or  $\Delta hemB$  SCV strains (Figure 5; Supplementary Figure 1). This type of cross-streaking experiment was previously shown to enable the identification of apparent SCV auxotrophies since *S. aureus* SCV isolates are able to exchange heme and menaquinone to support partial restoration of the respiring phenotype (Hammer et al., 2014).



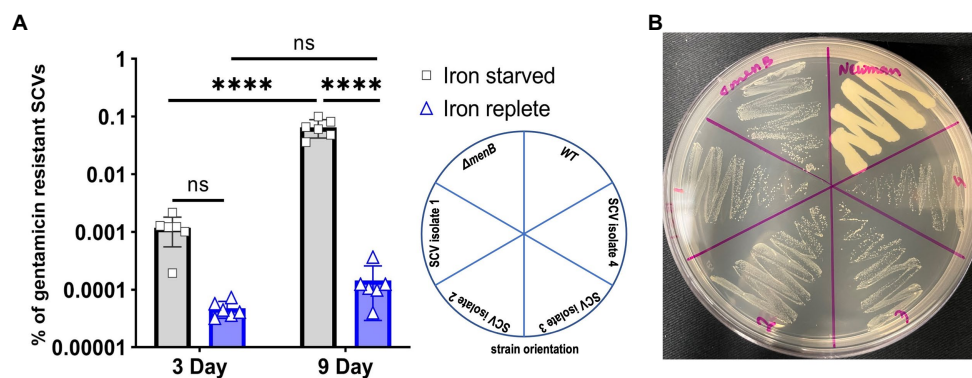


FIGURE 4

Iron starvation promotes increased rates of stable SCV emergence over time. (A) The respiring *Staphylococcus aureus* increased the number of gentamicin-resistant SCVs with the treatment by DIP. Each of the data points on the graph represents one of six replications that were passaged for 9 consecutive days. Error bars indicate the standard deviation. A one-way ANOVA (Tukey's multiple comparisons) was applied to determine the statistical significance of the differences ("\*\*\*\*" is the indication for  $p < 0.0001$ ; "NS" is for non-significant data). (B) Representative plate assessing the subcultured phenotype of the stable, iron starvation-induced SCV isolates when grown on untreated TSA plates.

TABLE 1 Phenotypic assessment of stable, iron starvation-induced SCV isolates.

Isolate #	Apparent SCV auxotrophy	Reversion of phenotype
1	Heme	—
2	Menaquinone	—
3	Heme	—
4	Menaquinone	—
5	Unknown	+
6	Menaquinone	—
7	Menaquinone	—
8	Heme	—
9	Menaquinone	—
10	Unknown	+
11	Unknown	+
12	Menaquinone	—
13	Heme	—
14	Unknown	+
15	Unknown	+
16	Unknown	+
17	Menaquinone	—
18	Menaquinone	—
19	Menaquinone	—
20	Unknown	—
21	Heme	—
22	Unknown	+
23	Unknown	—
24	Heme	—

An isolate was deemed a heme auxotroph if cross-streaking with the heme-producing, menaquinone-deficient  $\Delta menB$  SCV resulted in a growth rescue. An isolate was deemed a menaquinone auxotroph if cross-streaking with a menaquinone-producing, heme-deficient  $\Delta hemB$  SCV resulted in a growth rescue. Unknown auxotrophy indicates that either the isolate did not experience a growth rescue with either cross-streak or they were not a stable SCV that could be subjected to cross-streaking. An isolate in which the SCV phenotype reverted to normal growth upon passaging to fresh media is marked in with a "+" whereas a stable, non-revertible isolate is marked with "—".

Most of our isolates could be characterized as either menaquinone-deficient strains (9 of 24 isolates) when they achieved growth restoration in the presence of  $\Delta hemB$  but not  $\Delta menB$  or heme-deficient strains (6 of 24 isolates) when they achieved growth restoration in the presence of  $\Delta menB$  but not  $\Delta hemB$ . The remaining nine isolates were designated unknown auxotrophies, seven of which were unknown because the SCV phenotype reverted upon further subculture and two of which were unknown because they were not complemented in the presence of either SCV type (Table 1). In total, these data support the contention that brief periods of iron starvation promote a physiological, transient adoption of the SCV phenotype in respiration-capable *S. aureus*, but prolonged exposure to iron starvation can enable the emergence of stable, genetic SCVs within the population.

## Conclusion

The findings presented herein indicate that a major consequence of iron starvation in *S. aureus* is the inhibition of aerobic respiration. Additionally, we observe that the transient SCVs induced as a physiological response to iron starvation are equivalently resistant to aminoglycoside exposure as compared to genetic SCVs. Because certain niches of the host environment are known to be iron restricted *via* various mechanisms of nutritional immunity (Palmer and Skaar, 2016), these findings indicate that many *S. aureus* infections may contain inherently antibiotic-resistant cells even when SCVs are not found to be present during clinical laboratory investigation. Furthermore, experimental evolution demonstrated that SCV cells can emerge naturally within an *S. aureus* population at greater rates when cells are cultured under iron restriction. This finding could have implications during infection in that *S. aureus* chronically infecting iron-restricted host niches might become dominated by



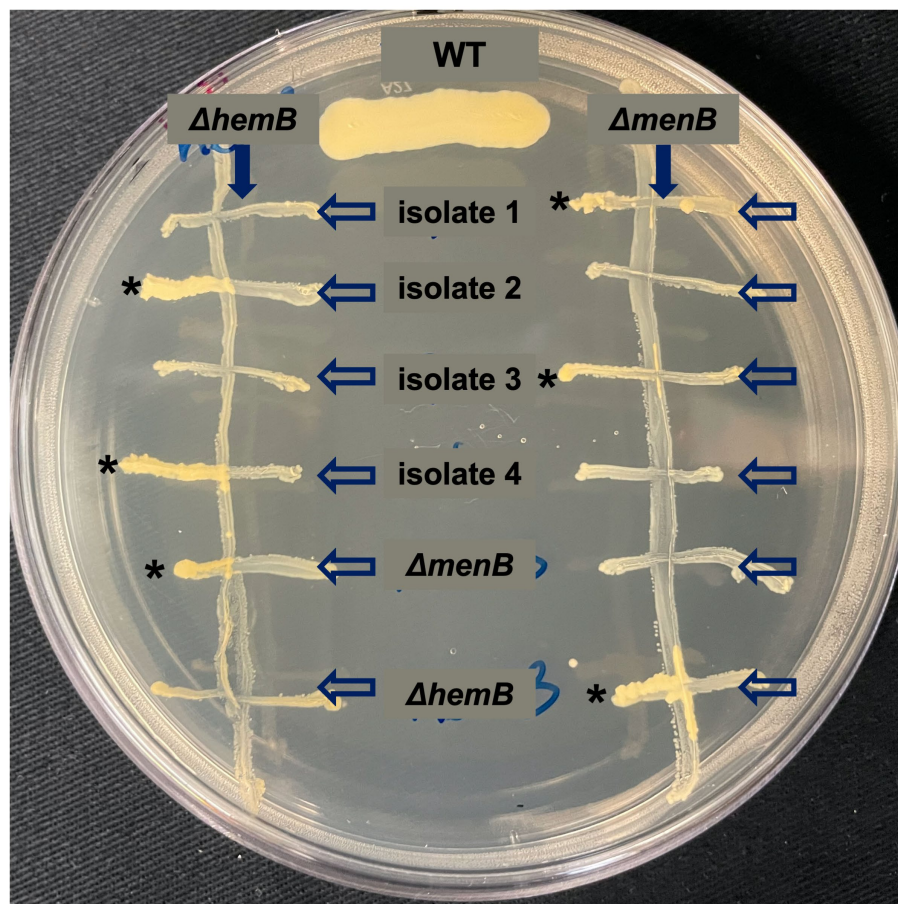


FIGURE 5

Cross-streaking experiments uncover apparent auxotrophies for stable, iron starvation-induced SCV isolates. Representative plate of cross-streaking experiment in which a menaquinone-deficient  $\Delta menB$  or a heme-deficient  $\Delta hemB$  were streaked vertically downward as a primary streak (direction of primary streak is emphasized by a solid arrow). SCV isolates and  $\Delta menB$  and  $\Delta hemB$  as controls were cross-streaked horizontally (direction of the secondary streak is emphasized by unfilled arrows). When a heme-deficient and menaquinone-deficient strain intersect, a moderate growth rescue resembling the WT control streak can be observed (marked with \*). In this way, apparent auxotrophies could be assigned for most SCVs that evolved during prolonged exposure during iron starvation. The raw photograph of this plate can be viewed in [Supplementary Figure 1](#).

naturally occurring genetic SCVs without prior antibiotic exposure.

## Data availability statement

The original contributions presented in the study are included in the article/[Supplementary material](#), further inquiries can be directed to the corresponding author.

## Author contributions

SI contributed to experimental design, data acquisition, data analysis, and manuscript drafting. AC and QH contributed to data acquisition and manuscript editing. CW contributed to project and experimental design, data analysis

and interpretation, and manuscript drafting and editing. All authors contributed to the article and approved the submitted version.

## Funding

Research in the Wakeman lab was supported by NIH/NIGMS (R15GM128072). Additionally, SI was awarded publication funding from Texas Tech American Society for Microbiology, received a long summer research assistantship from the Department of Biological Sciences at Texas Tech University, and was awarded the Doctoral Dissertation Completion Fellowship (DDCF) from Texas Tech Graduate School. AC received support from the Texas Tech University Centre for the Integration of STEM Education and Research (CISER).

## Acknowledgments

The authors would like to thank Kate C. O'Malley, Chukwuemeka C. Okafor, and Emmett Jamieson for their help in making bacterial growth media and some experimental assistance. Additionally, the authors would like to thank Hafij Al Mahmud, Jayc E. Waller, Md Siddiqur Rahman, Garrett N. Mathews, and Priyanka Mathews for their critical manuscript reading.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

- Baishya, J., and Wakeman, C. A. (2019). Selective pressures during chronic infection drive microbial competition and cooperation. *NPJ Biofilms Microbiomes* 5:16. doi: 10.1038/s41522-019-0089-2
- Bjorkman, J., Nagaev, I., Berg, O. G., Hughes, D., and Andersson, D. I. (2000). Effects of environment on compensatory mutations to ameliorate costs of antibiotic resistance. *Science* 287, 1479–1482. doi: 10.1126/science.287.5457.1479
- Boles, B. R., Thoendel, M., and Singh, P. K. (2004). Self-generated diversity produces “insurance effects” in biofilm communities. *Proc. Natl. Acad. Sci.* 101, 16630–16635. doi: 10.1073/pnas.0407460101
- Bui, L. M. G., Hoffmann, P., Turnidge, J. D., Zilm, P. S., and Kidd, S. P. (2015). Prolonged growth of a clinical *Staphylococcus aureus* strain selects for a stable small-colony-variant cell type. *Infect. Immun.* 83, 470–481. doi: 10.1128/IAI.02702-14
- Bui, L. M. G., and Kidd, S. P. (2015). A full genomic characterization of the development of a stable small colony variant cell-type by a clinical *Staphylococcus aureus* strain. *Infect. Genet. Evol.* 36, 345–355. doi: 10.1016/j.meegid.2015.10.011
- Bui, L. M. G., Turnidge, J. D., and Kidd, S. P. (2015). The induction of *Staphylococcus aureus* biofilm formation or small colony variants is a strain-specific response to host-generated chemical stresses. *Microbes Infect.* 17, 77–82. doi: 10.1016/j.micinf.2014.09.009
- Bullen, J. J., Griffiths, E., and Edmiston, C. E. (1999). *Iron and infection: Molecular, physiological and clinical aspects*. 2 Edn. United Kingdom: LWW.
- Calomiris, J. J., Armstrong, J. L., and Seidler, R. J. (1984). Association of metal tolerance with multiple antibiotic resistance of bacteria isolated from drinking water. *Appl. Environ. Microbiol.* 47, 1238–1242. doi: 10.1128/aem.47.6.1238-1242.1984
- Chareyre, S., Barras, F., and Mandin, P. (2019). A small RNA controls bacterial sensitivity to gentamicin during iron starvation. *PLoS Genet.* 15:e1008078. doi: 10.1371/journal.pgen.1008078
- Cohen, S., Sweeney, H. M., and Leitner, F. (1967). Relation between iron uptake, pH of growth medium, and penicillinase formation in *Staphylococcus aureus*. *J. Bacteriol.* 93, 1227–1235. doi: 10.1128/jb.93.4.1227-1235.1967
- Dassy, B., and Fournier, J.-M. (1996). Respiratory activity is essential for post-exponential-phase production of type 5 capsular polysaccharide by *Staphylococcus aureus*. *Infect. Immun.* 64, 2408–2414. doi: 10.1128/iai.64.7.2408-2414.1996
- Dastgheyb, S. S., and Otto, M. (2015). Staphylococcal adaptation to diverse physiologic niches: an overview of transcriptomic and phenotypic changes in different biological environments. *Future Microbiol.* 10, 1981–1995. doi: 10.2217/fmb.15.116
- Davies, J. A., Harrison, J. J., Marques, L. L. R., Foglia, G. R., Stremick, C. A., Storey, D. G., et al. (2007). The Gac S sensor kinase controls phenotypic reversion of small colony variants isolated from biofilms of *Pseudomonas aeruginosa* PA14. *FEMS Microbiol. Ecol.* 59, 32–46. doi: 10.1111/j.1574-6941.2006.00196.x
- Dean, M. A., Olsen, R. J., Wesley Long, S., Rosato, A. E., and Musser, J. M. (2014). Identification of point mutations in clinical *Staphylococcus aureus* strains that produce small-colony variants auxotrophic for menadione. *Infect. Immun.* 82, 1600–1605. doi: 10.1128/IAI.01487-13
- Duthie, E. S., and Lorenz, L. L. (1952). Staphylococcal coagulase: mode of action and antigenicity. *Microbiology* 6, 95–107. doi: 10.1099/00221287-6-1-2-95
- Edwards, M. J., Richardson, D. J., Paquette, C. M., and Clarke, T. A. (2020). Role of multiheme cytochromes involved in extracellular anaerobic respiration in bacteria. *Protein Sci.* 29, 830–842. doi: 10.1002/pro.3787
- Friedman, D. B., Stauff, D. L., Pishchany, G., Whitwell, C. W., Torres, V. J., and Skaar, E. P. (2006). *Staphylococcus aureus* redirects central metabolism to increase iron availability. *PLoS Pathog.* 2:e87. doi: 10.1371/journal.ppat.0020087
- Hammer, N. D., Cassat, J. E., Noto, M. J., Lojek, L. J., Chadha, A. D., Schmitz, J. E., et al. (2014). Inter- and intraspecies metabolite exchange promotes virulence of antibiotic-resistant *Staphylococcus aureus*. *Cell Host Microbe* 16, 531–537. doi: 10.1016/j.chom.2014.09.002
- Hammer, N. D., Reniere, M. L., Cassat, J. E., Zhang, Y., Hirsch, A. O., Indriati Hood, M., et al. (2013). Two heme-dependent terminal oxidases power *Staphylococcus aureus* organ-specific colonization of the vertebrate host. *MBio* 4, e00241–e00213. doi: 10.1128/mBio.00241-13
- Harrison, J. J., Ceri, H., and Turner, R. J. (2007). Multimetal resistance and tolerance in microbial biofilms. *Nat. Rev. Microbiol.* 5, 928–938. doi: 10.1038/nrmicro1774
- Herrin Brittany, E., Islam, S., Rentschler, K. N., Pert, L. H., Kopanski, S. P., and Wakeman, C. A. (2021). Haem toxicity provides a competitive advantage to the clinically relevant *Staphylococcus aureus* small colony variant phenotype. *Microbiology* 167:001044. doi: 10.1099/mic.0.001044
- Johnson, M., Cockayne, A., and Morrissey, J. A. (2008). Iron-regulated biofilm formation in *Staphylococcus aureus* Newman requires Ica and the secreted protein Emp. *Infect. Immun.* 76, 1756–1765. doi: 10.1128/IAI.01635-07
- Klevens, R. M., Morrison, M. A., Nadle, J., Petit, S., Gershman, K., Ray, S., et al. (2007). Invasive methicillin-resistant *Staphylococcus aureus* infections in the United States. *JAMA* 298, 1763–1771. doi: 10.1001/jama.298.15.1763
- Krishna, S., and Miller, L. S. (2012). Host–pathogen interactions between the skin and *Staphylococcus aureus*. *Curr. Opin. Microbiol.* 15, 28–35. doi: 10.1016/j.mib.2011.11.003
- Kuehnert, M. J., Kruszon-Moran, D., Hill, H. A., McQuillan, G., McAllister, S. K., Fosheim, G., et al. (2006). Prevalence of *Staphylococcus aureus* nasal colonization in the United States, 2001–2002. *J. Infect. Dis.* 193, 172–179. doi: 10.1086/499632
- Laupland, K. B., Lyytikäinen, O., Sgaard, M., Kennedy, K. J., Knudsen, J. D., Christian Ostergaard, J. C., et al. (2013). The changing epidemiology of *Staphylococcus aureus* bloodstream infection: a multinational population-based surveillance study. *Clin. Microbiol. Infect.* 19, 465–471. doi: 10.1111/j.1469-0691.2012.03903.x
- Ledala, N., Zhang, B., Seravalli, J., Powers, R., and Somerville, G. A. (2014). Influence of iron and aeration on *Staphylococcus aureus* growth, metabolism, and transcription. *J. Bacteriol.* 196, 2178–2189. doi: 10.1128/JB.01475-14
- Lopez, C. A., and Skaar, E. P. (2018). The impact of dietary transition metals on host-bacterial interactions. *Cell Host Microbe* 23, 737–748. doi: 10.1016/j.chom.2018.05.008
- Lung, W. F., Tania, E., and Prince, A. (2020). Consequences of metabolic interactions during *Staphylococcus aureus* infection. *Toxins* 12:581. doi: 10.3390/toxins12090581
- Masoud-Landgraf, L., Zarfel, G., Kaschnigg, T., Friedl, S., Feierl, G., Wagner-Eibel, U., et al. (2016). Analysis and characterization of *Staphylococcus aureus* small colony variants isolated from cystic fibrosis patients in Austria. *Curr. Microbiol.* 72, 606–611. doi: 10.1007/s00284-016-0994-z

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.978859/full#supplementary-material>

- McNamara, P. J., and Proctor, R. A. (2000). *Staphylococcus aureus* small colony variants, electron transport and persistent infections. *Int. J. Antimicrob. Agents* 14, 117–122. doi: 10.1016/S0924-8579(99)00170-3
- Minandri, F., Imperi, F., Frangipani, E., Bonchi, C., Visaggio, D., Facchini, M., et al. (2016). Role of iron uptake systems in *Pseudomonas aeruginosa* virulence and airway infection. *Infect. Immun.* 84, 2324–2335. doi: 10.1128/IAI.00098-16
- Nagaev, I., Björkman, J., Andersson, D. I., and Hughes, D. (2001). Biological cost and compensatory evolution in fusidic acid-resistant *Staphylococcus aureus*. *Mol. Microbiol.* 40, 433–439. doi: 10.1046/j.1365-2958.2001.02389.x
- Noble, W. C., Valkenburg, H. A., and Wolters, C. H. L. (1967). Carriage of *Staphylococcus aureus* in random samples of a normal population. *Epidemiol. Infect.* 65, 567–573. doi: 10.1017/S002217240004609X
- Onyango, L. A., Hugh Dunstan, R., Roberts, T. K., Macdonald, M. M., and Gottfries, J. (2013). Phenotypic variants of staphylococci and their underlying population distributions following exposure to stress. *PLoS One* 8:e77614. doi: 10.1371/journal.pone.0077614
- Painter, K. L., Strange, E., Parkhill, J., Bamford, K. B., Armstrong-James, D., and Edwards, A. M. (2015). *Staphylococcus aureus* adapts to oxidative stress by producing H<sub>2</sub>O<sub>2</sub>-resistant small-colony variants via the SOS response. *Infect. Immun.* 83, 1830–1844. doi: 10.1128/IAI.03016-14
- Palmer, L. D., and Skaar, E. P. (2016). Transition metals and virulence in bacteria. *Annu. Rev. Genet.* 50, 67–91. doi: 10.1146/annurev-genet-120215-035146
- Poole, K. (2012). Bacterial stress responses as determinants of antimicrobial resistance. *J. Antimicrob. Chemother.* 67, 2069–2089. doi: 10.1093/jac/dks196
- Poole, K., Krebes, K., McNally, C., and Neshat, S. H. A. D. I. (1993). Multiple antibiotic resistance in *Pseudomonas aeruginosa*: evidence for involvement of an efflux operon. *J. Bacteriol.* 175, 7363–7372. doi: 10.1128/jb.175.22.7363-7372.1993
- Proctor, R. A., Balwit, J. M., and Vesga, O. (1994). Variant subpopulations of *Staphylococcus aureus* as cause of persistent and recurrent infections. *Infect. Agents Dis.* 3, 302–312. PMID: 7889317
- Proctor, R. A., Kriegeskorte, A., Kahl, B. C., Becker, K., Löffler, B., and Peters, G. (2014). *Staphylococcus aureus* small Colony variants (SCVs): a road map for the metabolic pathways involved in persistent infections. *Front. Cell. Infect. Microbiol.* 4:99. doi: 10.3389/fcimb.2014.00099
- Proctor, R. A., Von Eiff, C., Kahl, B. C., Becker, K., McNamara, P., Herrmann, M., et al. (2006). Small colony variants: a pathogenic form of bacteria that facilitates persistent and recurrent infections. *Nat. Rev. Microbiol.* 4, 295–305. doi: 10.1038/nrmicro1384
- Richardson, A. R. (2019). Virulence and metabolism. *Microbiol. Spectr.* 7:7.2. 39. doi: 10.1128/microbiolspec.GPP3-0011-2018
- Rollin, G., Tan, X., Tros, F., Dupuis, M., Nassif, X., Charbit, A., et al. (2017). Intracellular survival of *Staphylococcus aureus* in endothelial cells: a matter of growth or persistence. *Front. Microbiol.* 8:1354. doi: 10.3389/fmicb.2017.01354
- Schaff, F., Bierbaum, G., Baumert, N., Bartmann, P., and Sahl, H.-G. (2003). Mutations are involved in emergence of aminoglycoside-induced small colony variants of *Staphylococcus aureus*. *Int. J. Med. Microbiol.* 293, 427–435. doi: 10.1078/1438-4221-00282
- Sifri, C. D., Baresch-Bernal, A., Calderwood, S. B., and von Eiff, C. (2006). Virulence of *Staphylococcus aureus* small colony variants in the *Caenorhabditis elegans* infection model. *Infect. Immun.* 74, 1091–1096. doi: 10.1128/IAI.74.2.1091-1096.2006
- Thammavongsa, V., Kim, H. K., Missiakas, D., and Schneewind, O. (2015). Staphylococcal manipulation of host immune responses. *Nat. Rev. Microbiol.* 13, 529–543. doi: 10.1038/nrmicro3521
- Trivier, D., and Courcol, R. J. (1996). Iron depletion and virulence in staphylococcus aureus. *FEMS Microbiol. Lett.* 141, 117–127. doi: 10.1111/j.1574-6968.1996.tb08373.x
- Tuchscher, L., Kreis, C. A., Hoerr, V., Flint, L., Hachmeister, M., Geraci, J., et al. (2016). *Staphylococcus aureus* develops increased resistance to antibiotics by forming dynamic small colony variants during chronic osteomyelitis. *J. Antimicrob. Chemother.* 71, 438–448. doi: 10.1093/jac/dkv371
- Wakeman, C. A., Hammer, N. D., Stauff, D. L., Attia, A. S., Anzaldi, L. L., Dikalov, S. I., et al. (2012). Menaquinone biosynthesis potentiates haem toxicity in *S. aureus*. *Mol. Microbiol.* 86, 1376–1392. doi: 10.1111/mmi.12063
- Wales, A. D., and Davies, R. H. (2015). Co-selection of resistance to antibiotics, biocides and heavy metals, and its relevance to foodborne pathogens. *Antibiotics* 4, 567–604. doi: 10.3390/antibiotics4040567
- Wilson, S. G., and Sanders, C. C. (1976). Selection and characterization of strains of *Staphylococcus aureus* displaying unusual resistance to aminoglycosides. *Antimicrob. Agents Chemother.* 10, 519–525. doi: 10.1128/AAC.10.3.519
- Wissenbach, U., Kroger, A., and Uden, G. (1990). The specific functions of menaquinone and demethylmenaquinone in anaerobic respiration with fumarate, dimethylsulfoxide, trimethylamine N-oxide and nitrate by *Escherichia coli*. *Arch. Microbiol.* 154, 60–66. doi: 10.1007/BF00249179
- Workentine, M. L., Harrison, J. J., Weljie, A. M., Tran, V. A., Stenroos, P. U., Tremaroli, V., et al. (2010). Phenotypic and metabolic profiling of colony morphology variants evolved from *Pseudomonas fluorescens* biofilms. *Environ. Microbiol.* 12, 1565–1577. doi: 10.1111/j.1462-2920.2010.02185.x



## OPEN ACCESS

## EDITED BY

Sabine Kleinstüber,  
Helmholtz Association of German Research  
Centres (HZ), Germany

## REVIEWED BY

James Ferry,  
The Pennsylvania State University,  
United States  
Barry Whitman,  
University of Georgia,  
United States

## \*CORRESPONDENCE

Nicole R. Buan  
nbuan@unl.edu

## SPECIALTY SECTION

This article was submitted to  
Microbial Physiology and Metabolism,  
a section of the journal  
Frontiers in Microbiology

RECEIVED 01 September 2022

ACCEPTED 28 October 2022

PUBLISHED 15 December 2022

## CITATION

Carr S and Buan NR (2022) Insights into the  
biotechnology potential of *Methanosarcina*.  
*Front. Microbiol.* 13:1034674.  
doi: 10.3389/fmicb.2022.1034674

## COPYRIGHT

© 2022 Carr and Buan. This is an open-  
access article distributed under the terms  
of the [Creative Commons Attribution  
License \(CC BY\)](#). The use, distribution or  
reproduction in other forums is permitted,  
provided the original author(s) and the  
copyright owner(s) are credited and that  
the original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use, distribution or  
reproduction is permitted which does not  
comply with these terms.

# Insights into the biotechnology potential of *Methanosarcina*

Sean Carr and Nicole R. Buan\*

Department of Biochemistry, University of Nebraska-Lincoln, Lincoln, NE, United States

Methanogens are anaerobic archaea which conserve energy by producing methane. Found in nearly every anaerobic environment on earth, methanogens serve important roles in ecology as key organisms of the global carbon cycle, and in industry as a source of renewable biofuels. Environmentally, methanogenic archaea play an essential role in the reintroducing unavailable carbon to the carbon cycle by anaerobically converting low-energy, terminal metabolic degradation products such as one and two-carbon molecules into methane which then returns to the aerobic portion of the carbon cycle. In industry, methanogens are commonly used as an inexpensive source of renewable biofuels as well as serving as a vital component in the treatment of wastewater though this is only the tip of the iceberg with respect to their metabolic potential. In this review we will discuss how the efficient central metabolism of methanoarchaea could be harnessed for future biotechnology applications.

## KEYWORDS

methanogens, archaea, biotechnology, methane, biofuel, terpenoid chemicals

## Methanogen ecology and diversity

Methanogens are single-celled organisms that conserve energy *via* the conversion of substrate carbon compounds into methane gas (Ferry, 2012). The majority of methanogens subsist in anaerobic environments by the reduction of one carbon (C1) compounds including carbon dioxide and carbon monoxide, methanol, methylamines, and methyl sulfides as well as the fermentation of acetate (Daniels et al., 1977; Rother and Metcalf, 2004; Buan, 2018). The gaseous methane they produce then bubbles back into the aerobic world where it is consumed by methanotrophic organisms and is returned to the carbon cycle. The methane produced by methanogens is of interest due to methane's ecological impact resulting from agricultural production by livestock (Johnson and Johnson, 1995) and rice cultivation (Schütz et al., 1989) as well as methane's benefits as a renewable source of natural gas (Luo and Angelidaki, 2012; Huang et al., 2017) which is a high energy fuel used for heat, electricity generation, and for transportation including as a propellant for rocket engines (Neill et al., 2009; Sheehan, 2021). In nature, methanogenic archaea have been identified in environments spanning the boundaries of life sustaining conditions, from acidic to alkaline (pH 3.0–10.2), thermophilic to psychrophilic (−2°C to 110°C), and including both fresh and saline aquatic environments (Martin and Sousa, 2016). In addition to these environments, methanogens are found symbiotically communing in a wide range of



single- and multi-cellular hosts ranging from amoebae (Holmes et al., 2014) and protozoa (Stumm and Zwart, 1986) to termites (Brune, 1998, 2018) bovines (Whitford et al., 2001), and humans (Fricke et al., 2006; Rajilić-Stojanović et al., 2007).

As more methanogen species are discovered, it is becoming evident that methanogens may be able to use a wider variety of substrates than previously known. Biomethane generation has been observed from subsurface coal beds (Ulrich and Bower, 2008; Mayumi et al., 2016) as well as oceanic oil sinks (Laso-Pérez et al., 2019; Zhou et al., 2022). Methanogens may form syntrophic partnerships with other microorganisms such as hydrocarbon-degrading bacteria, thereby indirectly facilitating the reintroduction of crude oil carbon into a bioavailable state (Zengler et al., 1999; Dolfig et al., 2008; Jones et al., 2008). However, methanogens are suspected of being capable of alkane oxidation independent of any other archaeal or bacterial partner (Borrel et al., 2019; Laso-Pérez et al., 2019). Ecological methane accumulation has been observed in correlation with colonized oil droplets at deep-sea oil seeps. These proposed alkane utilizing methanogens are not limited to short-chain alkanes; *Candidatus Methanoliparum* has been shown to degrade long-chain hydrocarbons with methanogenesis (Zhou et al., 2022). The mechanism by which these so far uncultured alkanotrophic methanogens are capable of utilizing hydrocarbons is still being investigated, though the phenomenon does not appear to be a rare occurrence. Alkane-degrading methanogens are widely distributed, (Zengler et al., 1999; Laso-Pérez et al., 2019; Zhou et al., 2022) indicating that methanogens are directly or indirectly involved in the bioconversion of crude oil to methane on a large scale and may serve a benefit to bioremediation efforts in anaerobic environments such as deep-sea sediments.

The ability for methanogens to thrive in these diverse environments is testament to their metabolic robustness. Regardless of the environment they inhabit, methanogens share a similar metabolic niche, the bioconversion of low-energy substrates into biomass and high-energy molecules with a high degree of efficiency. All cultured methanogens to date are strict obligate anaerobes and produce methane as an essential byproduct of metabolism (Daniels et al., 1977; Rother and Metcalf, 2004; Ferry, 2012; Buan, 2018). To grow on these energy poor substrates methanogens have adopted a highly efficient pathway for conserving energy called methanogenesis (Figure 1; Thauer, 2012; Gonnerman et al., 2013). In this review we discuss how methanogen metabolism allows these organisms to thrive under strict energetic conditions and how their special metabolic features could be utilized in biotechnology.

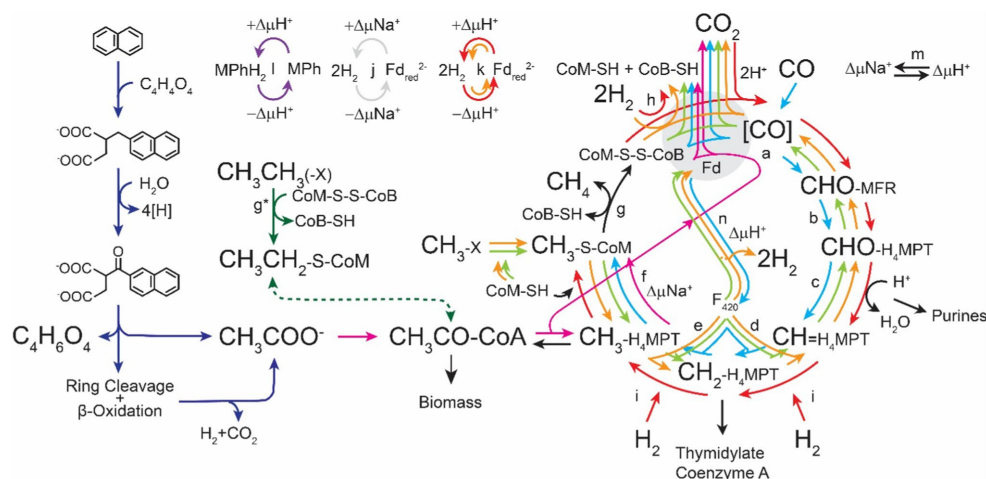
## Expanding the methanogenesis pathway

Despite their ability to live in a wide diversity of habitats, methanogens are united by their unique central metabolism. In the five characterized versions of the methanogenesis pathway,

substrates are reduced to methane while formate, primary alcohols/amines/thiols, or  $H_2$  are oxidized to  $CO_2$  or  $H_2O$  (Ermler et al., 1997; Buan, 2018). Redox cofactors associated with the methanogenesis pathway are regenerated through formation of a transmembrane ion gradient which is coupled to ATP synthesis via ATP synthase (Costa and Leigh, 2014; Diender et al., 2015). These reactions yield a small amount of energy for the methanogen amounting to between 0.5 and 2 moles of ATP per mole of substrate (Buan, 2018). A result of this low energy yield is a high relative flux through energy conservation pathways, with over 99% of the chemistry within the cell being directly tied to methanogenesis (Feist et al., 2006). The remaining 1–2% of carbon substrate is used to generate biomass for replication. The average macromolecular composition of a methanogen includes 63% protein, 0.1% fatty acid lipids, 5% isoprenoid lipids, 0.5% carbohydrates, 28% nucleic acids, and 4% metabolites and metabolic precursors (Gonnerman et al., 2013). The relatively high abundance of isoprenoid lipids and high protein concentration make them an appealing source of difficult-to-synthesize lipids and molecules from inexpensive C1 compounds or acetate, yields and titers of which could be further enhanced through genetic engineering. The unique properties of methanogenesis and highly efficient energy conservation mechanisms make methanoarchaea ideal organisms for the production of renewable biofuels as the vast majority of feed substrate is converted efficiently to methane.

It should be noted, however, that while methanogenesis is highly conserved and exceedingly efficient, it can also be modified to better serve biotechnological goals without necessarily undermining methanogenic growth. *Methanosarcina* in particular may be well-suited to metabolic engineering, as they can use multiple methanogenesis pathways and are genetically tractable (Metcalf et al., 1997; Ehlers et al., 2005).

Methanogenesis is inherently limited by substrate availability though this limitation can be overcome by expanding the carbon and energy sources available to methanoarchaea. *Methanosarcina acetivorans* has been successfully engineered to expand its substrate use and to enhance metabolic efficiency. *M. acetivorans* is a marine methanogen that can use methylotrophic and acetotrophic methanogenesis pathways, but unusually cannot use  $H_2$  for methanogenesis (Sowers et al., 1984; Guss et al., 2009). As a result, *M. acetivorans* appears to use very efficient intracellular redox balancing mechanisms, thus avoiding loss of  $H_2$  reducing equivalents by gas diffusion, which is a possibility for methanogens that use  $H_2$  cycling to generate transmembrane proton gradients (Kulkarni et al., 2009). Methylotrophic methanogenesis relies on substrate specific methyltransferases to convert substrates to  $CH_3$ -CoM for entry into the pathway. It has been demonstrated that heterologous expression of the bacterial broad-specificity esterase from *Pseudomonas veronii* in *M. acetivorans* increased esterase activity 80-fold and greatly enhanced growth on methyl acetate and methyl propionate substrates (Lessner et al., 2010). Once substrates have entered the methanogenesis pathway, energy conservation occurs by the regeneration of methanogenic cofactors (Thauer, 2012). Cofactor regeneration is catalyzed by membrane bound,



**FIGURE 1**  
Pathways for methanogenesis (adapted from [Buan, 2018](#)). The direction of arrows represents the direction of biochemical reactions. Reactions which are utilized in every methanogenic pathway are represented in black. Hydrogenotrophic methanogenesis (aka. The Wolfe Cycle) ([Thauer, 2012](#)) is represented in red. Methyl oxidation is represented in orange. Methylotrophic methanogenesis is represented in green. Acetotrophic methanogenesis is represented in fuchsia. Degradation of polyaromatic hydrocarbons is represented in dark blue ([Siegert et al., 2011](#)). Ethylene and long chain alkane reduction is represented in purple ([Lemaire and Wagner, 2022](#)). Carboxydrotrophic methanogenesis is represented in cyan. CoB-SH, coenzyme B thiol; CoM-S-S-CoB, coenzyme M-coenzyme B heterodisulfide; Fd, ferredoxin; Fd<sub>red</sub>, reduced ferredoxin; H<sub>4</sub>MPT, tetrahydromethanopterin; MFR, methanofuran; MPH, methanophenazine; MPH<sub>2</sub>, reduced methanophenazine. Enzymes involved in methanogenesis: **(a)** Formyl-methanofuran dehydrogenase (Fmd), **(b)** Formyl-methanofuran:H<sub>4</sub>MPT formyl transferase (Ftr), **(c)** Methylene-H<sub>4</sub>MPT cyclohydrolase (Mch), **(d)** F<sub>420</sub>-dependent Methylene-H<sub>4</sub>MPT dehydrogenase (Mtd), **(e)** F420-dependent Methylene-H<sub>4</sub>MPT reductase (Mer), **(f)** Methyl-H<sub>4</sub>MPT:coenzyme M methyltransferase (Mtr), **(g)** Methyl-coenzyme M reductase (Mcr), **(g\*)** Atypical methyl-coenzyme M reductase (Mcr), ([Wang et al., 2019](#)) **(h)** Electron-bifurcating hydrogenase:heterodisulfide reductase complex (Mvh:HdrABC), **(i)** F<sub>420</sub>-reducing hydrogenase (Frh), **(j)** Energy-converting sodium pumping ferredoxin hydrogenase, **(k)** Ferredoxin reducing hydrogenase (Eha/Ech), **(l)** Proton-translocating methanophenazine:heterodisulfide reductase (HdrED), **(m)** Sodium-proton antiporter (MrpA), **(n)** F<sub>420</sub> proton-pumping methanophenazine reductase (Fpo).

redox-driven enzyme complexes such as Rnf (Schlegel et al., 2012) and HdrED (Duszenko and Buan, 2017), which combine cofactor regeneration with ion transport or by cytoplasmic enzymes such as Fpo (Welte and Deppenmeier, 2011), or the terminal oxidase HdrABC (Catlett et al., 2015; Buckel and Thauer, 2018). By enhancing cofactor regeneration it is possible to stimulate increased methanogenesis. For example, when the cytoplasmic enzyme heterodisulfide reductase (HdrABC) is overexpressed, methane production on methanol is 30% faster without a detectable change in growth rate compared to the parent strain (Catlett et al., 2015). The exogenous addition of methanophenazine (MPh), an electron carrier found in methanogens which fulfils a similar role as quinones in other electron transport chains, was found to significantly increase growth in *Methanosarcina* spp (Duszenko and Buan, 2017). Additionally, it is possible that the regeneration of methanogenic cofactors could be achieved through pathway engineering (Aldridge et al., 2021). The reduction of the disulfide complex between coenzyme M and coenzyme B is the final step in all methanogenic pathways and is restricted to the heterodisulfide reductases HdrABC and HdrED (Buan and Metcalf, 2010; Yan and Ferry, 2018). Providing an alternative means of cofactor reduction would eliminate this metabolic bottleneck, freeing up cofactors at a greater rate (Aldridge et al., 2021). If a methanogen were engineered to produce a non-native metabolite which allows for the reduction of ferredoxin,  $F_{420}$ , coenzyme M, or coenzyme B then

production of that metabolite has the potential to increase the rate of methanogenesis while also synthesizing the desired product (Aldridge et al., 2021).

Due to the tight energetic restrictions methanogenesis is proposed to rely heavily on substrate channeling to minimize entropic effects (Costa et al., 2010; Matschiavelli et al., 2012; Catlett et al., 2015; Yan and Ferry, 2018; Watanabe et al., 2021). Substrate channeling allows methanogenesis to function efficiently but presents challenges for metabolic engineers as the metabolite pools for methanogenesis have limited availability outside of the channeled enzyme complexes. To overcome this metabolic obstacle metabolic engineers must choose products which draw from metabolites which are not directly channeled or incorporate the production of their products within methanogenesis. Table 1 lists potential strategies to increase substrate variety, optimize growth rates and culture conditions, or generate new metabolic products by engineering methanogenesis.

## Anaerobic oxidation of methane and reverse methanogenesis

Given the efficiency of methanogenesis and the abundance of anaerobic environments around the world, methanogens are distributed across every continent. Yet of the approximately 1 Gt

TABLE 1 Strategies for expanding the metabolic potential of the methanogenesis pathway.

Desired trait	Potential mechanism
Increased methanogenesis and methane production	<ul style="list-style-type: none"> <li>• <b>Overexpression of genes associated with methanogenesis or addition of parallel heterologous methanogenic pathways.</b></li> <li>• <b>Exogenous addition of metabolites or pathway engineering to supply limiting metabolites.</b></li> </ul> <p>Research has shown that overexpression of redox-active cofactors such as methanophenazine in <i>Methanosarcinales</i> relieves the metabolic bottleneck caused by cofactor regeneration and increases the production of methane (Catlett et al., 2015). Vitamins addition often stimulates growth (Tanner and Wolfe, 1988; Jin et al., 2017). Many methanogens are fully prototrophic, but some strains are dependent on exogenous addition of CoM or other vitamins (Catlett et al., 2022). Additionally, magnetite nanoparticles have been demonstrated to serve facilitate increased acetotrophic methanogenesis in cocultures between acetogens and methanogens (Tanner and Wolfe, 1988; Catlett et al., 2015; López Muñoz et al., 2015; Jin et al., 2017; Fu et al., 2019; Catlett et al., 2022).</p>
Increased substrate uptake rates	<ul style="list-style-type: none"> <li>• <b>By overexpressing endogenous or ortho/heterologous methyltransferases and hydrogenases more substrate carbon could enter methanogenesis.</b></li> </ul> <p>In methylotrophic methanogenesis entry point methanogenesis is limited by the substrate-specific methyltransferase whereas hydrogenotrophic methanogens rely upon membrane bound methyltransferase to conserve energy and maintain the methanogens sodium motive force (Kurth et al., 2020).</p>
Increased substrate diversity and mixotrophy	<ul style="list-style-type: none"> <li>• <b>Introduction of multiple substrate input pathways would allow more rapid substrate uptake, faster unitrophic growth, and mixotrophic growth.</b></li> </ul> <p>Substrate entry into methanogenesis is limited by substrate specific methyltransferases and whether the methanogen can directly utilize H<sub>2</sub> as an electron source. By introducing methyltransferases from different methanogens one can expand the substrates usable to the methanogen. Increasing extracellular-facing hydrogenases may allow increased rates of H<sub>2</sub> uptake and hydrogenotrophic methanogenesis. Upregulation of pyruvate ferredoxin oxidoreductase (<i>por</i>) in <i>M. barkeri</i> has been demonstrated to facilitate growth on pyruvate as a sole carbon and energy source (López Muñoz et al., 2015).</p>
Controlled energy conservation	<ul style="list-style-type: none"> <li>• <b>Selective uncoupling biomass from methanogenesis could allow maximal growth for bioreactor scale-up with a methanogenesis-only production phase.</b></li> </ul> <p>This could be accomplished by bypassing ATP synthesis, managing macromolecular accessibility, by adding protein synthesis inhibitors, or futile cycling for redox cofactors either chemically or genetically (Catlett et al., 2015).</p>
Increased stress resistance	<ul style="list-style-type: none"> <li>• <b>Increased stress tolerance could increase growth rate, improve expression of introduced enzymes, enable production of xenobiotic chemicals, and expand biorefining process parameters.</b></li> </ul> <p>All methanogens are strict anaerobes. Increased oxygen tolerance was observed in <i>Methanosarcina acetivorans</i> when gradually passed with increased O<sub>2</sub> concentrations over a course of 6 months (Jasso-Chávez et al., 2015). Transcripts from adapted <i>Methanosarcina</i> suggest the over expression of superoxide dismutase, catalase, and peroxidase will confer increased aerotolerance to other methanogens. Methanogens engineered to express the bacterial catalase EcKatG demonstrated increased tolerance of hydrogen peroxide, though no increase in resistance to O<sub>2</sub> was observed (Jennings et al., 2014). Though non-spore-forming, methanogens are capable of revival after desiccation with no significant loss of viability observed in aerobic environments (Anderson et al., 2012). Cocultivation with sulfate reducing bacteria has shown to mitigate heavy metal stress in methanogenic cultures (Paulo et al., 2015). The introduction or overexpression of the betaine transporter from <i>Methanosarcina thermophila</i> TM-1 increases internal ionic balance conferring protection against osmotic stress (Macario and Macario, 2003). Additionally, it has been noted that under high ammonia conditions which inhibits acetotrophic methanogenesis, the addition of magnetite reduces inhibition (Macario and Macario, 2003; Anderson et al., 2012; Horne and Lessner, 2013; Jennings et al., 2014; Jasso-Chávez et al., 2015; Paulo et al., 2015; Wang et al., 2020).</p>
Multiple	<ul style="list-style-type: none"> <li>• <b>Trait stacking for process optimization.</b></li> </ul> <p>By stacking the above traits may be possible to maximize methanogenic efficiency in mixed substrate environments such as the treatment of waste biomass or in process conditions that require multiple extremophilic conditions.</p>

Genetic methods in methanoarchaea are available in several species. Current tools have been recently reviewed or published (Tumbula et al., 1994; Pritchett et al., 2004; Ehlers et al., 2005; Guss et al., 2008; Atomi et al., 2012; Kohler and Metcalf, 2012; Nayak and Metcalf, 2017; Innard, 2018).

of methane produced by methanogens in the wild each year in anaerobic and microanaerobic environments, roughly half escapes into the aerobic carbon cycle (Conrad, 2009). It is estimated between 43–90% of biogenic methane is oxidized by

aerobic methanotrophs at the anaerobic/aerobic interface (Hao et al., 1988; Roslev and King, 1995; Le Mer and Roger, 2001; Conrad, 2009). The remainder of this methane is either trapped within anaerobic environments (as gas or methane gas hydrates)

or consumed by methanotrophic archaea and bacteria (Knittel and Boetius, 2009; Thauer, 2011). Given the estimated 70Gt of CO<sub>2</sub> fixed by photosynthesis into biomass, methane as a product of methanogenesis accounts for approximately 2% of the annual total carbon utilization (Thauer et al., 2008). Previously it was believed that the anaerobic oxidation of methane (AOM) was possible through the symbiotic exchange of metabolites and electrons between the methanotrophic archaea and sulfate reducers (Alperin and Hoehler, 2009; Summers et al., 2010). Within this process anaerobic methane-oxidizing archaea (ANME) consisting of Methanomicrobiales (ANME-1) and Methanosarcinales (ANME-2 and ANME-3) form granular aggregates with delta-proteobacteria in which electrons are transferred between organisms *via* multi-heme cytochromes (McGlynn et al., 2015). Metabolic modeling has suggested that iron and sulfate can be co-substrates in AOM (Riedinger et al., 2014) and 16S rRNA gene-sequences for *Candidatus* Methanoperedens correlated with increased AOM in sulfate-rich anoxic sediments suggesting the occurrence of AOM independent of a bacterial partner (Su et al., 2020). In laboratory conditions it was found that trace amounts of AOM was observed in *Methanothermobacter marburgensis* (Scheller et al., 2010) and *Methanosarcina acetivorans* (Moran et al., 2005) though it was not observed that these strains were able to use methane as the major source of carbon and energy for growth. However, by scouring the metagenomes of unculturable ANME-1 samples from aquatic regions with high amounts of AOM, a novel variant of methyl-coenzyme M reductase (Mcr) was discovered which correlated to AOM without the need for a syntrophic sulfate-reducing partner (Meyerdierks et al., 2010; Shima et al., 2012). When the uncultured ANME-1 Mcr was introduced into *M. acetivorans* it was found that isotope labeled methane was converted into acetate while also facilitating growth, (Soo et al., 2016). Furthermore, methanogen strains containing this ANME-1 Mcr gene can be utilized along with a consortia of microbes including *Geobacter sulfurreducens* to produce electricity in a microbial fuel cell utilizing only methane as a substrate (McAnulty et al., 2017). As every step of methanogenesis is reversible, reverse methanogenesis is theoretically possible for any methanogen though under most conditions these reactions are non-energy yielding (Thauer, 2012). These observations indicate that the bidirectionality of methanogenesis enables methane to be utilized as growth substrate for methanogens, particularly by *Methanosarcina* spp. For example: a *Methanosarcina* culture which has been engineered to produce a high-value terpenoid product is grown using methyl compounds until stationary phase is achieved and biomass accumulation is no longer necessary; this culture could then be induced to produce the terpenoids utilizing potentially any C1 compound or mixtures of compounds including CO, CO<sub>2</sub>, or CH<sub>4</sub> based on substrate availability. This potential extends beyond the production of secreted products, as the biomass of methanogens itself can be utilized as a source of valuable lipids.

## Potential for engineering the lipid membrane biosynthesis pathway as a valorization strategy

Methanogen membranes, like those found in all archaea, are distinct from those found in bacteria and eukarya. In bacterial and eukaryotic organisms lipid membrane structures are composed of fatty acid chains ester linked to glycerol-3-phosphate (G3P) (Koga and Morii, 2007). Archaeal lipids membranes instead utilize isoprenoid alkyl chains ether linked to glycerol-1-phosphate (G1P; Figure 2; Koga and Morii, 2007; Koga, 2012). This fundamental differentiation in membrane composition is the basis of the so called 'lipid divide' separating archaea from the other two domains of life (Villanueva et al., 2021). Given the high quantity and the molecular uniformity of lipid membranes, comprising on average 5% of total methanogen dry weight (Gonnerman et al., 2013), and the relatively high metabolic flux through the archaeal mevalonate lipid biosynthesis pathway, high-value isoprenoid lipids are attractive metabolic engineering targets. The isoprenoid lipids used by archaea allow them to tolerate a wide range of environmental stressors. The most abundant archaeal lipid structures are archeol, consisting of a pair of phytanyl chains ether linked to G1P and caldarcheol, a cyclic dimer of archeol. Caldarcheol is of particular biotechnological interest as the cyclized tetraether lipids maintain cellular homeostasis in the presence of extreme pH and thermal stress (Boyd et al., 2013; Siliakus et al., 2017). Archaeal ether linked lipids are more stable than ester linked membranes when exposed to extremes of pH and thermal conditions, and the unique monolayer structure of tetraether linked lipids imparts resistance to degradation to phospholipases (Jacquemet et al., 2009). These stable properties and the intrinsic monolayer formed by caldarcheol represents an enticing alternative to traditional phospholipids in liposome-based commercial applications. One such application is in the delivery of chemotherapeutic compounds *via* archaeal derived liposomes. It has been found that tetraether linked artificial liposomes reduce leakage of chemotherapeutic compounds by 9-fold compared to conventional eukaryotic derived liposomes, which results in a lower dose required for therapeutic effects (Leriche et al., 2017). The archaeal liposomes themselves also contribute therapeutic effects as archaeal liposomes utilized to transport vaccine components induce robust antigen specific humoral and cellular immune responses exceeding those found from traditional delivery mechanisms (Conlan et al., 2001; Patel and Chen, 2010; Haq et al., 2016; Landi et al., 2017).

In addition to the direct application of archaeal lipids, the high metabolic flux through the archaeal mevalonate pathway presents an opportunity for low-cost production of terpene compounds. Terpenes are the largest class of natural compounds and have a wide range of commercial applications. Odorant terpenes such as limonene, eucalyptol, and linalool are cornerstones of the \$29B flavor and fragrance industry (Markets, 2021). In addition to odorants, terpenes are often the active compound in pharmaceuticals including the anti-cancer drug paclitaxel and the



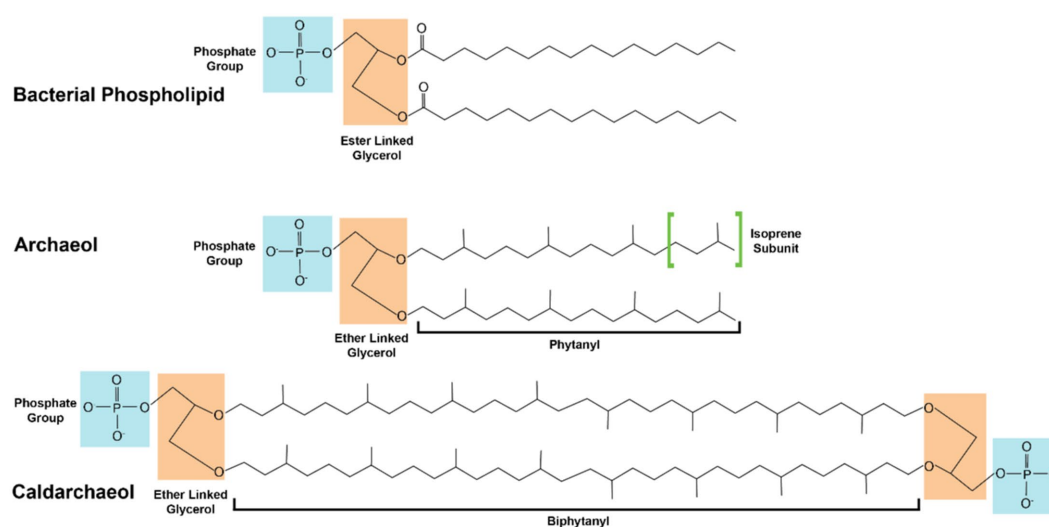


FIGURE 2

Comparison between structures of bacterial and archaeal lipids. Glycerol molecules are shaded in orange. Phosphate groups are shaded in cyan. The isoprenoid subunits which make up the archaeal lipids are highlighted in green brackets. Fully saturated lipids are shown; organisms may produce versions of unsaturated alkane lipids with multiple double bonds or hydroxyl moieties.

antimalarial artemisinin. Hundreds of natural terpenes have shown promising bioactivity (Gould, 1997; Sgadari et al., 2000; Friedman et al., 2002; Paduch et al., 2007; Ajikumar et al., 2008) yet are limited in application due to their availability. Many of these terpenes are currently harvested from their native plant, fungal, and marine producers which are limited by the endogenous expression levels which are prohibitively low (Long et al., 1998; Sills et al., 1998; Newman and Cragg, 2004) or non-renewably synthesized from petroleum precursors. Organically produced terpenes are primarily produced *via* compounds derived from one of two isoprenoid synthesis pathways, the mevalonate (MVA) pathway and the deoxyxyllose 5-phosphate (DXP) pathway (Lange et al., 2000). These pathways in non-archaeal organisms suffer low carbon flux and depletion of precursors towards non-target compounds (McGarvey and Croteau, 1995; Rodriguez-Concepcion and Boronat, 2002; Vranova et al., 2012). Archaea, however, synthesize the majority of membrane lipids through the mevalonate pathway, accounting for a higher flux as compared to eukaryal or bacterial organisms (Boucher et al., 2004; Jain et al., 2014; Villanueva et al., 2014). As such, there is a naturally higher abundance of metabolic precursors available for the synthesis of isoprenoid and terpene products using methanoarchaea. Concerns over the depletion of these membrane precursors have been alleviated by the synthesis of mono-isoprene from engineered strains of *M. acetivorans* and *Methanosarcina barkeri* (Aldridge et al., 2021; Carr et al., 2021). These strains demonstrate that methanogens are able to withstand the metabolic burden of membrane substrate depletion without a significant decrease in growth rate or final carrying capacity, opening the door for further isoprenoid products that could be produced by addition of relatively few genes (Table 2). Inducible promoters such as Ptet could also be used to drive expression of genes for terpenoid

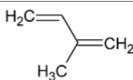
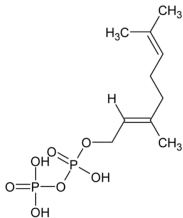
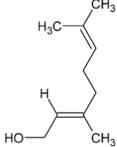
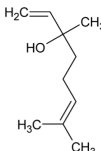
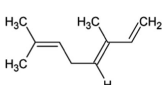
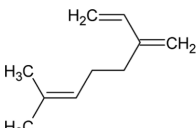
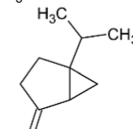

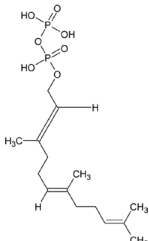
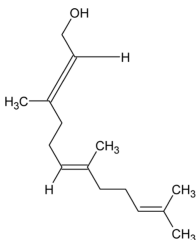
biosynthesis in two-stage fermentation processes to increase bioreactor carrying capacity and maximize terpenoid titer and yield (Urlinger et al., 2000; Loew et al., 2010). One challenge is that some terpenes require molecular oxygen for complete biosynthesis and this might be difficult for anaerobic organisms to achieve. However, *Methanosarcina acetivorans* is remarkably oxygen-tolerant and it is possible to further enhance resistance to oxidative stress through engineering or adaptation (Horne and Lessner, 2013; Jasso-Chávez et al., 2015). Therefore, it is theoretically feasible to use O<sub>2</sub> availability as a biosynthetic inducer during terpene fermentation with oxygen-tolerant methanogens.

## Benefits and challenges of methanogen biotechnologies

The use of methanogens in bioproduction is beneficial in a myriad of ways including ease of selection, low cost of media, and flexibility of products (Table 3). Methanogens have been shown to be an excellent source of metabolically active compounds such as coenzyme M (CoM) which acts as a potent chemotherapy adjuvant as the drug mesna (Shaw and Graham, 1987) as well as immune stimulating lipids for vaccine delivery (Patel and Chen, 2010; Haq et al., 2016). Due to their anaerobic metabolism which requires a lack of O<sub>2</sub>, they are able to produce novel chiral precursors which could later be tailored by chemists through custom oxidation steps and subsequent functionalization.

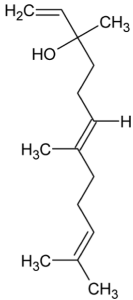
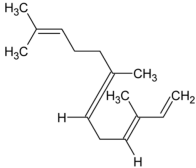
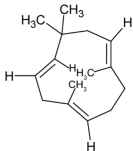
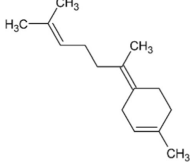
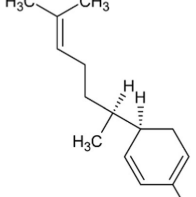
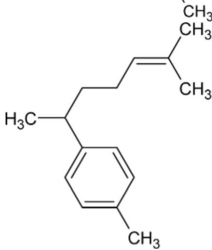
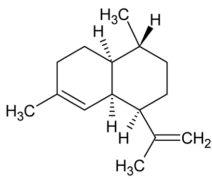
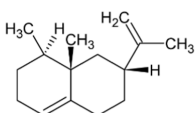
In large scale industrial fermentations pure aseptic environments are difficult to maintain, and often media and growth conditions are utilized to ensure continuous selection during the fermentation (Mosier and Ladisch, 2011; Doran, 2013). Methanogens circumvent this issue by growing in selective

TABLE 2 Potential terpenoids to be produced by methanogens based on category.

Name	Terpene class	Structure	Synthesis Enzyme	Substrate	Initial citation
Isoprene	Hemiterpene		Isoprene synthase (4.2.3.27)	Dimethylallyl pyrophosphate	<a href="#">Silver and Fall (1991)</a>
Geranyl pyrophosphate (GPP)	Monoterpene		(2E,6E)-farnesyl diphosphate synthase (2.5.1.10)	Dimethylallyl pyrophosphate	<a href="#">Cornforth et al. (1966)</a>
Geraniol	Monoterpene		Geraniol synthase (3.1.7.11)	Geranyl diphosphate	<a href="#">Iijima et al. (2004)</a>
Linalool	Monoterpene		S-linalool synthase (4.2.3.25)	Geranyl diphosphate	<a href="#">Pichersky et al. (1994)</a>
Ocimene	Monoterpene		(E)-beta-ocimene synthase (4.2.3.106)	Geranyl diphosphate	<a href="#">Bohlmann et al. (2000)</a>
Myrcene	Monoterpene		Myrcene synthase (4.2.3.15)	Geranyl diphosphate	<a href="#">Bohlmann et al. (1997)</a>
Sabinene	Bicyclic Monoterpenoid		(+)-sabinene synthase (4.2.3.110)	Geranyl diphosphate	<a href="#">Wise et al. (1998)</a>
Pinene	Bicyclic Monoterpenoid		Pinene synthase (4.2.3.14)	Geranyl diphosphate	<a href="#">Gambliel and Croteau (1984)</a>
Farnesyl diphosphate	Acyclic Sesquiterpenoid		Farnesyl diphosphate synthase (2.5.1.1)	Dimethylallyl diphosphate and isopentenyl diphosphate	<a href="#">Vandermoten et al. (2009)</a>
Farnesol	Acyclic Sesquiterpenoid		Farnesyl diphosphatase (3.1.7.6)	Farnesyl diphosphate	<a href="#">Meigs and Simoni (1997)</a>

(Continued)

TABLE 2 (Continued)

Name	Terpene class	Structure	Synthesis Enzyme	Substrate	Initial citation
Nerolidol	Acyclic Sesquiterpenoid		(3S,6E)-nerolidol synthase (4.2.3.48)	Farnesyl diphosphate	<a href="#">Donath and Boland (1995)</a>
Farnesene	Acyclic Sesquiterpenoid		Alpha-farnesene synthase (4.2.3.46) and beta-farnesene synthase (4.2.3.47)	Farnesyl diphosphate	<a href="#">Pechous and Whitaker (2004)</a>
Humulene	Monocyclic Sesquiterpenoid		Alpha-humulene synthase (4.2.3.104)	Farnesyl diphosphate	<a href="#">van Der Hoeven et al. (2000)</a>
Bisabolene	Monocyclic Sesquiterpenoid		Alpha-bisbolene synthase (4.2.3.38)	Farnesyl diphosphate	<a href="#">Bohlmann et al. (1998)</a>
Zingiberene	Monocyclic Sesquiterpenoid		Zingiberene synthase (4.2.3.65)	Farnesyl diphosphate	<a href="#">Zhuang et al. (2012)</a>
Curcumene	Monocyclic Sesquiterpenoid		Gamma-curcumene synthase (4.2.3.94)	Farnesyl diphosphate	<a href="#">Deguerry et al. (2006)</a>
Amorphadiene	Bicyclic Sesquiterpenoid		Amorpha-4,11-diene synthase (4.2.3.24)	Farnesyl diphosphate	<a href="#">Bouwmeester et al. (1999)</a>
Valencene	Bicyclic Sesquiterpenoid		Valencene synthase (4.2.3.73)	Farnesyl diphosphate	<a href="#">Sharon-Asa et al. (2003)</a>

Terpenoid classes requiring molecular oxygen for biosynthesis have been omitted. However, oxygen-tolerant *Methanosarcina* have potential to be used in this context.

environments free of oxygen using substrates that cannot be used by the majority of common contaminating factors such as lactic

acid bacteria and fungi ([Skinner and Leathers, 2004](#); [Beckner et al., 2011](#)). Methanogens are prototrophic organisms, able to

TABLE 3 Benefits and challenges of methanogen biotechnology.

Benefits	Challenges
Methanogens are some of the fastest-replicating organisms, particularly members of <i>Methanococcus</i> (Jones et al., 1983; Goyal et al., 2016; Long et al., 2017) and <i>Methanopyrus</i> (Takai et al., 2008) genus. (Jones et al., 1983; Takai et al., 2008; Goyal et al., 2016; Long et al., 2017).	Strain differences in growth rate and carrying capacity. Growth is flux-controlled depending on substrate feed rates. Gas-phase fermentation presents similar problems as oxygenation in traditional fermentations (Mosier and Ladisch, 2011; Chen, 2012; Luo and Angelidaki, 2012).
Methanogens can grow on inexpensive substrates including negative value substrates such as wastewater (Daniels et al., 1977; Schiraldi et al., 2002; McGenity, 2010; Ferry, 2012; Costa and Leigh, 2014; Borrel et al., 2016; Buan, 2018; Chadwick et al., 2022).	Process disfavors growth of aerobic pathogens. Co-product can be water ready for discharge to aquifers and waterways.
Methanogens already scaled up worldwide for water treatment and biogas production.	
Can be coupled directly or indirectly to electrodes for carbon capture by electrosynthesis or for electricity generation from biomass (Ragab et al., 2020).	Surface-to-area, substrate solubility, and other challenges commensurate with microbial fuel cell technologies.
Oxygenation not required. Can grow on non-gas substrates. No contamination by aerobic organisms.	Methanogens require specialized culture environments to maintain anaerobicity (Balch et al., 1979; Rouviere and Wolfe, 1988; Buan, 2018).
Mesophilic and thermophilic strains available to tailor to the desired product and process needs.	Methanogen chassis organisms may need different optimization strategies.
Novel metabolic pathways are constantly being discovered (Costa and Leigh, 2014; Guan et al., 2015; Borrel et al., 2016; Mayumi et al., 2016; Buan, 2018; Yan and Ferry, 2018; Chadwick et al., 2022; Zhou et al., 2022).	Methanogen genetics and biochemistry are less characterized than other model organisms.
Synthetic biology pathways often use archaeal or methanogen genes to improve yields and reduce feedback inhibition.	
Bacterial synthetic biology and genetic strategies have been successfully translated to methanogens.	
Methanogens have a high substrate to volume ratio with low accumulation of biomass relative to products (Thauer et al., 2008; Ferry, 2012; Buan, 2018).	High titers of intracellular products may be difficult to obtain unless accumulated into vacuoles or secreted extracellularly.
Multiple validated genetic tools available including tools for <i>Methanosarcina</i> spp., (Metcalf et al., 1997; Buan et al., 2011; Nayak and Metcalf, 2017) <i>Methanococcus maripaludis</i> , (Blank et al., 1995; Bao and Scheller, 2021) and <i>Methanothermobacter thermautotrophicus</i> (Buan et al., 2011; Sarmiento et al., 2011; Nayak and Metcalf, 2017; Bao and Scheller, 2021; Fink et al., 2021).	Variability in genome copy number can present challenges when performing chromosomal modifications (Hildenbrand et al., 2011; Aldridge et al., 2021).
The lack of cell wall and envelope in most methanogens ensures that products generated through methanogen fermentations are not contaminated with peptidoglycan or endotoxin (Jones et al., 1987; Claus and König, 2010).	Some methanogen species produce pseudomurein cell walls or extracellular polysaccharide capsules, although these are generally non-or weakly immunogenic (Sirohi et al., 2010; Subedi et al., 2021).
Methanoarchaea are non-pathogenic, though there have been studies suggesting a link between methanogens and other microbes in dysbiotic anaerobic abscesses (Drancourt et al., 2017; Sogodogo et al., 2019).	Not currently recognized as a GRAS (Generally Regarded as Safe) organism.

synthesize all vitamins and cofactors required for growth from inorganic material, allowing for additional selection by limiting available vitamins and nutrients required for contaminating growth by exclusion (Patil and Muskan, 2009; Thauer, 2012; Buan, 2018). While viral predation on methanogens has been observed (Park et al., 2007) there is little evidence that these methanophage/methanovirus particles have a substantial effect on methanogenic digester performance as viral titers did not correlate with a significant decrease in methane output and methanogen carrying capacity.

Another major challenge in industrial fermentations is the large amounts of fresh water required for *E. coli* or yeast (Chen, 2012). Methanogens, however, thrive in environments with high salt concentrations, allowing for the utilization of seawater in fermentations. Non-sterile hypersaline environments such as ocean water and hydraulic fracking fluids have been

demonstrated to select for methylotrophic methanogens such as *Methanohalophilus*, *Methanohalobium*, and *Methanosarcina* spp. while also presenting a high concentration of non-competitive substrates such as methylamines (McGenity, 2010; Guan et al., 2015). Methanogens are utilized worldwide for the production of renewable biogas in non-selective environments with high degrees of contamination such as municipal and agricultural wastewater treatment. In these environments methanogens are exposed to a wide variety of stressors including dramatic shifts in ammonia, osmotic shifts, and exposure to heavy metals (Yan et al., 2020). Many methanogens are natively capable of withstanding these stressors (Macario and Macario, 2003) though as stated above, using genetic tools it is possible to combine or “stack” desirable traits onto a single methanogen strain to gain the maximum benefit from a single organism.



## Conclusion

Methanogens are biologically important organisms with a wide-reaching impact both in ecological and biotechnological applications. Their extremely efficient central metabolism makes them an ideal source of renewable biofuels that can be captured through anaerobic digestion or fermentation processes. They are able to grow prototrophically with inexpensive feedstocks and can produce endotoxin-free protein, carbohydrates, and valuable isoprenoid lipids. Their unique membrane composition can be used to expand the biotechnological toolbox for the delivery of chemotherapeutics as well as source for novel terpene compounds previously not available *via* conventional extraction means. By continuing to investigate the molecular, genetic, and synthetic biology potential of these unique organisms, researchers may unlock a wide range of applications from environmental and ecological management, renewable energy, agriculture, chemical manufacturing, and pharmaceutical industries.

## Author contributions

SC conceived and wrote the manuscript. NRB conceived, wrote, and edited the manuscript. All authors contributed to the article and approved the submitted version.

## References

- Ajikumar, P. K., Tyo, K., Carlsen, S., Mucha, O., Phon, T. H., and Stephanopoulos, G. (2008). Terpenoids: opportunities for biosynthesis of natural product drugs using engineered microorganisms. *Mol. Pharm.* 5, 167–190. doi: 10.1021/mp700151b
- Aldridge, J., Carr, S., Weber, K. A., and Buan, N. R. (2021). Anaerobic production of isoprene by engineered *Methanosarcina* species archaea. *Appl. Environ. Microbiol.* 87, e02417–e02420. doi: 10.1128/AEM.02417-20
- Alperin, M. J., and Hoehler, T. M. (2009). Anaerobic methane oxidation by archaea/sulfate-reducing bacteria aggregates: 1. Thermodynamic and physical constraints. *Am. J. Sci.* 309, 869–957. doi: 10.2475/10.2009.02
- Anderson, K. L., Apolinario, E. E., and Sowers, K. R. (2012). Desiccation as a long-term survival mechanism for the archaeon *Methanosarcina barkeri*. *Appl. Environ. Microbiol.* 78, 1473–1479. doi: 10.1128/AEM.06964-11
- Atomi, H., Imanaka, T., and Fukui, T. (2012). Overview of the genetic tools in the archaea. *Front. Microbiol.* 3:337. doi: 10.3389/fmicb.2012.00337
- Balch, W., Fox, G., Magrum, L., Woese, C., and Wolfe, R. (1979). Methanogens: reevaluation of a unique biological group. *Microbiol. Rev.* 43:260. PMID: 281474
- Bao, J., and Scheller, S. (2021). Efficient CRISPR/Cas12a-based genome editing toolbox for metabolic engineering in *Methanococcus maripaludis*. bioRxiv. doi: 10.1101/2021.12.29.474413
- Beckner, M., Ivey, M. L., and Phister, T. G. (2011). Microbial contamination of fuel ethanol fermentations. *Lett. Appl. Microbiol.* 53, 387–394. doi: 10.1111/j.1472-765X.2011.03124.x
- Blank, C. E., Kessler, P. S., and Leigh, J. A. (1995). Genetics in methanogens: transposon insertion mutagenesis of a *Methanococcus maripaludis* *nifH* gene. *J. Bacteriol.* 177, 5773–5777.
- Bohlmann, J., Crock, J., Jetter, R., and Croteau, R. (1998). Terpeneoid-based defenses in conifers: cDNA cloning, characterization, and functional expression of wound-inducible (E)-alpha-bisabolene synthase from grand fir (*Abies grandis*). *Proc. Natl. Acad. Sci. U. S. A.* 95, 6756–6761. doi: 10.1073/pnas.95.12.6756
- Bohlmann, J., Martin, D., Oldham, N. J., and Gershenzon, J. (2000). Terpeneoid secondary metabolism in *Arabidopsis thaliana*: cDNA cloning, characterization, and functional expression of a myrcene/(E)-beta-ocimene synthase. *Arch. Biochem. Biophys.* 375, 261–269. doi: 10.1006/abbi.1999.1669
- Bohlmann, J., Steele, C. L., and Croteau, R. (1997). Monoterpene synthases from grand fir (*Abies grandis*). cDNA isolation, characterization, and functional expression of myrcene synthase, (-)-(4S)-limonene synthase, and (-)-(1S,5S)-pinene synthase. *J. Biol. Chem.* 272, 21784–21792. doi: 10.1074/jbc.272.35.21784
- Borrel, G., Adam, P. S., and Gribaldo, S. (2016). Methanogenesis and the Wood–Ljungdahl pathway: an ancient, versatile, and fragile association. *Genome Biol. Evol.* 8, 1706–1711. doi: 10.1093/gbe/evw114
- Borrel, G., Adam, P. S., McKay, L. J., Chen, L.-X., Sierra-García, I. N., Sieber, C. M., et al. (2019). Wide diversity of methane and short-chain alkane metabolisms in uncultured archaea. *Nat. Microbiol.* 4, 603–613. doi: 10.1038/s41564-019-0363-3
- Boucher, Y., Kamekura, M., and Doolittle, W. F. (2004). Origins and evolution of isoprenoid lipid biosynthesis in archaea. *Mol. Microbiol.* 52, 515–527. doi: 10.1111/j.1365-2958.2004.03992.x
- Bouwmeester, H. J., Wallaart, T. E., Janssen, M. H., van Loo, B., Jansen, B. J., Posthumus, M. A., et al. (1999). Amorpha-4,11-diene synthase catalyses the first probable step in artemisinin biosynthesis. *Phytochemistry* 52, 843–854. doi: 10.1016/S0031-9422(99)00206-x
- Boyd, E., Hamilton, T., Wang, J., He, L., and Zhang, C. (2013). The role of tetraether lipid composition in the adaptation of the thermophilic archaea to acidity. *Front. Microbiol.* 4:62. doi: 10.3389/fmicb.2013.00062
- Brune, A. (1998). Termite guts: the world's smallest bioreactors. *Trends Biotechnol.* 16, 16–21. doi: 10.1016/S0167-7799(97)00151-7
- Brune, A. (2018). In *(Endo) symbiotic Methanogenic Archaea 81–101* Berlin: Springer.
- Buan, N. R. (2018). Methanogens: pushing the boundaries of biology. *Emerg. Topics Life Sci.* 2, 629–646. doi: 10.1042/ETLS20180031

## Funding

This work was supported by grants from the National Science Foundation (IOS-1938948) and the Nebraska Center for Energy Sciences Research (Cycle 15). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the funding agencies.

## Conflict of interest

NRB has disclosed a significant financial interest in RollingCircle Biotech, LLC and Molecular Trait Evolution, LLC.

SC declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Buan, N., Kulkarni, G., and Metcalf, W. (2011). Genetic methods for *Methanosarcina* species. in *Methods in Enzymology* (Vol. 494, pp. 23–42). Academic Press.
- Buan, N. R., and Metcalf, W. W. (2010). Methanogenesis by *Methanosarcina acetivorans* involves two structurally and functionally distinct classes of heterodisulfide reductase. *Mol. Microbiol.* 75, 843–853. PMID: 19968794. doi: 10.1111/j.1365-2958.2009.06990.x
- Buckel, W., and Thauer, R. K. (2018). Flavin-based electron bifurcation, ferredoxin, Flavodoxin, and anaerobic respiration with protons (Ech) or NAD(+) (Rnf) as electron acceptors: a historical review. *Front. Microbiol.* 9. doi: 10.3389/fmicb.2018.00401
- Carr, S., Aldridge, J., and Buan, N. R. (2021). Isoprene production from municipal wastewater biosolids by engineered archaeon *Methanosarcina acetivorans*. *Appl. Sci.* 11:3342. doi: 10.3390/app11083342
- Catlett, J. L., Carr, S., Cashman, M., Smith, M. D., Walter, M., Sakka, Z., et al. (2022). R. Metabolic synergy between human symbionts *Bacteroides* and *Methanobrevibacter*. *Microbiol. Spectr.* 10:e0106722. doi: 10.1128/spectrum.01067-22 Epub 2022 May 10. PMID: 35536023; PMID: PMC9241691.
- Catlett, J. L., Ortiz, A. M., and Buan, N. R. (2015). Rerouting cellular electron flux to increase the rate of biological methane production. *Appl. Environ. Microbiol.* 81, 6528–6537. doi: 10.1128/AEM.01162-15
- Chadwick, G. L., Skennerton, C. T., Laso-Pérez, R., Leu, A. O., Speth, D. R., Yu, H., et al. (2022). Comparative genomics reveals electron transfer and syntrophic mechanisms differentiating methanotrophic and methanogenic archaea. *PLoS Biol.* 20:e3001508. PMID: 34986141; PMID: PMC9012536. doi: 10.1371/journal.pbio.3001508
- Chen, G.-Q. (2012). New challenges and opportunities for industrial biotechnology. *Microb. Cell Factories* 11, 1–3. PMID: 22905695; PMID: PMC3520750. doi: 10.1186/1475-2859-11-111
- Claus, H., and König, H. (2010). *In Prokaryotic Cell Wall Compounds* Berlin: Springer.
- Conlan, J. W., Krishnan, L., Willick, G. E., Patel, G. B., and Sprott, G. D. (2001). Immunization of mice with lipopeptide antigens encapsulated in novel liposomes prepared from the polar lipids of various Archaeobacteria elicits rapid and prolonged specific protective immunity against infection with the facultative intracellular pathogen *Listeria monocytogenes*. *Vaccine* 19, 3509–3517. PMID: 11348718. doi: 10.1016/s0264-410x(01)00041-x
- Conrad, R. (2009). The global methane cycle: recent advances in understanding the microbial processes involved. *Environ. Microbiol. Rep.* 1, 285–292. doi: 10.1111/j.1758-2229.2009.00038.x
- Cornforth, J., Cornforth, R. H., Popjak, G., and Yengoyan, L. (1966). Studies on the biosynthesis of cholesterol: XX. Steric course of decarboxylation of 5-pyrophosphomevalonate and of the carbon to carbon bond formation in the biosynthesis of farnesyl pyrophosphate. *J. Biol. Chem.* 241, 3970–3987. doi: 10.1016/S0021-9258(18)99800-5
- Costa, K. C., and Leigh, J. A. (2014). Metabolic versatility in methanogens. *Curr. Opin. Biotechnol.* 29, 70–75. PMID: 24662145. doi: 10.1016/j.copbio.2014.02.012
- Costa, K. C., Wong, P. M., Wang, T., Lie, T. J., Dodsworth, J. A., Swanson, I., et al. (2010). Protein complexing in a methanogen suggests electron bifurcation and electron delivery from formate to heterodisulfide reductase. *Proc. Natl. Acad. Sci. U. S. A.* 107, 11050–11055. PMID: 20534465; PMID: PMC2890747. doi: 10.1073/pnas.1003653107
- Daniels, L., Fuchs, G., Thauer, R., and Zeikus, J. (1977). Carbon monoxide oxidation by methanogenic bacteria. *J. Bacteriol.* 132, 118–126. PMID: 21159; PMID: PMC221834. doi: 10.1128/jb.132.1.118-126.1977
- Deguerry, F., Pastore, L., Wu, S., Clark, A., Chappell, J., and Schalk, M. (2006). The diverse sesquiterpene profile of patchouli, *Pogostemon cablin*, is correlated with a limited number of sesquiterpene synthases. *Arch. Biochem. Biophys.* 454, 123–136. doi: 10.1016/j.abb.2006.08.006
- Diender, M., Stams, A. J., and Sousa, D. Z. (2015). Pathways and bioenergetics of anaerobic carbon monoxide fermentation. *Front. Microbiol.* 6:1275. doi: 10.3389/fmicb.2015.01275
- Dolfing, J., Larter, S. R., and Head, I. M. (2008). Thermodynamic constraints on methanogenic crude oil biodegradation. *ISME J.* 2, 442–452. doi: 10.1038/ismej.2007.111
- Donath, J., and Boland, W. (1995). Biosynthesis of acyclic homoterpenes: enzyme selectivity and absolute configuration of the nerolidol precursor. *Phytochemistry* 39, 785–790. doi: 10.1016/0031-9422(95)00082-1
- Doran, P. M. ed. (2013). *Biotechnology of hairy root systems*. Springer.
- Drancourt, M., Nkanga, V. D., Lakhe, N. A., Régis, J.-M., Dufour, H., Fournier, P.-E., et al. (2017). Evidence of archaeal methanogens in brain abscess. *Clin. Infect. Dis.* 65, 1–5. PMID: 28379309. doi: 10.1093/cid/cix286
- Duszenko, N., and Buan, N. R. (2017). Physiological evidence for isopotential tunneling in the electron transport chain of methane-producing archaea. *Appl. Environ. Microbiol.* 83, e00950–e00917. PMID: 28710268; PMID: PMC5583484. doi: 10.1128/AEM.00950-17
- Ehlers, C., Weidenbach, K., Veit, K., Deppenmeier, U., Metcalf, W. W., and Schmitz, R. A. (2005). Development of genetic methods and construction of a chromosomal *glnK1* mutant in *Methanosarcina mazei* strain Gö1. *Mol. Gen. Genomics.* 273, 290–298. doi: 10.1007/s00438-005-1128-7
- Ermiler, U., Grabarse, W., Shima, S., Goubeaud, M., and Thauer, R. K. (1997). Crystal structure of methyl-coenzyme M reductase: the key enzyme of biological methane formation. *Science* 278, 1457–1462. PMID: 9367957. doi: 10.1126/science.278.5342.1457
- Feist, A. M., Scholten, J. C., Palsson, B. Ø., Brockman, F. J., and Ideker, T. (2006). Modeling methanogenesis with a genome-scale metabolic reconstruction of *Methanosarcina barkeri*. *Mol. Syst. Biol.* 2:2006.0004. PMID: 16738551; PMID: PMC1681478. doi: 10.1038/msb4100046
- Ferry, J. G. (2012). *Methanogenesis: Ecology, Physiology, Biochemistry & Genetics*. Berlin Springer Science & Business Media. ISBN: 9781461523918.
- Fink, C., Beblawy, S., Enkerlin, A. M., Mühling, L., Angenent, L. T., and Molitor, B. (2021). A shuttle-vector system allows heterologous gene expression in the thermophilic methanogen *Methanothermobacter thermautotrophicus* ΔH. *MBio* 12, e02766–e02721. PMID: 34809461; PMID: PMC8609365. doi: 10.1128/mBio.02766-21
- Fricke, W. F., Sedorf, H., Henne, A., Krüger, M., Liesegang, H., Hedderich, R., et al. (2006). The genome sequence of *Methanosphaera stadtmanae* reveals why this human intestinal archaeon is restricted to methanol and H<sub>2</sub> for methane formation and ATP synthesis. *J. Bacteriol.* 188, 642–658. PMID: 16385054; PMID: PMC1347301. doi: 10.1128/JB.188.2.642-658.2006
- Friedman, M., Henika, P. R., and Mandrell, R. E. (2002). Bactericidal activities of plant essential oils and some of their isolated constituents against *Campylobacter jejuni*, *Escherichia coli*, *Listeria monocytogenes*, and *Salmonella enterica*. *J. Food Protect.* 65, 1545–1560. PMID: 12380738. doi: 10.4315/0362-028x-65.10.1545
- Fu, L., Zhou, T., Wang, J., You, L., Lu, Y., Yu, L., et al. (2019). NanoFe<sub>3</sub>O<sub>4</sub> as solid electron shuttles to accelerate acetotrophic methanogenesis by *Methanosarcina barkeri*. *Front. Microbiol.* 10:388. doi: 10.3389/fmicb.2019.00388
- Gambliel, H., and Croteau, R. (1984). Pinene cyclases I and II. Two enzymes from sage (*Salvia officinalis*) which catalyze stereospecific cyclizations of geranyl pyrophosphate to monoterpene olefins of opposite configuration. *J. Biol. Chem.* 259, 740–748. PMID: 6693393.
- Gonnerman, M. C., Benedict, M. N., Feist, A. M., Metcalf, W. W., and Price, N. D. (2013). Genomically and biochemically accurate metabolic reconstruction of *Methanosarcina barkeri* Fusaro. *iMG746. Biotechnol. J.* 8, 1070–1079. doi: 10.1002/biot.201200266
- Gould, M. N. (1997). Cancer chemoprevention and therapy by monoterpenes. *Environ. Health Persp.* 105, 977–979. doi: 10.2307/3433313
- Goyal, N., Zhou, Z., and Karimi, I. A. (2016). Metabolic processes of *Methanococcus maripaludis* and potential applications. *Microb. Cell Factories* 15, 1–19. doi: 10.1186/s12934-016-0500-0
- Guan, Y., Hikmawan, T., Antunes, A., Ngugi, D., and Stingl, U. (2015). Diversity of methanogens and sulfate-reducing bacteria in the interfaces of five deep-sea anoxic brines of the Red Sea. *Res. Microbiol.* 166, 688–699. doi: 10.1016/j.resmic.2015.07.002
- Guss, A. M., Kulkarni, G., and Metcalf, W. W. (2009). Differences in hydrogenase gene expression between *Methanosarcina acetivorans* and *Methanosarcina barkeri*. *J. Bacteriol.* 191, 2826–2833. doi: 10.1128/JB.00563-08
- Guss, A. M., Rother, M., Zhang, J. K., Kulkarni, G., and Metcalf, W. W. (2008). New methods for tightly regulated gene expression and highly efficient chromosomal integration of cloned genes for *Methanosarcina* species. *Archaea* 2, 193–203. PMID: 19054746; PMID: PMC2685592. doi: 10.1155/2008/534081
- Hao, W., Scharffe, D., Crutzen, P., and Sanhueza, E. (1988). Production of N<sub>2</sub>O, CH<sub>4</sub>, and CO<sub>2</sub> from soils in the tropical savanna during the dry season. *J. Atmos. Chem.* 7, 93–105. doi: 10.1007/BF00048256
- Haq, K., Jia, Y., and Krishnan, L. (2016). Archaeal lipid vaccine adjuvants for induction of cell-mediated immunity. *Expert Rev. Vaccines* 15, 1557–1566. doi: 10.1080/14760584.2016.1195265
- Hildenbrand, C., Stock, T., Lange, C., Rother, M., and Soppa, J. (2011). Genome copy numbers and gene conversion in methanogenic archaea. *J. Bacteriol.* 193, 734–743. doi: 10.1128/JB.01016-10
- Holmes, D. E., Giloteaux, L., Orellana, R., Williams, K. H., Robbins, M. J., and Lovley, D. R. (2014). Methane production from protozoan endosymbionts following stimulation of microbial metabolism within subsurface sediments. *Front. Microbiol.* 5:366. doi: 10.3389/fmicb.2014.00366
- Horne, A. J., and Lessner, D. J. (2013). Assessment of the oxidant tolerance of *Methanosarcina acetivorans*. *FEMS Microbiol. Lett.* 343, 13–19. doi: 10.1111/1574-6968.12115
- Huang, Z., Sednek, C., Urynowicz, M. A., Guo, H., Wang, Q., Fallgren, P., et al. (2017). Low carbon renewable natural gas production from coalbeds and implications for carbon capture and storage. *Nat. Commun.* 8, 1–11. doi: 10.1038/s41467-017-00611-7

- Iijima, Y., Gang, D. R., Fridman, E., Lewinsohn, E., and Pichersky, E. (2004). Characterization of geraniol synthase from the peltate glands of sweet basil. *Plant Physiol.* 134, 370–379. PMID: 14657409; PMID: PMC316316. doi: 10.1104/pp.103.032946
- Innard, N. (2018). *Developing Genetic Tools in Methanococcus maripaludis*. Heslington: University of York. Available at: <https://etheses.whiterose.ac.uk/20486/>
- Jacquemet, A., Barbeau, J., Lemièvre, L., and Benvenne, T. (2009). Archaeal tetraether bipolar lipids: structures, functions and applications. *Biochimie* 91, 711–717. doi: 10.1016/j.biochi.2009.01.006
- Jain, S., Caforio, A., and Driessen, A. J. (2014). Biosynthesis of archaeal membrane ether lipids. *Front. Microbiol.* 5:641. doi: 10.3389/fmicb.2014.00641
- Jasso-Chávez, R., Santiago-Martínez, M. G., Lira-Silva, E., Pineda, E., Zepeda-Rodríguez, A., et al. (2015). Air-adapted *Methanosarcina acetivorans* shows high methane production and develops resistance against oxygen stress. *PLoS One* 10:e0117331. PMID: 25706146; PMID: PMC4338226. doi: 10.1371/journal.pone.0117331
- Jennings, M. E., Schaff, C. W., Horne, A. J., Lessner, F. H., and Lessner, D. J. (2014). Expression of a bacterial catalase in a strictly anaerobic methanogen significantly increases tolerance to hydrogen peroxide but not oxygen. *Microbiology* 160, 270–278. doi: 10.1099/mic.0.070763-0
- Jin, D., Kang, K., Wang, H., Wang, Z., Xue, B., Wang, L., et al. (2017). Effects of dietary supplementation of active dried yeast on fecal methanogenic archaea diversity in dairy cows. *Anaerobe* 44, 78–86. PMID: 28188879. doi: 10.1016/j.anaerobe.2017.02.007
- Johnson, K. A., and Johnson, D. E. (1995). Methane emissions from cattle. *J. Anim. Sci.* 73, 2483–2492. doi: 10.2527/1995.7382483x
- Jones, D., Head, I., Gray, N., Adams, J., Rowan, A., Aitken, C., et al. (2008). Crude-oil biodegradation via methanogenesis in subsurface petroleum reservoirs. *Nature* 451, 176–180. doi: 10.1038/nature06484
- Jones, W. J., Nagle, D., and Whitman, W. B. (1987). Methanogens and the diversity of archaeobacteria. *Microbiol. Rev.* 51, 135–177. PMID: 3104748; PMID: PMC373095. doi: 10.1128/mr.51.1.135-177.1987
- Jones, W. J., Paynter, M., and Gupta, R. (1983). Characterization of *Methanococcus maripaludis* sp. nov., a new methanogen isolated from salt marsh sediment. *Arch. Microbiol.* 135, 91–97. doi: 10.1007/BF00425213
- Knittel, K., and Boetius, A. (2009). Anaerobic oxidation of methane: progress with an unknown process. *Annu. Rev. Microbiol.* 63, 311–334. PMID: 19575572. doi: 10.1146/annurev.micro.61.080706.093130
- Koga, Y. (2012). Thermal adaptation of the archaeal and bacterial lipid membranes. *Archaea* 2012, 1–6. doi: 10.1155/2012/789652
- Koga, Y., and Morii, H. (2007). Biosynthesis of ether-type polar lipids in archaea and evolutionary considerations. *Microbiol. Mol. Biol. Rev.* 71, 97–120. PMID: 17347520; PMID: PMC1847378. doi: 10.1128/MMBR.00033-06
- Kohler, P. R., and Metcalf, W. W. (2012). Genetic manipulation of *Methanosarcina* spp. *Front. Microbiol.* 3:259. PMID: 22837755; PMID: PMC3403347. doi: 10.3389/fmicb.2012.00259
- Kulkarni, G., Kridelbaugh, D. M., Guss, A. M., and Metcalf, W. W. (2009). Hydrogen is a preferred intermediate in the energy-conserving electron transport chain of *Methanosarcina barkeri*. *Proc. Natl. Acad. Sci. U. S. A.* 106, 15915–15920. doi: 10.1073/pnas.0905914106
- Kurth, J. M., Op den Camp, H. J., and Welte, C. U. (2020). Several ways one goal—methanogenesis from unconventional substrates. *Appl. Microbiol. Biotechnol.* 104, 6839–6854. PMID: 32542472; PMID: PMC7374477. doi: 10.1007/s00253-020-10724-7
- Landi, A., Law, J., Hockman, D., Logan, M., Crawford, K., Chen, C., et al. (2017). Superior immunogenicity of HCV envelope glycoproteins when adjuvanted with cyclic-di-AMP, a STING activator or archaeosomes. *Vaccine* 35, 6949–6956. doi: 10.1016/j.vaccine.2017.10.072
- Lange, B. M., Rujan, T., Martin, W., and Croteau, R. (2000). Isoprenoid biosynthesis: the evolution of two ancient and distinct pathways across genomes. *Proc. Natl. Acad. Sci. U. S. A.* 97, 13172–13177. PMID: 11078528; PMID: PMC27197. doi: 10.1073/pnas.240454797
- Laso-Pérez, R., Hahn, C., van Vliet, D. M., Tegetmeyer, H. E., Schubotz, F., Smit, N. T., et al. (2019). Anaerobic degradation of non-methane alkanes by “*Candidatus Methanoliparia*” in hydrocarbon seeps of the Gulf of Mexico. *MBio* 10, e01814–e01819. PMID: 31431553; PMID: PMC6703427. doi: 10.1128/mBio.01814-19
- Le Mer, J., and Roger, P. (2001). Production, oxidation, emission and consumption of methane by soils: a review. *Eur. J. Soil Biol.* 37, 25–50. doi: 10.1016/S1164-5563(01)01067-6
- Lemaire, O. N., and Wagner, T. (2022). A structural view of alkyl-coenzyme M reductases, the first step of alkane anaerobic oxidation catalyzed by archaea. *Biochemistry* 61, 805–821. doi: 10.1021/acs.biochem.2c00135
- Lerliche, G., Cifelli, J. L., Sibucão, K. C., Patterson, J. P., Koyanagi, T., Gianneschi, N. C., et al. (2017). Characterization of drug encapsulation and retention in archaea-inspired tetraether liposomes. *Org. Biomol. Chem.* 15, 2157–2162. doi: 10.1039/C6OB02832B
- Lessner, D. J., Lhu, L., Wahal, C. S., and Ferry, J. G. (2010). An engineered methanogenic pathway derived from the domains bacteria and archaea. *MBio* 1, e00243–e00210. PMID: 21060738; PMID: PMC2975365. doi: 10.1128/mBio.00243-10
- Loew, R., Heinz, N., Hampf, M., Bujard, H., and Gossen, M. (2010). Improved Tet-responsive promoters with minimized background expression. *BMC Biotechnol.* 10, 1–13. PMID: 21106052; PMID: PMC3002914. doi: 10.1186/1472-6750-10-81
- Long, B. H., Carboni, J. M., Wasserman, A. J., Cornell, L. A., Casazza, A. M., Jensen, P. R., et al. (1998). Eleutherobin, a novel cytotoxic agent that induces tubulin polymerization, is similar to paclitaxel (Taxol (R)). *Cancer Res.* 58, 1111–1115. PMID: 9515790.
- Long, F., Wang, L., Lupa, B., and Whitman, W. B. (2017). A flexible system for cultivation of *Methanococcus* and other formate-utilizing methanogens. *Archaea* 2017, 1–12. doi: 10.1155/2017/7046026
- López Muñoz, M. M., Schönheit, P., and Metcalf, W. W. (2015). Genetic, genomic, and transcriptomic studies of pyruvate metabolism in *Methanosarcina barkeri* fusaro. *J. Bacteriol.* 197, 3592–3600. doi: 10.1128/JB.00551-15
- Luo, G., and Angelidaki, I. (2012). Integrated biogas upgrading and hydrogen utilization in an anaerobic reactor containing enriched hydrogenotrophic methanogenic culture. *Biotechnol. Bioeng.* 109, 2729–2736. doi: 10.1002/bit.24557
- Macario, E. C. D., and Macario, A. J. (2003). Molecular biology of stress genes in methanogens: potential for bioreactor technology. *Biomethanation* 1, 95–150. PMID: 12747562. doi: 10.1007/3-540-45839-5\_4
- Markets, M. A. (2021). Flavors & Fragrances Market by ingredients (natural, synthetic), end use (beverage, Savory & Snacks, bakery, dairy products, confectionery, consumer products, fine fragrances), and region (Asia Pacific, North America, Europe) – global forecast to 2026. Available at: <https://www.marketsandmarkets.com/Market-Reports/flavors-fragrance-market-175163912.html>
- Martin, W. F., and Sousa, F. L. (2016). Early microbial evolution: the age of anaerobes. *Cold Spring Harb. Perspect. Biol.* 8:a018127. doi: 10.1101/cshperspect.a018127
- Matschiavelli, N., Oelgeschläger, E., Cocchiariar, B., Finke, J., and Rother, M. (2012). Function and regulation of isoforms of carbon monoxide dehydrogenase/acetyl coenzyme a synthase in *Methanosarcina acetivorans*. *J. Bacteriol.* 194, 5377–5387. doi: 10.1128/JB.00881-12
- Mayumi, D., Mochimaru, H., Tamaki, H., Yamamoto, K., Yoshioka, H., Suzuki, Y., et al. (2016). Methane production from coal by a single methanogen. *Science* 354, 222–225. doi: 10.1126/science.aaf8821
- McAnulty, M. J., Poosarla, V., Kim, K.-Y., Jasso-Chávez, R., Logan, B. E., and Wood, T. K. (2017). Electricity from methane by reversing methanogenesis. *Nat. Commun.* 8, 1–8. doi: 10.1038/ncomms15419
- McGarvey, D. J., and Croteau, R. (1995). Terpenoid metabolism. *Plant Cell* 7, 1015–1026. doi: 10.1105/tpc.7.7.1015
- McGenity, T. J. (2010). “Methanogens and methanogenesis in hypersaline environments” in *Handbook of Hydrocarbon and Lipid Microbiology*. ed. K. N. Timmis Berlin, Heidelberg: Springer, 665–680.
- McGlynn, S. E., Chadwick, G. L., Kempes, C. P., and Orphan, V. J. (2015). Single cell activity reveals direct electron transfer in methanotrophic consortia. *Nature* 526, 531–535. doi: 10.1038/nature15512
- Meigs, T. E., and Simoni, R. D. (1997). Farnesol as a regulator of HMG-CoA reductase degradation: characterization and role of farnesyl pyrophosphatase. *Arch. Biochem. Biophys.* 345, 1–9. doi: 10.1006/abbi.1997.0200
- Metcalf, W. W., Zhang, J. K., Apolinario, E., Sowers, K. R., and Wolfe, R. S. (1997). A genetic system for archaea of the genus *Methanosarcina*: liposome-mediated transformation and construction of shuttle vectors. *Proc. Natl. Acad. Sci. U. S. A.* 94, 2626–2631. PMID: 9122246; PMID: PMC20139. doi: 10.1073/pnas.94.6.2626
- Meyerdierks, A., Kube, M., Kostadinov, I., Teeling, H., Glöckner, F. O., Reinhardt, R., et al. (2010). Metagenome and mRNA expression analyses of anaerobic methanotrophic archaea of the ANME-1 group. *Environ. Microbiol.* 12, 422–439. doi: 10.1111/j.1462-2920.2009.02083.x
- Moran, J. J., House, C. H., Freeman, K. H., and Ferry, J. G. (2005). Trace methane oxidation studied in several Euryarchaeota under diverse conditions. *Archaea* 1, 303–309. PMID: 15876563; PMID: PMC2685550. doi: 10.1155/2005/650670
- Mosier, N. S., and Ladisch, M. R. (2011). *Modern Biotechnology: Connecting Innovations in Microbiology and Biochemistry to Engineering Fundamentals*. Hoboken: John Wiley & Sons, ISBN: 1118210204.
- Nayak, D. D., and Metcalf, W. W. (2017). Cas9-mediated genome editing in the methanogenic archaeon *Methanosarcina acetivorans*. *Proc. Natl. Acad. Sci.* 114, 2976–2981. doi: 10.1073/pnas.1618596114



- Neill, T., Judd, D., Veith, E., and Rousar, D. (2009). Practical uses of liquid methane in rocket engine applications. *Acta Astronaut.* 65, 696–705. doi: 10.1016/j.actaastro.2009.01.052
- Newman, D. J., and Cragg, G. M. (2004). Marine natural products and related compounds in clinical and advanced preclinical trials. *J. Nat. Prod.* 67, 1216–1238. doi: 10.1021/np040031y
- Paduch, R., Kandefer-Szerszen, M., Trytek, M., and Fiedurek, J. (2007). Terpenes: substances useful in human healthcare. *Arch. Immunol. Ther. Ex.* 55, 315–327. doi: 10.1007/s00005-007-0039-1
- Park, M.-O., Ikenaga, H., and Watanabe, K. (2007). Phage diversity in a methanogenic digester. *Microb. Ecol.* 53, 98–103. PMID: 17186158. doi: 10.1007/s00248-006-9053-9
- Patel, G. B., and Chen, W. (2010). Archaeal lipid mucosal vaccine adjuvant and delivery system. *Expert Rev. Vaccines* 9, 431–440. PMID: 20370552. doi: 10.1586/erv.10.34
- Patil, U. K., and Muskan, K. (2009). *Essentials of Biotechnology*. Bengaluru: IK International Pvt Ltd. ISBN: 9380026528.
- Paulo, L. M., Stams, A. J., and Sousa, D. Z. (2015). Methanogens, sulphate and heavy metals: a complex system. *Rev. Environ. Sci. Biotechnol.* 14, 537–553. doi: 10.1007/s11157-015-9387-1
- Pechous, S. W., and Whitaker, B. D. (2004). Cloning and functional expression of an (E, E)- $\alpha$ -farnesene synthase cDNA from peel tissue of apple fruit. *Planta* 219, 84–94. doi: 10.1007/s00425-003-1191-4
- Pichersky, E., Raguso, R. A., Lewinsohn, E., and Croteau, R. (1994). Floral Scent Production in *Clarkia* (Onagraceae) (I. Localization and Developmental Modulation of Monoterpene Emission and Linalool Synthase Activity). *Plant Physiol.* 106, 1533–1540. doi: 10.1104/pp.106.4.1533
- Pritchett, M. A., Zhang, J. K., and Metcalf, W. W. (2004). Development of a markerless genetic exchange method for *Methanosarcina acetivorans* C2A and its use in construction of new genetic tools for methanogenic archaea. *Appl. Environ. Microbiol.* 70, 1425–1433. doi: 10.1128/AEM.70.3.1425-1433.2004
- Ragab, A., Shaw, D. R., Katuri, K. P., and Saikaly, P. E. (2020). Effects of set cathode potentials on microbial electrosynthesis system performance and biocathode methanogen function at a metatranscriptional level. *Sci. Rep.* 10, 1–15. doi: 10.1038/s41598-020-76229-5
- Rajilić-Stojanović, M., Smidt, H., and De Vos, W. M. (2007). Diversity of the human gastrointestinal tract microbiota revisited. *Environ. Microbiol.* 9, 2125–2136. doi: 10.1111/j.1462-2920.2007.01369.x
- Riedinger, N., Formolo, M. J., Lyons, T. W., Henkel, S., Beck, A., and Kasten, S. (2014). An inorganic geochemical argument for coupled anaerobic oxidation of methane and iron reduction in marine sediments. *Geobiology* 12, 172–181. doi: 10.1111/gbi.12077
- Rodriguez-Concepcion, M., and Boronat, A. (2002). Elucidation of the methylerythritol phosphate pathway for isoprenoid biosynthesis in bacteria and plastids. A metabolic milestone achieved through genomics. *Plant Physiol.* 130, 1079–1089. doi: 10.1104/pp.007138
- Roslev, P., and King, G. M. (1995). Aerobic and anaerobic starvation metabolism in methanotrophic bacteria. *Appl. Environ. Microbiol.* 61, 1563–1570. PMID: 16535004; PMCID: PMC1388422. doi: 10.1128/aem.61.4.1563-1570.1995
- Rother, M., and Metcalf, W. W. (2004). Anaerobic growth of *Methanosarcina acetivorans* C2A on carbon monoxide: an unusual way of life for a methanogenic archaeon. *Proc. Natl. Acad. Sci. U. S. A.* 101, 16929–16934. doi: 10.1073/pnas.0407486101
- Rouviere, P. E., and Wolfe, R. (1988). Novel biochemistry of methanogenesis. *J. Biol. Chem.* 263, 7913–7916. PMID: 3131330.
- Sarmiento, F. B., Leigh, J. A., and Whitman, W. B. (2011). *Methods in Enzymology* Amsterdam: Elsevier. PMID: 21402209. doi: 10.1016/B978-0-12-385112-3.00003-2
- Scheller, S., Goenrich, M., Boecher, R., Thauer, R. K., and Jaun, B. (2010). The key nickel enzyme of methanogenesis catalyses the anaerobic oxidation of methane. *Nature* 465, 606–608. doi: 10.1038/nature09015
- Schiraldi, C., Giuliano, M., and De Rosa, M. (2002). Perspectives on biotechnological applications of archaea. *Archaea* 1, 75–86. doi: 10.1155/2002/436561
- Schlegel, K., Welte, C., Deppenmeier, U., and Muller, V. (2012). Electron transport during acetoclastic methanogenesis by *Methanosarcina acetivorans* involves a sodium-translocating Rnf complex. *FEBS J.* 279, 4444–4452. doi: 10.1111/febs.12031
- Schütz, H., Seiler, W., and Conrad, R. (1989). Processes involved in formation and emission of methane in rice paddies. *Biogeochemistry* 7, 33–53. doi: 10.1007/BF00000896
- Sgadari, C., Toschi, E., Palladino, C., Barillari, G., Carlei, D., Cereseto, A., et al. (2000). Mechanism of paclitaxel activity in Kaposi's sarcoma. *J. Immunol.* 165, 509–517. PMID: 10861090. doi: 10.4049/jimmunol.165.1.509
- Sharon-Asa, L., Shalit, M., Frydman, A., Bar, E., Holland, D., Or, E., et al. (2003). Citrus fruit flavor and aroma biosynthesis: isolation, functional characterization, and developmental regulation of Cstps1, a key gene in the production of the sesquiterpene aroma compound valencene. *Plant J.* 36, 664–674. doi: 10.1046/j.1365-313x.2003.01910.x
- Shaw, I., and Graham, M. (1987). Mesna—a short review. *Cancer Treat. Rev.* 14, 67–86. PMID: 3119211. doi: 10.1016/0305-7372(87)90041-7
- Sheehan, S. W. (2021). Electrochemical methane production from CO<sub>2</sub> for orbital and interplanetary refueling. *Iscience* 24:102230. doi: 10.1016/j.isci.2021.102230
- Shima, S., Krueger, M., Weinert, T., Demmer, U., Kahnt, J., Thauer, R. K., et al. (2012). Structure of a methyl-coenzyme M reductase from Black Sea mats that oxidize methane anaerobically. *Nature* 481, 98–101. PMID: 22121022. doi: 10.1038/nature10663
- Siegert, M., Cichocka, D., Herrmann, S., Gründger, F., Feisthauer, S., Richnow, H.-H., et al. (2011). Accelerated methanogenesis from aliphatic and aromatic hydrocarbons under iron- and sulfate-reducing conditions. *FEMS Microbiol. Lett.* 315, 6–16. doi: 10.1111/j.1574-6968.2010.02165.x
- Siliakus, M. F., van der Oost, J., and Kengen, S. W. (2017). Adaptations of archaeal and bacterial membranes to variations in temperature, pH and pressure. *Extremophiles* 21, 651–670. doi: 10.1007/s00792-017-0939-x
- Sills, A. K., Williams, J. L., Tyler, B. M., Epstein, D. S., Sipsos, E. P., Davis, J. D., et al. (1998). Squalamine inhibits angiogenesis and solid tumor growth in vivo and perturbs embryonic vasculature. *Cancer Res.* 58, 2784–2792. PMID: 9661892.
- Silver, G. M., and Fall, R. (1991). Enzymatic synthesis of isoprene from dimethylallyl diphosphate in aspen leaf extracts. *Plant Physiol.* 97, 1588–1591. PMID: 16668590; PMCID: PMC1081206. doi: 10.1104/pp.97.4.1588
- Sirohi, S., Pandey, N., Singh, B., and Puniya, A. (2010). Rumen methanogens: a review. *Indian J. Microbiol.* 50, 253–262. doi: 10.1007/s12088-010-0061-6
- Skinner, K. A., and Leathers, T. D. (2004). Bacterial contaminants of fuel ethanol production. *J. Ind. Microbiol. Biotechnol.* 31, 401–408. doi: 10.1007/s10295-004-0159-0
- Sogodogo, E., Drancourt, M., and Grine, G. (2019). Methanogens as emerging pathogens in anaerobic abscesses. *Eur. J. Clin. Microbiol. Infect. Dis.* 38, 811–818. doi: 10.1007/s10096-019-03510-5
- Soo, V. W., McAnulty, M. J., Tripathi, A., Zhu, F., Zhang, L., Hatzakis, E., et al. (2016). Reversing methanogenesis to capture methane for liquid biofuel precursors. *Microb. Cell Factories* 15, 1–14. doi: 10.1186/s12934-015-0397-z
- Sowers, K. R., Baron, S. F., and Ferry, J. G. (1984). *Methanosarcina acetivorans* sp. nov., an Acetotrophic methane-producing bacterium isolated from marine sediments. *Appl. Environ. Microbiol.* 47, 971–978. doi: 10.1128/aem.47.5.971-978.1984
- Stumm, C., and Zwart, K. (1986). Symbiosis of protozoa with hydrogen-utilizing methanogens. *Microbiol. Sci.* 3, 100–105. PMID: 3153149.
- Su, G., Zopfi, J., Yao, H., Steinle, L., Niemann, H., and Lehmann, M. F. (2020). Manganese/iron-supported sulfate-dependent anaerobic oxidation of methane by archaea in lake sediments. *Limnol. Oceanogr.* 65, 863–875. doi: 10.1002/lno.11354
- Subedi, B. P., Martin, W. F., Carbone, V., Duin, E. C., Cronin, B., Sauter, J., et al. (2021). Archaeal pseudomurein and bacterial murein cell wall biosynthesis share a common evolutionary ancestry. *FEMS Microb. Lett.* 424, 1–12. doi: 10.1093/femsmc/xtab012
- Summers, Z. M., Fogarty, H. E., Leang, C., Franks, A. E., Malvankar, N. S., and Lovley, D. R. (2010). Direct exchange of electrons within aggregates of an evolved syntrophic coculture of anaerobic bacteria. *Science* 330, 1413–1415. doi: 10.1126/science.1196526
- Takai, K., Nakamura, K., Toki, T., Tsunogai, U., Miyazaki, M., Miyazaki, J., et al. (2008). Cell proliferation at 122°C and isotopically heavy CH<sub>4</sub> production by a hyperthermophilic methanogen under high-pressure cultivation. *Proc. Natl. Acad. Sci. U. S. A.* 105, 10949–10954. doi: 10.1073/pnas.0712334105
- Tanner, R. S., and Wolfe, R. S. (1988). Nutritional requirements of *Methanococcus mobilis*. *Appl. Environ. Microbiol.* 54, 625–628. PMID: 3377488; PMCID: PMC202515. doi: 10.1128/aem.54.3.625-628.1988
- Thauer, R. K. (2011). Anaerobic oxidation of methane with sulfate: on the reversibility of the reactions that are catalyzed by enzymes also involved in methanogenesis from CO<sub>2</sub>. *Curr. Opin. Microbiol.* 14, 292–299. doi: 10.1016/j.mib.2011.03.003
- Thauer, R. K. (2012). The Wolfe cycle comes full circle. *Proc. Natl. Acad. Sci. U. S. A.* 109, 15084–15085. doi: 10.1073/pnas.1213193109
- Thauer, R. K., Kaster, A.-K., Seedorf, H., Buckel, W., and Hedderich, R. (2008). Methanogenic archaea: ecologically relevant differences in energy conservation. *Nat. Rev. Microbiol.* 6, 579–591. PMID: 18587410. doi: 10.1038/nrmicro193
- Tumbula, D. L., Makula, R. A., and Whitman, W. B. (1994). Transformation of *Methanococcus maripaludis* and identification of a Pst I-like restriction system. *FEMS Microbiol. Lett.* 121, 309–314. doi: 10.1111/j.1574-6968.1994.tb07118.x
- Ulrich, G., and Bower, S. (2008). Active methanogenesis and acetate utilization in Powder River Basin coals, United States. *Int. J. Coal Geol.* 76, 25–33. doi: 10.1016/j.coal.2008.03.006



- Urlinger, S., Baron, U., Thellmann, M., Hasan, M. T., Bujard, H., and Hillen, W. (2000). Exploring the sequence space for tetracycline-dependent transcriptional activators: novel mutations yield expanded range and sensitivity. *Proc. Natl. Acad. Sci.* 97, 7963–7968. doi: 10.1073/pnas.130192197
- van Der Hoeven, R. S., Monforte, A. J., Breeden, D., Tanksley, S. D., and Steffens, J. C. (2000). Genetic control and evolution of sesquiterpene biosynthesis in *Lycopersicon esculentum* and *L. hirsutum*. *Plant Cell* 12, 2283–2294. doi: 10.1105/tpc.12.11.2283
- Vandermoten, S., Santini, S., Haubruge, E., Heuze, F., Francis, F., Brasseur, R., et al. (2009). Structural features conferring dual geranyl/farnesyl diphosphate synthase activity to an aphid prenyltransferase. *Insect Biochem. Mol. Biol.* 39, 707–716. doi: 10.1016/j.ibmb.2009.08.007
- Villanueva, L., Damste, J. S., and Schouten, S. (2014). A re-evaluation of the archaeal membrane lipid biosynthetic pathway. *Nat. Rev. Microbiol.* 12, 438–448. doi: 10.1038/nrmicro3260
- Villanueva, L., von Meijenfildt, F. B., Westbye, A. B., Yadav, S., Hopmans, E. C., Dutilh, B. E., et al. (2021). Bridging the membrane lipid divide: bacteria of the FCB group superphylum have the potential to synthesize archaeal ether lipids. *ISME J.* 15, 168–182. PMID: 32929208; PMCID: PMC7852524. doi: 10.1038/s41396-020-00772-2
- Vranova, E., Coman, D., and Gruissem, W. (2012). Structure and dynamics of the isoprenoid pathway network. *Mol. Plant* 5, 318–333. doi: 10.1093/mp/sss015
- Wang, C., Wang, C., Liu, J., Xu, Q., Han, Z., Xu, X., et al. (2020). Tolerance of acetoclastic methanogenesis enhanced by magnetite under the condition of ammonia stress. *ACS Sustain. Chem. Eng.* 8, 1417–1426. doi: 10.1021/acssuschemeng.9b05585
- Wang, Y., Wegener, G., Hou, J., Wang, F., and Xiao, X. (2019). Expanding anaerobic alkane metabolism in the domain of archaea. *Nat. Microbiol.* 4, 595–602. doi: 10.1038/s41564-019-0364-2
- Watanabe, T., Pfeil-Gardiner, O., Kahnt, J., Koch, J., Shima, S., and Murphy, B. J. (2021). Three-megadalton complex of methanogenic electron-bifurcating and CO<sub>2</sub>-fixing enzymes. *Science* 373, 1151–1156. PMID: 34516836. doi: 10.1126/science.abg5550
- Welte, C., and Deppenmeier, U. (2011). Re-evaluation of the function of the F<sub>420</sub> dehydrogenase in electron transport of *Methanosarcina mazei*. *FEBS J.* 278, 1277–1287. doi: 10.1111/j.1742-4658.2011.08048.x
- Whitford, M. E., Teather, R. M., and Forster, R. J. (2001). Phylogenetic analysis of methanogens from the bovine rumen. *BMC Microbiol.* 1, 1–5. PMID: 11384509; PMCID: PMC32158. doi: 10.1186/1471-2180-1-5
- Wise, M. L., Savage, T. J., Katahira, E., and Croteau, R. (1998). Monoterpene synthases from common sage (*Salvia officinalis*). cDNA isolation, characterization, and functional expression of (+)-sabinene synthase, 1,8-cineole synthase, and (+)-bornyl diphosphate synthase. *J. Biol. Chem.* 273, 14891–14899. doi: 10.1074/jbc.273.24.14891
- Yan, Z., and Ferry, J. G. (2018). Electron bifurcation and confurcation in methanogenesis and reverse methanogenesis. *Front. Microbiol.* 9:1322. doi: 10.3389/fmicb.2018.01322
- Yan, M., Treu, L., Zhu, X., Tian, H., Basile, A., Fotidis, I. A., et al. (2020). Insights into ammonia adaptation and methanogenic precursor oxidation by genome-centric analysis. *Environ. Sci. Technol.* 54, 12568–12582. doi: 10.1021/acs.est.0c01945
- Zengler, K., Richnow, H. H., Rosselló-Mora, R., Michaelis, W., and Widdel, F. (1999). Methane formation from long-chain alkanes by anaerobic microorganisms. *Nature* 401, 266–269. doi: 10.1038/45777
- Zhou, Z., Zhang, C.-J., Liu, P.-F., Fu, L., Laso-Pérez, R., Yang, L., et al. (2022). Non-syntrophic methanogenic hydrocarbon degradation by an archaeal species. *Nature* 601, 257–262. doi: 10.1038/s41586-021-04235-2
- Zhuang, X., Köllner, T. G., Zhao, N., Li, G., Jiang, Y., Zhu, L., et al. (2012). Dynamic evolution of herbivore-induced sesquiterpene biosynthesis in sorghum and related grass crops. *Plant J.* 69, 70–80. doi: 10.1111/j.1365-3113.2011.04771.x



## OPEN ACCESS

## EDITED BY

Graciela L. Lorca,  
University of Florida,  
United States

## REVIEWED BY

Gregory B. Whitfield,  
Université de Montréal,  
Canada  
Kaylie Allyson Padgett-Pagliai,  
University of Florida,  
United States

## \*CORRESPONDENCE

Catherine A. Wakeman  
✉ catherine.wakeman@ttu.edu

## †PRESENT ADDRESS

Karishma Bisht,  
Department of Molecular Biology,  
Princeton University, Princeton, NJ, United  
States

## SPECIALTY SECTION

This article was submitted to  
Microbial Physiology and Metabolism,  
a section of the journal  
Frontiers in Microbiology

RECEIVED 30 August 2022

ACCEPTED 07 December 2022

PUBLISHED 06 January 2023

## CITATION

Bisht K, Luecke AR and Wakeman CA (2023)  
Temperature-specific adaptations and  
genetic requirements in a biofilm formed  
by *Pseudomonas aeruginosa*.  
*Front. Microbiol.* 13:1032520.  
doi: 10.3389/fmicb.2022.1032520

## COPYRIGHT

© 2023 Bisht, Luecke and Wakeman. This is  
an open-access article distributed under  
the terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,  
distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in  
this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Temperature-specific adaptations and genetic requirements in a biofilm formed by *Pseudomonas aeruginosa*

Karishma Bisht<sup>†</sup>, Alex R. Luecke and Catherine A. Wakeman<sup>\*</sup>

Department of Biological Sciences, Texas Tech University, Lubbock, TX, United States

*Pseudomonas aeruginosa* is a gram-negative opportunistic pathogen often associated with nosocomial infections that are made more severe by this bacterium's ability to form robust biofilms. A biofilm is a microbial community encompassing cells embedded within an extracellular polymeric substrate (EPS) matrix that is typically secreted by the encased microbial cells. Biofilm formation is influenced by several environmental cues, and temperature fluctuations are likely to be an important stimulus in the lifecycle of *P. aeruginosa* as it transitions between life in aquatic or soil environments to sites of infection in the human host. Previous work has demonstrated that human body temperature can induce a shift in the biofilm EPS relative to room temperature growth, resulting in an incorporation of a filamentous phage coat protein into the biofilm EPS. In this study, we sought to identify adaptations enabling biofilm formation at room temperature or temperatures mimicking the natural environment of *P. aeruginosa* (23°C and 30°C) relative to temperatures mimicking life in the human host (37°C and 40°C). We identified higher biofilm: biomass ratios at lower temperatures on certain substrates, which correlated with a higher relative abundance of apparent polysaccharide EPS content. However, the known genes for EPS polysaccharide production in *P. aeruginosa* PA14 did not appear to be specifically important for temperature-dependent biofilm adaptation, with the *pelB* gene appearing to be generally important and the *algD* gene being generally expendable in all conditions tested. Instead, we were able to identify two previously uncharacterized hypothetical proteins (*PA14\_50070* and *PA14\_67550*) specifically required for biofilm formation at 23°C and/or 30°C relative to temperatures associated with the human host. These unstudied contributors to biofilm integrity may have been previously overlooked since most *P. aeruginosa* biofilm studies tend to use 37°C growth temperatures. Overall, our study demonstrates that temperature shifts can have dramatic impacts on biofilm structure and highlights the importance of studying environment-specific adaptations in biofilm physiology.

## KEYWORDS

biofilm, thermal adaptation, *Pseudomonas aeruginosa*, environment biofilm, host biofilm, EPS matrix

## Introduction

Many bacterial species can form biofilms, resulting in a group of bacteria attached to a surface and surrounded by an extracellular polymeric substance (EPS) matrix which is composed of a mixture of macromolecules such as polysaccharides, proteins, carbohydrates and extracellular DNA (Donlan, 2002; Bjarnsholt, 2013). A biofilm often has a complex structure composed of differentiated groups of cells and a heightened antibiotic resistance (Donlan, 2001; Hall-Stoodley et al., 2004). *Pseudomonas aeruginosa*, a common nosocomial pathogen, forms a biofilm in the lungs of cystic fibrosis patients, in burn wounds, and on indwelling medical devices, causing infections that are largely resistant to antibiotics and to human immune defenses (Mulcahy et al., 2014; Maurice et al., 2018; Bisht et al., 2020). The EPS matrix is an important component of the biofilm since it not only provides structural stability to the biofilm but also offers functional benefits that can lead to both enhanced virulence and antimicrobial tolerance (Flemming et al., 2007; Dragoš and Kovács, 2017; Karygianni et al., 2020). Additionally, the EPS matrix is known to protect the bacterial cells from environmental stress factors like pH and osmotic stress (Hwang et al., 2016; Yan et al., 2017).

Another environmental factor that can influence a pathogen's survival requirements is temperature. Since *P. aeruginosa* is a ubiquitous microbe that can not only infect immunocompromised humans, but can also survive in the soil, in plants, and in within streams, where temperature is different from that of a human host (Schmidt et al., 1996; Hardalo and Edberg, 1997; Mena and Gerba, 2009; Silby et al., 2011; Balcázar et al., 2015; Radó et al., 2017), studying the effect of temperature on the physiology of this important pathogen is crucial. Additionally, this pathogen can form biofilms on different surface materials ranging from medical equipment to industrial pipes, in both hospital as well as industrial settings (de Abreu et al., 2014; Bédard et al., 2016). The impact of temperature on the biofilm formation has been studied in several microbes including *P. aeruginosa* (Townsend and Yildiz, 2015; Townsend et al., 2016; Plumley et al., 2017; Kim et al., 2020; Almblad et al., 2021; del Peso Santos et al., 2021). Temperature is also known to cause stress in mesophilic pathogens found in cold water (Yi et al., 2022). Previous research has shown that temperature can indeed influence the gene expression and protein profile of *P. aeruginosa* at environmental versus host-associated temperatures (Bisht et al., 2021). The temperature fluctuations can also result in physiological adaptation and structural changes in the biofilm, which eventually help *P. aeruginosa* adapt better during its transition from the outside environment into a host.

Our goal in this study was to elucidate the effect of temperature on the EPS matrix of the biofilm formed by *P. aeruginosa* PA14 at 23°C, 30°C, 37°C and 40°C. The strain was selected as it has fewer potentially partially redundant mechanisms to promote EPS formation which could help uncover as yet unexplored biofilm EPS adaptations. Additionally, since light/dark cycling has been shown to impact PA14 biofilm gene expression, affecting both metabolic pathways and redox homeostasis, all biofilms were kept

in dark for our experimental purposes (Kahl Lisa et al., 2022). The temperature range in our study was chosen to simulate conditions relevant to both industrial/environmental and medical settings, ranging from the standard room temperature of a hospital setting to temperatures associated with human fever. We identified EPS adaptations of *Pseudomonas aeruginosa* and specific genetic requirements that support biofilm formation at temperatures lower than the human host. These adaptations can serve as potential targets to further differentiate treatment methods in clinical and environmental settings, which could assist in limiting the emergence of resistance in environment reservoirs to the same antibiotics that are used in humans.

## Results and discussion

### Environmental temperature biofilm is associated with higher biomass

*Pseudomonas aeruginosa* is known to form biofilm in the different niches and undergoes a temperature shift when transitioning from an outside environment into the human body environment. To investigate the role of temperature in biofilm formation, we used *P. aeruginosa* lab strain UCBPP-PA14 (PA14) to grow biofilms at the four temperatures in a static condition. We used an established crystal violet (CV assay) staining method for biofilm and biomass measurement (O'Toole, 2011) and found that the biomass was higher at 23°C compared to the other temperatures (Figure 1A). As the temperature rose, there was a trend toward decreasing biofilm biomass. Of note, at 40°C we observed more variability in biofilm formation, with some replications displaying unusually high biomass and others displaying unusually low biofilm-associated biomass. This variation was consistent in all the rounds that we performed using the CV assay. Our observation of a higher biomass at lower temperature is consistent with the findings of other research groups (Townsend and Yildiz, 2015; Çam and Brinkmeyer, 2020; Kim et al., 2020).

To assess whether the temperature-dependent trends in biofilm formation were due to overall differences in growth, we measured viable cell numbers at 48 h for each temperature (Figure 1B) in addition to performing a full growth curve over this period (Supplementary Figure S1). These analyses revealed that PA14 grows to the highest density at 30°C with 37°C providing the second most robust growth of this strain. The 40°C temperature is suboptimal compared to 30°C with 37°C, but 23°C resulted in the lowest amount of overall growth. Therefore, we can confidently state that overall growth levels achieved at the different temperatures are not the driver of the biofilm formation trends since 23°C and 30°C displayed the most robust biofilms even though 23°C displayed the lowest overall growth while 30°C displayed the highest overall growth. Thus, we conclude that the differences in overall biofilm at the different temperatures represent a true physiological adaptation to temperature shifts.

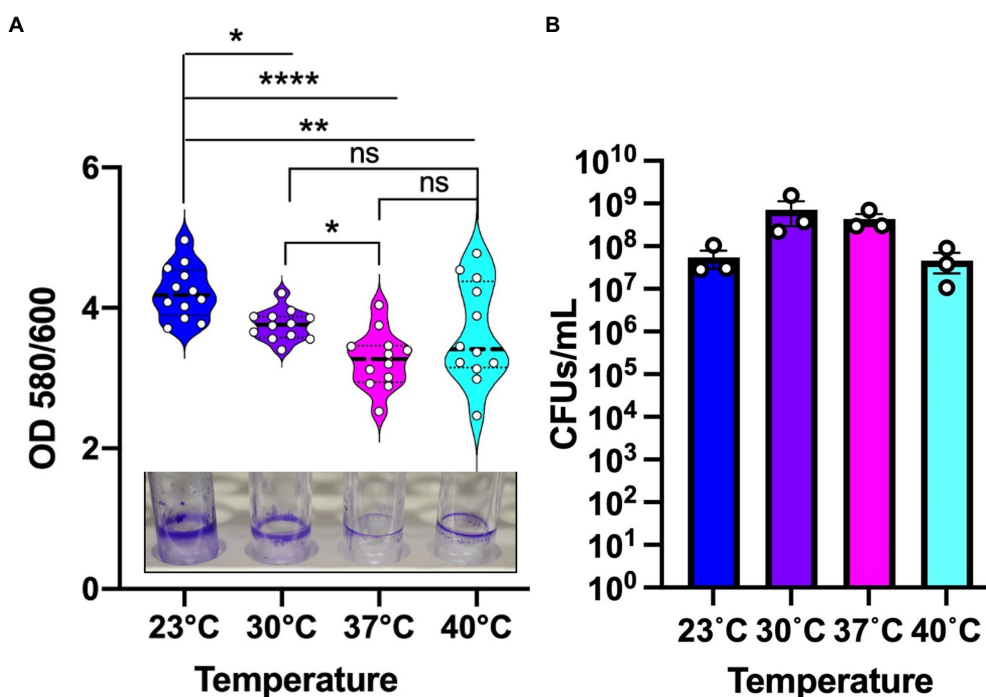


FIGURE 1

Environmental temperature biofilm is associated with higher biomass on polystyrene surfaces. (A) The biomass of the biofilm formed at the four temperatures. Violin plots represent means of three experiments with four replicates per sample. Crystal violet staining revealed subtle temperature-dependent architectural differences at the four temperatures. The center black line represents the interquartile range. Unpaired t-test (two-tailed) was used to measure statistical significance. ns: not significant, \* $p \leq 0.05$ , \*\* $p \leq 0.01$ , and \*\*\*\* $p \leq 0.0001$ . (B) CFU count of PA14 WT planktonic cells at environment versus host temperature. After 48 h of growth, the cultures were serially diluted in sterile 1X PBS and plated on LA plates. Plates were grown overnight at 37°C for sufficient colony formation. Error bars represent the standard error of mean of three biological replicates. The mean of each biological replicate was based on eight technical replicates after outlier analysis.

## Temperature-driven biofilm adaptations occur independently of nutrient sources and promote differential adherence to surface types

We next sought to determine whether these temperature dependent biofilm growth trends were dependent on the nutrient composition of the media. Therefore, we tested temperature dependent biofilm formation in M9 minimal media with either glycerol, malonate, or glucose as the sole carbon source. All these growth substrates are known to play a role in influencing the metabolism of *P. aeruginosa* as reported in previous studies (Frimmersdorf et al., 2010; She et al., 2019; Elmassry et al., 2021). In all cases, the temperature-dependent biofilm trends were similar to those observed in LB (Figure 2A). Additionally, since *P. aeruginosa* can adhere to different surface materials, including both biotic and abiotic surfaces (Martin et al., 1993; Pericolini et al., 2018), we wanted to assess if differences in the surface material can have any impact on the biofilm formation at the four temperatures. For our experiments we used polystyrene and borosilicate glass material, the two most widely used materials for biofilm growth in lab environment as well as disposable urinary catheters made from polyvinyl chloride to represent a

surface encountered in the clinic (Figure 2B). For both the polystyrene and borosilicate glass, we observed trends of more robust biofilm formation at environmental temperatures rather than host-associated temperatures. Interestingly, on the urinary catheters, overall biofilm biomass was greatly increased in all conditions with a slight biofilm growth preference observed at human body temperature. These results indicate that temperature adaptation in biofilms can occur independently of nutrient source and that the adaptations can lead to differential adherence properties to different substrates. Since EPS matrix is a vital component of the biofilm for surface adherence, we next wanted to study the temperature-specific EPS adaptations in *P. aeruginosa* biofilms.

## Biofilm architecture and EPS composition is modulated by temperature

To investigate the role of temperature on biofilm architecture, we first utilized microscopy tools. We grew *P. aeruginosa* strain PA14 biofilms at both environmental (23°C and 30°C) and host temperatures (37°C and 40°C) and visualized the biofilm using



both scanning electron microscopy (SEM) and confocal laser scanning microscopy (CLSM). The SEM micrographs revealed a striking difference in EM-visible matrix production at environmental versus host temperatures. At 23°C and 30°C, the matrix appeared to be more scattered while at 37°C and 40°C, it was denser with cells entangled within the matrix (Figure 3). We also observed a few circular vesicle-like structures only in the 40°C biofilm (Supplementary Figure S2). Several studies have reported that high-temperature stress can change the composition of the outer membrane resulting in the formation of membrane

vesicles (Cooke et al., 2019, 2020; Mozaheb and Mingeot-Leclercq, 2020). Also, a recent paper reported that explosive cell lysis in a subpopulation of cells under stress conditions could lead to the formation of membrane vesicles and biofilms (Turnbull et al., 2016). We believe these vesicle-like structures represent a similar phenomenon for our 40°C-grown biofilm.

Since SEM data revealed higher levels of EM-visible EPS in glass-associated biofilms at higher temperatures even though we observed lower overall biofilm, we theorized that there are other components contributing to the robust biofilms forming at

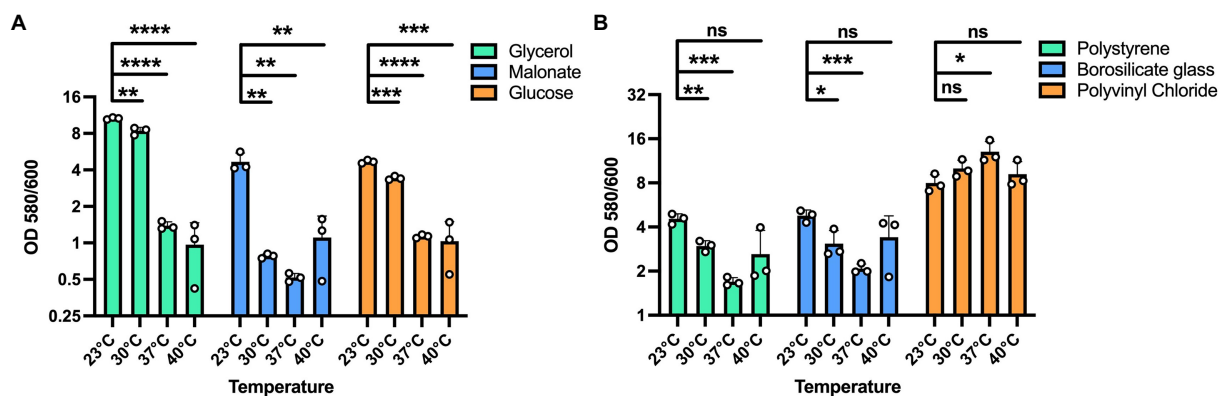


FIGURE 2

Effect of different substrates and materials on temperature-dependent biofilm formation. (A) The overall biofilm trend was same for all the carbon sources used. A 100 mM concentration of each carbon source in M9 media was used for biofilm growth at each temperature. In all cases, biofilm formation was most robust at 23°C and displayed decreasing levels with increasing temperatures. (B) The trend of higher biofilm formation at lower temperatures is observed on polystyrene and borosilicate glass. However, on polyvinyl chloride urinary catheters, biofilm formation was generally robust at all temperatures with a statistically significant advantage at 37°C. Bars represent the mean of three biological replicates performed on different days. The mean of each biological replicate was based on three technical replicates. Error bars represent the standard error of mean of the biological replicates. Unpaired t-test (two-tailed) was used to measure statistical significance. Y-axis is scaled by log2. ns: not significant, \* $p \leq 0.05$ , \*\* $p \leq 0.01$ , \*\*\* $p \leq 0.001$ , and \*\*\*\* $p \leq 0.0001$ .

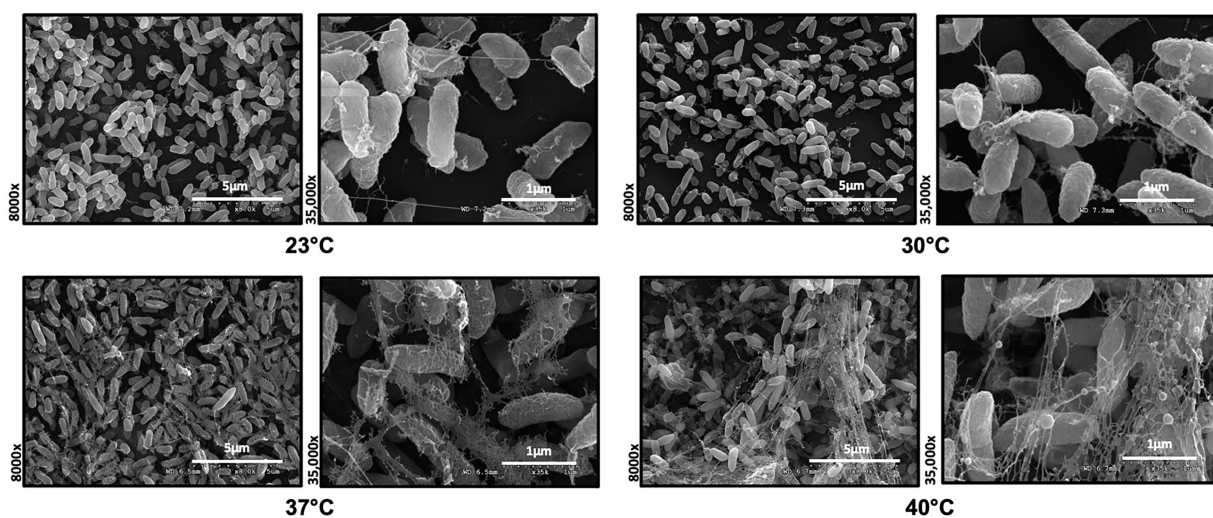


FIGURE 3

Microscopy reveals temperature-induced structural changes in *P. aeruginosa* biofilms. Comparison of the Scanning Electron Microscopy images of *P. aeruginosa* wild type biofilms grown for 48 h at 23°C, 30°C, 37°C and 40°C. Images show 8,000Å- and 35,000Å-magnification and are representative of three independent experiments.

23 and 30°C that might not be visible in our SEM micrographs, perhaps due to being washed away during sample processing steps. We therefore switched to CLSM to study the architecture of biofilms grown at different temperatures. To accomplish this, we used the Live/Dead Biofilm viability kit to observe the live versus dead cells as well as the overall biofilm architecture of *P. aeruginosa* at all the four temperatures (Deng et al., 2020). Also, since the propidium iodide (PI) which usually stains the dead cells, is known to stain the eDNA component of the matrix, we were able to observe some of the features of the matrix as well. Qualitatively, the overall architecture of the biofilm at 23°C and 30°C appeared slightly smoother with fewer textured features than biofilms grown at 37°C or 40°C (Figure 4A). On further quantifying certain parameters of the biofilm using COMSTAT2 and ImageJ, we observed the same trend of inverse correlation between biofilm thickness and temperature (Figure 4B). The surface area to biovolume ratio was also significantly higher in 23°C biofilms (Figure 4C). These trends did not appear to correlate with higher numbers of dead/lysed bacterial cells as the 23°C

biofilms displayed the highest live to dead ratio of all the other conditions (Figure 4D). Therefore, it is likely not eDNA released from lysed cells that is contributing to the robust biofilms at 23°C.

Since polysaccharide is one of the important components of the extracellular matrix, we next wanted to investigate its presence at different temperatures. For this purpose, we used calcofluor white, a polysaccharide-binding dye, which has been previously used to visualize the extracellular matrix (Neut et al., 2005; Wakeman et al., 2016; Figure 5A). This dye can bind to a variety of polysaccharides, with a strong interaction with glycosidic bonds, mainly  $\beta$ 1-3 and  $\beta$ 1-4 polysaccharide (Alemayehu et al., 2012; Soler-Arango et al., 2019). On quantifying the polysaccharide content relative to overall biofilm cell density, we observed that the room temperature biofilm had a slightly higher polysaccharide content per cell compared to the rest of the temperatures (Figure 5B). Further support of the temperature-associated changes in polysaccharide content in the EPS was provided by growth on Congo Red plates (Figure 5C; Madsen et al., 2015). This assay utilizes a Congo Red dye that binds to the

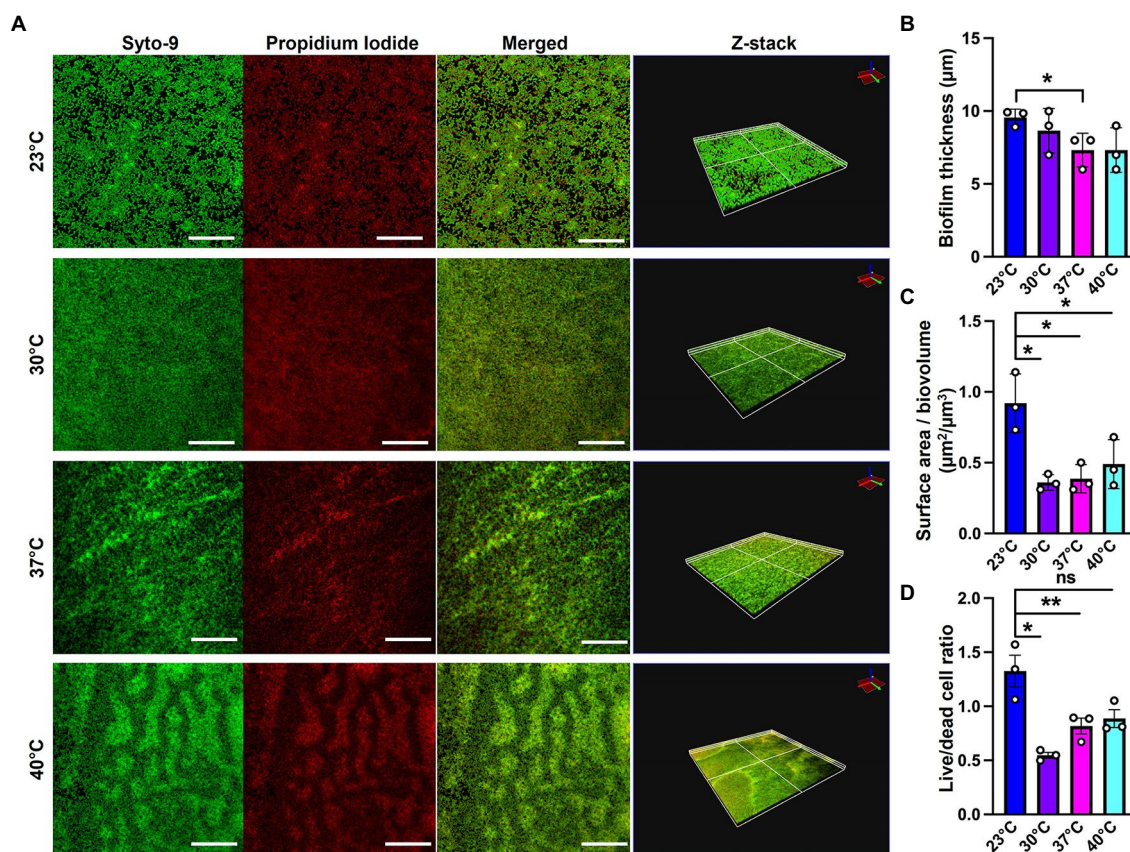


FIGURE 4

Microscopy using EPS-specific stains reveals temperature-induced structural changes in *P. aeruginosa* biofilms. (A) The live/dead staining results of *P. aeruginosa* PA14 biofilm after 48 h of growth at 23°C, 30°C, 37°C and 40°C. Biofilm was grown for 48 h on a microscope glass slide at 23°C, 30°C, 37°C and 40°C and stained with FilmTracer LIVE/DEAD Biofilm Viability kit; SYTO 9 shows live cells in green and propidium iodide shows dead cells in red. Scale bar: 50 μm. Z-stacks were collected for biofilms at the four temperatures using CLSM and analyzed by COMSTAT2 to determine. (B) The biofilm thickness and (C) the surface area to biovolume. (D) Live/dead cell ratio was analyzed using ImageJ. Images are representative of three independent experiments. Unpaired t-test (two-tailed) was used to measure statistical significance. ns: not significant, \* $p \leq 0.05$ , and \*\* $p \leq 0.01$ .



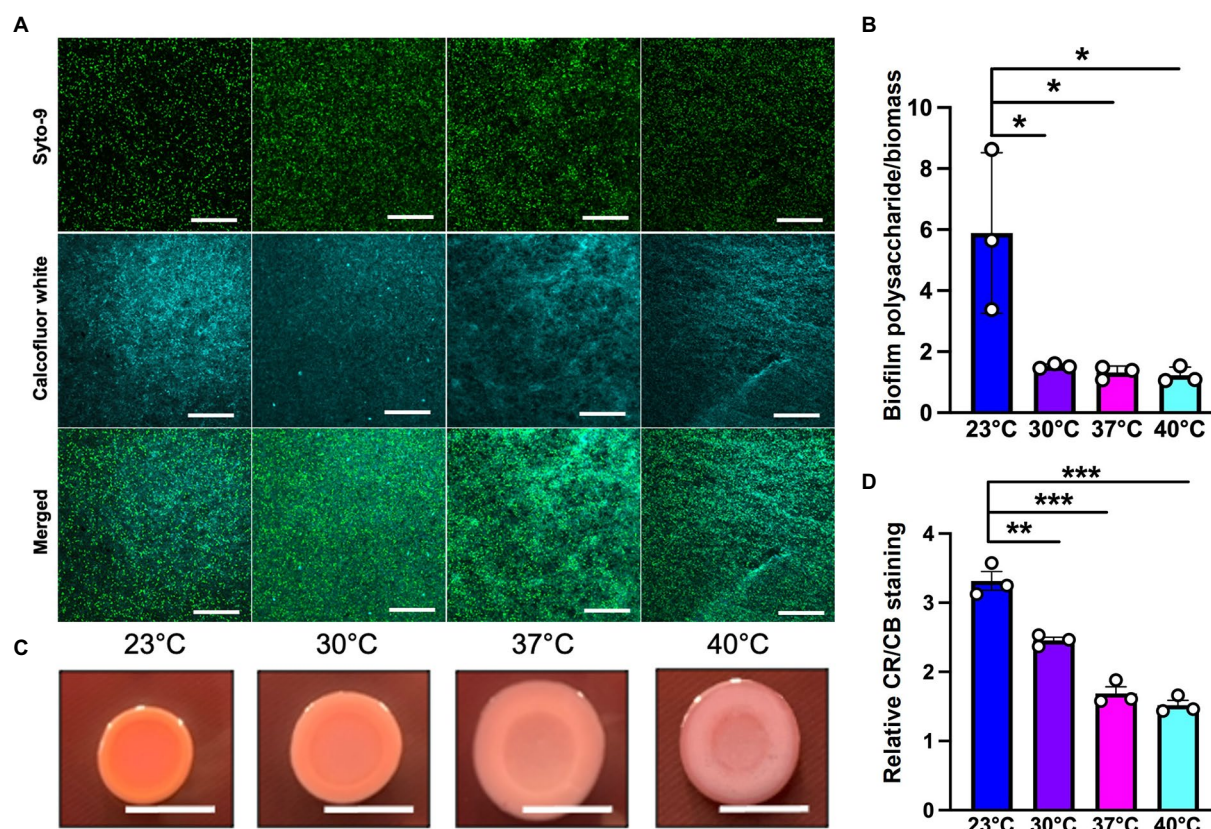


FIGURE 5

Microscopy using EPS-specific stains reveals temperature-induced increase in EPS carbohydrates in *P. aeruginosa* biofilms grown at 23°C. (A) The biofilm staining results of *P. aeruginosa* PA14 biofilm after 48 h of growth at 23°C, 30°C, 37°C and 40°C. Biofilm was grown for 48 h on a microscope glass slide at 23°C, 30°C, 37°C and 40°C and stained with SYTO 9 (live cells in green) and Calcofluor white (stains the polysaccharide component of the matrix cyan). Scale bar: 50  $\mu$ m. (B) Quantification of biofilm polysaccharide per biomass in *P. aeruginosa* PA14 biofilm after 48 h of growth at 23°C, 30°C, 37°C and 40°C using ImageJ. (C) Congo red binding assay reveals temperature specific difference in colony morphology. Extracellular matrix production by the wildtype was evaluated on tryptone agar plates containing Congo Red and Coomassie brilliant blue G after incubation at the four temperatures for 72 h. Representative images of the colony morphologies of PA14 are shown. Scale bar: 1 cm. (D) Quantification of CR and CB stain was performed using ImageJ. CR=Congo Red and CB=Coomassie Blue. Unpaired t-test (two-tailed) was used to measure statistical significance. \* $P \leq 0.05$ , \*\* $P \leq 0.01$ , and \*\*\* $P \leq 0.001$ .

polysaccharide component of the matrix and a Coomassie Blue dye that binds to the protein component of the matrix. The PA14 colonies looked morphologically different at all the four temperatures, with greater apparent Congo Red dye absorption and less apparent Coomassie Blue dye absorption at 23°C compared to the other three temperatures (Figures 5C,D). Previous work by our group has shown that there is a global proteomic difference in biofilms associated with 23°C and 37°C temperatures (Bisht et al., 2021), which further bolsters our present finding where we observe a different level of Coomassie dye absorption by the matrix at all four temperatures. Finally, the results in this study are consistent with previous findings stating the relation between temperature and polysaccharide production (Frølund et al., 1996; Sakuragi and Kolter, 2007). In total, these findings suggest that the polysaccharide component of the matrix could be contributing to biofilm formation at room temperature.

Previous studies have observed that *pel* and alginate carbohydrates genes increase at 20–23°C relative to 37°C in *P. aeruginosa* (Sakuragi and Kolter, 2007; Kim et al., 2020).

Therefore, it is possible that this phenomenon is occurring in our biofilms. However, we did not observe any striking differences in gene expression for *pel* or alginate biosynthetic genes (Supplementary Table S1) when we analyzed our previously published RNA sequencing data set (Bisht et al., 2021). However, these data do not exclude the possibility that temperature changes are influencing protein and/or activity levels of these polysaccharides. Therefore, we directly assessed the biofilm forming capacity of mutants lacking *pelB* and *algD* to demonstrate that neither *pel* nor alginate serve to specifically bolster biofilms grown at lower temperatures. The *pelB* mutant was found to be generally defective at all temperatures, while the *algD* mutant was generally expendable at all temperatures (Supplementary Figure S3). This finding highlights the general rather than temperature-specific importance of *pel* as an essential component of the matrix in the PA14 *P. aeruginosa* strain. Overall, our results are indicative of the presence of more EM-visible matrix and greater biofilm texture at higher temperatures and a flatter but generally thicker biofilm architecture at lower temperatures. We also see a slightly higher

polysaccharide content at lower temperature, consistent with the findings of some other research groups, which could be contributing toward the higher biofilm biomass observed at this temperature. However, future work focusing on the characterization of polysaccharide components at low temperature biofilm is required, especially since no temperature-specific polysaccharide requirements were identified for PA14.

## Genetic requirements to establish a robust biofilm at environmentally relevant temperatures

Because specific components contributing to the robust biofilm formation at room temperature remained elusive, a preliminary screening was performed on a commercially available transposon mutant library of *P. aeruginosa* comprising of over 5,500 mutant genes (Liberati et al., 2006) to identify temperature specific genetic requirements. We performed the established microtiter dish biofilm formation (crystal violet) assay to assess biofilm biomass of wild type and mutants at all the four temperatures (O'Toole, 2011). While most of the potential hits in this screen remain to be validated and thus are not included in the present manuscript, we turned our focus to hypothetical/uncharacterized proteins that might be specifically contributing to biofilm growth and adaptation at 23°C to better understand the odd phenomenon of temperature specific biofilm increases with no currently characterized genetic contributors to this structure.

The mutants in genes *PA14\_50070* and *PA14\_67750* were among the hits showing a significant reduction in biofilm formation at 23°C when compared to wild-type (Figure 6; Supplementary Figure S3). *PA14\_50070* which encodes for a hypothetical protein showed a defect in biofilm formation at 23°C and 30°C. Further characterization of this protein in the future can help us gain some more information about its role in low-temperature biofilm formation and adaptation. Another mutant of gene *PA14\_67750*, encoding for a rhodanese-like domain-containing protein, also showed a defective biofilm: biomass ratio at 23°C and a trend toward a defect at 30°C (Figure 6; Supplementary Figure S3). Because this gene is found in an apparent operon, we also performed the *PA14\_67750* complementation to confirm if the defect was due to the *PA14\_67750* gene and not the *PA14\_67740* (*grx*) gene that lies next to it. Our complementation experiment was able to recapitulate our findings, thus strengthening our observation that *PA14\_67750* indeed plays a key role in biofilm formation specifically at 23°C (Supplementary Figure S4). Additionally, the Congo red binding assay showed subtle changes in the colony morphology between the WT and the *PA14\_50070* and *PA14\_67750* deletion mutants (Supplementary Figure S5). For example, there is a trend toward a slight decrease in the level of Congo Red staining in these strains at 23°C and 30°C. We therefore believe that these hypothetical proteins represent previously uncharacterized temperature-specific contributors to *P. aeruginosa* biofilm formation and could be potential gene targets to eradicate biofilms growing at environment-associated temperature.

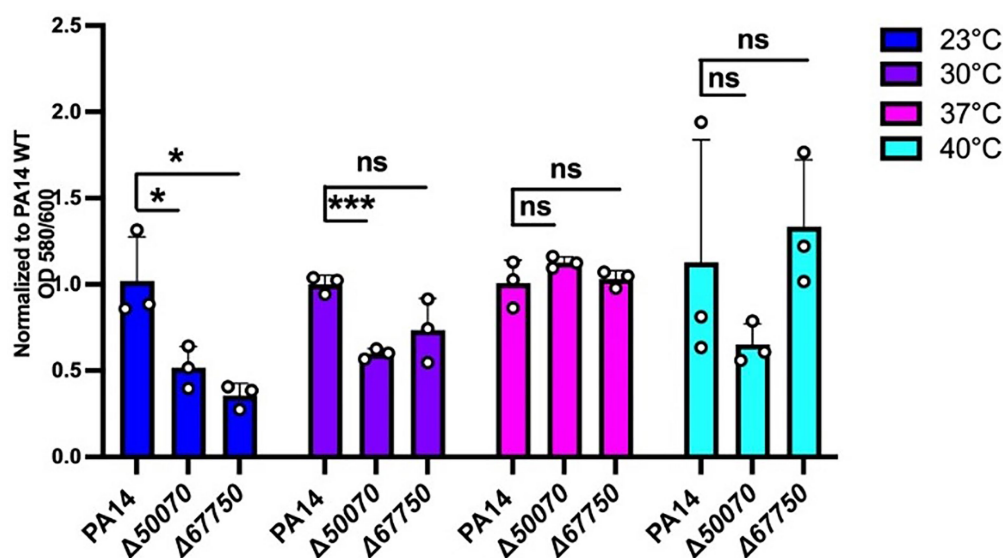


FIGURE 6

Growth requirements for biofilm formation at environmental versus host temperatures. The biofilm biomass (ratio of OD580/OD600), at each temperature for the two mutants was normalized relative to WT PA14 at the different temperatures. The biomass of the biofilm was defective at 23°C for both mutants and at 30°C for *PA14\_50070* when compared to PA14. Bars represent the mean of three biological replicates performed on different days. The mean of each biological replicate was based on three technical replicates. Error bars represent the standard error of mean of the biological replicates. Unpaired t-test (two-tailed) was used to measure statistical significance between the wildtype and each mutant. ns: not significant, and \* $p \leq 0.05$ , and \*\*\* $p \leq 0.001$ .



Since *P. aeruginosa* PA14 uses different strategies to form biofilms than other strains of *P. aeruginosa*, such as not having the *psl* operon, PA14 is considered somewhat of an outlier in species group. For this reason, we wanted to explore the relative conservation of our genes of interest. We downloaded all of the completed *P. aeruginosa* strains from NCBI database (accessed October 2022), 494 in total and ran them through Roary (Page et al., 2015) to assign the genes to ortholog groups. We found that *algD* and *pelB* were found in 98.78 and 36.03% of the strains analyzed, respectively. Conversely, all strains contained a gene which was assigned to the ortholog groups containing either *PA14\_50070* and *PA14\_67750*, (Supplementary Table S2). Additionally, we observed potential synteny or conservation of the sequence of genes in the region surrounding both genes. (Supplementary Figure S6) As a result of this data, we believe that our observations made regarding both hypothetical proteins to be relevant to the species group as a whole. However, future studies will be required to further verify this contention.

## Conclusion

Temperature is an important external stimulus that can tremendously impact the biofilm formation changing both gene and protein expression, as well as the metabolite levels (Bisht et al., 2021; Li et al., 2022). In this study we wanted to observe the effect of temperature on biofilm architecture at both environment and human host associated temperatures. We also sought to discover genes contributing to novel *P. aeruginosa* biofilm adaptations at room temperature rather than human body temperature. Overall, the data presented herein indicate that temperature can induce EPS-specific adaptations in *P. aeruginosa* by affecting the matrix composition and promoting differential surface association. We also report that despite an apparent increase in biofilm carbohydrates at 23°C, no known carbohydrates of strain PA14 were specifically required for biofilm formation at this temperature, with *pel* being generally required and alginate being generally expendable. Instead, previously uncharacterized hypothetical proteins appear to be playing a role in the temperature-specific biofilm adaptation. An in-depth understanding of the biofilm matrix of *P. aeruginosa* in its varied habitats is thus essential. By finding drug targets specific to particular niches (environmental versus host), we may be able to limit the spread of antibiotic resistance *via* environmental reservoirs.

## Materials and methods

### Bacterial strains, media, and growth conditions

*P. aeruginosa* strain UCBPP-PA14, a highly virulent strain of *P. aeruginosa* originally isolated from a wound infection, was used in all experiments unless otherwise stated (He et al., 2004). Mutants used for the screening are mentioned in Table 1 (Liberati

et al., 2006). Strains were routinely grown overnight and maintained at 37°C in Luria-Bertani (LB) broth. Gentamicin was added at 15 µg/ml to maintain the transposon in the mutants. Also, 100 mM of glucose, malonate, and glycerol in M9 media were used to assess biofilm biomass.

### Crystal violet assay

To quantify and study biofilm formation at each temperature we used the previously established microtiter biofilm formation assay (O'Toole, 2011). Wild-type *Pseudomonas aeruginosa* PA14 and the mutant strain were grown overnight in a 96 well round bottom plate in 150 microliters of LB broth with shaking at 220 rpm and 37°C overnight incubation. Next day, five microliters of the overnight culture were transferred to a fresh 96 well plate with 145 µl of LB media. This was done for 3 replicates at 3 different days. The plates were incubated for 48 h at 23°C, 30°C, 37°C and 40°C. An absorbance reading at 600-nm wavelength was taken after 48 h using a Synergy Hi5 Microplate Reader, Biotek. Planktonic cells were then aspirated out and the remaining biofilm was washed three times with 300 microliters of PBS. This step helps remove unattached cells and media components that can be stained in the next step and significantly lowers background staining. Next, 200 microliters of 100% ethanol were added to the wells and incubated for 15 min. The ethanol was then aspirated out completely and the plates are flipped upside down and left for drying. After the ethanol dried, 200 µl of a 1% solution of crystal violet (CV) was added to each well of the microtiter plate. The microtiter plate was then incubated for 15 min at room temperature followed by rinsing it 3–4 times with water

TABLE 1 Mutants selected to assess their role in biofilm formation at the four temperatures.

PA14 with MAR2xT7 mariner transposon insertion within the specific gene; Gm <sup>r</sup> (Liberati et al., 2006)			
Gene locus	Gene name	Product description	Functional category
PA14_18580	<i>algD</i>	GDP-mannose 6-dehydrogenase AlgD	Secreted Factors (toxins, enzymes, alginate)
PA14_24490	<i>pelB</i>	Hypothetical protein	Hypothetical, unclassified, unknown
PA14_50070	Hypothetical protein	Hypothetical protein	Hypothetical, unclassified, unknown
PA14_67750		Rhodanese-like domain-containing protein	Transport of small molecules

by submerging the plate in a tub of water and blotting vigorously on a stack of paper towels to get rid of the excess water. The microtiter plate was then left to dry for 1–2 h. Finally, 150 microliters of 30% acetic acid solution were added to each well of the microtiter plate to solubilize the crystal violet. After an hour incubation at room temperature, an absorbance reading was taken at 580 nm. Using the biomass baseline, this reading was quantified and analyzed to produce readable data. For data comparing biofilm formation on polystyrene versus borosilicate glass, cells were grown in 5 ml culture tubes composed of these two substrates each containing 0.5 ml of culture. Sample processing mimicked that of the microtiter plates. Prior to analysis, outliers were removed by determining the Q1 and Q3 values of each test group and then values outside of 1.5 times the range of Q1 and Q3 were removed.

## Biofilm formation on urinary catheters

Four overnight PA14 wild type (WT) isolates on LA plates were inoculated into 200 µl of LB media in a 96-well round bottom plate and grown overnight at 37°C. 145 µl of LB media was added into eight wells in a 96-well plate and 5 µl of each overnight culture was added into 2 of those wells. One of the duplicates acted as a control (no catheter) and to another, a 1 cm long piece of sterile urinary catheter (Cure Medical) was added. This inoculation was repeated three more times, so that there would be four plate copies: one for each temperature (23°, 30°, 37° and 40°C). Plates were placed in their respective incubators for 48 h. The fixing, staining and OD readings mentioned in “Crystal Violet Assay” method was followed with minor changes. After fixing and staining, the catheters were moved to clean wells in between each DI wash and the solubilized CV in acetic acid from the catheter wells had to be diluted to 1:10 ratio prior to taking the OD580 reading.

## Bacterial growth curve

Frozen cultures of mutants and PA14 WT were streaked out on LA plates and incubated overnight at 37°C. Isolated colonies were selected from both the WT and mutants which were then inoculated in 200 µl of LB within a 96-well round bottom plate. The plate was incubated for 18 h at 37°C in shaking conditions. In a 96-well flat bottom plate, 145 µl of LB and 5 µl of overnight grown cultures were added into each well. This step was repeated 3 more times to have a plate for each temperature (23°, 30°, 37° and 40°C). The 0-h OD 600 reading was taken, and plates were placed into their respective static incubators. Additional OD 600 readings were taken at 2, 6, 12, 24, 30, 36, and 48 h. Blank LB was used for background subtraction. The protocol was done on two independent days. The OD readings for all the time points were plotted using GraphPad Prism 8.

## Cell viability assay

A frozen culture of PA14 wild type (WT) was streaked on an LA plate and grown at 37°C overnight. Eight isolated colonies were picked and inoculated into 200 µl of LB within a 96-well round bottom plate. The culture plate was incubated for 18 h at 37°C in shaking conditions. 5 µl of overnight culture was inoculated into 145 µl of fresh LB within a 96-well flatbottom plate. This step was done a total of 4 times, one plate for each temperature (23°, 30°, 37° and 40°C). The plates were placed in their respective incubators in static conditions for 48 h. Afterwards, 1:10 serial dilutions of the cultures were done in filtered sterilized 1X PBS (a total of eight dilutions steps were needed). 5 µl of each dilution was spot plated onto LA plates and incubated overnight at 37°C. This was repeated on three independent replicates. Once all of the CFUs were counted, replicates were subjected to outlier analysis as mentioned previously and the averages of each independent replicate was calculated, which were plotted using GraphPad Prism 8.

## Congo red binding assay and quantification

To phenotypically assess the different components of the EPS matrix, a colony-morphology assay was performed as described previously (Madsen et al., 2015). Extracellular matrix production by *P. aeruginosa* PA14 and the mutants was evaluated on tryptone agar plates containing Congo Red and Coomassie brilliant blue G after incubation at 23°C, 30°C, 37°C and 40°C for 24, 48 and 72 h. Five microliters of overnight precultures were spotted on Petri plates containing 20 ml of the assay medium (1% tryptone, Congo red dye (40 µg/ml), Coomassie brilliant blue G dye (20 µg/ml) and 1% agar). Colonies were grown at 23°C, 30°C, 37°C and 40°C for 24, 48 and 72 h. Images of the colonies were taken daily using a Nikon camera. For quantification purpose, the images were first imported into ImageJ, after that the red (Congo red) and blue (Coomassie blue) channels were used to calculate the area and mean intensity. Normalization of all the mean integrated density values were done with respect to the area and the background. Apparent Congo red to Coomassie blue ratio was then calculated and plotted using GraphPad Prism.

## Electron microscopy

For scanning electron microscopy (SEM), static cultures were grown in 50-ml conical tubes with circular glass coverslips semi-submerged in 2 ml of LB broth. Coverslip samples were handled and processed as described previously (Gaddy et al., 2009). Upon removal of the culture medium, the coverslips were immediately flooded with 2.5% glutaraldehyde in 0.05 M sodium cacodylate and incubated at room temperature for 1 h. The fixative was then removed and replaced immediately with 0.05 M sodium

cacodylate to prevent sample dehydration. Next, the coverslips were incubated in osmium tetroxide for 15 min followed by dehydration with increasing concentrations of ethanol, ranging from 25 to 100%, and CO<sub>2</sub> critical point drying. Samples were carbon-coated and visualized with a Hitachi S-4300 scanning electron microscope.

## Biofilm imaging and quantification

Biofilms grown on glass slides for 48 h were imaged by confocal laser scanning microscopy (CLSM) (Olympus; FV3000). 20 µl of the overnight grown culture of PA14 were added into a fresh 20 ml of LB broth, in a 50-mL Falcon tube along with a microscope glass slide, and grown for 48 h. The biofilm was then washed with PBS followed by staining with FilmTracer LIVE/DEAD Biofilm Viability kit (ThermoFisher). Next, biofilm was incubated with SYTO™ 9 and propidium iodide stains for 20 min in the dark at room temperature as per the manufacturer's protocol. After that, the biofilm was washed with PBS and the slides were imaged using a OlympusFV3000 microscope with an objective lens of ×60 (oil). For another experiment, to visualize bacterial cells we used SYTO™ 9 and to visualize polysaccharide component of the matrix, calcofluor white (blue; Sigma-Aldrich) was used. For each glass slide, five image stacks were taken with a z-step size of 0.7 µm. Biofilm quantification at all four temperatures were done using COMSTAT2 plugin in ImageJ (Heydorn et al., 2000). ImageJ was used to quantify the live/dead cell ratio. Images were analyzed by first splitting each image into RGB channels and then collecting the fluorescence integrated density values for SYTO™ 9 stain from the green channel and propidium iodide from the red channel. Normalization of all the fluorescence integrated density values were done with respect to the area and the background. Finally, a ratio of SYTO™ 9 relative to propidium iodide signal was calculated using the average density values of each image and plotted. Similarly, ImageJ was also used to quantify polysaccharide/biofilm biomass content for which images were analyzed by first splitting each image into RGB channels and then collecting the fluorescence integrated density values for calcofluor white stain from the blue channel and for SYTO™ 9 stain from the green channel. Normalization of all the fluorescence integrated density values were done with respect to the area and the background. Finally, a ratio of calcofluor white signal relative to SYTO™ 9 signal was calculated using the average density values of each image and plotted using GraphPad Prism. Each experiment included three independent biological replicates, and at least three images were taken for each replicate.

## Plasmid construction

To design constructs for the complementation assay, we used the previously established method with slight modifications (Hmelo et al., 2015). First, DNA was extracted from PA14 WT overnight grown culture, followed by amplification of the gene loci *PA14\_67750* using forward and reverse primers. The primers were designed for

*PA14\_67750* with minor modifications upstream of the gene to add HindIII and EcoRI restriction sites. The primers *PA14\_67750-F* (5'-CCGGCAGAATTCCGCCCTGGATATCCGTCACC-3') and *PA14\_67750-R* (5'-GATCACAAGCTTCGGCATGTCACTTCACCAGC-3') targeting *PA14\_67750* were designed for our experimental use. PCR was conducted using the Q5 High-Fidelity Master Mix (New England BioLabs) followed by PCR purification (QIAGEN PCR purification kit). pHERD20T-specific primer sequences were obtained from a previously published study (Qiu et al., 2008). Isolated pHERD20T plasmid and the purified PCR products were digested by HINDIII and EcoRI, followed by ligation of the digested products using a quick ligation mix (New England BioLabs). The ligated products were then chemically transformed into *Escherichia coli* S17 and incubated on Luria Agar (LA) [with carbenicillin (50 µg/ml)] plates. Isolated colonies were selected, and plasmid was confirmed by colony PCR and gel electrophoresis. Confirmed colonies were grown in 5 ml of LB at 37°C overnight and stored. Transposon mutants  $\Delta$ *PA14\_67750* was streaked on LA plate and the *E. coli* S17 mutant was grown on LA [with carbenicillin (50 µg/ml)] at 37°C overnight. Isolated colonies were grown in 5 ml of LB at 37°C (carbenicillin was added to *E. coli* S17 mutants to have a final concentration of 50 µg/ml). 1 ml of the overnight cultures of *E. coli* S17 mutant (donor) was pelleted, and supernatant was discarded. Simultaneously, 1 ml of  $\Delta$ *PA14\_67750* (recipient) was added to the donor pellet and resuspended. Next, 400 µl of the donor/recipient mixture and donor only mixture was spotted on LA plates and grown for 5 h at 37°C. The spotted cultures were collected and suspended in 1 ml of sterile 1X PBS, pelleted and washed with 1 ml of sterile 1X PBS. 1:10 dilutions of conjugates were plated on Vogel-Bonner minimal medium (VBMM) salt plates [with carbenicillin (300 µg/ml)] and grown at 37°C overnight. Overnight cultures were then streaked on LA [with carbenicillin (300 µg/ml)] plates to get isolated colonies and grown at 37°C. Colonies were selected and constructs were confirmed by colony PCR and gel electrophoresis. Confirmed colonies were grown in 5 ml of LB [with carbenicillin (300 µg/ml)] overnight at 37°C and stored.

## Biofilm assay using complementation construct

Tn-*PA14\_67750::67750*, Tn-*PA14\_67750*:pHERD, and PA14 WT:pHERD was streaked on LA [with carbenicillin (200 µg/ml)] plates and grown overnight at 37°C. Isolated colonies from each plate were selected and inoculated in a 96-well plate with 200 µl of LB [with carbenicillin (200 µg/ml)]. The plate was grown at 37°C overnight in shaking conditions. Next day, 145 µl LB [with carbenicillin (200 µg/ml) and 1.5% arabinose] was added into wells of four sterile 96-well plates and then inoculated with 5 µl of overnight culture. Each plate was wrapped in plastic wrap and placed at 23°C in dark and grown for 48 h. Previously established microtiter biofilm formation assay, as discussed earlier, was then performed. Each experiment included four independent biological replicates and was performed on three different days.

## Orthologous analysis of *Pseudomonas aeruginosa* strains

The FASTA files of the complete genomes for *P. aeruginosa* were downloaded from NCBI. The genomes were ran through Prokka software to annotate (Seemann, 2014). The annotated GTF files were then compared for orthologs using the default settings of Roary (Page et al., 2015). The output file “gene\_presence\_absence.csv” was used to determine the number of genes by strain to each ortholog group.

## Statistical analysis

Statistical analyses were performed using GraphPad Prism 8.0 (GraphPad Software, Inc., San Diego, CA). Unpaired t-test (two tailed) was used to calculate the statistical significance.

## Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

## Author contributions

KB contributed to the project design, data acquisition, interpretation of data, and manuscript writing. AL contributed to the complementation assay, data acquisition and interpretation of the biofilm assay as well as manuscript editing. CW contributed to project design, data acquisition, interpretation of data, and editing of the paper. All authors contributed to the article and approved the submitted version.

## Funding

This research work in the Wakeman lab was supported by NIH/NIGMS (R15GM128072). KB was supported by the Doctoral

Dissertation Completion Fellowship granted from Texas Tech University Graduate School. CAW received publication funding through TTU Open Access Fund. ARL received publication funding through TechASM.

## Acknowledgments

We would like to thank Alan Hauser for generously providing us with the pHERD20T plasmid for complementation assays, and Moamen Elmassry for thoughtful discussion during the preparation of the manuscript and help with COMSTAT. We would also like to thank Austin Brown, Jake Sellers, Jeanette Rimbey, Kiddist Yihunie, Jennifer Herrera, Peyton Webster, Andrew Salazar, Jayson Long, and Temitayo Ahmed for assistance with the biofilm high throughput screening. We thank Darian Winkle, Gretta Poole, Adolph Salinas, Priyanka Mathews for their help in accessing phenotypes of different mutants. We thank the College of Arts & Sciences Microscopy, Texas Tech University, for use of facilities.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.1032520/full#supplementary-material>

## References

- Alemayehu, D., Casey, P. G., McAuliffe, O., Guinane, C. M., Martin, J. G., Shanahan, F., et al. (2012). Bacteriophages  $\phi$ MR299-2 and  $\phi$ NH-4 can eliminate *Pseudomonas aeruginosa* in the murine lung and on cystic fibrosis lung airway cells. *MBio* 3, e00029–e00012. doi: 10.1128/mBio.00029-12
- Almblad, H., Randall, T. E., Liu, F., Leblanc, K., Groves, R. A., Kittichotirat, W., et al. (2021). Bacterial cyclic diguanylate signaling networks sense temperature. *Nat. Commun.* 12:1986. doi: 10.1038/s41467-021-22176-2
- Balcázar, J. L., Subirats, J., and Borrego, C. M. (2015). The role of biofilms as environmental reservoirs of antibiotic resistance. *Front. Microbiol.* 6:1216. doi: 10.3389/fmicb.2015.01216
- Bédard, E., Prévost, M., and Déziel, E. (2016). *Pseudomonas aeruginosa* in premise plumbing of large buildings. *Microbiology* 5, 937–956. doi: 10.1002/mbo3.391
- Bisht, K., Baishya, J., and Wakeman, C. A. (2020). *Pseudomonas aeruginosa* polymicrobial interactions during lung infection. *Curr. Opin. Microbiol.* 53, 1–8. doi: 10.1016/j.mib.2020.01.014
- Bisht, K., Moore, J. L., Caprioli, R. M., Skaar, E. P., and Wakeman, C. A. (2021). Impact of temperature-dependent phage expression on *Pseudomonas aeruginosa* biofilm formation. *NPJ Biofilms Microbiomes* 7:22. doi: 10.1038/s41522-021-00194-8



- Bjarnsholt, T. (2013). The role of bacterial biofilms in chronic infections. *APMIS Suppl.* 121, 1–58. doi: 10.1111/apm.12099
- Çam, S., and Brinkmeyer, R. (2020). The effects of temperature, pH, and iron on biofilm formation by clinical versus environmental strains of *Vibrio vulnificus*. *Folia Microbiol.* 65, 557–566. doi: 10.1007/s12223-019-00761-9
- Cooke, A. C., Florez, C., Dunshee, E. B., Lieber, A. D., Terry, M. L., Light, C. J., et al. (2020). *Pseudomonas* quinolone signal-induced outer membrane vesicles enhance biofilm dispersion in *Pseudomonas aeruginosa*. *mSphere* 5, e01109–e01120. doi: 10.1128/mSphere.01109-20
- Cooke, A. C., Nello, A. V., Ernst, R. K., and Schertzer, J. W. (2019). Analysis of *Pseudomonas aeruginosa* biofilm membrane vesicles supports multiple mechanisms of biogenesis. *PLoS One* 14:e0212275. doi: 10.1371/journal.pone.0212275
- de Abreu, P. M., Farias, P. G., Paiva, G. S., Almeida, A. M., and Morais, P. V. (2014). Persistence of microbial communities including *Pseudomonas aeruginosa* in a hospital environment: a potential health hazard. *BMC Microbiol.* 14:118. doi: 10.1186/1471-2180-14-118
- del Peso Santos, T., Alvarez, L., Sit, B., Irazoki, O., Blake, J., Warner, B. R., et al. (2021). BipA exerts temperature-dependent translational control of biofilm-associated colony morphology in *Vibrio cholerae*. *Elife* 10:e60607. doi: 10.7554/eLife.60607
- Deng, B., Ghatak, S., Sarkar, S., Singh, K., Ghatak, P. D., Mathew-Steiner, S. S., et al. (2020). Novel bacterial diversity and fragmented eDNA identified in Hyperbiofilm-forming *Pseudomonas aeruginosa* rugose small Colony variant. *iScience* 23:100827. doi: 10.1016/j.isci.2020.100827
- Donlan, R. M. (2001). Biofilm formation: a clinically relevant microbiological process. *Clin. Infect. Dis.* 33, 1387–1392. doi: 10.1086/322972
- Donlan, R. M. (2002). Biofilms: microbial life on surfaces. *Emerg. Infect. Dis.* 8, 881–890. doi: 10.3201/eid0809.020063
- Dragoš, A., and Kovács, Á. T. (2017). The peculiar functions of the bacterial extracellular matrix. *Trends Microbiol.* 25, 257–266. doi: 10.1016/j.tim.2016.12.010
- Elmassry, M. M., Bisht, K., Colmer-Hamood, J. A., Wakeman, C. A., San Francisco, M. J., and Hamood, A. N. (2021). Malonate utilization by *Pseudomonas aeruginosa* affects quorum-sensing and virulence and leads to formation of mineralized biofilm-like structures. *Mol. Microbiol.* 116, 516–537. doi: 10.1111/mmi.14729
- Flemming, H.-C., Neu, T. R., and Wozniak, D. J. (2007). The EPS matrix: the “house of biofilm cells”. *J. Bacteriol.* 189, 7945–7947. doi: 10.1128/JB.00858-07
- Frimmersdorf, E., Horatzek, S., Pelnikevich, A., Wiehlmann, L., and Schomburg, D. (2010). How *Pseudomonas aeruginosa* adapts to various environments: a metabolomic approach. *Environ. Microbiol.* 12, 1734–1747. doi: 10.1111/j.1462-2920.2010.02253.x
- Frolund, B., Palmgren, R., Keiding, K., and Nielsen, P. H. (1996). Extraction of extracellular polymers from activated sludge using a cation exchange resin. *Water Res.* 30, 1749–1758. doi: 10.1016/0043-1354(95)00323-1
- Gaddy, J. A., Tomaras, A. P., and Actis, L. A. (2009). The *Acinetobacter baumannii* 19606 OmpA protein plays a role in biofilm formation on abiotic surfaces and in the interaction of this pathogen with eukaryotic cells. *Infect. Immun.* 77, 3150–3160. doi: 10.1128/IAI.00096-09
- Hall-Stoodley, L., Costerton, J. W., and Stoodley, P. (2004). Bacterial biofilms: from the natural environment to infectious diseases. *Nat. Rev. Microbiol.* 2, 95–108. doi: 10.1038/nrmicro821
- Hardalo, C., and Edberg, S. C. (1997). *Pseudomonas aeruginosa*: assessment of risk from drinking water. *Crit. Rev. Microbiol.* 23, 47–75. doi: 10.3109/10408419709115130
- He, J., Baldini, R. L., Déziel, E., Saucier, M., Zhang, Q., Liberati, N. T., et al. (2004). The broad host range pathogen *Pseudomonas aeruginosa* strain PA14 carries two pathogenicity islands harboring plant and animal virulence genes. *Proc. Natl. Acad. Sci. U. S. A.* 101, 2530–2535. doi: 10.1073/pnas.0304622101
- Heydorn, A., Nielsen, A. T., Hentzer, M., Sternberg, C., Givskov, M., Ersbøll, B. K., et al. (2000). Quantification of biofilm structures by the novel computer program COMSTAT. *Microbiology (Reading)* 146, 2395–2407. doi: 10.1099/00221287-146-10-2395
- Hmelo, L. R., Borlee, B. R., Almlad, H., Love, M. E., Randall, T. E., Tseng, B. S., et al. (2015). Precision-engineering the *Pseudomonas aeruginosa* genome with two-step allelic exchange. *Nat. Protoc.* 10, 1820–1841. doi: 10.1038/nprot.2015.115
- Hwang, G., Liu, Y., Kim, D., Sun, V., Aviles-Reyes, A., Kajfasz, J. K., et al. (2016). Simultaneous spatiotemporal mapping of in situ pH and bacterial activity within an intact 3D microcolony structure. *Sci. Rep.* 6:32841. doi: 10.1038/srep32841
- Kahl Lisa, J., Eckart, K. N., Morales, D. K., Price-Whelan, A., Dietrich, L. E. P., et al. (2022). Light/dark and temperature cycling modulate metabolic electron flow in *Pseudomonas aeruginosa* biofilms. *MBio* 13, e01407–e01422. doi: 10.1128/mbio.01407-22
- Karygianni, L., Ren, Z., Koo, H., and Thurnheer, T. (2020). Biofilm Matrixome: extracellular components in structured microbial communities. *Trends Microbiol.* 28, 668–681. doi: 10.1016/j.tim.2020.03.016
- Kim, S., Li, X. -H., Hwang, H. -J., and Lee, J. -H. (2020). Thermoregulation of biofilm formation in *Pseudomonas aeruginosa*. *Appl. Environ. Microbiol.* 86, 01584–01520. doi: 10.1128/AEM.01584-20
- Li, W., Siddique, M. S., Graham, N., and Yu, W. (2022). Influence of temperature on biofilm formation mechanisms using a gravity-driven membrane (GDM) system: insights from microbial community structures and metabolomics. *Environ. Sci. Technol.* 56, 8908–8919. doi: 10.1021/acs.est.2c01243
- Liberati, N. T., Urbach, J. M., Miyata, S., Lee, D. G., Drenkard, E., Wu, G., et al. (2006). An ordered, nonredundant library of *Pseudomonas aeruginosa* strain PA14 transposon insertion mutants. *Proc. Natl. Acad. Sci. U. S. A.* 103, 2833–2838. doi: 10.1073/pnas.0511100103
- Madsen, J. S., Lin, Y. C., Squyres, G. R., Price-Whelan, A., de Santiago Torio, A., Song, A., et al. (2015). Facultative control of matrix production optimizes competitive fitness in *Pseudomonas aeruginosa* PA14 biofilm models. *Appl. Environ. Microbiol.* 81, 8414–8426. doi: 10.1128/AEM.02628-15
- Martin, D. W., Schurr, M. J., Mudd, M. H., Govan, J. R., Holloway, B. W., and Deretic, V. (1993). Mechanism of conversion to mucoidy in *Pseudomonas aeruginosa* infecting cystic fibrosis patients. *Proc. Natl. Acad. Sci. U. S. A.* 90, 8377–8381. doi: 10.1073/pnas.90.18.8377
- Maurice, N. M., Bedi, B., and Sadikot, R. T. (2018). *Pseudomonas aeruginosa* biofilms: host response and clinical implications in lung infections. *Am. J. Respir. Cell Mol. Biol.* 58, 428–439. doi: 10.1165/rcmb.2017-0321TR
- Mena, K. D., and Gerba, C. P. (2009). Risk assessment of *Pseudomonas aeruginosa* in water. *Rev. Environ. Contam. Toxicol.* 201, 71–115. doi: 10.1007/978-1-4419-0032-6\_3
- Mozaheb, N., and Mingeot-Leclercq, M.-P. (2020). Membrane vesicle production as a bacterial defense against stress. *Front. Microbiol.* 11:3120. doi: 10.3389/fmicb.2020.600221
- Mulcahy, L. R., Isabella, V. M., and Lewis, K. (2014). *Pseudomonas aeruginosa* biofilms in disease. *Microb. Ecol.* 68, 1–12. doi: 10.1007/s00248-013-0297-x
- Neut, D., Hendriks, J. G. E., van Horn, J. R., van der Mei, H. C., and Busscher, H. J. (2005). *Pseudomonas aeruginosa* biofilm formation and slime excretion on antibiotic-loaded bone cement. *Acta Orthop.* 76, 109–114. doi: 10.1080/00016470510030427
- O'Toole, G. A. (2011). Microtiter dish biofilm formation assay. *J. Vis. Exp.* 47:2437. doi: 10.3791/2437
- Page, A. J., Cummins, C. A., Hunt, M., Wong, V. K., Reuter, S., Holden, M. T. G., et al. (2015). Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31, 3691–3693. doi: 10.1093/bioinformatics/btv421
- Pericolini, E., Colombari, B., Ferretti, G., Iseppi, R., Ardizzone, A., Girardis, M., et al. (2018). Real-time monitoring of *Pseudomonas aeruginosa* biofilm formation on endotracheal tubes in vitro. *BMC Microbiol.* 18:84. doi: 10.1186/s12866-018-1224-6
- Plumley, B. A., Martin, K. H., Borlee, G. I., Marlenee, N. L., Burtneck, M. N., Brett, P. J., et al. (2017). Thermoregulation of biofilm formation in *Burkholderia pseudomallei* is disrupted by mutation of a putative Diguanylate cyclase. *J. Bacteriol.* 199, e00780–e00716. doi: 10.1128/JB.00780-16
- Qiu, D., Damron, F. H., Mima, T., Schweizer, H. P., and Yu, H. D. (2008). PBAD-based shuttle vectors for functional analysis of toxic and highly regulated genes in *Pseudomonas* and *Burkholderia* spp. and other bacteria. *Appl. Environ. Microbiol.* 74, 7422–7426. doi: 10.1128/AEM.01369-08
- Radó, J., Kaszab, E., Petrovics, T., Pászti, J., Kriszt, B., and Szoboszlai, S. (2017). Characterization of environmental *Pseudomonas aeruginosa* using multilocus sequence typing scheme. *J. Med. Microbiol.* 66, 1457–1466. doi: 10.1099/jmm.0.000589
- Sakuragi, Y., and Kolter, R. (2007). Quorum-sensing regulation of the biofilm matrix genes (pel) of *Pseudomonas aeruginosa*. *J. Bacteriol.* 189, 5383–5386. doi: 10.1128/JB.00137-07
- Schmidt, K. D., Tümmler, B., and Römmling, U. (1996). Comparative genome mapping of *Pseudomonas aeruginosa* PAO with *P. aeruginosa* C, which belongs to a major clone in cystic fibrosis patients and aquatic habitats. *J. Bacteriol.* 178, 85–93. doi: 10.1128/jb.178.1.85-93.1996
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153
- She, P., Wang, Y., Liu, Y., Tan, F., Chen, L., Luo, Z., et al. (2019). Effects of exogenous glucose on *Pseudomonas aeruginosa* biofilm formation and antibiotic resistance. *Microbiologyopen* 8:e933. doi: 10.1002/mbo3.933
- Silby, M. W., Winstanley, C., Godfrey, S. A. C., Levy, S. B., and Jackson, R. W. (2011). *Pseudomonas* genomes: diverse and adaptable. *FEMS Microbiol. Rev.* 35, 652–680. doi: 10.1111/j.1574-6976.2011.00269.x

- Soler-Arango, J., Figoli, C., Muraca, G., Bosch, A., and Brelles-Mariño, G. (2019). The *Pseudomonas aeruginosa* biofilm matrix and cells are drastically impacted by gas discharge plasma treatment: a comprehensive model explaining plasma-mediated biofilm eradication. *PLoS One* 14:e0216817. doi: 10.1371/journal.pone.0216817
- Townsley, L., Sison Mangus, M. P., Mehic, S., and Yildiz, F. H. (2016). Response of *Vibrio cholerae* to low-temperature shifts: CspV regulation of type VI secretion, biofilm formation, and association with zooplankton. *Appl. Environ. Microbiol.* 82, 4441–4452. doi: 10.1128/AEM.00807-16
- Townsley, L., and Yildiz, F. H. (2015). Temperature affects c-di-GMP signalling and biofilm formation in *Vibrio cholerae*. *Environ. Microbiol.* 17, 4290–4305. doi: 10.1111/1462-2920.12799
- Turnbull, L., Toyofuku, M., Hynen, A. L., Kurosawa, M., Pessi, G., Petty, N. K., et al. (2016). Explosive cell lysis as a mechanism for the biogenesis of bacterial membrane vesicles and biofilms. *Nat. Commun.* 7:11220. doi: 10.1038/ncomms11220
- Wakeman, C. A., Moore, J. L., Noto, M. J., Zhang, Y., Singleton, M. D., Prentice, B. M., et al. (2016). The innate immune protein calprotectin promotes *Pseudomonas aeruginosa* and *Staphylococcus aureus* interaction. *Nat. Commun.* 7:11951. doi: 10.1038/ncomms11951
- Yan, J., Nadell, C. D., Stone, H. A., Wingreen, N. S., and Bassler, B. L. (2017). Extracellular-matrix-mediated osmotic pressure drives *Vibrio cholerae* biofilm expansion and cheater exclusion. *Nat. Commun.* 8:327. doi: 10.1038/s41467-017-00401-1
- Yi, X., Chen, Y., Cai, H., Wang, J., Zhang, Y., Zhu, Z. Q., et al. (2022). The temperature-dependent expression of type II secretion system controls extracellular product secretion and virulence in mesophilic *Aeromonas salmonicida* SRW-OG1. *Front. Cell. Infect. Microbiol.* 12. doi: 10.3389/fcimb.2022.945000



## OPEN ACCESS

## EDITED BY

Graciela L. Lorca,  
University of Florida,  
United States

## REVIEWED BY

Barbara Drigo,  
University of South Australia,  
Australia  
Xiaoyong Yang,  
Technical University of Denmark, Denmark

## \*CORRESPONDENCE

Niculina Musat  
✉ niculina.musat@ufz.de

## SPECIALTY SECTION

This article was submitted to  
Microbial Physiology and Metabolism,  
a section of the journal  
Frontiers in Microbiology

RECEIVED 30 September 2022

ACCEPTED 10 January 2023

PUBLISHED 25 January 2023

## CITATION

Haenelt S, Wang G, Kasmanas JC, Musat F,  
Richnow HH, da Rocha UN, Müller JA and  
Musat N (2023) The fate of sulfonamide  
resistance genes and anthropogenic pollution  
marker *int1* after discharge of wastewater into  
a pristine river stream.  
*Front. Microbiol.* 14:1058350.  
doi: 10.3389/fmicb.2023.1058350

## COPYRIGHT

© 2023 Haenelt, Wang, Kasmanas, Musat,  
Richnow, da Rocha, Müller and Musat. This is  
an open-access article distributed under the  
terms of the [Creative Commons Attribution  
License \(CC BY\)](#). The use, distribution or  
reproduction in other forums is permitted,  
provided the original author(s) and the  
copyright owner(s) are credited and that the  
original publication in this journal is cited, in  
accordance with accepted academic practice.  
No use, distribution or reproduction is  
permitted which does not comply with these  
terms.

# The fate of sulfonamide resistance genes and anthropogenic pollution marker *int1* after discharge of wastewater into a pristine river stream

Sarah Haenelt<sup>1</sup>, Gangan Wang<sup>1</sup>, Jonas Coelho Kasmanas<sup>2</sup>,  
Florin Musat<sup>1,3</sup>, Hans Hermann Richnow<sup>1,4</sup>,  
Ulisses Nunes da Rocha<sup>2</sup>, Jochen A. Müller<sup>2,5</sup> and Niculina Musat<sup>1\*</sup>

<sup>1</sup>Department of Isotope Biogeochemistry, Helmholtz Centre for Environmental Research, Leipzig, Germany,

<sup>2</sup>Department of Environmental Microbiology, Helmholtz Centre for Environmental Research, Leipzig, Germany, <sup>3</sup>Department of Molecular Biology and Biotechnology, Faculty of Biology and Geology, Babeş-Bolyai University, Cluj-Napoca, Romania, <sup>4</sup>Isodetect Umweltmonitoring GmbH, Leipzig, Germany, <sup>5</sup>Institute for Biological Interfaces (IBG 5), Karlsruhe Institute of Technology, Eggenstein-Leopoldshafen, Germany

**Introduction:** Currently there are sparse regulations regarding the discharge of antibiotics from wastewater treatment plants (WWTP) into river systems, making surface waters a latent reservoir for antibiotics and antibiotic resistance genes (ARGs). To better understand factors that influence the fate of ARGs in the environment and to foster surveillance of antibiotic resistance spreading in such habitats, several indicator genes have been proposed, including the integrase gene *int1* and the sulfonamide resistance genes *sul1* and *sul2*.

**Methods:** Here we used quantitative PCR and long-read nanopore sequencing to monitor the abundance of these indicator genes and ARGs present as class 1 integron gene cassettes in a river system from pristine source to WWTP-impacted water. ARG abundance was compared with the dynamics of the microbial communities determined via 16S rRNA gene amplicon sequencing, conventional water parameters and the concentration of sulfamethoxazole (SMX), sulfamethazine (SMZ) and sulfadiazine (SDZ).

**Results:** Our results show that WWTP effluent was the principal source of all three sulfonamides with highest concentrations for SMX (median 8.6ng/l), and of the indicator genes *sul1*, *sul2* and *int1* with median relative abundance to 16S rRNA gene of 0.55, 0.77 and 0.65%, respectively. Downstream from the WWTP, water quality improved constantly, including lower sulfonamide concentrations, decreasing abundances of *sul1* and *sul2* and lower numbers and diversity of ARGs in the class 1 integron. The riverine microbial community partially recovered after receiving WWTP effluent, which was consolidated by a microbiome recovery model. Surprisingly, the relative abundance of *int1* increased 3-fold over 13km of the river stretch, suggesting an internal gene multiplication.

**Discussion:** We found no evidence that low amounts of sulfonamides in the aquatic environment stimulate the maintenance or even spread of corresponding ARGs. Nevertheless, class 1 integrons carrying various ARGs were still present 13km downstream from the WWTP. Therefore, limiting the release of ARG-harboring microorganisms may be more crucial for restricting the environmental spread of antimicrobial resistance than attenuating ng/L concentrations of antibiotics.

## KEYWORDS

class 1 integron, sulfamethoxazole, sulfonamide resistance, *sul1*, *sul2*, *int1*, river ecosystem, one health

## Introduction

The environmental fate of antimicrobial resistance is a crucial component of the One Health concept (McEwen and Collignon, 2018; The World Health Organization, The Food and Agriculture Organization of the United Nations, and The World Organisation for Animal Health, 2019). Surface waters are among the environmental reservoirs and potential dissemination avenues for antibiotic resistant bacteria and antimicrobial resistance genes (ARGs; Larsson and Flach, 2022). The close connection between such water bodies, animals, and humans entails the risk of cross-transmission, which ultimately jeopardizes the successful treatment of infectious diseases with antibiotics (Collignon et al., 2018). Therefore, monitoring the fate of ARGs in surface waters is an important component in assessing the risk of antimicrobial resistance spread in the environment to human and animal health (Berendonk et al., 2015; Wang et al., 2021).

The fate of ARGs in surface waters and other environmental compartments is likely influenced by their frequent association with mobile genetic elements such as transposons and plasmids (Yao et al., 2022). Such elements may promote ARG transfer from antibiotic resistant bacteria stemming from humans and warm-blooded animals to riverine microorganisms, eventually leading to the spread of resistance among aquatic bacterial communities (Koczura et al., 2016; Chaturvedi et al., 2021a). The mobility of ARGs can be further enhanced when they are associated with integrons, which are powerful tools for microorganisms to adapt quickly to changes in their environment (Mazel, 2006). Integrons can be located on mobile genetic elements or on the chromosome. They can acquire ARGs, heavy metal and disinfectant resistance genes (Gillings, 2018). The class 1 integron, the best-studied integron variant, consists of two conserved segments flanking a variable region (Supplementary Figure S1). The 5' conserved segment (5'CS) contains the *intI1* gene, encoding for an integrase enzyme, an *attI* recombination site, and a promoter governing the expression of genes cassettes in the variable region. Gene cassettes or any DNA fragment carrying an *attC* recombination site can be inserted into or excised from the variable region by the integrase. The insertion mechanism can lead to gene cassette arrays, making the class 1 integron a carrier of multi-resistance (Cambray et al., 2010). The 3' conserved segment (3'CS) harbors the *qacEΔ1* gene, encoding for resistance against quaternary ammonium compounds, and the *sul1* gene, conferring sulfonamide resistance. The relative abundance of the so-called clinical *intI1* variant of class 1 integron has increased in the natural environment worldwide due to human activities, including wastewater discharge and selection pressure from antibiotic use (Gillings et al., 2008). The *in situ* copy number of this gene is used as a proxy for anthropogenic pollution in the environment (Gillings et al., 2015). Given the co-occurrence of *intI1* and *sul1* in the class 1 integron, the abundance of these genes shows frequently a strong positive correlation in impacted environments (Gillings et al., 2015). Like *intI1*, *sul1* and its close relative *sul2* are used as indicators for the level of pollution with antimicrobial agents in the environment (Pruden et al., 2006; Adelowo et al., 2018; Lee et al., 2021; Chaturvedi et al., 2021b). Both *sul* genes are often located on mobile genetic elements such as plasmids, which may promote their dissemination in the environment (Chaturvedi et al., 2021b).

Prominent interfaces between human activities and surface waters are wastewater treatment plants (WWTPs), whose effluents increase the amounts of ARGs and antibiotic resistant bacteria in receiving waterbodies (Hernández et al., 2019; Wang et al., 2021). WWTPs also discharge a measurable amount of antibiotics into the aquatic

environment. Sulfonamides like sulfamethoxazole (SMX) are commonly detected in WWTP effluents with average concentrations ranging from tens to hundreds of ng/L, while peak concentrations can even be in the low µg/L range (Carvalho and Santos, 2016). In the absence of extreme environmental conditions, they are chemically rather stable in surface waters (Straub, 2016). Together with their continuous presence in WWTP discharge, this leads to typical mean concentrations in receiving water bodies in the low ng/L range, with occasionally low µg/L concentrations (Martin-Laurent et al., 2019). It is an important question whether continuous sulfonamide contamination affects the spread of the corresponding ARGs and microbial community structure in riverine ecosystems and at what concentrations such effects are detectable. According to the European Committee on Antimicrobial Susceptibility Testing (EUCAST), the lowest minimal inhibitory concentration MIC for SMX in any microbial species is around 1 mg/l, thus  $10^4$ – $10^6$  times higher than typical mean concentrations in surface waters (Bengtsson-Palme and Larsson, 2016). However, recent studies have also evaluated the predicted no-effect concentration (PNEC) for selection of sulfonamide resistance. The PNEC varies between 0.6 and 16 µg/l, with the first one being at the high end of SMX concentrations found in WWTP effluents (Bengtsson-Palme and Larsson, 2016; Le Page et al., 2017). In a recent study conducted by Borsetto et al. (2021), the effect of SMX on the microbiome of a WWTP-impacted river was tested over 24 days using an artificial river water flume system. The addition of 4 µg/l SMX did not lead to the spread of sulfonamide resistance genes in this particular experimental set-up. However, a potential draw-back of such experiments with a microbial community already influenced by WWTP effluent containing sulfonamides and other antimicrobials is, that short-term concentration increases may not generate detectable effects. Therefore, it remains uncertain whether SMX has an impact on riverine microbial communities at typical surface water concentrations, including whether it promotes the spread of ARGs in the environment (Cairns et al., 2018).

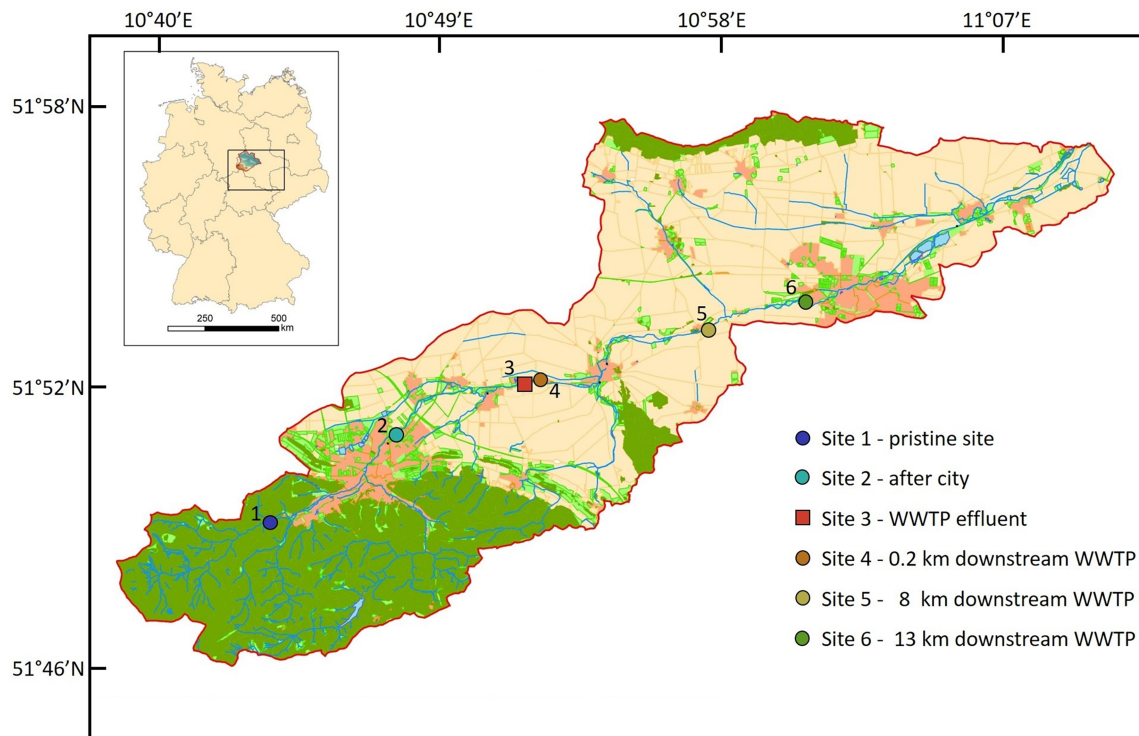
Here, we investigated how the microbial community of a pristine river is affected by WWTP effluent containing sulfonamides and the corresponding ARGs. As a model site, we selected the Holtemme river, which flows from the Harz mountains to the central German plain and has been well characterized within the Terrestrial Environmental Observatories (TERENO) (Zacharias et al., 2011; Wollschläger et al., 2017). We profiled the microbial community, monitored the abundance of *sul1*, *sul2* and *intI1*, and determined sulfonamide concentrations along the gradient of anthropogenic impact on the river. Furthermore, we carried out long-read nanopore sequencing of gene cassettes of the clinical class 1 integron in river samples to better understand how the integron contributes to the local prevalence of ARGs.

## Materials and methods

### Study region

This study was conducted on the Holtemme river in Saxony-Anhalt (Germany), which is serving as a landscape model for studies on the effects of anthropogenic pollution on riverine ecosystem functioning (Inostroza et al., 2016; Beckers et al., 2018; Brase et al., 2018). The third order river is 47 km long, originating from the low mountain range of Harz (862 m a.s.l), and drains a catchment area of 278 km<sup>2</sup>. After leaving the Harz national park, the river unites with a small stream called Zillierbach and passes through the city of Wernigerode (32,000





**FIGURE 1**  
Location of all six sites in the catchment of the Holtemme river in the Harz mountain region, Germany. Map created from Authoritative Topographic–Cartographic Information System (ATKIS).

inhabitants), farmed plains, some smaller villages and the city of Halberstadt (39,000 inhabitants). There are two activated sludge-based, tertiary WWTPs, treating urban wastewater and discharging into the Holtemme, one just downstream Wernigerode (WWTP Silstedt) and one downstream Halberstadt. Two small brooks enter the Holtemme upstream and three more downstream from the WWTP Silstedt, all of which contribute very little to the water volume of Holtemme river. Downstream the mountainous region, the Holtemme is often rectified and canalized (Kärthe et al., 2017). The land use in the river catchment is dominated by agriculture (60% of the catchment area), followed by forests (30%), and urban areas (10%). The above mentioned features render the Holtemme an excellent model system for studying the effect of anthropogenic pollutions on a pristine river ecosystem (Wollschläger et al., 2017).

## Water sampling and sample preparation

Surface water samples were collected at six different Sites along the river stream (Figure 1). Site 1 (51°49′00.9″N 10°43′28.6″E) in the mountain region is a pristine Site with minor human impact. Site 2 (51°50′49.6″N 10°47′29.3″E) is located in the city of Wernigerode, Site 3 (51°51′59.4″N 10°51′43.3″E) is the discharge water of WWTP Silstedt, whereas Site 4 (51°51′59.0″N 10°51′55.5″E) is located 250 m downstream the WWTP. Site 5 (51°53′05.8″N 10°57′47.5″E) and Site 6 (51°53′46.5″N 11°01′03.5″E) are 8 and 13 km downstream Site 3 before the Holtemme enters the city of Halberstadt. Site 1–5 were sampled on December 7, 2020, and weekly from January 4 through February 1, 2021. Additional daily samples were collected from Site 1 to 6 from February 24 through

March 4, 2021. Apart from eight unusually cold days in February 2021, precipitation and temperature patterns were representative for a winter season in this region (Supplementary Figure S2). No sampling was conducted during the unusually cold period in February 2021. For Sites 1 and 5, water flow rates are available from official measurement points of the state Saxony-Anhalt. At Site 1, rates varied between 0.063 and 0.119 m<sup>3</sup>/s with an average of 0.1 m<sup>3</sup>/s, and at Site 5 the average flow rate was 1 m<sup>3</sup>/s. At Site 3 (WWTP effluent) it ranged between 0.104 and 0.19 m<sup>3</sup>/s during the sampling period and values were provided by the WWTP Silstedt.

Water samples were collected in sterile 1 l glass flasks after pre-rinsing with river water by immersing the bottles into the river at a water depth of approximately 20 cm. Water samples were collected in triplicates at each Site. Afterwards, the samples were transported to the laboratory within 3 h in a thermostable box. At Sites 1 and 5, physicochemical parameters pH, water temperature [°C], conductivity [mS/cm] and NO<sub>3</sub><sup>-</sup>-N [mg/L] were obtained in the MOBICOS monitoring stations located at the river (Fink et al., 2020). Physicochemical parameters from Site 3 (WWTP effluent) were kindly provided by the WWTP Silstedt.<sup>1</sup> At Sites 2, 4, and 6, pH and temperature were measured using a SenTix41 probe (Xylem Analytics, Germany). Immediately after arriving in the laboratory, samples were prepared for further analyses. A volume of 500 ml of water was filtered (PSE PALL filters, diameter 47 mm, pore size 0.22 μm, Pall Corporation, NY United States) using a vacuum pump. Three filters per sample were

<sup>1</sup> <https://www.wahb.eu/>

prepared. One filter was used for DNA extraction and two were frozen at  $-20^{\circ}\text{C}$  as back-up. For each sample, 100 ml of flow-through were collected and stored at  $4^{\circ}\text{C}$  overnight for the measurement of sulfonamide concentration.

## Solid phase extraction (SPE) and HPLC-MS analysis

Sulfamethoxazole (SMX), sulfadiazine (SDZ) and sulfamethazine (SMZ) concentrations were measured in seven replicates per Site (Supplementary Table S1). To concentrate the water samples, solid phase extraction was carried out using Oasis HLB 6cc-500 mg columns (Waters, CT United States). Cartridges were conditioned with 4 ml of methanol (Chemsolute, Germany) and equilibrated with 4 ml of  $\text{H}_2\text{O}$  (Chemsolute, Germany). Filtered river samples (100 ml) were loaded onto the column at a flow rate of 1 ml/min using a vacuum pump. After sample loading was completed, cartridges were dried for 5 min under vacuum. Analytes were eluted with 4 ml of methanol. Eluents were evaporated under a gentle stream of pure nitrogen until complete dryness. The analytes were then re-dissolved in 900  $\mu\text{l}$  of  $\text{H}_2\text{O}$  with 0.1% (v/v) formic acid (Serva, Germany) and 100  $\mu\text{l}$  of methanol and transferred to HPLC vials. Vials were kept at  $-20^{\circ}\text{C}$  until measurement at HPLC-MS.

For HPLC-MS analysis a stock solution of 10 mg/l mixed standard with SMX (Sigma-Aldrich, MO United states), SDZ (Alfa Aesar, MA USA) and SMZ (Sigma-Aldrich, MO, United states) was prepared in pure methanol and stored at  $-20^{\circ}\text{C}$ . For the calibration curve, working concentrations of 0.01, 0.025, 0.05, 0.075, 0.1, 0.25, 0.5, 0.75, 1, and 2  $\mu\text{g/l}$  in 900  $\mu\text{l}$   $\text{H}_2\text{O}$  with 0.1% formic acid and 100  $\mu\text{l}$  of methanol were used. The chromatographic separation was done using a Zorbax Eclipse Plus Rapid Resolution HT-C18 (100 mm  $\times$  3 mm, 1.8  $\mu\text{m}$ ) column on a 1,260 Infinity II HPLC (Agilent Technologies, CA, United states). Mobile phase A was  $\text{H}_2\text{O}$  with 0.1% formic acid, while mobile phase B consisted of methanol with 0.1% formic acid. The gradient was: 90% A from 0 to 2 min, linear decrease to 40% A from 2 to 3 min, linear decrease to 0% A from 3 to 8 min, 0% A from 8 to 9 min, increase to the original condition of 90% A from 9 to 9.1 min, 90% A from 9.1 to 16 min. The injection volume was 50  $\mu\text{l}$ , HPLC was conducted under a constant flow rate of 0.4 ml/min and column temperature kept constant at  $30^{\circ}\text{C}$ .

The HPLC system was coupled to a QTRAP 6500+ MS/MS with linear ion trap (AB Sciex, United Kingdom). Analysis was performed in multiple reaction monitoring mode (MRM) with positive ion mode and electrospray ionization (ESI). Ion source parameters were curtain gas CUR = 30 psi, ion source gas 1 GS1 = 60 psi, ion source gas 2 GS2 = 60 psi, ion spray voltage IS = 5,000 V, collision gas CAD = medium and temperature TEM =  $450^{\circ}\text{C}$ . Dwell time was 40 ms. MRM transitions and optimized MS parameters can be found in Supplementary Table S2. All samples and standards were measured in three technical replicates.

## DNA extraction and 16S rRNA gene amplicon sequencing

For DNA extraction, one filter per sample was halved and then cut into smaller slices. Each half of the filter was placed into a separate 2 ml tube containing approx. 0.15 g of 1 mm diameter zirconium beads (Carl Roth, Germany) and 200  $\mu\text{l}$  BE buffer (Macherey Nagel, Düren, Germany). After agitating for 20 s at 4 m/s (FastPrep 24, MP Biomedicals,

Germany), tubes were centrifuged in a bench-top centrifuge (neoLab, Germany) and the liquid phase from both tubes was pooled. Additional 100  $\mu\text{l}$  BE buffer were added to the tubes containing filters, vortexed and centrifuged. The liquid phases were combined to get maximum DNA yield. DNA extraction was further conducted following the manufacturers protocol (NucleoSpin Microbial DNA Kit, Macherey Nagel, Düren, Germany). DNA extracts were used for quantitative PCR (qPCR), as well as for sequencing of 16S rRNA gene amplicons and *intI1* gene cassettes.

To determine the diversity of the microbial community alongside the river flow, a 16S rRNA gene amplicon library was prepared for each Site from five sampling days each (Supplementary Table S1). Primers sequences for PCR-amplification of the variable region V3 of the 16S rRNA gene are given in Supplementary Table S3. For the first PCR, 1  $\mu\text{l}$  of template (pure to 1:100 dilutions in  $\text{ddH}_2\text{O}$ ) was used in 12.5  $\mu\text{l}$  of repliQa HiFi ToughMix (Quantabio, MA, United states), 0.75  $\mu\text{l}$  of each primer 27F and 1492R (both at 10  $\mu\text{M}$ ) and 10  $\mu\text{l}$   $\text{ddH}_2\text{O}$ . The PCR program was [ $98^{\circ}\text{C}$  for 10 s,  $55^{\circ}\text{C}$  for 5 s,  $68^{\circ}\text{C}$  for 8 s]  $\times$  30,  $8^{\circ}\text{C}$  hold. PCR was carried out on a peqSTAR 2  $\times$  Universal gradient (VWR, PA, United states). For the nested PCR, 1.5  $\mu\text{l}$  PCR product was used in 25  $\mu\text{l}$  NEBNext Ultra II Q5 Master Mix (New England Biolabs, Germany), 2.5  $\mu\text{l}$  of each primer 341F-TS and 518R-TS (both at 10  $\mu\text{M}$ ) and 20  $\mu\text{l}$   $\text{ddH}_2\text{O}$ . The PCR program was  $98^{\circ}\text{C}$  for 30 s, [ $98^{\circ}\text{C}$  for 20 s,  $51^{\circ}\text{C}$  for 30 s,  $72^{\circ}\text{C}$  for 30 s]  $\times$  35,  $72^{\circ}\text{C}$  for 30 s,  $8^{\circ}\text{C}$  hold. All PCR products were cleaned using 0.9x AmpureXP Beads (Beckmann Coulter, CA, United states) according to the manufacturers protocol and re-dissolved in 25  $\mu\text{l}$  of 5 mM TE buffer. The samples were sequenced on an Illumina NextSeq 550 machine using the NextSeq 500/550 High Output Kit v2.5. For data analysis with the software package QIIME 2 (Bolyen et al., 2019; RRID: SCR\_021258), all fastq files were imported as paired-end sequences, primers were trimmed and reads merged using DADA2 (Callahan et al., 2016). Classification was done in QIIME 2 using the pre-trained classifier “Silva 138 99% OTU full-length sequences” (Bokulich et al., 2018; Robeson et al., 2021). Afterwards, data was exported and manually curated in R using phyloseq (McMurdie and Holmes, 2013; RRID: SCR\_013080). Before further analysis, one outlying dataset was removed for Sites 1–4 (Supplementary Figure S3). Rarefaction curves before and after rarefying the reads can be found in Supplementary Figure S4. Alpha diversity was assessed by computing the Observed Richness and Shannon diversity indices. For beta diversity, a non-metric multidimensional scaling (NMDS) plot was built from Bray-Curtis dissimilarities using the R package vegan (Oksanen et al. (2022), RRID: SCR\_011950). Statistical analyses were performed using Kruskal-Wallis and Dunn’s tests in R. In addition, the beta diversity differences from all Sites were measured against the pristine Site 1. To do so, we first calculated the centroid from the NMDS beta diversity plot for all Sites and then calculated the distance between the centroid of each Site and the pristine Site 1. Furthermore, we applied the microbiome recovery model described by Shaw et al. (2019). In short, we used equation 1 to model the beta-diversity distance to Site 1.

$$x(d) = \frac{De^{\Phi_1}e^{\Phi_2}}{e^{\Phi_2} - e^{\Phi_1}} \times \left( e^{-e^{\Phi_1}d} - e^{-e^{\Phi_2}d} \right) + \text{Asym} * \left( 1 - e^{-e^{\Phi_1}d} \right) \quad (1)$$

Here,  $d$  refers to the geographical distance from each Site to Site 1 in km,  $x(d)$  is the calculated distance from the centroid in the NMDS plot to the pristine centroid,  $e$  is the Euler’s number,  $D$  is the magnitude of displacement,  $\Phi$  refers to variables derived by the damping on the system and the strength of the restoring force, and *Asym* is the value

where the equilibrium asymptotically tends to stabilize. A Bayesian framework was applied to fit Equation 1, using the R package *rstan* (Stan Development Team. 2022. Stan Modeling Language Users Guide and Reference Manual, v.2.21.7.)<sup>2</sup> as described previously (Shaw et al., 2019). Aside from  $d$  and  $x(d)$ , all variables were adjusted by the model. We used a uniformly distributed prior for  $D$ , uniform priors for  $\Phi_1$ ,  $\Phi_2$ , and uniform distribution for  $Asym$ , as shown in equations 2.

$$D \sim \text{uniform}(0,10) \quad (2.1)$$

$$\Phi_1 \sim \text{uniform}(-0.99,0.99) \quad (2.2)$$

$$\Phi_2 \sim \text{uniform}(-1,1) \quad (2.3)$$

$$Asym \sim \text{uniform}(0,1) \quad (2.4)$$

We further investigated the phylotypes with >1% frequency on a family level to identify phylotypes specific for each individual Site.

## Quantitative PCR

The absolute abundances of *sul1*, *sul2*, *intI1* and the bacterial 16S rRNA gene were quantified by SYBR Green-based qPCR with established primers (Supplementary Table S3) and three technical replicates per sample for all sampling dates and Sites essentially as described before (Adelowo et al., 2020). As calibration standards we used a 16S rRNA gene fragment from *Pseudomonas putida* KT2440, inserted into plasmid pGEM (Invitrogen) and cloned in *Escherichia coli* JM109, and purified PCR products obtained from *Citobacter* sp. strain EC35 for *sul1* and *sul2*, and from *Thauera aromatica* K172 for *intI1* (both strains from UFZ culture collection). The PCR mixture contained 6.25 µl KAPA SYBR® FAST (Sigma-Aldrich, MO, United States), 4.75 µl ddH<sub>2</sub>O, 0.25 µl each of forward and reverse primer (both at 10 µM) and 1 µl of template. To compensate for PCR biases, 1:10 dilutions of each sample in ddH<sub>2</sub>O were prepared in transparent 96 Fast PCR-Plates (Sarstedt, Germany). Measurements were done on a StepOnePlus Real Time PCR System with software version 2.1 (AB Applied Biosystems, MA, United States). Absolute abundances were normalized to copy numbers/100 ml. The relative abundances of ARGs and the *intI1* gene were calculated as absolute abundance divided by absolute abundance of the 16S rRNA gene.

## Sequencing of class 1 integron gene cassettes

The long-read nanopore sequencing method (Oxford Nanopore Technology—ONT; Oxford, United Kingdom) was chosen to analyze the whole length of gene cassette inserts in the class 1 integron. The

region between 5'CS and 3'CS of class 1 integron in the extracted DNA was inspected *via* PCR followed by long-read sequencing with ONT. For each Site, extracted DNA from nine different sampling dates was chosen (Supplementary Table S1). A volume of 2 µl of DNA template was used in 22.5 µl RedTaq 1.1 × (VWR, PA, United States), 0.5 µl each of primers 5'CS and 3'CS (both at 10 µM, Supplementary Table S3). The PCR program was: 95°C for 2 min, [95°C for 30 s, 55°C for 60 s, 72°C for 60 s] × 35, 72°C for 5 min. For Sites 1 and 2, no PCR product was visible in agarose gel electrophoresis. Therefore, these samples were not processed further. For Sites 3 to 6, band patterns observed in agarose gel were identical between all sampling dates, hence replicates were pooled before the clean-up step. PCR products were cleaned with 1.8× AmpureXP Beads, re-dissolved in 25 µl 1 mM Tris buffer and prepared for ONT sequencing using the Rapid Barcoding kit (ONT, United Kingdom) with AMPure XT beads (Beckman Coulter, CA, United States) following the manufacturer's protocol. Flongle flow cells (ONT, United Kingdom) were loaded with 15 fmol of the total sequencing library after priming with the Flongle sequencing expansion kit (ONT, United Kingdom) following the manufacturer's protocol. ONT sequencing was carried out on a MinION MK1b with MinKNOW software (19.06.8) for 24 h with standard settings applied. Base-calling and de-multiplexing for converting ONT fast5 to fastq files were conducted using Guppy (v3.6.0; ONT, United Kingdom). Data analysis was handled in Geneious Prime (2022.0.1, Geneious, New Zealand). Barcodes were removed by trimming 110 bp on the 5' end of each read. The Comprehensive Antibiotic Resistance Database CARD (Alcock et al., 2020) was used as a reference database from which available ARG nucleotide sequences were imported. Furthermore, we imported the sequences of 5'CS and 3'CS from the INTEGRALL database (Moura et al., 2009). All trimmed reads were then annotated from the in-house reference database choosing standard settings ("Index Length" 10 and "Best Match Criteria" 75%) and a sequence similarity of 70%, to account for the error rate of ~10–15% of nanopore sequencing. Non-annotated reads were checked manually using blastn with standard settings. Reads with query coverage <40% were defined as "low quality or potential artefact" and removed (7.7% of total reads). For a graphical overview of the whole workflow, refer to Supplementary Figure S5.

## Results

### Physicochemical parameters

This study was carried out in the winter season 2020/21. The average water temperature at sampling Sites varied from 3.3°C at Sites 1 and 2 located upstream from the WWTP discharge point to 10.7°C at Site 3, the WWTP effluent, and 5°C at downstream Sites 4 to 6. The pH value at Sites 1 and 2 was on average 8.3 and 8.0, respectively. The WWTP effluent (pH 7.0) only slightly altered the pH of the river at downstream Sites 4, 5 and 6 to an average of 7.7. Nitrate concentration upstream the WWTP effluent ranged between 0.9 to 3.9 mg/l (Site 1) while downstream varied between 2.3 and 3.8 mg/l (Site 5). The nitrate concentration in the WWTP effluent was on average 3.9 mg/l, and did not affect significantly the nitrate concentration in the river (Supplementary Figure S6). The nitrate concentration in the Holtemme river were similar to those reported previously for the winter season (Brase et al., 2018), indicating that there was no substantial pollution of river water by agricultural runoff. Ammonium concentrations were low in the WWTP effluent throughout the sampling period (0.016–0.044 mg/l).

<sup>2</sup> <https://mc-stan.org>



The concentration of SMX was highest in the WWTP effluent (Site 3) with a median concentration of 8.6 ng/l and a peak value of 27.3 ng/l (Figure 2), which was in the low range of previous SMX quantifications in effluent of other WWTPs (Carvalho and Santos, 2016; Paumelle et al., 2020). Downstream Sites 4–6 showed a steady decrease of SMX concentrations reaching below 3 ng/l at Site 6. The SMX concentration at Site 4 was about 35% of the SMX concentration at Site 3. Thus we estimate that the wastewater contributed approximately one third to the river water in the immediate vicinity (200 m downstream), which fits well with calculations done by Krauss et al. (2019), who estimated a fraction of wastewater of 27% in the Holtemme river. SDZ and SMZ concentrations were likewise highest in the WWTP effluent, however at considerably lower concentrations than SMX with median values <1 ng/l (Supplementary Figure S7). Previously, SMX and SMZ as well as sulfapyridine were quantified ~1 km downstream from our Site 4. Median SMX concentrations were about double than those in our study, while SMZ was not detectable. Sulfapyridine, which was not measured in our study, was found at a median concentration of 65.3 ng/l (Tousova et al., 2017).

## Microbial community structure

Microbial community profiling by 16S rRNA gene amplicon sequencing showed the lowest species richness according to the number of observed amplicon sequence variants (ASVs) for Site 1, with an average of 757. This value was constantly increasing along the flow path up to a number of 1,658 ASVs for Site 6 (Figure 3A). The increase in species richness was highest at Sites 3 to 4 suggesting a mixture of distinct microbial communities from WWTP and river waters. The Shannon index estimates the entropy of a system by combining species

richness and evenness (Allaby, 2009). Comparable to the number of observed ASVs, the Shannon index of the river communities (excluding Site 3) was constantly increasing from Site 1 (Shannon index 6) to Site 6 (Shannon index 6.6). The average Shannon index was lowest in the wastewater effluent with a value of 5.6 (Figure 3B).

We compared the microbial diversity at all Sites using NMDS of Bray-Curtis dissimilarities. A clear separation was observed between the pristine Sites 1 and 2 on one hand, and the WWTP-impacted Sites 3 to 6 on the other hand (Figure 3C). Furthermore, downstream the discharge point of WWTP, beta diversity tended to return to the pre-discharge state. When comparing Bray-Curtis dissimilarities to those of Site 1, the differences in beta diversity were significant (Dunn's test,  $p \leq 0.05$ ). Compared to Site 2, differences in beta diversity were significant for all Sites except Sites 5 and 6 (Supplementary Figure S8). These results indicate that microbial community partially recovered to its pristine state after discharge of wastewater into the river. To test this hypothesis, we visualized the microbial community's recovery after discharge of WWTP effluent into the pristine river using a microbiome recovery model (Shaw et al., 2019). The model also showed that the microbial community was recovering after wastewater discharge yet did not fully revert to its original composition (Figure 3D).

The analysis of the taxonomic composition revealed a coherent picture of a river system with a gradient from pristine to effluent-impacted water. The family of *Caulobacteraceae*, which include many oligotrophic members, was the most abundant phylotype at pristine Site 1. Other highly abundant phylotypes at this Site were affiliated with the *Comamonadaceae* and *Sphingomonadaceae*, typical inhabitants of natural river ecosystems. Phylotypes specific for Site 1 belonged to the *Caulobacteraceae*, *Nocardiaceae* and *Rhizobiales*. The increased occurrence of phylotypes associated with the *Clostridia* taxonomic class,

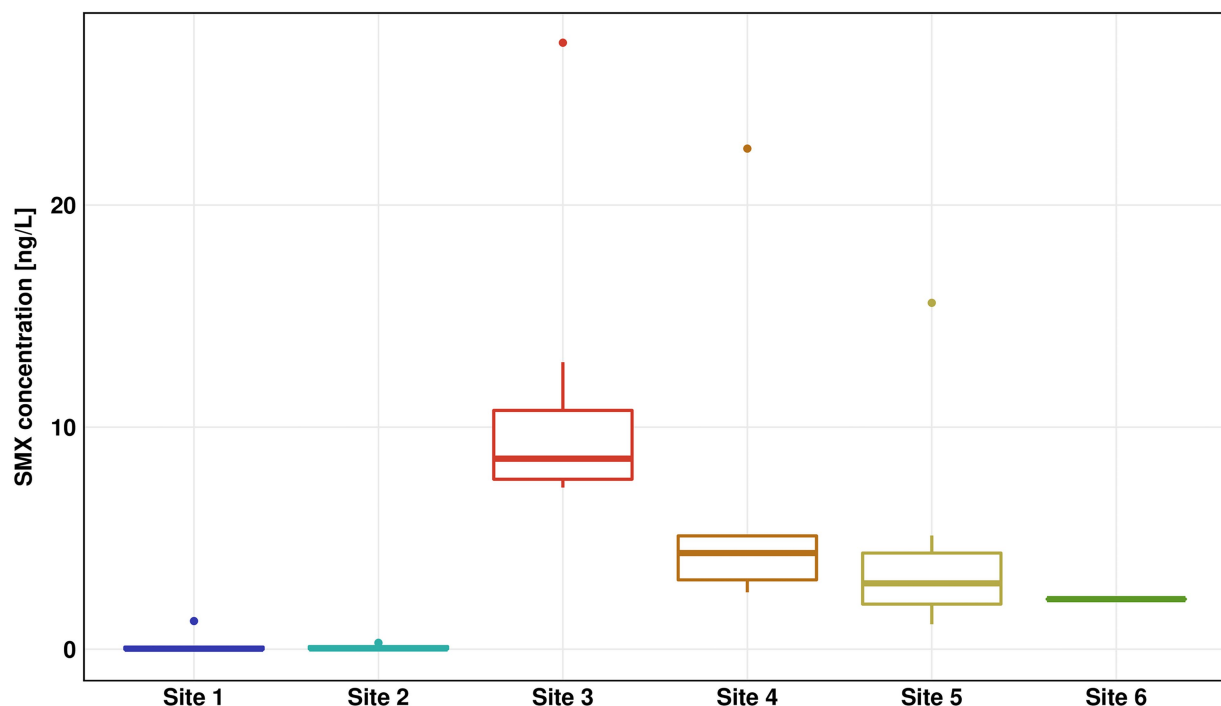
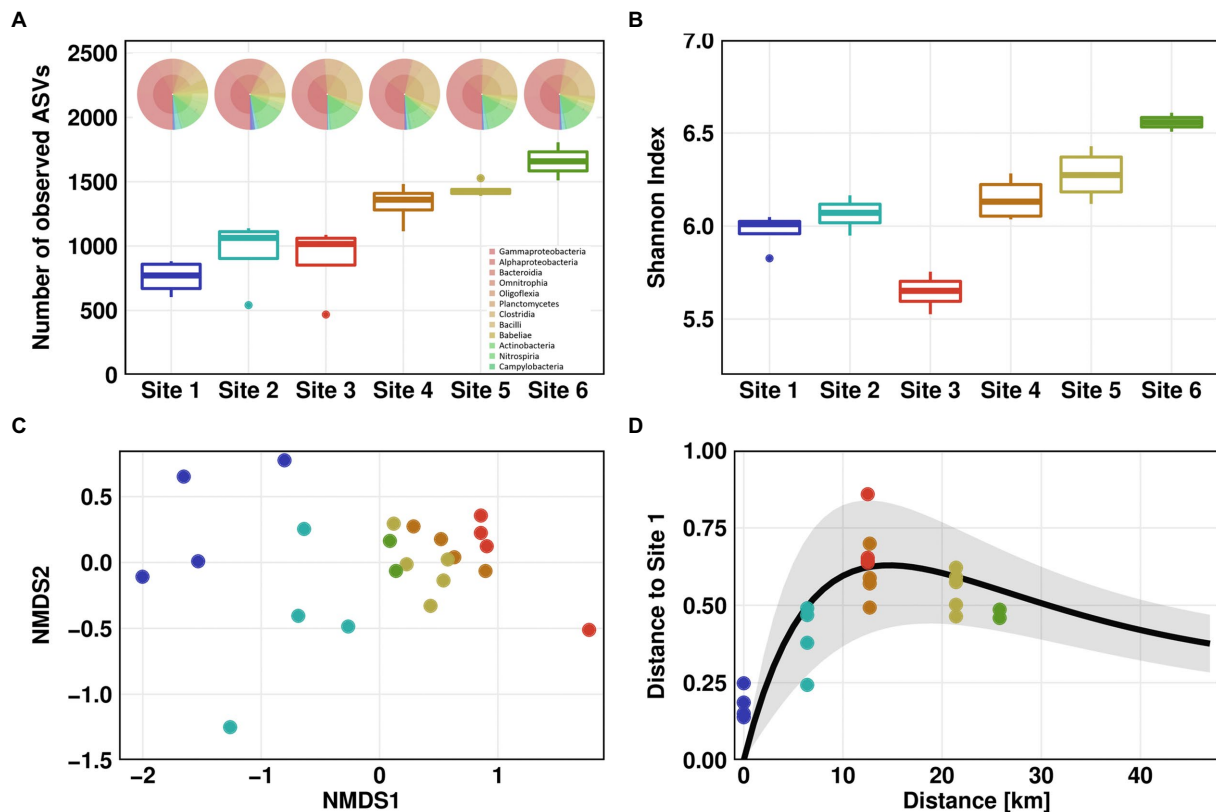


FIGURE 2  
Sulfamethoxazole (SMX) concentration for all six sites measured by high performance liquid chromatography–mass spectrometry (HPLC-MS)/mass spectrometry (MS). The results of seven replicates are shown using box-whiskers-plots with the median represented by a horizontal line. Colors are representing the different sites (Site 1 in blue, Site 2 in turquoise, Site 3 in red, Site 4 in orange, Site 5 in yellow, and Site 6 in green).





**FIGURE 3**  
Microbial community structure. Microbial community alpha diversity based on 16S rRNA amplicon sequencing data using Observed (A) and Shannon measurement (B). KRONA charts showing microbial community structure on genus level are displayed on top of (A) for February 26th, 2021. Microbial community beta diversity based on 16S rRNA amplicon sequencing data, ordination by non-metric multidimensional scaling (NMDS) of Bray–Curtis dissimilarities (C). Impulse model showing the displacement from the pristine state at Site 1 (D). For Sites 1–4, one outlier was removed in each dataset before analysis.

mainly members of *Ruminococcaceae* and *Lachnospiraceae* family, indicates a slight urban impact at Site 2, with *Flavobacteriaceae* being a phylotype specific for Site 2. Most abundant phylotypes in the effluent at Site 3 were affiliated with the genus *Trichococcus*, which is frequently found in other WWTPs (Kristensen et al., 2020). The families specific for Site 3 were *Rhodobacteraceae*, *Diplorickettsiaceae* and *Saprospiraceae*. *Peptostreptococcaceae* related phylotypes were detected at the downstream Sites and could not be found upstream from the discharge point. The abundance of phylotypes associated with *Comamonadaceae* and *Sphingomonadaceae* was gradually increasing downstream the WWTP discharge point (Supplementary Table S4).

## Antibiotic resistance gene abundance

The absolute copy numbers of the 16S rRNA gene per 100 mL did not exceed  $9 \times 10^6$  in river water and  $2.5 \times 10^6$  in the WWTP effluent. In general, the 16S rRNA gene as well as *int11*, *sul1* and *sul2* showed the same trends with low copy numbers for pristine Site 1 and the slightly impacted Site 2, highest copy numbers in WWTP effluent and decreasing copy numbers downstream the WWTP discharge point (Figure 4). The decrease in absolute abundance of the 16S rRNA gene from Sites 3 to 4 was ~38%, which is in line with the decrease of the SMX concentration and supports the calculated dilution factor of WWTP effluent discharged into the river water of about one third.

The relative abundance of *sul1* and *sul2* vs. 16S rRNA showed a similar pattern to the absolute gene abundances and measured concentrations of SMX. Surprisingly, the relative abundance of *int11* was continuously increasing downstream the WWTP discharge point. The increase in relative abundance is statistically significant (Dunn's test,  $p \leq 0.05$ ) when comparing Site 6 to Sites 3 and 4 (Figure 5).

## Gene cassette inserts

The long-read sequencing generated 340 to 684 reads per Site, with about 92% retained for further analysis after quality check. Sequence analysis revealed a high percentage of class 1 integrons without inserts in the variable region, ranging from 23% at Site 3 and up to 36% at Site 6. An even higher percentage of reads (65, 57, 64 and 63%, respectively) carried gene cassette inserts without known function. These were either not annotated previously or consisted of non-coding DNA fragments. According to GenBank queries, most of these sequences have been found previously in plasmids or integrons. Comparative analyses of sequence alignments from each individual Site revealed a decreasing trend in diversity from Sites 3 to 6.

During the class 1 integron sequence analysis, 7.3, 6.3, 1.5 and 0.6% of the reads could be annotated as ARGs for the Sites 3–6, respectively (Figure 6A). Annotated genes showed different types of ARGs, conferring resistance against aminoglycosides, beta-lactams,

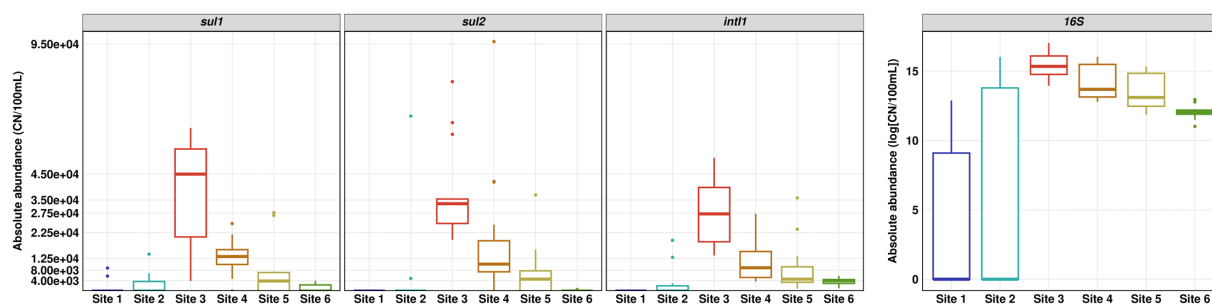


FIGURE 4  
Absolute abundance of *sul1*, *sul2*, *int1* and 16S rRNA gene determined by quantitative real-time PCR. Number of replicates ( $n$ )=15.

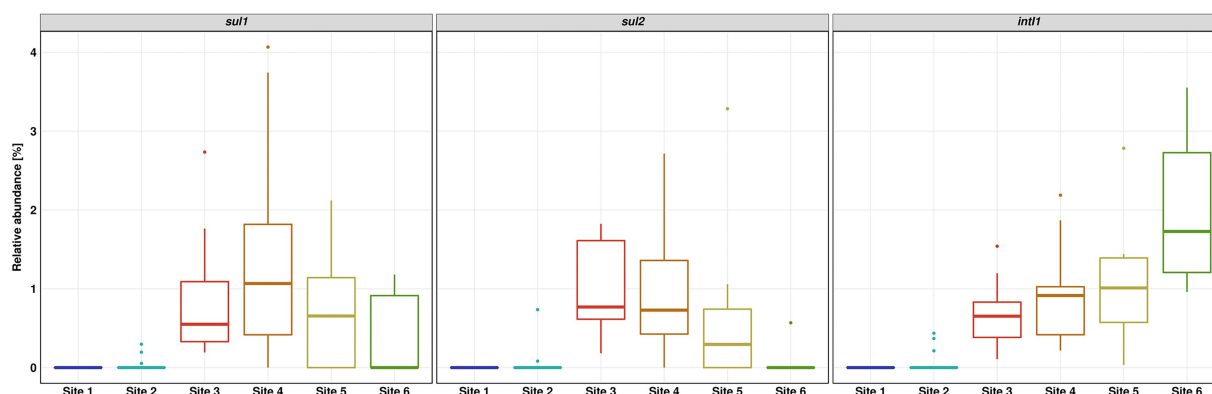


FIGURE 5  
Relative abundance of *sul1*, *sul2* and *int1* over 16S rRNA gene determined by quantitative real-time PCR. Number of replicates ( $n$ )=15.

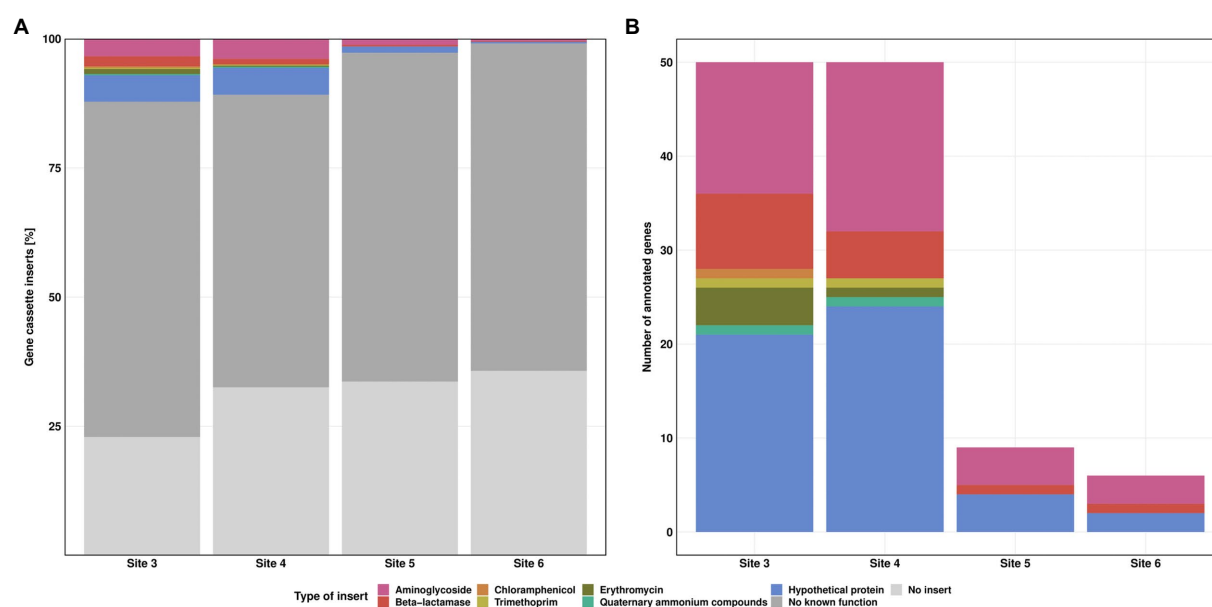


FIGURE 6  
Class 1 integron gene cassette inserts. Relative frequency of insert types for each Site (A). Absolute counts of annotated gene cassette inserts for each Site (B). Resistance against streptomycin (pink), beta-lactamase (red), chloramphenicol (orange), trimethoprim (yellow), erythromycin (green) and quaternary ammonium compounds (turquoise). "No known function" (dark grey) refers to reads without known annotation in GenBank database. "No insert" (light grey) refers to reads where 5'CS and 3'CS were directly linked, without a variable region in between.

chloramphenicol, trimethoprim, erythromycin and quaternary ammonium compounds. In addition, we also identified 4.9, 4.6, 1.2, and 0.3% genes for hypothetical proteins, whereby seven out of eight hypothetical proteins were previously reported in integron gene cassettes from hospital effluents (Stalder et al., 2014).

The highest diversity in inserts was observed at Site 3 with eight different types of ARGs, followed by Site 4 with seven different types. The number of annotated genes was similar for Sites 3 and 4. Sites 5 and 6 showed a clear decrease in both, diversity and number of ARGs. We detected streptomycin, beta lactamase resistance genes, and genes coding for hypothetical proteins for these Sites (Figure 6B). Multiple gene cassettes in one integron were identified only at Site 3, with one read each containing *bla*<sub>BEL-1</sub>|*qacL* and *aadA23*|*bla*<sub>OXA-4</sub>. For streptomycin resistance, we detected mainly *aadA2*, *aadA11* and *aadA6/aadA10*. In addition, we also found *aadA*, *aadA1*, *aadA3*, *aadA4*, *aadA6*, *aadA7*, *aadA22*, *aadA23*, *aadA24*, and *aadA25*. Most of the resistance genes against beta-lactam antibiotics were of the *bla*<sub>OXA</sub> type (*bla*<sub>OXA-2</sub>, *bla*<sub>OXA-4</sub>, *bla*<sub>OXA-10</sub>, *bla*<sub>OXA-33</sub>, *bla*<sub>OXA-36</sub>, *bla*<sub>OXA-129</sub>, *bla*<sub>OXA-392</sub>, and *bla*<sub>OXA-824</sub>, each of them found at most once per Site). For the *bla*<sub>BEL</sub> type we found the *bla*<sub>BEL-1</sub> and *bla*<sub>GES-11</sub> gene. For erythromycin resistance we detected only *ereA2*. Furthermore, we also found *cmlA5*, *dfrA1*, *dfrA7* and *qacL* genes. Of particular concern are *bla*<sub>OXA-4</sub>, *bla*<sub>OXA-10</sub>, *bla*<sub>GES-11</sub>, *cmlA5* and *dfrA1* as they are clinically relevant and have the greatest potential to contribute to multidrug resistance in pathogens. The *dfrA1* gene is listed as high-risk ARG by the World Health Organisation (Zhang et al., 2021). None of these high-risk genes was detected at the downstream Sites 5 and 6. Detailed information regarding the gene cassette inserts can be found in Supplementary Table S5.

## Discussion

Wastewater treatment plants (WWTP) effluents are key source of antibiotics and ARGs in riverine ecosystems (Harnisz et al., 2020; Hutinel et al., 2021; Larsson and Flach, 2022). Currently, there are conflicting results about the extent to which such contaminants spread over long distances and affect microbial communities at downstream Sites (Lee et al., 2021; Wang et al., 2021). A better understanding of the fate of ARGs in rivers will be helpful to regulatory bodies in deciding where and how best to act to curtail the spread of antibiotic resistance through environmental compartments (Vikesland et al., 2017).

Overall, our results revealed a coherent picture regarding the pollution pattern in the investigated stretch of the Holtemme river, showing the highest SMX concentration and abundance of sulfonamide resistance genes *sul1* and *sul2* in the WWTP effluent and at the nearby downstream Site 4, with a steep decrease at further downstream Sites located at 8 and 13 km distance from the discharge point. The observed pollution pattern is in line with previous studies conducted at the Holtemme, that focused on monitoring anthropogenic micropollutants in the river (Krauss et al., 2019; Beckers et al., 2020; Švara et al., 2021). We found no evidence that SMX at ambient low concentrations impacted the abundance of ARGs *sul1* and *sul2* in the river water. Most likely they shared the same source, in this case the WWTP (Larsson and Flach, 2022). Lee et al. (2021) have drawn similar conclusions, arguing that the majority of wastewater-origin resistance determinants are not persistent in rivers after discharge from WWTP due to dilution as well as additional removal mechanisms. Concerning the fate of SMX, we hypothesize that mainly physicochemical processes such as dilution, sedimentation and sorption were responsible for the concentration

decrease with increasing distance to the WWTP discharge point (Luo et al., 2010).

The changes in microbial community structure reflected the pollution pattern in the Holtemme. The microbiome recovery model provided quantitative support for this ecological assumption. A pairwise diversity comparison helped identify the individual difference in the microbiome state. However, it could not provide general information on the microbiome dynamics. Our impulse model, trained based on the work from Shaw et al. (2019), gave additional insight into the long and short-term impact of wastewater treatment on the evolution of the aquatic microbiome diversity. The result of our microbiome recovery model showed the recovery of the riverine microbial community through a transition state. Both, microbial community analysis as well as our model consolidate the hypothesis that microorganisms from the wastewater may not actively thrive in the riverine environment over a more extended period, suggesting a high probability of the natural river community to recover from the displacement by effluent discharge. At the same time, it is important to notice that the community seems to be stabilizing to a different state than before the disturbance. The die-off of wastewater derived microorganisms in the river water might also explain the decrease in ARG abundance. Further investigations are needed to understand potential causal relations between the apparent stabilization of the microbial community and ARG prevalence in riverine hosts (Spencer et al., 2016). We would like to note, that our last sampling Site 6 was just upstream from the city of Halberstadt, with its own WWTP discharging. Consequently, it is likely that the microbial community structure in the Holtemme river downstream from the first WWTP will never return to its original composition. On the contrary, further WWTP effluent discharge will continuously alter the microbial community structure.

The decrease in ARG abundance downstream the WWTP, which we followed closely for *sul1* and *sul2*, was also observed for the class 1 integron. Furthermore, the diversity of gene cassette inserts was decreasing with increasing distance to the discharge point. It seems unlikely that host species carrying larger class 1 integrons were outcompeted due to the higher metabolic burden (San Millan and Craig MacLean, 2022). This suggests that some of the inserts were excised by the *intI1* gene product, potentially due to higher integrase activity triggered by cellular stress (Guerin et al., 2009; Strugeon et al., 2016; Knecht et al., 2022). Due to the missing selective pressure from high antibiotic concentrations, the loss of gene cassette inserts did not result in an evolutionary disadvantage, leading to an overall decrease in class 1 integron inserts. To our knowledge, little is known about the underlying mechanisms which are driving the integrase to either insert or excise DNA fragments from the gene cassette.

We detected some genes with high clinical relevance near the discharge point of the WWTP, like *bla*<sub>OXA-4</sub>, *bla*<sub>OXA-10</sub>, *bla*<sub>GES-11</sub>, *cmlA5* and *dfrA1* (Zhang et al., 2021), but none of them was found at the downstream Sites. Aminoglycoside and beta-lactam resistance genes were the most prevalent gene cassette inserts in the class 1 integron, consistent with a recent study by Yang et al. (2021), as well as a survey of the distribution of integrons in sequenced genomes (Zhang et al., 2018). Meanwhile trimethoprim resistance, which is also frequently detected in comparable studies (Wu et al., 2012; Adelowo et al., 2018; Chaturvedi et al., 2021a), was rare at our study sites. These findings highlight the diversity and complexity of class 1 integron arrays depending on environmental conditions as well as facilities located upstream from the WWTP (e.g., hospital, livestock). Despite the high sequencing error rate, nanopore sequencing of the clinical class 1

integron can improve our understanding of environmental dynamics at comparably low cost and with sufficient sensitivity and sequencing accuracy (Kasuga et al., 2022). We anticipate that this approach will facilitate future studies on mobilized ARGs and their dissemination through the class 1 integron.

Considering the pollution pattern of the Holtemme, the application of *intI1* as biomarker for anthropogenic pollution, and the frequent association of *intI1* and *sul1* as part of the class 1 integron, the continuous increase of *intI1* relative abundance downstream the WWTP was surprising (Gillings et al., 2015; Zhang et al., 2018). Recently, Lee et al. (2021) investigated two rivers in Switzerland up- and downstream from their respective first WWTP and observed an occasional increase in the abundance of *intI1* that did not match the pollution level of the river, but they also measured a parallel increase in the abundance of *sul1*. They hypothesized that either additional sources or an in-system growth could be plausible causes. Similarly, a study on treatment wetland found an increase of *intI1* and *sul1* that occurred independent of the presence of sulfonamides (Knecht et al., 2022). In our study, the continuous increase of *intI1* relative abundance may result either from an intracellular multiplication of the gene or from the proliferation of the host microorganisms within the riverine microbial community. Interestingly, Koczura et al. (2016) likewise observed a higher relative abundance of *intI1* in river water during the cold season, which did not correlate with the abundance of *sul1*.

As our study sites are near agricultural fields, it could be hypothesized that field run-offs from agricultural lands may have introduced a variety of new phylotypes and therefore *intI1* genes to the river system. The same accounts for small tributaries of the Holtemme. These hypotheses are however not supported by the diversity analyses (Figure 3), which did not reveal a sudden emergence of new phylotypes. The potential role of other stressors, like heavy metals or a secondary source of antibiotics, as drivers for the increase in *intI1* relative abundance also appears unlikely since long-read sequencing indicated that the integrase was rather excising inserts out of the gene cassette. A remaining plausible explanation, which prompts for future investigations, is an increase in the relative abundance of *intI1* by intracellular gene multiplication.

## Conclusion

In conclusion, our monitoring study shows a partial recovery of the microbial community after wastewater discharge. SMX concentration as well as *sul1* and *sul2* abundance decreased steadily downstream from the WWTP, and the number and diversity of ARG cassettes in the class 1 integron declined. Apparently, low ng/l concentrations of sulfonamides did not stimulate maintenance or even proliferation of the corresponding ARGs in the river water. Nevertheless, copies of class 1 integron harboring ARGs against aminoglycosides and beta-lactams were still present at the most downstream Site 6, just before the river flows into the next city. Furthermore, the *intI1* relative abundance increased 3-fold over a course of 13 km downstream from the WWTP. Even though the underlying mechanism remains unclear, the observed increase was statistically significant. It seems that factors other than antimicrobials affect the environmental dynamics of *intI1* and potentially the class 1 integron. If this is the case, then limiting the release of ARG-harboring bacteria into the aqueous environment may be more prudent than attenuating ng/L concentrations of sulfonamides when aiming for restricting the environmental spread of antimicrobial resistance.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: <https://www.ncbi.nlm.nih.gov/>, PRJNA858676. <https://figshare.com/>; <https://doi.org/10.6084/m9.figshare.20310003.v1>.

## Author contributions

NM, JM, and HHR conceived the study and the experimental design. SH and GW performed the sampling. SH executed the sample preparation, solid phase extraction (SPE), HPLC-MS, qPCR, and preparation for sequencing. FM supported with chemical analysis. SH, JK, and UR performed the microbial community analysis. SH, JM, and NM wrote the manuscript with contribution from all authors and all authors approved the submitted version.

## Funding

JK was supported by São Paulo Research Foundation (FAPESP; grant 2019/03396-9 and 2022/03534-5).

## Acknowledgments

We would like to thank Florian Tschernikl for his support in solid phase extraction. We appreciate the support of Ines Mäusezahl for DNA extraction and qPCR, extend special thanks to David Thiele and Florian Lenk for library preparation and Illumina sequencing, and to Patrick Fink, Mario Brauns and Florian Zander for their continuous effort maintaining the MOBICOS stations.

## Conflict of interest

HHR is employed by Isodetect Umweltmonitoring GmbH (Leipzig, Germany).

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2023.1058350/full#supplementary-material>



## References

- Adelowo, O. O., Helbig, T., Knecht, C., Reincke, F., Mäusezahl, I., and Müller, J. A. (2018). High abundances of class 1 integrase and sulfonamide resistance genes, and characterisation of class 1 integron gene cassettes in four urban wetlands in Nigeria. *PLoS One* 13:e0208269. doi: 10.1371/journal.pone.0208269
- Adelowo, O. O., Ikimiukor, O. O., Knecht, C., Vollmers, J., Bhatia, M., Kaster, A.-K., et al. (2020). A survey of extended-spectrum beta-lactamase-producing Enterobacteriaceae in urban wetlands in southwestern Nigeria as a step towards generating prevalence maps of antimicrobial resistance. *PLoS One* 15:e0229451. doi: 10.1371/journal.pone.0229451
- Alcock, B. P., Raphenya, A. R., Lau, T. T. Y., Tsang, K. K., Bouchard, M., Edalatmand, A., et al. (2020). CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res.* 48, D517–D525. doi: 10.1093/nar/gkz935
- Allaby, M. (2009). *A Dictionary of Zoology*. Oxford: Oxford University Press.
- Beckers, L.-M., Brack, W., Dann, J. P., Krauss, M., Müller, E., and Schulze, T. (2020). Unraveling longitudinal pollution patterns of organic micropollutants in a river by non-target screening and cluster analysis. *Sci. Total Environ.* 727:138388. doi: 10.1016/j.scitotenv.2020.138388
- Beckers, L.-M., Busch, W., Krauss, M., Schulze, T., and Brack, W. (2018). Characterization and risk assessment of seasonal and weather dynamics in organic pollutant mixtures from discharge of a separate sewer system. *Water Res.* 135, 122–133. doi: 10.1016/j.watres.2018.02.002
- Bengtsson-Palme, J., and Larsson, D. G. J. (2016). Concentrations of antibiotics predicted to select for resistant bacteria: proposed limits for environmental regulation. *Environ. Int.* 86, 140–149. doi: 10.1016/j.envint.2015.10.015
- Berendonk, T. U., Manaia, C. M., Merlin, C., Fatta-Kassinos, D., Cytryn, E., Walsh, F., et al. (2015). Tackling antibiotic resistance: the environmental framework. *Nat. Rev. Microbiol.* 13, 310–317. doi: 10.1038/nrmicro3439
- Bokulich, N. A., Kaehler, B. D., Rideout, J. R., Dillon, M., Bolyen, E., Knight, R., et al. (2018). Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* 6:90. doi: 10.1186/s40168-018-0470-z
- Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 37, 852–857. doi: 10.1038/s41587-019-0209-9
- Borsetto, C., Raguideau, S., Travis, E., Kim, D. W., Lee, D. H., Bottrill, A., et al. (2021). Impact of sulfamethoxazole on a riverine microbiome. *Water Res.* 201:117382. doi: 10.1016/j.watres.2021.117382
- Brase, L., Sanders, T., and Dähnke, K. (2018). Anthropogenic changes of nitrogen loads in a small river: external nutrient sources vs. internal turnover processes. *Isot. Environ. Health Stud.* 54, 168–184. doi: 10.1080/10256016.2018.1428580
- Cairns, J., Ruokolainen, L., Hultman, J., Tamminen, M., Virta, M., and Hiltunen, T. (2018). Ecology determines how low antibiotic concentration impacts community composition and horizontal transfer of resistance genes. *Commun. Biol.* 1, 35–38. doi: 10.1038/s42003-018-0041-7
- Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016). DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. Methods* 13, 581–583. doi: 10.1038/nmeth.3869
- Cambray, G., Guerout, A.-M., and Mazel, D. (2010). Integrons. *Annu. Rev. Genet.* 44, 141–166. doi: 10.1146/annurev-genet-102209-163504
- Carvalho, I. T., and Santos, L. (2016). Antibiotics in the aquatic environments: a review of the European scenario. *Environ. Int.* 94, 736–757. doi: 10.1016/j.envint.2016.06.025
- Chaturvedi, P., Shukla, P., Giri, B. S., Chowdhary, P., Chandra, R., Gupta, P., et al. (2021a). Prevalence and hazardous impact of pharmaceutical and personal care products and antibiotics in environment: a review on emerging contaminants. *Environ. Res.* 194:110664. doi: 10.1016/j.envres.2020.110664
- Chaturvedi, P., Singh, A., Chowdhary, P., Pandey, A., and Gupta, P. (2021b). Occurrence of emerging sulfonamide resistance (sul1 and sul2) associated with mobile integrons-integrase (intl1 and intl2) in riverine systems. *Sci. Total Environ.* 751:142217. doi: 10.1016/j.scitotenv.2020.142217
- Collignon, P., Beggs, J. J., Walsh, T. R., Gandra, S., and Laxminarayan, R. (2018). Anthropological and socioeconomic factors contributing to global antimicrobial resistance: a univariate and multivariable analysis. *Lancet Planet. Health* 2, e398–e405. doi: 10.1016/S2542-5196(18)30186-4
- Fink, P., Norf, H., Anlanger, C., Brauns, M., Kamjunke, N., Risse-Buhl, U., et al. (2020). Streamside mobile mesocosms (MOBICOS): a new modular research infrastructure for hydro-ecological process studies across catchment-scale gradients. *Int. Rev. Hydrobiol.* 105, 63–73. doi: 10.1002/iroh.201902009
- Gillings, M. R. (2018). DNA as a pollutant: the clinical class 1 Integron. *Curr. Pollut. Rep.* 4, 49–55. doi: 10.1007/s40726-018-0076-x
- Gillings, M., Boucher, Y., Labbate, M., Holmes, A., Krishnan, S., Holley, M., et al. (2008). The evolution of class 1 integrons and the rise of antibiotic resistance. *J. Bacteriol.* 190, 5095–5100. doi: 10.1128/JB.00152-08
- Gillings, M. R., Gaze, W. H., Pruden, A., Smalla, K., Tiedje, J. M., and Zhu, Y.-G. (2015). Using the class 1 integron-integrase gene as a proxy for anthropogenic pollution. *ISME J.* 9, 1269–1279. doi: 10.1038/ismej.2014.226
- Guerin, E., Cambray, G., Sanchez-Alberola, N., Campoy, S., Erill, I., Da Re, S., et al. (2009). The SOS response controls integron recombination. *Science* 324:1034. doi: 10.1126/science.1172914
- Harnisz, M., Kiedrzyńska, E., Kiedrzyński, M., Korzeniewska, E., Czatkowska, M., Koniuszewska, I., et al. (2020). The impact of WWTP size and sampling season on the prevalence of antibiotic resistance genes in wastewater and the river system. *Sci. Total Environ.* 741:140466. doi: 10.1016/j.scitotenv.2020.140466
- Hernández, F., Calisto-Ulloa, N., Gómez-Fuentes, C., Gómez, M., Ferrer, J., González-Rocha, G., et al. (2019). Occurrence of antibiotics and bacterial resistance in wastewater and sea water from the Antarctic. *J. Hazard. Mater.* 363, 447–456. doi: 10.1016/j.jhazmat.2018.07.027
- Hutinel, M., Fick, J., Larsson, D. G. J., and Flach, C.-F. (2021). Investigating the effects of municipal and hospital wastewaters on horizontal gene transfer. *Environ. Pollut.* 276:116733. doi: 10.1016/j.envpol.2021.116733
- Inostroza, P. A., Vera-Escalona, I., Wicht, A.-J., Krauss, M., Brack, W., and Norf, H. (2016). Anthropogenic stressors shape genetic structure: insights from a model freshwater population along a land use gradient. *Environ. Sci. Technol.* 50, 11346–11356. doi: 10.1021/acs.est.6b04629
- Karthe, D., Lin, P.-Y., and Westphal, K. (2017). Instream coliform gradients in the Holtemme, a small headwater stream in the Elbe River Basin, Northern Germany. *Front. Earth Sci.* 11, 544–553. doi: 10.1007/s11707-017-0648-x
- Kasuga, I., Nagasawa, K., Suzuki, M., Kurisu, F., and Furumai, H. (2022). High-throughput screening of antimicrobial resistance genes and their association with class 1 Integrons in urban Rivers in Japan. *Front. Environ. Sci.* 10, 12–13. doi: 10.3389/fenvs.2022.825372
- Knecht, C. A., Krüger, M., Kellmann, S., Mäusezahl, I., Möder, M., Adelowo, O. O., et al. (2022). Cellular stress affects the fate of microbial resistance to folate inhibitors in treatment wetlands. *Sci. Total Environ.* 845:157318. doi: 10.1016/j.scitotenv.2022.157318
- Koczura, R., Mokracka, J., Taraszewska, A., and Łopacińska, N. (2016). Abundance of class 1 Integron-Integrase and Sulfonamide resistance genes in river water and sediment is affected by anthropogenic pressure and environmental factors. *Microb. Ecol.* 72, 909–916. doi: 10.1007/s00248-016-0843-4
- Krauss, M., Hug, C., Bloch, R., Schulze, T., and Brack, W. (2019). Prioritising site-specific micropollutants in surface water from LC-HRMS non-target screening data using a rarity score. *Environ. Sci. Eur.* 31:9. doi: 10.1186/s12302-019-0231-z
- Kristensen, J. M., Nierychlo, M., Albertsen, M., and Nielsen, P. H. (2020). Bacteria from the genus *Arcobacter* are abundant in effluent from wastewater treatment plants. *Appl. Environ. Microbiol.* 86:e03044-19. doi: 10.1128/AEM.03044-19
- Larsson, D. G. J., and Flach, C.-F. (2022). Antibiotic resistance in the environment. *Nat. Rev. Microbiol.* 20, 257–269. doi: 10.1038/s41579-021-00649-x
- Le Page, G., Gunnarsson, L., Snape, J., and Tyler, C. R. (2017). Integrating human and environmental health in antibiotic risk assessment: a critical analysis of protection goals, species sensitivity and antimicrobial resistance. *Environ. Int.* 109, 155–169. doi: 10.1016/j.envint.2017.09.013
- Lee, J., Ju, F., Maile-Moskowitz, A., Beck, K., Maccagnan, A., McArdell, C. S., et al. (2021). Unraveling the riverine antibiotic resistome: the downstream fate of anthropogenic inputs. *Water Res.* 197:117050. doi: 10.1016/j.watres.2021.117050
- Luo, Y., Mao, D., Rysz, M., Zhou, Q., Zhang, H., Xu, L., et al. (2010). Trends in antibiotic resistance genes occurrence in the Haihe River, China. *Environ. Sci. Technol.* 44, 7220–7225. doi: 10.1021/es100233w
- Martin-Laurent, F., Topp, E., Billet, L., Batisson, I., Malandain, C., Besse-Hoggan, P., et al. (2019). Environmental risk assessment of antibiotics in agroecosystems: ecotoxicological effects on aquatic microbial communities and dissemination of antimicrobial resistances and antibiotic biodegradation potential along the soil-water continuum. *Environ. Sci. Pollut. Res. Int.* 26, 18930–18937. doi: 10.1007/s11356-019-05122-0
- Mazel, D. (2006). Integrons: agents of bacterial evolution. *Nat. Rev. Microbiol.* 4, 608–620. doi: 10.1038/nrmicro1462
- McEwen, S. A., and Collignon, P. J. (2018). Antimicrobial resistance: a one health perspective. *Microbiol. Spectr.* 6. doi: 10.1128/microbiolspec.ARBA-0009-2017
- McMurdie, P. J., and Holmes, S. (2013). Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8:e61217. doi: 10.1371/journal.pone.0061217
- Moura, A., Soares, M., Pereira, C., Leitão, N., Henriques, I., and Correia, A. (2009). INTEGRALL: a database and search engine for integrons, integrases and gene cassettes. *Bioinformatics* 25, 1096–1098. doi: 10.1093/bioinformatics/btp105
- Oksanen, J., Simpson, G., Blanchet, F., Kindt, R., Legendre, P., Minchin, P., et al. (2022). Vegan: Community Ecology Package. R package version 2.6-2. Available at: <https://CRAN.R-project.org/package=vegan>
- Paumelle, M., Donnadiu, F., Joly, M., Besse-Hoggan, P., and Artigas, J. (2020). Effects of sulfonamide antibiotics on aquatic microbial community composition and functions. *Environ. Int.* 146:106198. doi: 10.1016/j.envint.2020.106198
- Pruden, A., Pei, R., Storteboom, H., and Carlson, K. H. (2006). Antibiotic resistance genes as emerging contaminants: studies in Northern Colorado. *Environ. Sci. Technol.* 40, 7445–7450. doi: 10.1021/es060413l

- Robeson, M. S., O'Rourke, D. R., Kaehler, B. D., Ziemski, M., Dillon, M. R., Foster, J. T., et al. (2021). RESCRIPT: Reproducible sequence taxonomy reference database management. *PLoS Comput. Biol.* 17:e1009581. doi: 10.1371/journal.pcbi.1009581
- San Millan, A., and Craig MacLean, R. (2022). Fitness costs of plasmids: a limit to plasmid transmission. *Microbiol. Spectr.* 5. doi: 10.1128/9781555819743.ch4
- Shaw, L. P., Bassam, H., Barnes, C. P., Walker, A. S., Klein, N., and Balloux, F. (2019). Modelling microbiome recovery after antibiotics using a stability landscape framework. *ISME J.* 13, 1845–1856. doi: 10.1038/s41396-019-0392-1
- Spencer, S. J., Tamminen, M. V., Preheim, S. P., Guo, M. T., Briggs, A. W., Brito, I. L., et al. (2016). Massively parallel sequencing of single cells by epicPCR links functional genes with phylogenetic markers. *ISME J.* 10, 427–436. doi: 10.1038/ismej.2015.124
- Stalder, T., Barraud, O., Jové, T., Casellas, M., Gaschet, M., Dagot, C., et al. (2014). Quantitative and qualitative impact of hospital effluent on dissemination of the integron pool. *ISME J.* 8, 768–777. doi: 10.1038/ismej.2013.189
- Straub, J. O. (2016). Aquatic environmental risk assessment for human use of the old antibiotic sulfamethoxazole in Europe. *Environ. Toxicol. Chem.* 35, 767–779. doi: 10.1002/etc.2945
- Strugeon, E., Tilloy, V., Ploy, M.-C., and Da Re, S. (2016). The stringent response promotes antibiotic resistance dissemination by regulating Integron Integrase expression in biofilms. *mBio* 7:7. doi: 10.1128/mBio.00868-16
- Švara, V., Krauss, M., Michalski, S. G., Altenburger, R., Brack, W., and Luckenbach, T. (2021). Chemical pollution levels in a river explain site-specific sensitivities to micropollutants within a genetically homogeneous population of freshwater amphipods. *Environ. Sci. Technol.* 55, 6087–6096. doi: 10.1021/acs.est.0c07839
- The World Health Organization, The Food and Agriculture Organization of the United Nations, and The World Organisation for Animal Health (2019). Taking a Multisectoral, One Health Approach: A Tripartite Guide to Addressing Zoonotic Diseases in Countries.
- Tousova, Z., Oswald, P., Slobodnik, J., Blaha, L., Muz, M., Hu, M., et al. (2017). European demonstration program on the effect-based and chemical identification and monitoring of organic pollutants in European surface waters. *Sci. Total Environ.* 601–602, 1849–1868. doi: 10.1016/j.scitotenv.2017.06.032
- Vikesland, P. J., Pruden, A., Alvarez, P. J. J., Aga, D., Bürgmann, H., Li, X.-D., et al. (2017). Toward a comprehensive strategy to mitigate dissemination of environmental sources of antibiotic resistance. *Environ. Sci. Technol.* 51, 13061–13069. doi: 10.1021/acs.est.7b03623
- Wang, R., Ji, M., Zhai, H., Guo, Y., and Liu, Y. (2021). Occurrence of antibiotics and antibiotic resistance genes in WWTP effluent-receiving water bodies and reclaimed wastewater treatment plants. *Sci. Total Environ.* 796:148919. doi: 10.1016/j.scitotenv.2021.148919
- Wollschläger, U., Attinger, S., Borchardt, D., Brauns, M., Cuntz, M., Dietrich, P., et al. (2017). The bode hydrological observatory: a platform for integrated, interdisciplinary hydro-ecological research within the TERENO Harz/Central German Lowland Observatory. *Environ. Earth Sci.* 76:29. doi: 10.1007/s12665-016-6327-5
- Wu, K., Wang, F., Sun, J., Wang, Q., Chen, Q., Yu, S., et al. (2012). Class 1 integron gene cassettes in multidrug-resistant gram-negative bacteria in southern China. *Int. J. Antimicrob. Agents* 40, 264–267. doi: 10.1016/j.ijantimicag.2012.05.017
- Yang, Y., Zhang, A.-N., Che, Y., Liu, L., Deng, Y., and Zhang, T. (2021). Underrepresented high diversity of class 1 integrons in the environment uncovered by PacBio sequencing using a new primer. *Sci. Total Environ.* 787:147611. doi: 10.1016/j.scitotenv.2021.147611
- Yao, Y., Maddamsetti, R., Weiss, A., Ha, Y., Wang, T., Wang, S., et al. (2022). Intra- and interpopulation transposition of mobile genetic elements driven by antibiotic selection. *Nat. Ecol. Evol.* 6, 555–564. doi: 10.1038/s41559-022-01705-2
- Zacharias, S., Bogen, H., Samaniego, L., Mauder, M., Fuß, R., Pütz, T., et al. (2011). A network of terrestrial environmental observatories in Germany. *Vadose Zone J.* 10, 955–973. doi: 10.2136/vzj2010.0139
- Zhang, A.-N., Gaston, J. M., Dai, C. L., Zhao, S., Poyet, M., Groussin, M., et al. (2021). An omics-based framework for assessing the health risk of antimicrobial resistance genes. *Nat. Commun.* 12:4765. doi: 10.1038/s41467-021-25096-3
- Zhang, A. N., Li, L.-G., Ma, L., Gillings, M. R., Tiedje, J. M., and Zhang, T. (2018). Conserved phylogenetic distribution and limited antibiotic resistance of class 1 integrons revealed by assessing the bacterial genome and plasmid collection. *Microbiome* 6:130. doi: 10.1186/s40168-018-0516-2



## OPEN ACCESS

EDITED AND REVIEWED BY  
Graciela L. Lorca,  
University of Florida, United States

\*CORRESPONDENCE  
Niculina Musat  
✉ niculina.musat@ufz.de

RECEIVED 19 April 2023  
ACCEPTED 17 May 2023  
PUBLISHED 30 May 2023

## CITATION

Haenelt S, Wang G, Kasmanas JC, Musat F, Richnow HH, da Rocha UN, Müller JA and Musat N (2023) Corrigendum: The fate of sulfonamide resistance genes and anthropogenic pollution marker *intI1* after discharge of wastewater into a pristine river stream. *Front. Microbiol.* 14:1208555. doi: 10.3389/fmicb.2023.1208555

## COPYRIGHT

© 2023 Haenelt, Wang, Kasmanas, Musat, Richnow, da Rocha, Müller and Musat. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Corrigendum: The fate of sulfonamide resistance genes and anthropogenic pollution marker *intI1* after discharge of wastewater into a pristine river stream

Sarah Haenelt<sup>1</sup>, Gangan Wang<sup>1</sup>, Jonas Coelho Kasmanas<sup>2</sup>, Florin Musat<sup>1,3</sup>, Hans Hermann Richnow<sup>1,4</sup>, Ulisses Nunes da Rocha<sup>2</sup>, Jochen A. Müller<sup>2,5</sup> and Niculina Musat<sup>1\*</sup>

<sup>1</sup>Department of Isotope Biogeochemistry, Helmholtz Centre for Environmental Research, Leipzig, Germany, <sup>2</sup>Department of Environmental Microbiology, Helmholtz Centre for Environmental Research, Leipzig, Germany, <sup>3</sup>Department of Molecular Biology and Biotechnology, Faculty of Biology and Geology, Babeş-Bolyai University, Cluj-Napoca, Romania, <sup>4</sup>Isodetect Umweltmonitoring GmbH, Leipzig, Germany, <sup>5</sup>Institute for Biological Interfaces (IBG 5), Karlsruhe Institute of Technology, Eggenstein-Leopoldshafen, Germany

## KEYWORDS

class 1 integron, sulfamethoxazole, sulfonamide resistance, *sul1*, *sul2*, *intI1*, river ecosystem, one health

## A corrigendum on

The fate of sulfonamide resistance genes and anthropogenic pollution marker *intI1* after discharge of wastewater into a pristine river stream

by Haenelt, S., Wang, G., Kasmanas, J. C., Musat, F., Richnow, H. H., da Rocha, U. N., et al. (2023). *Front. Microbiol.* 14:1058350. doi: 10.3389/fmicb.2023.1058350

In the original article, there was an error in [Figure 4](#) as published. The absolute abundances of *sul1*, *sul2*, *intI1* and 16S rRNA gene were calculated incorrectly.

The corrected [Figure 4](#) and its caption “[Figure 4](#). Absolute abundance of *sul1*, *sul2*, *intI1* and 16S rRNA gene determined by quantitative real time PCR. Number of replicates (*n*) = 15.” appears below.

In the original article, there was an error in the results section. The absolute abundances of *sul1*, *sul2*, *intI1* and 16S rRNA gene were calculated incorrectly.

A correction has been made to **Results**, *ARG abundance*, Paragraph 1. This sentence previously stated:

“The absolute copy numbers of the 16S rRNA gene per 100 mL did not exceed  $2 \times 10^5$  in river water and  $5 \times 10^5$  in the WWTP effluent.”

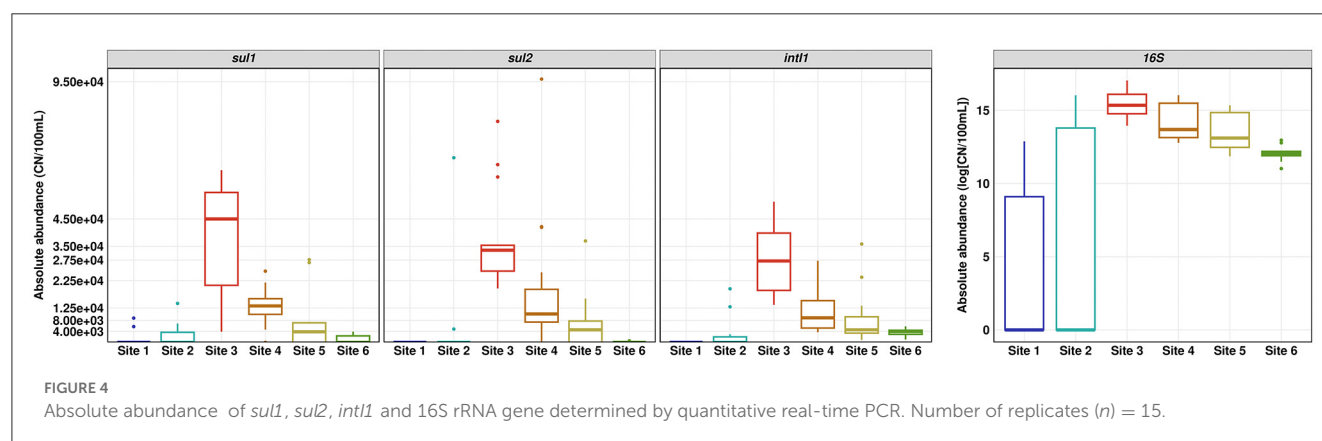
The corrected sentence appears below:

“The absolute copy numbers of the 16S rRNA gene per 100 mL did not exceed  $9 \times 10^6$  in river water and  $2.5 \times 10^6$  in the WWTP effluent.”

The author apologizes for these errors and state that this does not change the scientific conclusions of the article in any way. The original article has been updated.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.







## OPEN ACCESS

## EDITED BY

Graciela L. Lorca,  
University of Florida,  
United States

## REVIEWED BY

Ying Yang,  
Sun Yat-sen University,  
China  
Florentin Constancias,  
ETH Zürich,  
Switzerland

## \*CORRESPONDENCE

Stefanía Magnúsdóttir  
✉ stefania.magnusdottir@ufz.de  
Ulisses Nunes da Rocha  
✉ ulisses.rocha@ufz.de

## SPECIALTY SECTION

This article was submitted to  
Microbial Physiology and Metabolism,  
a section of the journal  
Frontiers in Microbiology

RECEIVED 06 September 2022

ACCEPTED 02 January 2023

PUBLISHED 25 January 2023

## CITATION

Magnúsdóttir S, Saraiva JP, Bartholomäus A,  
Soheili M, Toscan RB, Zhang J, Nunes da  
Rocha U and CLUE-TERRA consortium (2023)  
Metagenome-assembled genomes indicate  
that antimicrobial resistance genes are highly  
prevalent among urban bacteria and multidrug  
and glycopeptide resistances are ubiquitous in  
most taxa.  
*Front. Microbiol.* 14:1037845.  
doi: 10.3389/fmicb.2023.1037845

## COPYRIGHT

© 2023 Magnúsdóttir, Saraiva, Bartholomäus,  
Soheili, Toscan, Zhang, Nunes da Rocha and  
CLUE-TERRA consortium. This is an open-  
access article distributed under the terms of the  
[Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).  
The use, distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in this  
journal is cited, in accordance with accepted  
academic practice. No use, distribution or  
reproduction is permitted which does not  
comply with these terms.

# Metagenome-assembled genomes indicate that antimicrobial resistance genes are highly prevalent among urban bacteria and multidrug and glycopeptide resistances are ubiquitous in most taxa

Stefanía Magnúsdóttir<sup>1\*</sup>, Joao Pedro Saraiva<sup>1</sup>,  
Alexander Bartholomäus<sup>2</sup>, Majid Soheili<sup>1</sup>, Rodolfo Brizola Toscan<sup>1</sup>,  
Junya Zhang<sup>3,4</sup>, Ulisses Nunes da Rocha<sup>1\*</sup> and CLUE-TERRA  
consortium

<sup>1</sup>Department of Environmental Microbiology, Helmholtz Centre for Environmental Research-UFZ, Leipzig, Germany, <sup>2</sup>GFZ German Research Centre for Geosciences, Section Geomicrobiology, Potsdam, Germany, <sup>3</sup>Department of Isotope Biogeochemistry, Helmholtz Centre for Environmental Research-UFZ, Leipzig, Germany, <sup>4</sup>Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing, China

**Introduction:** Every year, millions of deaths are associated with the increased spread of antimicrobial resistance genes (ARGs) in bacteria. With the increasing urbanization of the global population, the spread of ARGs in urban bacteria has become a more severe threat to human health.

**Methods:** In this study, we used metagenome-assembled genomes (MAGs) recovered from 1,153 urban metagenomes in multiple urban locations to investigate the fate and occurrence of ARGs in urban bacteria. Additionally, we analyzed the occurrence of these ARGs on plasmids and estimated the virulence of the bacterial species.

**Results:** Our results showed that multidrug and glycopeptide ARGs are ubiquitous among urban bacteria. Additionally, we analyzed the deterministic effects of phylogeny on the spread of these ARGs and found ARG classes that have a non-random distribution within the phylogeny of our recovered MAGs. However, few ARGs were found on plasmids and most of the recovered MAGs contained few virulence factors.

**Discussion:** Our results suggest that the observed non-random spreads of ARGs are not due to the transfer of plasmids and that most of the bacteria observed in the study are unlikely to be virulent. Additional research is needed to evaluate whether the ubiquitous and widespread ARG classes will become entirely prevalent among urban bacteria and how they spread among phylogenetically distinct species.

## KEYWORDS

Antimicrobial resistance, urban, prevalence, metagenome-assembled genome, bacteria, virulence factor, plasmid

## Introduction

Every year, millions of human deaths can be associated with the spread of antimicrobial resistance genes (ARGs) in microbes, and antibiotic resistance will become the leading cause of death worldwide within the next decade (United Nations, 2016). Every year, the proportion of the global population living in urban environments increases, and to this date, more than half of the world population lives in urban environments (United Nations, 2018). People living in urban

environments have more common surface areas that can be viewed as hot spots for microbial spread (e.g., public transit infrastructures, supermarkets, and park benches). People touch these common areas, leave, and pick up microbes through those contacts.

Antimicrobial resistance is a natural defense mechanism in bacteria and the spread of ARGs among bacteria is a naturally occurring phenomenon observed even in pristine environments. ARGs have, for example, been found in bacteria isolated from 2,000-year-old glacial ice and water samples (Dancer et al., 1997) and in metagenomic samples from over 10,000-year-old sediments (Song et al., 2005). However, because of the everyday use of antibiotics in healthcare and food production (United Nations, 2016), antimicrobial resistance is spreading among the microbes in our environment at an alarming rate.

A recent study showed that most samples from urban environments have some ARG levels (Danko et al., 2021). However, these studies did not report ARG spread among urban bacterial taxa. In this study, we aimed to investigate the spread of different types and classes of ARGs in urban bacteria. We gathered publicly available metagenomic sequences from studies that collected samples from urban environments. We then recovered metagenome-assembled genomes (MAGs) from each sample and searched for the presence of known ARGs in these genomes. Current protein alignment methods fail to identify over a third of microbial proteins (Bileschi et al., 2022). We, therefore, opted to use DeepARG (Arango-Argoty et al., 2018) to identify ARGs in urban MAGs. DeepARG is a deep learning model specifically trained to identify ARGs belonging to over 30 different classes in bacterial genomes.

Our findings show that ARGs are present at all branches in the phylogenetic tree of our recovered MAGs, and several ARG classes have already spread throughout the entire phylogenetic tree. We also see many ARG classes that are prevalent among urban bacteria but have not yet fully spread throughout the phylogenetic tree. These ARG classes need further investigation to evaluate how quickly we must act before these ARGs spread across the entire phylogenetic tree.

## Materials and methods

### Urban metagenome selection

We extracted metagenomic samples belonging to urban environments (TMDB biome: “urban”) from the Terrestrial Metagenome Database (Corrêa et al., 2019), which is a part of the Collaborative Multi-domain Exploration of Terrestrial metagenomes (CLUE-TERRA) consortium.<sup>1</sup> Metagenomic samples in the CLUE-TERRA consortium have previously been filtered based on the following criteria: (i) Because non-metagenomic libraries in the Sequence Read Archive (SRA) can be wrongfully annotated as metagenomic, only true whole genome shotgun (WGS) libraries were kept. This was achieved using PARTIE (Torres et al., 2017), using default parameters. (ii) Metagenomes with sequence quality scores below 70%, determined using SRA-Tinder<sup>2</sup> with default parameters, were discarded. (iii) To allow for comparative studies, only metagenomes sequenced using the Illumina sequencing platform and with a minimum of 8 million paired-end reads per library were kept. (iv) Given the CLUE-TERRA consortium’s focus on terrestrial

environments, all libraries containing coordinates or terms for sea environments were excluded. After filtering, metagenomes from 1,023 SRA experiments remained from four different studies: (i) “Urban waterways sediment Metagenome” (BioProject: PRJNA267173, SRA: SRP051069) (Saxena et al., 2018), (ii) “New York City MTA subway samples Metagenome” (BioProject: PRJNA271013, SRA: SRP051511) (Afshinnikoo et al., 2015), (iii) “Metagenomics based spatiotemporal study of Chicago River microbiome” (BioProject: PRJNA336577, SRA: SRP080963), and (iv) “Antimicrobial resistance of urban water samples” (BioProject: PRJNA400857, SRA: SRP116665) (Supplementary Table 1). We downloaded the SRA run tables using the SRA run selector<sup>3</sup> and the SRA accession IDs.

### Pre-processing and library assembly

Metagenome-assembled genomes (MAGs) were recovered using the Multi-Domain Genome Recovery tool (MuDoGeR) (Nunes da Rocha et al., 2022). In short, the raw reads were quality-controlled using metaWrap (Uritskiy et al., 2018) with default parameters. Trimming of raw reads was performed using TrimGalore<sup>4</sup> with the default settings. High-quality reads (using default Phred scores from TrimGalore) were aligned to potential host genomes using bmtagger (Rotmistrovsky and Agarwala, 2011) using default parameters and the human build 38 patch release 12 database (GRCh38.p12). This alignment aims to remove human DNA contamination and read pairs with only a single aligned read from the metagenomic libraries. We used metaSpades (Nurk et al., 2017) to assemble the reads using default parameters. We then binned the assembled contigs into MAGs using Metabat2 (Kang et al., 2019), Maxbin2 (Wu et al., 2016), and CONCOCT (Alneberg et al., 2014). The recovered bins were refined and dereplicated (Nunes da Rocha et al., 2022). All assembled bins were quality-checked using CheckM (Parks et al., 2015) and filtered for metagenome-assembled genomes (MAGs) based on the following criteria: at least 50% completeness and less than 10% contamination based on CheckM results (medium and high-quality MAGs), and a quality score higher or equal to 50, where quality score = completeness-5\*contamination” (Parks et al., 2017; Supplementary Table 2).

### Genomic operational taxonomic unit clustering and taxonomy classification

We used the OTU picking script from Module 5 in the MuDoGeR tool (Nunes da Rocha et al., 2022) with default parameters to cluster the 4,281 MAGs into genomic operational taxonomic units (gOTUs). In short, the tool first filtered out MAGs that did not meet the following criteria: at least 50% completeness, a quality score higher than 50, and an N50 higher or equal to 10,000. It then divided the 2,073 remaining MAGs into taxonomic groups based on their GTDB-tk classifications (Chaumeil et al., 2019). Finally, it used FastANI (Jain et al., 2018) to cluster the MAGs into gOTUs based on an average nucleotide identity (ANI) of 95, which can be used as a proxy for species (Jain et al., 2018). To each gOTU, we assigned a “best bin” that represents the cluster. The “best bin” was defined as the bin with the highest quality score. In the

<sup>1</sup> <https://www.ufz.de/index.php?en=47300>

<sup>2</sup> [https://github.com/NCBI-Hackathons/SRA\\_Tinder](https://github.com/NCBI-Hackathons/SRA_Tinder)

<sup>3</sup> <https://trace.ncbi.nlm.nih.gov/Traces/study/>

<sup>4</sup> <https://github.com/FelixKrueger/TrimGalore>

event of a tie, we selected the bin with the lowest number of contigs, then the highest N50, and finally, the lowest strain heterogeneity. GTDB taxon names are often appended with an alphabetic, which means that the taxon is either not monophyletic in the GTDB reference tree or its placement in the tree is unstable (Parks et al., 2022). For simplicity, we have removed these suffixes in our analysis, but all unmodified GTDB-tk classification names are listed in [Supplementary Table 3](#).

## Antimicrobial gene annotation

We used DeepARG-LS (Arango-Argoty et al., 2018) to identify ARGs (ARGs) in each MAG. We first translated the MAG genome sequences to amino acid sequences using the faTrans tool from KentUtils (Kent, 2022). We defined ARGs as present in our MAGs when they met the following DeepARG-LS output criteria: equal or higher than 80% probability, an e-value lower than  $1 \times 10^{-10}$ , and percent identity of 35% or higher, as done by Wicaksono et al. in a recent study (Wicaksono et al., 2021; [Supplementary Table 4](#)). We chose the 35% identity cutoff to utilize the novel ARG prediction power of the DeepARG-LS model. A cutoff of 50% identity or higher allows for the prediction of high-quality ARGs (Arango-Argoty et al., 2018). However, because we are working with MAGs, which largely reflect the uncultured, and thus less known, proportion of urban prokaryotes, it is fitting to allow for the prediction of novel ARGs within our data. In order to test whether our findings resulted from an overestimation of ARGs, we repeated our analyses with the default DeepARG-LS 50% identity cutoff ([Supplementary Figures 1–4](#)). We converted all ARG identifiers to upper case to standardize the IDs and removed the differentiation between gene and protein identifiers. The correlation between the numbers of ARGs in an ARG class and the numbers of gOTUs containing at least one MAG with at least one ARG from an ARG class was calculated using Kendall's rank correlation. Kendall's tau can be used to evaluate the correlation of non-linear monotonic continuous data as an alternative to Pearson correlation, which assumes linearity, an assumption that our data does not fulfill (Puth et al., 2015).

## Predicting plasmid sequences and virulence factors

We used PlasFlow (Krawczyk et al., 2018) with default parameters and a threshold of 0.7 to predict the sequence types of all sequences that were at least 1,000 base pairs long in all of our MAGs ([Supplementary Table 5](#)). Additionally, we aligned the MAG nucleotide sequences to all amino acid sequences belonging to the Victors database of virulence factors (Sayers et al., 2019) using blastx. Virulence factors with an e-value lower than  $1 \times 10^{-10}$  and a percent identity above 80 were considered ([Supplementary Table 6](#)). The number of virulence factors per gOTU was calculated by counting the unique Victors virulence factors aligned to at least one of the MAGs belonging to the gOTU.

## Prevalence and randomness of ARGs in different taxonomic levels

Prevalences of ARGs and ARG classes in gOTUs were calculated based on the presence and absence of each ARG in each MAG. For each gOTU, the prevalence was calculated as the number of MAGs containing

the ARG or ARG belonging to an ARG class, divided by the number of MAGs belonging to the respective cluster, similar to the prevalence definition by Danko et al. (2021) ([Supplementary Table 7](#)). Weighted average prevalence (WAP) was calculated per ARG class on the phylum level by multiplying each phylum's prevalence with the number of gOTU species in that phylum and dividing the result with the total number of gOTUs, and finally adding up all the values:

$$WAP = \sum_{i=1}^p \frac{\text{prevalence}_i \times \text{number of gOTUs}_i}{\text{total number of gOTUs}}, \text{ where } p \text{ is the number}$$

of unique phyla. For each ARG class, we calculated the normalized mutual information (NMI) per taxonomic class using the taxonomic order level to categorize the data using the "aricode" package in R (Chiquet et al., 2020). In information theory, the mutual information (MI) of two variables is a measure that can compute the mutual dependency between those variables (Press et al., 1992). In other words, this measure quantifies the amount of information obtained about a variable by observing the other random variable. The value of the MI can be between  $[0, \infty]$ , where a zero value means that there is no dependency between two variables. The regular version of the MI is known as NMI, where values have been normalized to the range  $[0, 1]$ , and bias to the variable with a larger number of bins is removed (Singh et al., 2014). To test if the differences we observed using NMI were significant, we used Pearson's chi-squared test, which can be used to assess if there is a statistically significant difference between two categorical variables (Xia and Sun, 2017). When the chi-squared test's significance value was less than 0.05, we assigned a significant difference between the two variables used in the NMI calculation ([Supplementary Table 8](#)).

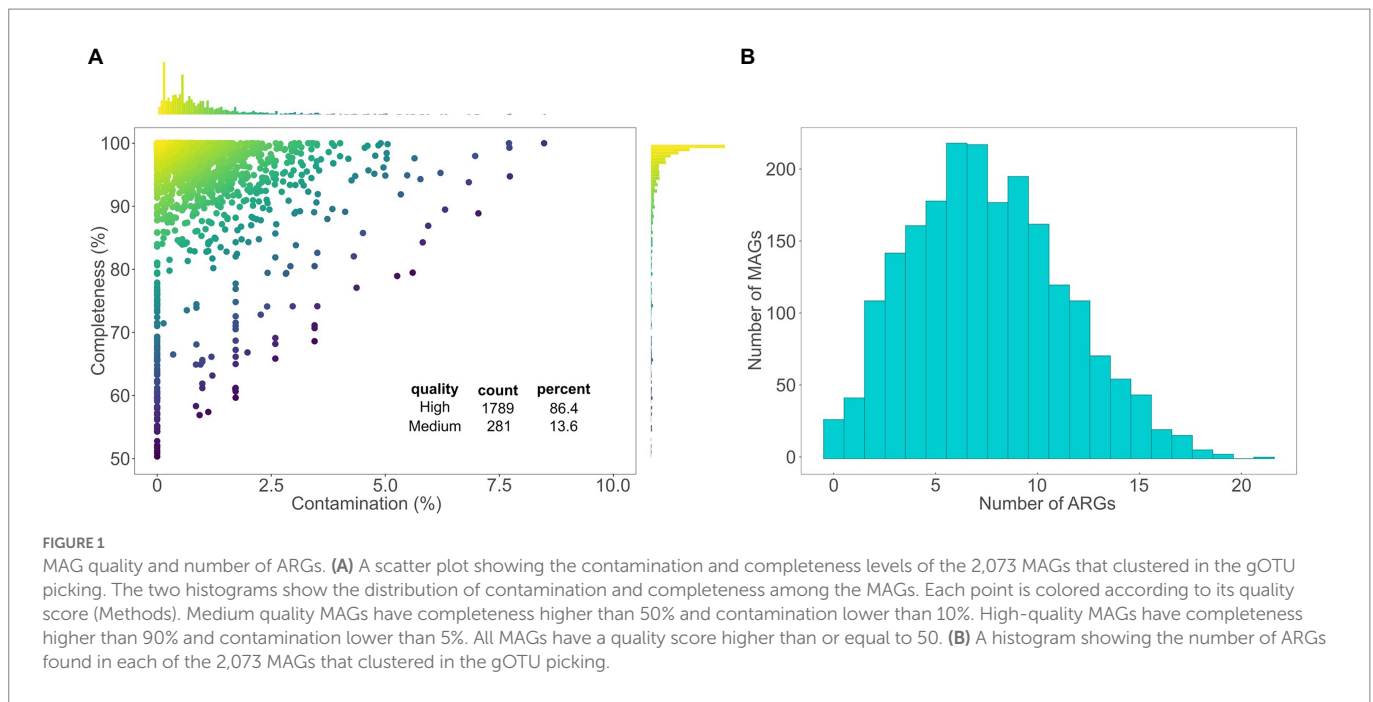
## Phylogenetic tree

Phylogenetic trees were created using the GTDB-tk taxonomic classifications of each MAG or gOTU. The GTDB-tk taxon levels were uploaded to the phyloT (v2) webpage,<sup>5</sup> (accessed 01.08.2022) to generate the tree (Letunic, 2022). As input, we used the lowest taxonomic level provided by the GTDB-tk classification of each gOTU's best MAG ([Supplementary Table 3](#)). However, not all identifiers were accepted as input in the phyloT tool. In these cases, we first used the phyloT search function to check for an alternate identifier for the same taxonomic level. When that did not result in a valid identifier, we instead used the next higher taxonomic level provided by GTDB-tk classification. At the lowest accepted levels, the 317 gOTUs belonged to 305 GTDB identifiers that were accepted as input into phyloT, out of which 283 are represented as leaves in the final tree. In the tree, gOTUs belonging to the same GTDB identifier cannot be differentiated. The tree and heatmap were plotted using the R package "gtree" (Guangchuang, 2020).

## Data availability statement

Metagenome-assembled genome sequences are available in the NCBI BioProject database (BioProject accession number PRJNA850115).

<sup>5</sup> <https://phyloT.biobyte.de/>



All analyzed data sets are available as [Supplementary material](#) on the publisher's page, and the code is available on GitHub.<sup>6</sup>

## Results

### The recovered urban MAGs are of high quality and taxonomically diverse

We recovered 1,396 medium (completeness >50, contamination <10) and 2,885 high (completeness >90, contamination <5) quality MAGs (4,281 in total) from 1,023 metagenomic samples of four different studies that contained urban terrestrial metagenomes from diverse surfaces ([Supplementary Table 1](#); [Supplementary Figure 5](#)). According to the GTDB-tk classification ([Chaumeil et al., 2019](#)), the 4,281 recovered MAGs belonged to 17 different phyla, [Supplementary Figure 6](#); [Supplementary Table 3](#). We performed genomic operational taxonomic unit (gOTU) clustering on the 4,281 MAGs. After filtering, 2,073 MAGs were divided into 317 gOTUs (i.e., distinct species; [Supplementary Table 3](#)), where each gOTU contains MAGs that belong to the same species. Of the 2,073 MAGs, 13.6% were of medium quality and 86.4% of high-quality ([Figure 1](#)). The 317 gOTUs were assigned to 259 known GTDB-tk species and 58 potentially novel species belonging to 13 phyla. The majority (292) of the gOTUs belonged to the four phyla most commonly found in urban metagenomic samples ([Danko et al., 2021](#)): *Actinobacteria* (15 gOTUs; 11 known species), *Bacteroidota* (34 gOTUs; 25 known species), *Firmicutes* (53 gOTUs; 26 known species), and *Proteobacteria* (190 gOTUs; 172 known species). Other phyla included *Cyanobacteria* (5 gOTUs; 2 known species) and *Patescibacteria* (12 gOTUs, 1 known species). We identified the presence of ARGs in each MAG using DeepARG-LS ([Arango-Argoty et al., 2018](#)). ARGs with

equal or higher than 80% probability, an e-value lower than  $1 \times 10^{-10}$ , and percent identity of 35% or higher ([Wicaksono et al., 2021](#)) were considered present. Because we are working with MAGs that largely reflect the uncultured, and thus less known, proportion of urban prokaryotes, it is fitting to allow for the prediction of novel ARGs within our data. Therefore, we chose the 35% percent identity cutoff to utilize the novel ARG prediction power of the DeepARG-LS model. Out of the 2,073 MAGs assigned to gOTUs, 2,046 (98.7%) contained at least one ARG, with an average of 8 ARGs per MAG ([Figure 1B](#)).

### Multiple antimicrobial resistance gene classes are highly prevalent in all detected branches of the urban bacterial phylogeny

We found that the ARGs showing resistance to multidrug and glycopeptides had already spread among most urban bacterial gOTUs ([Figure 2A](#)). We also saw that ARGs of macrolide-lincosamide-streptogramin (MLS), beta-lactam, and bacitracin were already present in two-thirds of our gOTUs. Beta-lactam, bacitracin, MLS, and unclassified ARGs followed the multidrug and glycopeptide ARGs, being present in over half of all gOTUs ([Figure 2A](#)). To investigate the spread of these ARGs in urban bacteria, we calculated the prevalence of all ARG classes among the MAGs within each gOTU ([Figure 2B](#)). We defined four categories based on ARG prevalence: i) ubiquitous (prevalence >0.75), ii) widespread ( $0.5 < \text{prevalence} \leq 0.75$ ), iii) common ( $0.25 < \text{prevalence} \leq 0.5$ ), and iv) sparse ( $0 \leq \text{prevalence} \leq 0.25$ ). Based on the dendrogram in [Figure 2B](#), we observed that, based on their prevalence on the phylum level, the 29 ARG classes could be divided into four groups, each group corresponding to one of the four categories based on the weighted average of the prevalence (WAP) among the four phyla. The first group mainly contained sparse classes, the second group contained the two ubiquitous classes, glycopeptide and multidrug, the third group contained mostly common classes, and the fourth group contained widespread classes.

<sup>6</sup> <https://github.com/mdsufz/UrbanARGs>



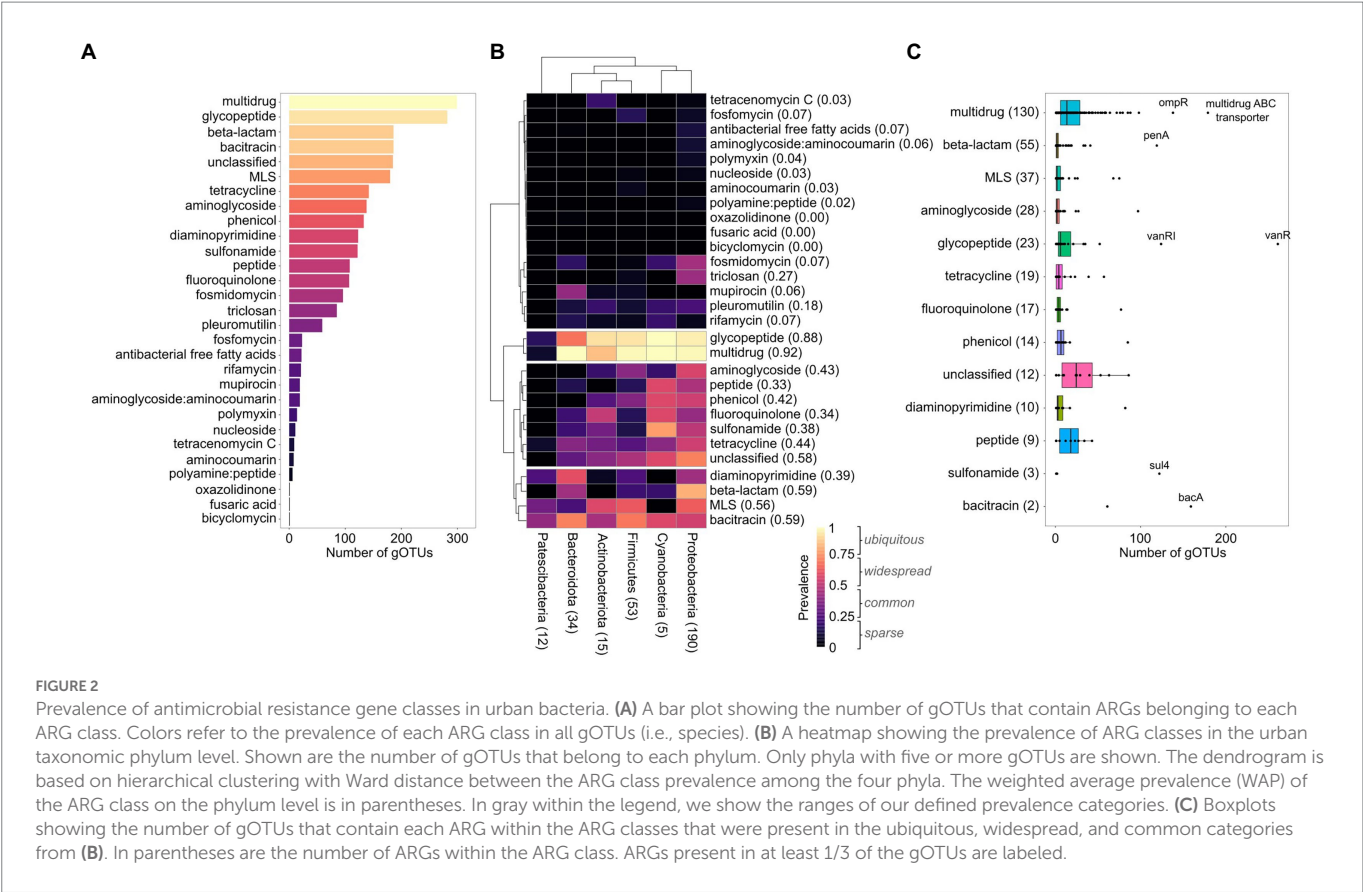


TABLE 1 Resistance mechanisms of the ARGs that were found in more than 100 gOTUs.

ARG	ARG class	Resistance mechanism
Multidrug ABC transporter	Multidrug	Antibiotic efflux*
ompR	Multidrug	Not available in CARD
penA	Beta-lactam	Antibiotic target alteration
vanR	Glycopeptide	Antibiotic target alteration
vanRI	Glycopeptide	Antibiotic target alteration
sul4	Sulfonamide	Antibiotic target replacement
bacA	Bacitracin	Antibiotic target alteration

Resistance mechanisms were taken from the CARD database (Alcock et al., 2019). \*Not available by name in CARD, but because the names include the word “transporter,” we assign the antibiotic efflux mechanism.

In total, we found 13 ARG classes that were present in 100 or more gOTUs (Figure 2A). However, we also investigated the prevalence of individual ARGs among the 317 gOTUs and found that 5 out of the 30 ARG classes contained a total of seven individual ARGs that were found in 100 or more gOTUs (Figure 2C). This means that the ARGs within some ARG classes had spread across different gOTUs. In the next section, we dive deeper into the spread of ARGs in the urban prokaryotic tree of life by testing whether ARGs were randomly spread across different taxonomic levels. According to the CARD database, most of the highly prevalent ARGs have target-altering or efflux pump resistance mechanisms (Table 1; Alcock et al., 2019).

Our analysis also showed that the number of ARGs within an ARG class was not indicative of the ARG class prevalence (Figure 2C). The

multidrug class was the most prevalent ARG class in our data and had the most individual ARGs. However, *vanR* was the most prevalent ARG in our data and belonged to the glycopeptide class, which encapsulates only a fifth of the number of ARGs in the multidrug class. In addition, the third most prevalent ARG (*bacA*) belongs to the bacitracin class, which contained only two ARGs in our data. To test this observation, we calculated the correlation between the number of unique ARGs per ARG class and its prevalence in our observed gOTUs (Kendall’s tau 0.69). However, this correlation was not observed when looking only at ARG classes belonging to the ubiquitous and widespread categories (Kendall’s tau 0.27).

To test if our results came from an overestimation of ARGs, we repeated our analyses with a percent identity cutoff of 50%, the universal cutoff according to DeepARG (Arango-Argoty et al., 2018), which removed the novel predicted ARGs from the data and allowed us to focus on the high-confidence ARGs. Despite the fewer ARGs (Supplementary Figure 1), we saw a widespread prevalence of the multidrug class and a common spread of the glycopeptide, bacitracin, beta-lactam, MLS, fosmidomycin, and aminoglycoside classes, which was largely driven by the prevalence of these classes among *Proteobacteria* and *Firmicutes* (Supplementary Figures 2, 3). When inspecting the individual high-confidence ARGs, we saw a similar distribution of ARGs among the gOTUs as seen with the results that included the novel predicted ARGs (Figure 2C; Supplementary Figure 4), except for a large reduction in predicted multidrug ABC transporters, which was one of the most abundant ARGs when including novel predicted ARGs. However, the fact that we still saw a widespread prevalence of the multidrug class when removing the novel ARGs indicated that our result did not stem from an overestimation of ARGs.

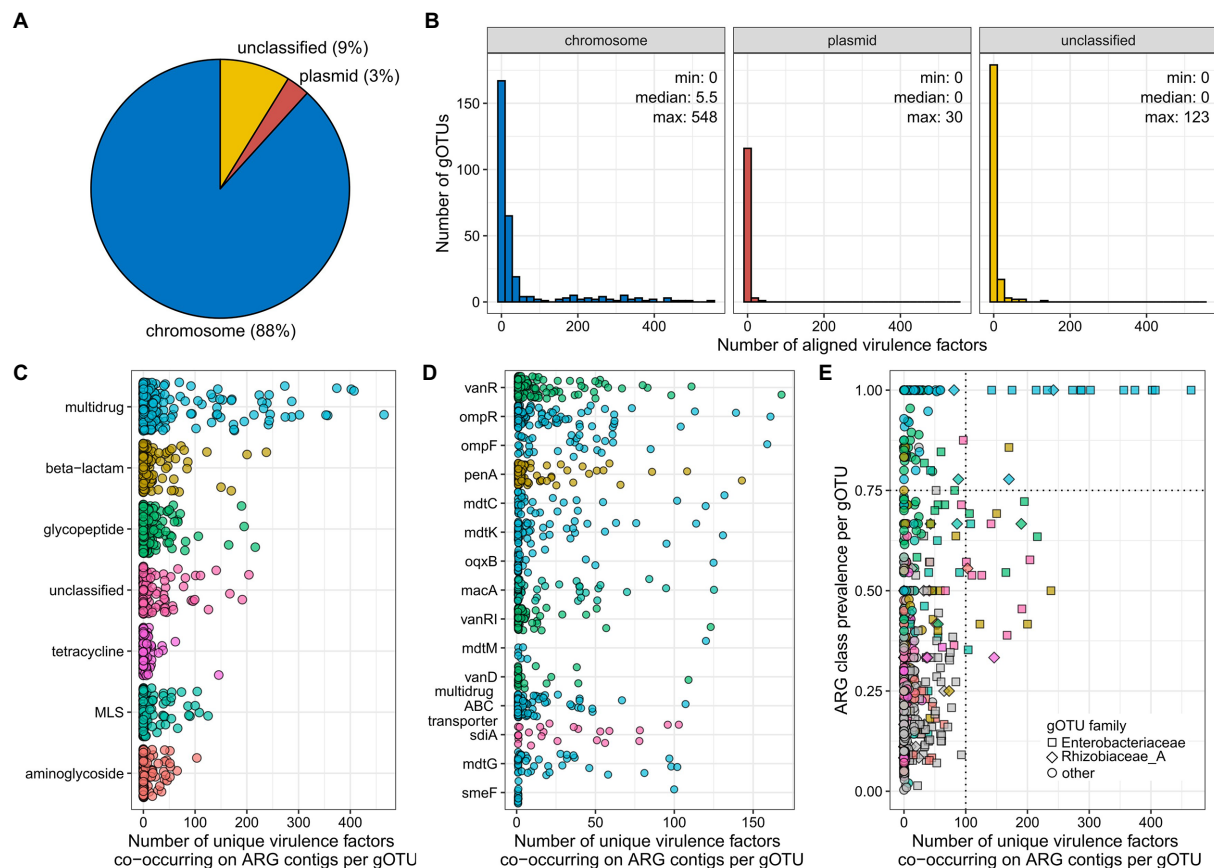


FIGURE 3

Occurrence of ARGs on plasmids and co-occurrence with virulence factors. (A) Percentages of ARGs found on chromosomes, plasmids, or unclassified reads in our data, as determined by PlasFlow (Krawczyk et al., 2018). (B) Histograms showing the number of aligned virulence factors (Sayers et al., 2019) per gOTU per read type (chromosome, plasmid, unclassified). (C) Number of unique virulence factors that co-occur on contigs with ARGs belonging to each ARG class, per gOTU. Show are only those ARG classes that had more than 100 co-occurring virulence factors in at least one gOTU. (D) Number of unique virulence factors that co-occur on contigs with ARGs belonging to the multidrug ARG class, per gOTU. Show are only those ARGs that had more than 100 co-occurring virulence factors in at least one gOTU. (E) ARG class prevalence per gOTU as a function of the number of unique virulence factors that co-occur on contigs with ARGs belonging to each ARG class, per gOTU. Highlighted are gOTUs that have both ubiquitous ARG classes and contain more than 100 co-occurring virulence factors for ARGs in the ARG class. The colors correspond to the ARG classes and the shapes to the gOTU taxonomic family. Shown are only those gOTUs that contained more than five MAGs.

We found that multidrug and glycopeptide ARGs were highly prevalent in all branches of the urban phylogenetic tree. Several other ARG classes showed similar trends as they were widespread in the urban phylogenetic tree.

## Virulence potential and ARG horizontal gene transfer among urban bacteria

In order to estimate the horizontal gene transfer potential of ARGs among urban bacteria, we identified the read type of all contigs in our MAGs as chromosome, plasmid, or unclassified, using PlasFlow (Krawczyk et al., 2018). Our results show that the majority of the urban ARGs in our study were found on chromosomes (91%), and only 2% were found on plasmids (Figure 3A). Additionally, we estimated the virulence of all MAGs by aligning their reads to the Victors database (Sayers et al., 2019), which contains the protein sequences of known bacterial virulence factors. We found that, on average,  $41 \pm 98$  (median: 0) virulence

factors aligned to each of the gOTUs, ranging from 0 to 550 virulence factors per gOTU (Figure 3B). Interestingly, the ARGs that were found on plasmids did not belong to the highly virulent gOTUs (Figure 3B). However, when we inspect the number of virulence factors per gOTU that co-occur with at least one ARG per ARG class, we saw that the ubiquitous and widespread ARG classes are frequently present in gOTU species that align to more virulence factors (Figure 3C) compared to the other ARG classes (Supplementary Figure 7). We found that gOTUs with ARGs belonging to the multidrug class have a significantly higher number of virulence factors than the other ARG classes (Wilcoxon Rank Sum test,  $p$ -value range  $4 \times 10^{-14}$  to  $8 \times 10^{-3}$ , Supplementary Table 9). We also found that five of the seven ARGs that were present in at least 100 gOTUs (Figure 2C) were co-occurring with more than 100 virulence factors in multiple gOTUs (Figure 3D). Finally, we observed several gOTUs from the *Enterobacteriaceae* and *Rhizobiaceae* families that contain ARGs from the ubiquitous multidrug or beta-lactam ARG classes in combination with having over 100 virulence factors aligned (Figure 3E; Table 2).

**TABLE 2** List of the gOTUs that had more than 100 virulence factors co-occurring with ARGs in ARG classes that were widespread within the gOTU, i.e., ARG class prevalence equal to or higher than 0.75 (Figure 3E).

Family	gOTU	Species	Number of MAGs in gOTU	ARG class	Prevalence	Virulence factors
Enterobacteriaceae	group-160-1	<i>Atlantibacter hermannii</i>	72	Multidrug	1.0	356
Enterobacteriaceae	group-52-1	<i>Atlantibacter</i> sp002358165	6	Multidrug	1.0	175
Enterobacteriaceae	group-164-1	<i>Enterobacter</i> _D sp000116015	13	Multidrug	1.0	374
Enterobacteriaceae	group-20-1	<i>Escherichia coli</i>	11	Multidrug	1.0	465
Enterobacteriaceae	group-255-1	<i>Klebsiella oxytoca</i>	8	Multidrug	1.0	287
Enterobacteriaceae	group-108-1	<i>Klebsiella variicola</i>	7	Multidrug	1.0	273
Enterobacteriaceae	group-108-1	<i>Klebsiella variicola</i>	7	Beta-lactam	0.86	170
Enterobacteriaceae	group-48-1	<i>Kosakonia cowanii</i>	13	Multidrug	1.0	214
Enterobacteriaceae	group-70-1	<i>Leclercia adecarboxylata</i>	156	Multidrug	1.0	407
Enterobacteriaceae	group-133-1	<i>Leclercia adecarboxylata</i> _C	12	Multidrug	1.0	402
Enterobacteriaceae	group-92-1	<i>Leclercia</i> sp002902985	10	Multidrug	1.0	232
Enterobacteriaceae	group-82-1	<i>Mixta calida</i>	12	Multidrug	1.0	142
Enterobacteriaceae	group-65-1	<i>Pseudoescherichia</i> sp002918705	11	Multidrug	1.0	302
Enterobacteriaceae	group-125-1	UBA7405 sp000755535	7	Multidrug	1.0	287
Rhizobiaceae_A	group-264-1	<i>Ochrobactrum intermedium</i>	9	Multidrug	0.78	170
Rhizobiaceae_A	group-256-1	<i>Ochrobactrum anthropi</i>	12	Multidrug	1.0	242

## Non-randomly spread ARG classes within the urban phylogenetic tree

Seeing the prevalence of ARG classes in urban bacteria, we wanted to test whether the ARGs within these classes were randomly distributed throughout the bacterial taxonomic levels or if there are deterministic events within the phylogenetic tree that affect the distributions of ARGs. We calculated each ARG class's normalized mutual information (NMI) at every taxonomic level in the phylogenetic tree to assess randomness. We found multiple nodes in the phylogenetic tree where the ARG distributions were flagged as non-random (Figure 4; Table 3). Non-random distributions of ARGs within ARG classes were present in the three most populous phyla; Proteobacteria, Firmicutes, and Bacteroidota. Most ARG classes whose ARGs show non-random distributions within the phylogenetic tree belong to the common and widespread prevalence categories (Figure 2B). The presence of individual ARGs within the taxons flagged as non-random are shown in Supplementary Figure 8.

Multiple ARG classes show non-random distributions among the urban bacterial taxonomic levels. It is possible that the spread of these classes is determined by bacterial phylogeny and that they will not reach the high prevalence of multidrug and glycopeptide resistance genes in urban bacteria. When specifically looking at ARGs that were aligned to predicted plasmids, we could not identify any non-random distributions of any ARG classes among the different taxonomic levels (Supplementary Table 10).

## Discussion

Our results show that two ARG classes are highly prevalent among urban bacteria. The taxonomic diversity of the MAGs we recovered and analyzed on a gOTU (i.e., species) level mirrored previously reported

urban taxonomic diversity (Danko et al., 2021). In this study, we identified multidrug and glycopeptide ARGs in nearly all urban bacterial species. One reason for the prevalence of the multidrug class is the spread of multidrug ABC transporters (Figure 2C; Table 1), which are among the most common resistance mechanisms in bacteria (Allen et al., 2010) and are commonly found even in pristine environments such as Antarctica (Van Goethem et al., 2018). Additionally, our data showed an alarming ubiquitous prevalence of glycopeptide ARGs and widespread prevalence of MLS, bacitracin, and beta-lactam ARGs. However, despite the high prevalence of these ARG classes in urban bacteria, our analysis suggests that the incidence of horizontal gene transfer among these bacteria is low and that the urban bacteria identified in this study are mostly non-virulent. However, we see that a subset of the putative virulent bacteria contained ARGs that belong to the widespread and ubiquitous ARG classes. However, further studies should be carried out to confirm if the putative virulent bacteria identified in our study are actually virulent. Therefore, we cannot confirm the risk assessment of these bacteria and indicate whether our results are cause for concern.

Our study has some limitations which need addressing. At the time of analysis, the database used to train the DeepARG models had not been updated since April 2020. Therefore, ARGs may be missing in our study. Homology searches for ARGs using the CARD database, updated every three months, are possible but have lower accuracy than deep learning algorithms (Bileschi et al., 2022). Urban microbiomes from similar environments across the globe have different microbial compositions and distributions of ARG classes (Danko et al., 2021). Our samples are largely represented by samples from the United States, and it is thus important to keep in mind that other geographical locations may differ in both taxonomic and ARG diversity and prevalence.

Our results showed that most of the urban bacteria in our study had few virulence factors, which suggests that they are unlikely to be pathogenic to humans. However, we observed many gOTUs with



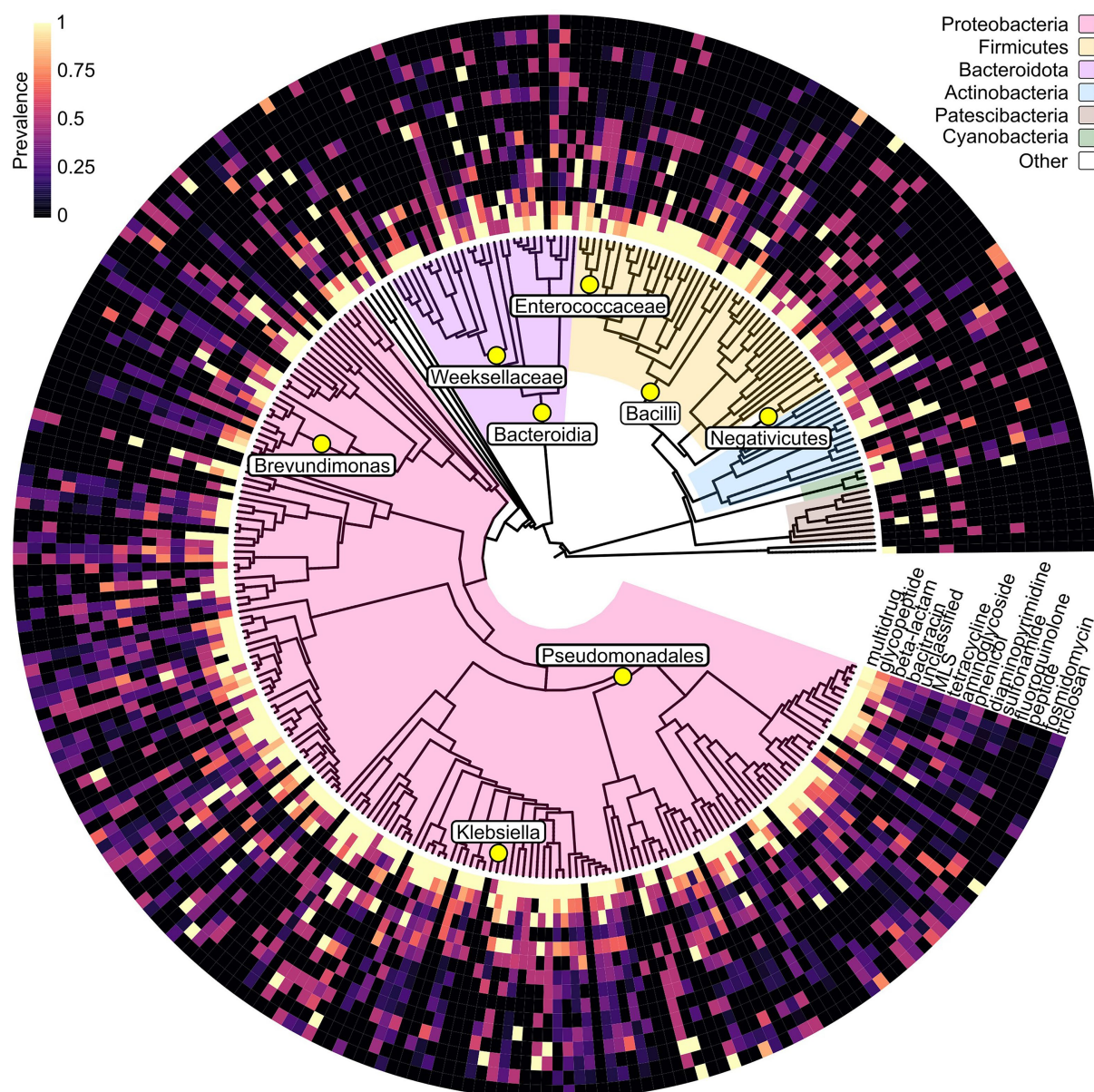


FIGURE 4

A phylogenetic tree of the urban gOTUs based on their GTDB-tk taxonomy. Tree branches are colored based on their phyla. The surrounding heatmap shows per cluster the prevalence of ARGs belonging to the 15 most prevalent ARG classes in the MAGs belonging to each gOTU. Some gOTUs were identified by GTDB-tk to belong to the same taxonomic category, and the prevalence was then calculated based on all MAGs belonging to all the affected gOTUs. Yellow circles show the presence of non-randomly distributed ARG classes at the taxonomic level displayed in the tree, and the labels show the corresponding taxons. Non-randomness was determined with Normalized Mutual Information (NMI)  $\geq 0.5$  and chi-squared  $p$ -value  $< 0.05$ . Table 3 lists the taxons with non-random distributions of ARGs within the listed ARG class, along with the NMI and  $p$ -values.

hundreds of virulence factors in addition to highly prevalent ARGs, and the combination of these two factors suggests the presence of antimicrobial-resistant pathogenic bacteria. We have highlighted bacterial species that have high numbers of virulence factors, in combination with the highly prevalent ARG classes, and suggest that further research is needed to assess the potential risk from these species to humans. We identified 15 gOTUs that contained a high number of virulence factors and were ubiquitous in multidrug or beta-lactam ARG classes (Table 2), most of which belonged to the Enterobacteriaceae family, which is known hosts several human pathogens. In addition, some of these identified species and genera are recorded as emerging human pathogens, which further highlights the need for risk assessment

of urban microbes (Rodríguez-Medina et al., 2019; Zayet et al., 2021; Jeyaraman et al., 2022).

Most of the ARG classes identified in this study were randomly distributed among the gOTUs (i.e., different species), indicating that there were no deterministic factors on the phylogenetic level that determined the acquiring of ARGs in urban bacteria. This is in line with the results from a recent study where they could not find strong links between phylogeny and ARG diversity in urban sewage (Hendriksen et al., 2019). However, we did see several non-random distributions of ARGs in the urban bacterial phylogenetic tree, which indicates that phylogenetic traits possibly determine the distribution of some ARGs. The question is whether all ARG classes will follow the trends of



**TABLE 3** Normalized mutual information (NMI) value and chi-squared *p*-value based on the presence of ARGs within each ARG class within the different taxons starred in [Figure 4](#).

Phylum	Taxon	ARG class	NMI	<i>p</i> -value	<i>n</i>
Firmicutes	Bacilli	Phenicol	0.82	$7.7 \times 10^{-5}$	15
Firmicutes	Bacilli	Diaminopyrimidine	0.55	$3.5 \times 10^{-3}$	13
Bacteroidota	Bacteroidia	MLS	0.65	$8.3 \times 10^{-3}$	11
Bacteroidota	Weeksellaceae	Beta-lactam	0.63	$4.6 \times 10^{-2}$	15
Proteobacteria	Negativicutes	MLS	0.52	$5.0 \times 10^{-2}$	6
Proteobacteria	Pseudomonadales	Antibacterial_free_fatty_acids	1.00	$1.1 \times 10^{-2}$	10
Proteobacteria	Klebsiella	MLS	0.51	$4.4 \times 10^{-2}$	8
Proteobacteria	Brevundimonas	Beta-lactam	0.63	$4.6 \times 10^{-2}$	12
Proteobacteria	Enterobacteriaceae	Aminoglycoside	0.84	$1.4 \times 10^{-2}$	8

Also listed are the taxon phyla and the number of MAGs within the taxon that contained at least one ARG within the ARG class. Here we list only those taxons with an NMI above 0.5 and a *p*-value lower than 0.05. For the complete results, see [Supplementary Table 8](#).

multidrug and glycopeptide resistance gene spread or if they are determined by phylogeny and will not spread to all urban bacteria. Because antimicrobial resistance in human pathogens is an increasing global problem and that increasing proportions of the global population live in urban environments, antimicrobial resistance genes must not spread fully throughout all urban bacteria. We identified two categories where this is already the case, but we may still be able to prevent other categories from reaching the same level of spreading.

This study includes a large data set of MAGs recovered from over 1,000 urban samples. However, our study covers only a fraction of the possible urban locations worldwide, even with a high number of samples. It would be interesting to perform similar analyses as those we describe in this study on a larger data set (with 10–50 thousand samples) representing different and balanced urban environments. Further research will be required to expand the analysis to more urban metagenomic samples. Another question is whether we see similar antimicrobial resistance gene spreads in other environments (e.g., natural and pristine ecosystems). The urban environment is highly anthropogenic, and it would be interesting to compare our results to further analyses of other highly anthropogenic environments and pristine environments that humans have impacted less.

With this study, we would like to raise awareness that urban ARGs are highly prevalent when analyzing metagenome-assembled genomes and emphasize the need for risk assessment studies at both local and global scales, as these are environments that most of the global population interacts with on a daily basis.

In conclusion, our results showed a high prevalence of antimicrobial resistance genes within the urban bacterial phylogenetic tree. ARGs belonging to two commonly used antimicrobial classes, multidrug and glycopeptide, were ubiquitous among urban bacteria, and multiple other ARG classes follow a similar trend. Additionally, we also saw multiple ARG classes with non-random distributions among urban bacteria. However, we found little evidence that these non-randomly distributed ARGs were transmitted *via* plasmids. Further analyses of additional data sets, more diverse environments with a balanced number of samples, and ARG transmission among bacteria are needed to fully explore the deterministic effects of phylogeny on the distribution of ARGs among urban bacteria.

## Data availability statement

The data availability statement needs to be changed to “Metagenome-assembled genome sequences are available in the NCBI BioProject database (BioProject accession numbers PRJNA843551 and PRJNA850115)”.

## Author contributions

SM: investigation, formal analysis, visualization, and writing. JS and AB: methodology and critical review. MS and RT: methodology. JZ: critical review. UR: conceptualization, supervision, and critical review. All authors reviewed and agreed to the content of the manuscript.

## Funding

This work was funded by the NFDI4Microbiota consortium, funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project number 460129525. UR, RT, and JS were financed by the Helmholtz Young Investigator grant VH-NG-1248 Micro “Big Data”.

## Acknowledgments

The scientific results were computed at the High-Performance Computing (HPC) Cluster EVE, a joint effort of both the Helmholtz Centre for Environmental Research - UFZ (<http://www.ufz.de/>) and the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig (<http://www.idiv-biodiversity.de/>). We would like to thank the administration and support staff of EVE who keep the system running and support us with our scientific computing needs: Thomas Schnicke, Ben Langenberg, Guido Schramm, Toni Harzendorf, Tom Strempel and Lisa Schurack from the UFZ, and Christian Krause from iDiv. We would like to thank the CLUE-TERRA consortium (<https://www.ufz.de/index.php?en=47300>) for making available the Terrestrial Metagenome Database (<https://webapp.ufz.de/tmdb/>). Additional members of the

CLUE-TERRA consortium are: Alexander Bartholomäus, Alexandre Soares Rosado, Ana-Maria Fiore-Donno, André Carlos Ponce de Leon Ferreira de Carvalho, Cecile Gubry-Rangin, Daniel Machado, Danilo S. Sanches, Dirk Wagner, Gabriele Berg, Ines Mundic Mulec, Marie Muehe, Michael Bonkowski, Newton Gomes, Raquel Peixoto, Rodrigo Costa, Sabine Kleinstuber, Simonetta Gribaldo, Tina Keller Costa, and Vivian Pellizari.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

- Afshinnekoo, E., Meydan, C., Chowdhury, S., Jaroudi, D., Boyer, C., Bernstein, N., et al. (2015). Geospatial resolution of human and bacterial diversity with City-scale Metagenomics. *Cell Syst.* 1, 72–87. doi: 10.1016/j.cels.2015.01.001
- Alcock, B. P., Raphenya, A. R., Lau, T. T. Y., Tsang, K. K., Bouchard, M., Edalatmand, A., et al. (2019). CARD 2020: antibiotic resistance surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res.* 48, D517–D525. doi: 10.1093/nar/gkz935
- Allen, H. K., Donato, J., Wang, H. H., Cloud-Hansen, K. A., Davies, J., and Handelsman, J. (2010). Call of the wild: antibiotic resistance genes in natural environments. *Nat. Rev. Microbiol.* 8, 251–259. doi: 10.1038/nrmicro2312
- Alneberg, J., Bjarnason, B. S., de Bruijn, I., Schirmer, M., Quick, J., Ijaz, U. Z., et al. (2014). Binning metagenomic contigs by coverage and composition. *Nat. Methods* 11, 1144–1146. doi: 10.1038/nmeth.3103
- Arango-Argoty, G., Garner, E., Pruden, A., Heath, L. S., Vikesland, P., and Zhang, L. (2018). DeepARG: a deep learning approach for predicting antibiotic resistance genes from metagenomic data. *Microbiome* 6:23. doi: 10.1186/s40168-018-0401-z
- Bileschi, M. L., Belanger, D., Bryant, D. H., Sanderson, T., Carter, B., Sculley, D., et al. (2022). Using deep learning to annotate the protein universe. *Nat. Biotechnol.* 40, 932–937. doi: 10.1038/s41587-021-01179-w
- Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P., and Parks, D. H. (2019). GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. *Bioinformatics* 36, 1925–1927. doi: 10.1093/bioinformatics/btz848
- Chiquet, J., Rigail, G., and Sundqvist, M. (2020). Aricode: efficient computations of standard clustering comparison measures. Available at: <https://CRAN.R-project.org/package=aricode> (Accessed August 27, 2022).
- Corrêa, F. B., Saraiva, J. P., Stadler, P. F., and Nunes da Rocha, U. (2019). TerrestrialMetagenomeDB: a public repository of curated and standardized metadata for terrestrial metagenomes. *Nucleic Acids Res.* 48, D626–D632. doi: 10.1093/nar/gkz994
- Dancer, S. J., Shears, P., and Platt, D. J. (1997). Isolation and characterization of coliforms from glacial ice and water in Canada's high Arctic. *J. Appl. Microbiol.* 82, 597–609. doi: 10.1111/j.1365-2672.1997.tb03590.x
- Danko, D., Bezdan, D., Afshin, E. E., Ahsanuddin, S., Bhattacharya, C., Butler, D. J., et al. (2021). A global metagenomic map of urban microbiomes and antimicrobial resistance. *Cells* 10:317. doi: 10.1016/j.cell.2021.05.002
- Guangchuang, Y. (2020). Using ggtree to visualize data on tree-like structures. *Curr. Protoc. Bioinformatics* 69:e96. doi: 10.1002/cpbi.96
- Hendriksen, R. S., Munk, P., Njage, P., van Bunnik, B., McNally, L., Lukjancenko, O., et al. (2019). Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. *Nat. Commun.* 10:1124. doi: 10.1038/s41467-019-08853-3
- Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., and Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* 9:5114. doi: 10.1038/s41467-018-07641-9
- Jeyaraman, M., Muthu, S., Sarangan, P., Jeyaraman, N., and Packiyarathinam, R. P. (2022). Ochrobactrum anthropi - An emerging opportunistic pathogen in musculoskeletal disorders - a case report and review of literature. *J. Orthop. Case Rep.* 12, 85–90. doi: 10.13107/jocr.2022.v12.i03.2730
- Kang, D. D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., et al. (2019). MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* 7:e7359. doi: 10.7717/peerj.7359
- Kent, J. (2022). Genome browser and blat application binaries built for standalone command-line use on various supported Linux and UNIX platforms. Available at: <http://hgdownload.soe.ucsc.edu/admin/exe/> (Accessed August 27, 2022).
- Krawczyk, P. S., Lipinski, L., and Dziembowski, A. (2018). PlasFlow: predicting plasmid sequences in metagenomic data using genome signatures. *Nucleic Acids Res.* 46:e35. doi: 10.1093/nar/gkx1321
- Letunic, I. (2022). phyloT: phylogenetic tree generator v2. Available at: [phylot.biobyte.de](http://phylot.biobyte.de) (Accessed September 03, 2022).
- Nunes da Rocha, U., Kasmanas, J. C., Kallies, R., Saraiva, J. P., Toscan, R. B., Štefanić, P., et al. (2022). MuDoGeR: Multi-domain genome recovery from metagenomes made easy. *bioRxiv* 496983. doi: 10.1101/2022.06.21.496983
- Nurk, S., Meleshko, D., Korobeynikov, A., and Pevzner, P. A. (2017). metaSPAdes: a new versatile metagenomic assembler. *Genome Res.* 27, 824–834. doi: 10.1101/gr.213959.116
- Parks, D. H., Chuvochina, M., Rinke, C., Mussig, A. J., Chaumeil, P.-A., and Hugenholtz, P. (2022). GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res.* 50, D785–D794. doi: 10.1093/nar/gkab776
- Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25, 1043–1055. doi: 10.1101/gr.186072.114
- Parks, D. H., Rinke, C., Chuvochina, M., Chaumeil, P.-A., Woodcroft, B. J., Evans, P. N., et al. (2017). Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nat. Microbiol.* 2, 1533–1542. doi: 10.1038/s41564-017-0012-7
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1992). "Numerical recipes in C" in *The Art of Scientific Computing*. 2nd. ed. W. H. Press (Cambridge: Cambridge University Press)
- Putth, M.-T., Neuhäuser, M., and Ruxton, G. D. (2015). Effective use of Spearman's and Kendall's correlation coefficients for association between two measured traits. *Anim. Behav.* 102, 77–84. doi: 10.1016/j.anbehav.2015.01.010
- Rodríguez-Medina, N., Barrios-Camacho, H., Duran-Bedolla, J., and Garza-Ramos, U. (2019). Klebsiella variicola: an emerging pathogen in humans. *Emerg. Microb. Infect.* 8, 973–988. doi: 10.1080/22221751.2019.1634981
- Rotmistrovsky, K., and Agarwala, R. (2011). BMTagger: Best match tagger for removing human reads from metagenomics datasets. Available at: <http://www.mmnt.net/db/0/0/ftp.ncbi.nlm.nih.gov/pub/agarwala/bmtagger> (Accessed August 11, 2022).
- Saxena, G., Mitra, S., Marzinelli, E. M., Xie, C., Wei, T. J., Steinberg, P. D., et al. (2018). Metagenomics reveals the influence of land use and rain on the benthic microbial communities in a tropical urban waterway. *mSystems* 3, e00136–e00117. doi: 10.1128/mSystems.00136-17
- Sayers, S., Li, L., Ong, E., Deng, S., Fu, G., Lin, Y., et al. (2019). Victors: a web-based knowledge base of virulence factors in human and animal pathogens. *Nucleic Acids Res.* 47, D693–D700. doi: 10.1093/nar/gky999
- Singh, B., Kushwaha, N., and Vyas, O. P. (2014). A feature subset selection technique for high dimensional data using symmetric uncertainty. *JDAIP* 02, 95–105. doi: 10.4236/jdaip.2014.24012
- Song, J. S., Jeon, J. H., Lee, J. H., Jeong, S. H., Jeong, B. C., Kim, S.-J., et al. (2005). Molecular characterization of TEM-type beta-lactamases identified in cold-seep sediments of Edison seamount (south of Lihir Island, Papua New Guinea). *J. Microbiol. Seoul Korea* 43, 172–178.
- Torres, P. J., Edwards, R. A., and McNair, K. A. (2017). PARTIE: a partition engine to separate metagenomic and amplicon projects in the sequence read archive. *Bioinformatics* 33, 2389–2391. doi: 10.1093/bioinformatics/btx184
- United Nations (2016). Political declaration of the high-level meeting of the general assembly on antimicrobial resistance. United Nations, New York. Available at: [https://digitallibrary.un.org/record/842813/files/A\\_71\\_L-2-EN.pdf?ln=en](https://digitallibrary.un.org/record/842813/files/A_71_L-2-EN.pdf?ln=en) (Accessed July 26 2022).
- United Nations (2018). World economic situation and prospects 2018. United Nations, New York. Available at: <https://population.un.org/wup/Publications/Files/WUP2018-Report.pdf> (Accessed July 26, 2022).

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2023.1037845/full#supplementary-material>

- Uritskiy, G. V., DiRuggiero, J., and Taylor, J. (2018). MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome* 6:158. doi: 10.1186/s40168-018-0541-1
- Van Goethem, M. W., Pierneef, R., Bezuidt, O. K. I., Van De Peer, Y., Cowan, D. A., and Makhallanyane, T. P. (2018). A reservoir of 'historical' antibiotic resistance genes in remote pristine Antarctic soils. *Microbiome* 6:40. doi: 10.1186/s40168-018-0424-5
- Wicaksono, W. A., Kusstatscher, P., Erschen, S., Reisenhofer-Graber, T., Grube, M., Cernava, T., et al. (2021). Antimicrobial-specific response from resistance gene carriers studied in a natural, highly diverse microbiome. *Microbiome* 9:29. doi: 10.1186/s40168-020-00982-y
- Wu, Y.-W., Simmons, B. A., and Singer, S. W. (2016). MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* 32, 605–607. doi: 10.1093/bioinformatics/btv638
- Xia, Y., and Sun, J. (2017). Hypothesis testing and statistical analysis of microbiome. *Genes Dis.* 4, 138–148. doi: 10.1016/j.gendis.2017.06.001
- Zayet, S., Lang, S., Garnier, P., Pierron, A., Plantin, J., Toko, L., et al. (2021). Leclercia adecarboxylata as emerging pathogen in human infections: clinical features and antimicrobial susceptibility testing. *Pathogens* 10:1399. doi: 10.3390/pathogens10111399



## OPEN ACCESS

## EDITED BY

Graciela L. Lorca,  
University of Florida, United States

## REVIEWED BY

Christoph Engl,  
Queen Mary University of London,  
United Kingdom  
Roshni R. Kharadi,  
Corteva Agriscience, United States

## \*CORRESPONDENCE

Emily E. Weinert  
✉ emily.weinert@psu.edu

RECEIVED 30 December 2022

ACCEPTED 26 June 2023

PUBLISHED 07 July 2023

## CITATION

Fekete FJ, Marotta NJ, Liu X and  
Weinert EE (2023) An O<sub>2</sub>-sensing diguanylate  
cyclase broadly affects the aerobic  
transcriptome in the phytopathogen  
*Pectobacterium carotovorum*.  
*Front. Microbiol.* 14:1134742.  
doi: 10.3389/fmicb.2023.1134742

## COPYRIGHT

© 2023 Fekete, Marotta, Liu and Weinert. This  
is an open-access article distributed under the  
terms of the [Creative Commons Attribution  
License \(CC BY\)](#). The use, distribution or  
reproduction in other forums is permitted,  
provided the original author(s) and the  
copyright owner(s) are credited and that the  
original publication in this journal is cited, in  
accordance with accepted academic practice.  
No use, distribution or reproduction is  
permitted which does not comply with these  
terms.

# An O<sub>2</sub>-sensing diguanylate cyclase broadly affects the aerobic transcriptome in the phytopathogen *Pectobacterium carotovorum*

Florian J. Fekete<sup>1</sup>, Nick J. Marotta<sup>2</sup>, Xuanyu Liu<sup>1</sup> and  
Emily E. Weinert<sup>1,3\*</sup>

<sup>1</sup>Department of Biochemistry and Molecular Biology, Penn State University, University Park, PA, United States, <sup>2</sup>Graduate Program in Molecular, Cellular, and Integrative Biosciences, Penn State University, University Park, PA, United States, <sup>3</sup>Department of Chemistry, Penn State University, University Park, PA, United States

*Pectobacterium carotovorum* is an important plant pathogen responsible for the destruction of crops through bacterial soft rot, which is modulated by oxygen (O<sub>2</sub>) concentration. A soluble globin coupled sensor protein, *Pcc* DgcO (also referred to as *Pcc*GCS) is one way through which *P. carotovorum* senses oxygen. DgcO contains a diguanylate cyclase output domain producing c-di-GMP. Synthesis of the bacterial second messenger c-di-GMP is increased upon oxygen binding to the sensory globin domain. This work seeks to understand regulation of function by DgcO at the transcript level. RNA sequencing and differential expression analysis revealed that the deletion of DgcO only affects transcript levels in cells grown under aerobic conditions. Differential expression analysis showed that DgcO deletion alters transcript levels for metal transporters. These results, followed by inductively coupled plasma–mass spectrometry showing decreased concentrations of six biologically relevant metals upon DgcO deletion, provide evidence that a globin coupled sensor can affect cellular metal content. These findings improve the understanding of the transcript level control of O<sub>2</sub>-dependent phenotypes in an important phytopathogen and establish a basis for further studies on c-di-GMP-dependent functions in *P. carotovorum*.

## KEYWORDS

cyclic-di-GMP, oxygen, *Pectobacterium*, globin, diguanylate cyclase

## 1. Introduction

Bacteria sense and respond to environmental signals such as nutrient availability, temperature, signals from other organisms, and oxygen (O<sub>2</sub>) (Miller and Bassler, 2001; Toth and Birch, 2005; Charkowski, 2006). Obligate and facultatively aerobic organisms can use molecular O<sub>2</sub> as a terminal electron acceptor, while many obligate anaerobe species are inhibited by its presence (Taylor, 1983). O<sub>2</sub> availability can influence numerous bacterial functions, such as motility, virulence factor expression, and host colonization (Miller and Bassler, 2001; Edwards et al., 2006; Walker et al., 2017). O<sub>2</sub> can be sensed directly by ligand binding or indirectly through redox sensors (Edwards et al., 2006; Vinogradov et al., 2013). Globin coupled sensor (GCS) proteins are a wide-spread family of direct gas sensing proteins that have been identified in



hundreds of bacterial genomes and are also predicted in some archaea and lower eukaryotes (Vinogradov et al., 2013; Walker et al., 2017). GCS proteins consist of a heme-containing globin domain, which binds diatomic ligands (O<sub>2</sub>, NO, CO) at the iron center, linked by a variable middle domain to a variety of output domains, such as phosphodiesterases, methyl-accepting chemotaxis proteins, adenylate cyclases, and diguanylate cyclases (Vinogradov et al., 2013; Walker et al., 2017).

Commonly found in prokaryotes, GGDEF domains are responsible for synthesizing the second messenger bis-(3'-5')-cyclic diguanylate (c-di-GMP) (Römling et al., 2013; Jenal et al., 2017; Hengge, 2021), which controls functions such as motility, virulence, and biofilm formation (Boyd and O'Toole, 2012; Jenal et al., 2017; Hengge, 2021). C-di-GMP levels can regulate cellular functions by directly interacting with target proteins or through regulating the activity of riboswitches (Sudarsan et al., 2008; Boyd and O'Toole, 2012; Jenal et al., 2017; Hengge, 2021). Evidence suggests that, besides the overall cellular c-di-GMP pool, local concentrations of c-di-GMP also affect protein activity, and thus cellular function, in a specific, localized manner (Hengge, 2021). These interactions can involve GGDEF proteins closely interacting with their targets, tightly regulating local c-di-GMP concentrations and therefore protein activity (Hengge, 2021). While numerous targets and ways of regulation have been identified, the full extent of bacterial c-di-GMP signaling remains uncharacterized (Hengge, 2010; Boyd and O'Toole, 2012; Römling et al., 2013; Jenal et al., 2017; Hengge, 2021).

The phytopathogen *Pectobacterium carotovorum* subsp. *carotovorum* (*Pcc*) is a leading cause of bacterial soft rot and is responsible for crop losses worth millions of US dollars each year (Toth and Birch, 2005; Charkowski, 2006; Kim et al., 2009). Upon infecting the plant host, *Pcc* secretes plant cell wall degrading exoenzymes (PCWDEs); this varied group of pectate lyases, cellulases, and polysaccharide lyases is responsible for the rotting of plant tissue (Toth and Birch, 2005; Charkowski, 2006; Babujee et al., 2012; Pöllumaa et al., 2012). *Pcc* is known to exhibit increased virulence under low oxygen concentrations, which is often caused by flooding of fields (Babujee et al., 2012). *Pcc* encodes for one GCS protein, previously known as *PccGCS*, which we hereby propose to rename *PccDgcO* based on the conserved GGDEF motif and the previously established systematic nomenclature of *E. coli* GGDEF domain-containing proteins and based on the similarity of *PccGCS* to the *E. coli* GCS protein *EcDgcO* (also known as *EcDosC*) (Tuckerman et al., 2009; Burns et al., 2014; Tarnawski et al., 2015; Hengge et al., 2016; Burns et al., 2017; Walker et al., 2017). Biochemical characterization of *PccDgcO* was demonstrated diguanylate cyclase activity in response to O<sub>2</sub> binding to the sensor globin and a number of phenotypes associated with *PccDgcO* have been studied in the *Pcc* strain WPP14 (Kim et al., 2009; Burns et al., 2014, 2017; Walker et al., 2017; Rivera et al., 2018). *PccDgcO* is known to regulate virulence, motility and extracellular levels of the quorum sensing molecule N-acyl homoserine lactone (AHL) under aerobic conditions, but has not been associated with phenotypes anaerobically (Miller and Bassler, 2001; Burns et al., 2017).

Transcriptomics, proteomics, phenotypic assays, and inductively coupled plasma-mass spectrometry (ICP-MS) were used to investigate the full role of *PccDgcO* in regulating cellular function. Results of these studies provide new insights into how *PccDgcO* regulates previously described phenotypes, identifies additional protein

interactors of *PccDgcO*, and provides evidence of *PccDgcO*-dependent control of cellular metal homeostasis. Overall, this work provides insight into the *PccDgcO*-dependent regulation of cellular oxygen response in *P. carotovorum*.

## 2. Materials and methods

### 2.1. Bacterial strains and general information

Wild-type (WT) *Pcc* WPP14 and the *PccDgcO* deletion strain ( $\Delta dgcO$ ) strains were used in all assays, unless otherwise noted. Construction of the  $\Delta dgcO$  strain was performed by Creative Biogene, Inc. (Shirley, NY, United States). Briefly, the *dgcO* gene was PCR amplified from WT *Pcc* and cloned into a pLP12 suicide vector conferring resistance to chloramphenicol (Luo et al., 2015), and transformed into *E. coli* DH5 $\alpha$  cells. Plasmids extracted from positive clones were then electroporated into *E. coli* strain  $\beta$ 2163 cells, which was then used to conjugate the deletion sequence into WT *Pcc* cells to delete the *dgcO* gene through homologous recombination. The deletion region was then verified by PCR amplification of the target sequence, followed by sequencing to verify the presence of a scar sequence (Supplementary file S1), as commonly used in the literature (Mole et al., 2010; Marquez-Villavicencio et al., 2011; Bowden et al., 2013; Koiv et al., 2013; Ernst et al., 2020). Primer sequences for amplification and sequencing of deletion region (5'  $\rightarrow$  3'): GATCGCCATCAGTGCCAAGG (anneals bp. 1781101–1781120), CATCGTTGGAAGAGCAAGCAG (anneals bp. 1783954–1783974). Primers annealing within *dgcO* were designed and used to verify the knockout (5'  $\rightarrow$  3'): GAACGATCATGAATTC (anneals bp. 1782149–1782164), GCTGACGCGTTTAC (anneals bp. 1783158–1783172). All coordinate references from the published *Pcc* WPP14 genome under NCBI GenBank accession number CP051652.1 with the primary assembly GCF\_013488025 (Liu et al., 2020). Additionally, the deletion and the lack of further insertions was verified using 200 Mbp Illumina whole genome sequencing by SeqCenter, LLC (Pittsburgh, PA, United States).

All chemicals and reagents were purchased and used without further purification.

For all assays, bacteria were grown on Luria-Bertani (LB) (Research Products International) agar (Sigma-Aldrich) plates at 30°C, isolated colonies from these plates were used to inoculate liquid cultures specified in each experiment. Data represents  $n = 3$  biological replicates, unless stated otherwise. Statistical analyses were performed as described at each experiment, with a  $p$ -value  $< 0.05$  being considered significant for all assays.

### 2.2. RNA sequencing and differential expression analysis

To quantify mRNA levels, WT and  $\Delta dgcO$  *Pcc* were grown overnight on Luria-Bertani (LB) (RPI) agar plates. Three isolated colonies of each strain were inoculated into 3 mL LB broth and grown overnight at 30°C with 225 rpm shaking in 15 mL plastic culture tubes (VWR). LB was chosen as *Pcc* grows well in LB at 30°C, and as it has been used for transcriptomic analyses for *Pectobacterium* species in

the past (Gorshkov et al., 2017; Wang et al., 2018; Fan et al., 2020). From these cultures, 50  $\mu$ L was inoculated into 5 mL fresh LB and grown at 30°C with 225 rpm orbital shaking in 15 mL culture tubes either aerobically (grown to  $OD_{600} \sim 0.68$ ) or anaerobically in an anaerobic glovebag (Coy; grown to  $OD_{600} \sim 0.2$ ). RNA extraction for all cultures was performed identically, under aerobic conditions. 1 mL of each culture was pelleted and then RNA extraction was performed using a New England Biolabs Monarch RNA Miniprep Kit (T2010S), according to the procedures outlined by the manufacturer. RNA samples were flash-frozen in liquid  $N_2$  and stored at  $-80^\circ\text{C}$  until submission for analysis by SeqCenter, LLC (Pittsburgh, PA) on dry ice, where RNA sequencing and differential expression analysis was performed. Briefly, Illumina Stranded RNA library preparation with RiboZero Plus rRNA depletion was performed, with 25 M reads. Quality control and adapter trimming was performed with bcl2fastq (version: 2.20.0.445 with default parameters) (Illumina, 2019). Read mapping was performed with HISAT2 (version: 2.2.0 with default parameters + “-very-sensitive”) (Kim et al., 2019). Read quantification was performed using Subread’s featureCounts (version: 2.0.1 with default parameters + “-Q 20”) functionality (Liao et al., 2014). Read counts were loaded into R (version: 4.0.2 with default parameters) and were normalized using edgeR’s (version: 1.14.5 with default parameters) Trimmed Mean of M values (TMM) algorithm (Robinson et al., 2010). Subsequent values were then converted to counts per million (cpm). Differential expression analysis was performed using edgeR’s Quasi Linear *F*-test. Significance of differential expression was established by using  $|\log_2 FC| > 1$  and adjusted *p*-values of  $< 0.05$ , calculated by the Benjamini-Hochberg method, with the false detection rate being  $< 0.05$ . The NCBI GenBank accession number CP051652.1 with the primary assembly GCF\_013488025.1 was used to annotate the genomes and perform the analyses (Liu et al., 2020). The *dgcO* gene is annotated by the LocusTag HER17\_08095. Annotated transcripts were manually classified in Microsoft Excel using online tools, such as the NCBI Gene database, NCBI BLAST, and Uniprot (Altschul et al., 1990; Doğan et al., 2016). Sequencing data was submitted to the NCBI Gene Expression Omnibus (GEO) under the accession: GSE214075.

## 2.3. Quantification of c-di-GMP

WT and  $\Delta dgcO$  *Pcc* were grown overnight at 30°C on LB agar plates. Three isolated colonies of each strain were inoculated into 3 mL LB broth overnight at 30°C with 225 rpm shaking in 15 mL plastic culture tubes (VWR). For quantification of c-di-GMP from aerobically growing *Pcc*, overnight cultures were diluted 1:100 into 15 mL LB in 50 mL plastic culture tubes (Falcon) and grown at 30°C with 225 rpm shaking to an  $OD_{600}$  of  $\sim 0.68$ . For quantification of c-di-GMP produced anaerobically, the overnight cultures were diluted 1:100 into 15 mL LB in 50 mL conicals (Falcon) in an anaerobic glovebag (Coy). Cultures were grown at 30°C with 230 rpm shaking to  $OD_{600} \sim 0.18$ . All further steps were performed aerobically for all cultures. 10 mL of each culture were pelleted at  $3,000 \times g$  for 15 min at 4°C, the supernatant decanted, the pellets flash frozen in liquid  $N_2$  and stored at  $-80^\circ\text{C}$  until use. Extraction of c-di-GMP was done as described previously, with minor modifications (Fontaine et al., 2018). Pellets were resuspended in 1 mL cold extraction solvent/5 mL of original culture, consisting of 2/2/1 v/v/v MeCN/MeOH/ $H_2O$ . Resuspended

cells were then lysed by sonication in a cup horn (Qsonica Q500, part #431C2) at 85% amplitude, 15 s on/15 s off, for a total sonication time of 6 min. Sonicated samples were then centrifuged at  $3,500 \times g$  for 10 min at 4°C to pellet debris. After centrifugation, supernatants were dried by lyophilization and stored at  $-80^\circ\text{C}$  until analysis by mass spectrometry (MS). Prior to analysis, samples were resuspended in 250  $\mu$ L HPLC grade water containing 0.5  $\mu$ M 8-bromoadenosine 3',5'-cyclic monophosphate (8-Br 3',5'-cAMP) (Millipore Sigma) as an internal standard and centrifuged at  $12,000 \times g$  at 4°C for 15 min to further remove debris. 50  $\mu$ L of supernatant was then submitted for analysis to the Penn State Metabolomics Core Facility (University Park, PA, United States) in LC-MS autosampler vials.

Samples were analyzed by Vanquish UHPLC system coupled with a TSQ Quantis Plus mass spectrometer in positive ion mode using a H-ESI<sup>TM</sup> ion source (all Thermo Fisher Scientific) with a Waters (Milford, MA) CORTECS C18+ column (2.1  $\times$  100 mm, 1.6  $\mu$ m particle size). The injection volume was 2  $\mu$ L. The mobile phase was composed of HPLC grade water with 0.1% formic acid (solvent A) and/or acetonitrile with 0.1% formic acid (solvent B). The flow rate of the mobile phase was 0.3 mL/min with column temperature of 40°C. The solvent gradient was as follows: 0% B from 0 to 4 min, 1.5% B from 4 to 15 min, 15% B from 20 to 25 min, increasing to 85% B over 25 to 28 min, then 100% B from 28 to 35 min. Mobile phase was held at 100% B until 39 min, then switched to 0% B (100% A) until 50 min. During the analysis, an ion spray voltage of 4,000 V, sheath gas of 40 psi, aux. Gas of 12 psi, sweep gas of 1, ion transfer tube temperature of 325°C and vaporizer temperature of 300°C were applied. C-di-GMP was detected based on a precursor *m/z* of 691.135 with fragments at 135.21 (collision energy [CE] 90 V), 152.118 (CE 60 V), and 248.168 (CE 40 V). 8-Br 3',5'-cAMP was detected based on a precursor of 408.03 *m/z* and fragments of 134.92 (CE 53 V), and 213.83 (CE 25 V).

Peak integration was performed using Freestyle version 1.8.63.0. For c-di-GMP, retention time and fragmentation were compared to an authentic standard and the fragment at 152.118 *m/z* was used for quantification. For 8-Br 3',5'-cAMP, 213.83 was used for quantification. C-di-GMP peak area was divided by 8-Br 3',5'-cAMP peak area to normalize for ionization efficiency between samples. This peak area ratio was then normalized to cell pellet mass.

## 2.4. Aerobic to anaerobic growth transition

To assess transition from aerobic to anaerobic growth, *Pcc* WT and  $\Delta dgcO$  were grown overnight on LB agar plates at 30°C. Four isolated colonies of each strain were inoculated into 3 mL LB broth in 15 mL plastic culture tubes (VWR), and grown overnight at 30°C with 225 rpm shaking. 980  $\mu$ L LB or M9 + LLL media were added to each well of a 24 well plate, and the overnight cultures were diluted in fresh LB and were added to the wells to a normalized 0.005  $OD_{600}$  starting concentration ( $\sim 20 \mu$ L for each well). Plates were then covered by BreathEasy<sup>TM</sup> (Diversified Bioscience) film and placed into a Biotek Epoch 2 plate reader for 315 min at 30°C with 228 cpm continuous orbital shaking.  $OD_{600}$  was measured every 15 min. After 315 min, the plate was removed from the plate reader, and holes were poked above each well on the film using a sterile needle. The lid was placed back onto the plate, and the plate was then cycled into an anaerobic chamber (Coy). After cycling into the chamber, the lid was removed and the plate was covered with a layer of film to prevent contamination

and evaporation. The plate was then placed into a Biotek Epoch 2 plate reader and grown anaerobically for 24 h under identical parameters to the aerobic portion of this assay.

To assess transition from anaerobic to aerobic growth, overnight cultures of WT and  $\Delta dgcO$  *Pcc* were inoculated in an identical way as for measuring aerobic to anaerobic transition, except that the overnight cultures were set up in an anaerobic chamber (Coy), in de-gassed LB media. Layout and preparation of 24 well plate was all performed in the glovebag, with steps identical to aerobic to anaerobic transition. The plate was placed into a Biotek Epoch 2 plate reader for 360 min at 30°C with 228 cpm continuous orbital shaking. OD<sub>600</sub> was measured every 15 min. The plate was then cycled out of the chamber and placed into a Biotek Epoch 2 plate reader and grown aerobically for 24 h under identical parameters to the anaerobic portion of the assay.

After the growths were completed, means and standard deviations for each strain were calculated, and the OD<sub>600</sub> values were plotted as a function of time in Igor Pro 6.10.

## 2.5. Identification of proteins interacting with *PccDgcO*

Pull-down and identification of proteins interacting with *PccDgcO* was performed as previously described (Burns et al., 2017). Briefly, *PccDgcO* was expressed and purified as previously described (Burns et al., 2014). WT *Pcc* WPP14 5 mL overnight LB cultures were grown at 30°C in a shaking incubator and then spun down using an Allegra X-30R centrifuge. Novagen BugBuster® Protein Extraction Reagent protocol was then followed to lyse the cell pellets. Total cell lysate was then spun down at 12,000 × g and supernatant (~200–500 µL) mixed with an equal amount of purified Fe<sup>II</sup>-O<sub>2</sub> *Pcc* DgcO, and mixed using an orbital mixer for 2 h at 4°C. Nickel-NTA resin was prepared using a 20 mM imidazole buffer (50 mM Tris, 20 mM Imidazole, 300 mM NaCl, 5% glycerol, pH = 7.0). After mixing for 2 h, the mixture was loaded onto 1 mL of prepared resin and flow through was collected. Then 1 mL of 20 mM imidazole buffer was added, allowed to drip through the column, and collected as the wash. In 1 mL increments, the buffers containing increasing amounts of imidazole (50, 100, and 200 mM imidazole buffers) were loaded onto the resin and flow-through collected. All fractions were then flash frozen in liquid N<sub>2</sub> and stored at –80°C until used for protein digestion. The pull-downs were repeated twice, with one biological replicate in the first run and three biological replicates in the second run.

Tryptic and chymotryptic digestions were performed on each fraction using the In-solution Tryptic Digestion and Guanidination Kit (Thermo Scientific) and Mass Spectrometry-Grade Chymotrypsin Endoproteinase (Thermo Scientific) according to their respective kit's instructions. In brief, 10.5 µL of each fraction was mixed with the respective digestion buffer. Samples were heated at 95°C, cooled, alkylated using iodoacetamide (IAA), and then 1 µL of the respective protease was added to the mixture and allowed to react overnight at room temperature before 3 µL of trifluoroacetic acid (TFA) was added to stop the reaction. Digestion samples were loaded into mass spectrometer vials.

The mass spectrometry experiments were performed on an LTQ-FTMS Ultra from Thermo Fisher equipped with a Thermo Ultimate 3,000 HPLC dual pump controlled through Chromeleon

software and a Shimadzu SIL-ACHT autosampler controlled by a Shimadzu CBM-20A controller. The control of the mass spectrometer and the synchronization of the system was accomplished in the Xcalibur software supplied by Thermo Fisher.

One of the dual pumps was used to load the sample injected from the autosampler onto a Micro Bioresource Captrap and washed using 2% formic acid in LC/MS water at a flow rate of 50 µL/min for 20 min. The Captrap was switched in the flow from the second pump into the picochip [New Objective (1PCH7515-105H354-NV)] using the divert valve loaded on to the capillary. The flow from the pump to the chip was set at 0.150 mL/min but was split using a homemade splitter just before the diverter valve; based on pressure obtained and specifications of the microchip, it is estimated that the flow rate to the chip was approximately 1 µL/min. The chromatography profile was as follows: 98% A (0.1% formic acid in H<sub>2</sub>O) and 2% B (0.01% formic acid in acetonitrile) for the 20 min; followed by a gradient to 60% A/40% B over 90 min; followed by a gradient to 2% A/98% B over 40 min; followed by a gradient to 98% A/2% B over 40 min.

The LTQ FTMS Ultra acquired a full scan FTMS from 200 to 2,000 m/z at 50,000 resolution followed by 5 data dependent (5 biggest peaks) ion trap MS/MS at 35 V normalize collision energy with an isolation window of 2 m/z. The data dependent used dynamic exclusion to allow for obtaining MS/MS on more ions. The following parameters were used for the source: ESI voltage, 2 KV; capillary temperature, 200°C; capillary voltage, 40 V; and tube lens voltage, 150 V. Data analysis was performed using Thermo Proteome Discoverer 1.3. As identified proteins can be false hits or due to non-specific interactions, the analyses from all the biological replicates were compared to identify proteins that were found in more than one biological replicate. The list of proteins identified multiple times was then sorted based on the Proteome Discoverer Peptide Sorting Match (PSM) to rank hits.

## 2.6. Quantification of cellular metals

Quantification of cellular metals was performed as described previously (Duggal et al., 2020). Briefly, from overnight LB-agar plates grown at 30°C, *n* = 5 isolated colonies of WT and  $\Delta dgcO$  *Pcc* were each inoculated into 3 mL of M9 minimal media [0.4% glucose (Research Products International) with 0.2% casamino acids (Becton, Dickinson and Co.), made using metal-free glassware (washed with 6 M HCl) and supplemented with various metals (Sigma-Aldrich) in 15 mL plastic conical tubes (VWR), then incubated overnight at 30°C with 225 rpm orbital shaking]. The list of metals and their concentrations can be found in Table 1. 15 mL of the same medium was inoculated with 100 µL of the overnight culture in 50 mL plastic conical tubes and grown to an OD<sub>600</sub> ~ 0.8 at 30°C with 225 rpm orbital shaking. 10 mL of each culture was then transferred into pre-weighed 15 mL ICP-MS Teflon sample tubes and pelleted at 4000 × g for 10 min at 4°C. Supernatants were removed and the pellets weighed. Tubes were then flash frozen in liquid N<sub>2</sub> and submitted for analysis by the Penn State Laboratory for Isotopes and Metals in the Environment (LIME) (University Park, PA, United States). Sample digestion and analysis was performed by LIME; briefly, bacterial pellets were digested using nitric acid, then analyzed using a Thermo Fisher Scientific Icap RQ ICP-MS instrument. Statistical significance was established using a 2-tailed Student's *t*-test in Igor Pro 6.10 (Wavemetrics, 2009), metal content (in parts per billion) was normalized to cell pellet mass.



TABLE 1 Metals added to M9 media for quantification by ICP-MS.

Metal	Concentration in media
CaCl <sub>2</sub> •2H <sub>2</sub> O	4 μM
MnCl <sub>2</sub> •4H <sub>2</sub> O	2 μM
ZnSO <sub>4</sub> •7H <sub>2</sub> O	2 μM
CoCl <sub>2</sub> •6H <sub>2</sub> O	0.4 μM
CuCl <sub>2</sub>	0.4 μM
NiCl <sub>2</sub>	0.4 μM
NaMoO <sub>4</sub> •2H <sub>2</sub> O	0.4 μM
Na <sub>2</sub> SeO <sub>3</sub>	0.4 μM
H <sub>3</sub> BO <sub>3</sub>	0.4 μM
FeCl <sub>3</sub>	10 μM

TABLE 2 List of *Vibrio* reporter and control strains used in the quantification of extracellular signaling molecules.

Strain	Genotype	Role	Source
<i>V. fischeri</i> JHK007	ES114 $\Delta$ ainS $\Delta$ luxR-luxI, P <sub>luxI</sub> -luxCDABEG	Nonluminescent, negative control for AHL production	(Colton et al., 2015)
<i>V. fischeri</i> DJ01	ES114 $\Delta$ ainS $\Delta$ luxR-luxI, luxR <sup>ES114</sup> , P <sub>luxI</sub> -luxCDABEG	3-oxo-C6-HSL reporter	(Colton et al., 2015)
<i>V. fischeri</i> DC22	ES114 $\Delta$ ainS $\Delta$ luxR-luxI, mutant luxR <sup>B</sup> (MJ1-T33A, R67M, S116A, M135I), P <sub>luxI</sub> -luxCDABEG	3-oxo-C8-HSL reporter	(Colton et al., 2015)
<i>V. campbellii</i> BB120 Wild-type		Positive control for AI-2, AHL production	(Colton et al., 2015)
<i>V. campbellii</i> MM32	luxN::Cm, luxS::Tn5Kan	AI-2 reporter	(Miller et al., 2004)
<i>V. campbellii</i> BB151	luxA::Tn5lac	Nonluminescent, negative control for AI-2 production	(Surette et al., 1999)

## 2.7. Bacterial growth with additives

To assess growth in the presence of Cu, Mn and Zn, 1 M stocks of CuCl<sub>2</sub>, MnCl<sub>2</sub>•4H<sub>2</sub>O and ZnSO<sub>4</sub>•7H<sub>2</sub>O were made in H<sub>2</sub>O. These were then serially diluted in H<sub>2</sub>O to 0.001 M and added to M9 media (supplemented with 0.4% glucose and 0.2% casamino acids) to concentrations identical to those in ICP-MS experiment (Table 1). Additionally, the three 1 M stock solutions were diluted into LB to final working concentrations of 0 mM (LB with no additional metal), 3 mM, 9 mM, and 12 mM.

To assess carbon utilization, a 50% (v/v) stock solution of glycerol (Sigma-Aldrich) and a 25% (w/w) stock solution of arabinose (Acros Organics) were made in H<sub>2</sub>O. Appropriate volumes of these stock solutions were added to M9 media containing no carbon sources for final concentrations of 2 and 4% for glycerol, and 1 and 3% for arabinose.

For growth assays in the presence of Cu, Mn, Zn, and the sole carbon source utilization assays, *Pcc* WT and  $\Delta$ dgcO were grown

overnight on LB agar plates at 30°C. Four isolated colonies of each strain were inoculated into 3 mL LB broth in 15 mL plastic culture tubes (VWR), and grown overnight at 30°C with 225 rpm shaking. 196 μL of the appropriate media was added to wells of a 96 well plate, and the overnight cultures were diluted in fresh LB and were added to the wells to a normalized 0.005 OD<sub>600</sub> starting concentration (~4 μL for each well).

For growth in the presence of FeCl<sub>3</sub>, *Pcc* WT and  $\Delta$ dgcO were grown overnight on LB agar plates at 30°C. Three isolated colonies of each strain were inoculated into 3 mL LB broth in 15 mL plastic culture tubes (VWR), and grown overnight at 30°C with 225 rpm shaking. 980 μL of M9 media (supplemented with 0.4% glucose) with the appropriate concentration of FeCl<sub>3</sub> (1 M stock solution made in H<sub>2</sub>O) was added to wells of a 24 well plate, and the overnight cultures were diluted in fresh LB to a normalized OD<sub>600</sub> = 1.0, 20 μL of these normalized cultures was then added to the wells of the 24 well plate.

For all experiments, plates were covered by BreathEasy™ (Diversified Biocience) film and placed into a Biotek Epoch 2 plate reader for 24 h at 30°C with 228 cpm continuous orbital shaking, with OD<sub>600</sub> being measured every 15 min. After the growth assay was completed, means and standard deviations for each strain were calculated, and the OD<sub>600</sub> values were plotted as a function of time in Igor Pro 6.10.

## 2.8. Quantification of extracellular signaling molecules

Production of extracellular signaling molecules was measured using the methods described previously, with modifications (Surette et al., 1999). The AHLs 3-oxo-C6-HSL and 3-oxo-C8-HSL, were quantified, along with autoinducer-2 (AI-2). A list of control and reporter strains, alongside their role in these experiments, can be found in Table 2.

Briefly, cell free supernatants of WT and  $\Delta$ dgcO *Pcc* were obtained by pelleting 1 mL cells grown in LB or M9 + Lettuce Leaf Lysate (LLL) media (3 mL inoculated with a 1:100 dilution of an overnight culture, grown at 30°C, 225 rpm orbital shaking) at the appropriate OD<sub>600</sub> (OD<sub>600</sub> = ~0.5, ~0.75, ~1.0, or ~1.2), followed by 0.22 μm filter sterilization of supernatant.

LLL was prepared as described previously (Burns et al., 2017). Briefly, 610.12 g of organic romaine lettuce hearts were washed then macerated using a manual juicer (Lexen). Lysate was then spun at 35,000 rpm (11 pprox. 100,000 × g) for 45 min in a Beckman-Coulter Optima XE-90 ultracentrifuge. The supernatants were filter sterilized using 0.22 μm filters before use. LLL was stored at 4°C for short-term, or at -80°C for long-term storage. M9 + LLL media was made by adding 12 μL LLL to 3 mL M9 media (0.4% glucose (Research Products International) with 0.2% casamino acids (Becton, Dickinson and Co.)) immediately before inoculation with bacteria (Burns et al., 2017).

Cell-free supernatants of the *V. campbellii* BB120 and BB151 control strains were prepared in a similar way to those of *Pcc*, but by streaking on LB-Salt (LBS) plates and growing bacteria in Autoinducer Bioassay (AB) (per liter: 17.5 g NaCl, 12.3 g MgSO<sub>4</sub>, 2.0 g Casamino acids, 10 mL 1 M pH 7.0 Potassium phosphate buffer, 10 mL 0.1 M L-arginine, 10 mL glycerol, all in H<sub>2</sub>O) medium to a 0.65 final OD<sub>600</sub> (Surette et al., 1999).

All *Vibrio* reporter strains were grown overnight on LBS plates at 30°C, then a single colony was inoculated into 3 mL AB media at 30°C



with 225 rpm shaking in 15 mL plastic culture tubes. These cultures were serially diluted to 1:5,000 with fresh AB, and used immediately after preparation.

To measure extracellular signaling molecule production, 40  $\mu$ L 1:5,000 of the appropriate reporter culture was added to 160  $\mu$ L cell-free supernatant of *Pcc* experimental or *V. campbellii* control strains in 96 well white/clear bottom plates (Thermo Scientific, 165306). For each biological replicate of *Pcc*, 2 technical replicates were used. Plates were covered using BreatheEasy film (Diversified Bioscience). Luminescence and OD<sub>600</sub> were measured over 12 h using a Biotek Cytation5 plate reader 30°C with 180 cpm orbital shaking.

Signaling molecule production was quantified by comparing specific luminescence (LUM/OD<sub>600</sub>) values of wells containing the appropriate reporter strains grown in the presence of cell-free spent media from WT or  $\Delta$ *dgcO* *Pcc* strains.

## 3. Results

### 3.1. WT vs. $\Delta$ *dgcO* *Pcc* exhibit differentially regulated transcripts when grown aerobically

RNA sequencing and differential expression analysis revealed the differential expression of 372 transcripts between WT and  $\Delta$ *dgcO* *Pcc* strains grown aerobically, but no significantly differentially expressed genes between WT and  $\Delta$ *dgcO* grown anaerobically. The 372 aerobically differentially expressed transcripts correspond to approximately 9% of the genome (Figure 1). It is important to note that a large percentage of transcripts (330 transcripts, 88.7% of the total) have a subtle,  $1 < |\log_2\text{FC}| < 2$ , change in their expression levels, with 237 transcripts (63.7% of total) having a change in expression of  $1 < |\log_2\text{FC}| < 1.5$ . The majority (258, 69.4%) of differentially expressed transcripts had lower while 114 (30.6%) had higher expression levels in the  $\Delta$ *dgcO* strain. A breakdown of the fold changes for each condition can be found in Table 3. The functions with the largest number of transcripts showing a decrease in abundance include the type 6 secretion system (T6SS) (15 transcripts), flagellar elements (8), virulence factors (5), and transcripts related to iron homeostasis and transport (4). Various transporters (20), methyl-accepting chemotaxis proteins (6), and transcripts related to carbohydrate metabolism (17) have increased transcript abundance in  $\Delta$ *dgcO* compared to WT. Figure 2 shows a comparison between the numbers and identities of genes expressed between WT and  $\Delta$ *dgcO*. A full list of quantified transcripts can be found in Supplementary Data Sheet 2, while a table classifying transcripts differentially expressed between WT and  $\Delta$ *dgcO* *Pcc* grown aerobically can be found in Supplementary Table S1 (hypothetical proteins are excluded). All further Log<sub>2</sub>FC values refer to the differential expression of genes between WT and  $\Delta$ *dgcO* *Pcc* grown aerobically, except where otherwise noted in Section 3.3.

### 3.2. Quantification of c-di-GMP

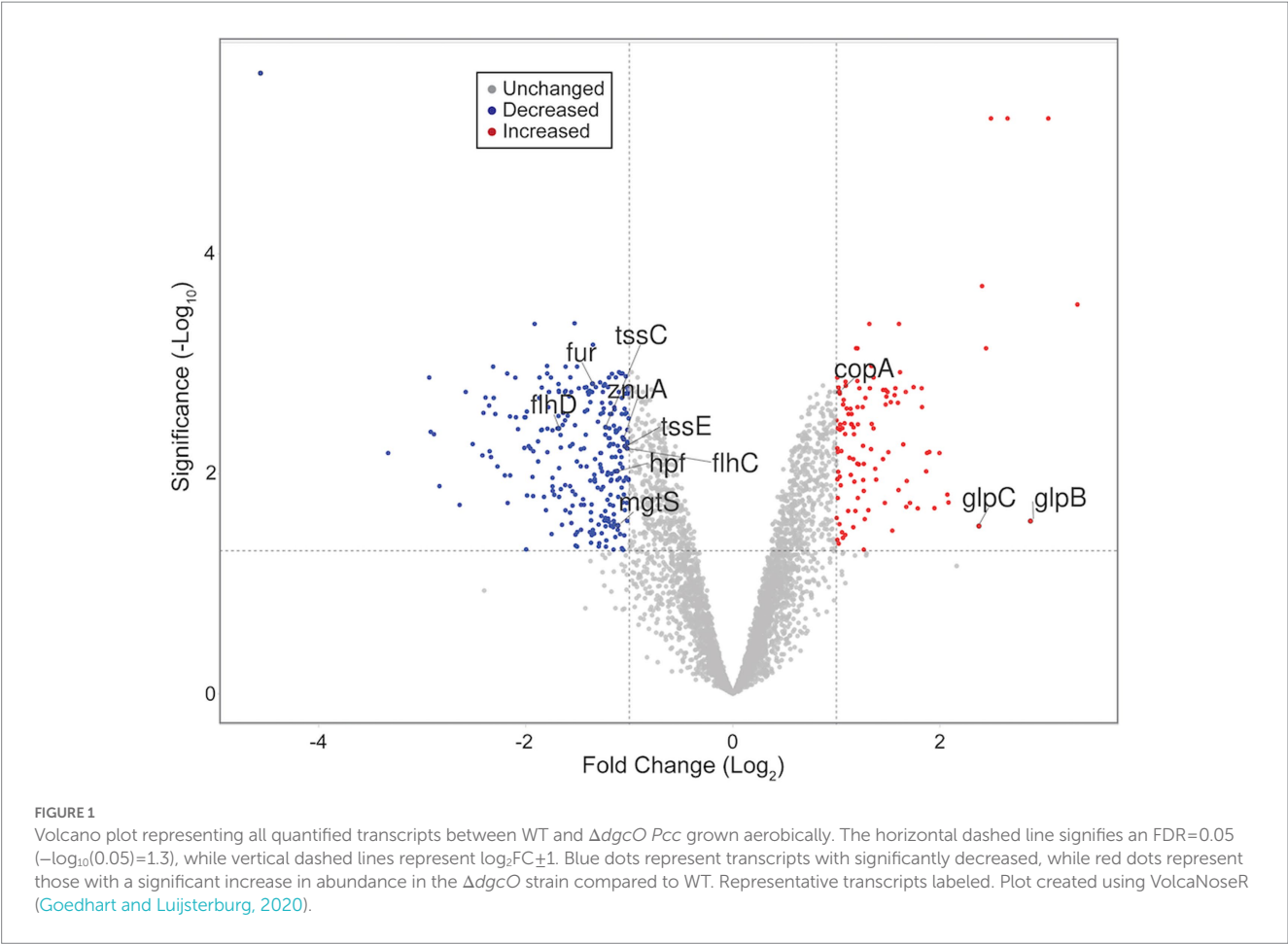
The subtle transcript level changes between WT and  $\Delta$ *dgcO* *Pcc* suggest that most regulation by DgcO happens through localized c-di-GMP signaling and not through major changes in the cellular c-di-GMP pool. To test this hypothesis, c-di-GMP was quantified

from WT and  $\Delta$ *dgcO* *Pcc* grown under aerobic and anaerobic conditions. Aerobically, the levels of c-di-GMP were found to be low for both strains, with c-di-GMP concentration in the  $\Delta$ *dgcO* strain being under the limit of quantification. Overall, aerobic c-di-GMP concentrations show no significant difference between WT and  $\Delta$ *dgcO* *Pcc* (Figure 3). When comparing c-di-GMP levels between WT and  $\Delta$ *dgcO* *Pcc* grown anaerobically, the WT strain was found to have a higher c-di-GMP concentration compared to the  $\Delta$ *dgcO* strain (Figure 3). When comparing c-di-GMP levels of the same strain between the two conditions, anaerobically grown WT and  $\Delta$ *dgcO* *Pcc* both had a significant increase in their c-di-GMP levels compared to cells of the same strain grown aerobically (Figure 3).

### 3.3. $\Delta$ *dgcO* Strain exhibits altered effects on transcripts in response to anaerobic conditions

The differences between *Pcc* WT transcript levels for cells grown aerobically vs. anaerobically were found to vary when compared to those identified in a comparison of *Pcc*  $\Delta$ *dgcO* grown aerobically vs. anaerobically. Differences were found in the number of transcripts with altered levels (789 vs. 639, respectively), gene identities, and Log<sub>2</sub>FC's when comparing the WT and  $\Delta$ *dgcO* strains. For example, there are no differences in transcript levels for genes encoding ribosomal proteins or elongation factors in *Pcc* WT, but *Pcc*  $\Delta$ *dgcO* exhibits 1–1.5 Log<sub>2</sub> fold changes in 18 transcripts encoding ribosomal proteins and elongation factors. Of the 639 altered transcripts identified in *Pcc*  $\Delta$ *dgcO* grown under anaerobic vs. aerobic conditions, 89 are also altered when comparing aerobically grown *Pcc*  $\Delta$ *dgcO* vs. WT strains. In all cases, transcript levels that have higher expression levels in aerobic *Pcc*  $\Delta$ *dgcO* vs. WT also have higher expression levels in *Pcc*  $\Delta$ *dgcO* aerobic vs. anaerobic datasets. The remaining 550 transcripts that were identified in the *Pcc*  $\Delta$ *dgcO* aerobic vs. anaerobic comparison suggest that *Pcc* does not transition properly between aerobic and anaerobic environments in the absence of DgcO. All of the 50 most differentially expressed transcripts identified when comparing *Pcc*  $\Delta$ *dgcO* grown aerobically vs. anaerobically were also identified in the comparison of *Pcc* WT aerobic vs. anaerobic RNAseq. Of this set of transcripts, 75 were found to share either increase or decrease in transcript levels between *Pcc* WT and  $\Delta$ *dgcO*, although the magnitude of the Log<sub>2</sub>FC between aerobically and anaerobically grown cells varied (Supplementary Table S2). The remaining 25 transcripts exhibited opposing changes in transcript levels, with the 16 transcripts with the largest decrease in abundance in *Pcc*  $\Delta$ *dgcO* all showing increased expression in *Pcc* WT. While six of the genes encode for putative or hypothetical proteins, genes encoding parts of the type VI secretion system (HER17\_04885, *tssM*, *tssH*, *tssA*) and a sigma-54-dependent Fis family transcriptional regulator (HER17\_04895), which may be involved in regulation of virulence, have decreased levels in *Pcc*  $\Delta$ *dgcO* but have increased abundance in *Pcc* WT. These differences may be involved in the previously observed decrease in virulence factor excretion and virulence in a potato host by *Pcc*  $\Delta$ *dgcO*, as compared to *Pcc* WT (Burns et al., 2017).

Based on these discrepancies, the transition of WT and  $\Delta$ *dgcO* *Pcc* from aerobic to anaerobic growth was studied. Results showed that while there is no difference between adaptation to anaerobic or aerobic conditions in LB (Supplementary Figures S1A,B), when grown in M9 + LLL media, the  $\Delta$ *dgcO* strain exhibits a longer lag when moved



**TABLE 3** Number and percentage of differentially expressed transcripts between WT and  $\Delta dgcO$  *Pcc* grown aerobically broken down based on level of differential expression.

Change	Lower expression in $\Delta dgcO$	Higher expression in $\Delta dgcO$
$1 <  \log_2FC  < 1.5$	156 (60.5%)	81 (71.1%)
$1.5 \leq  \log_2FC  < 2$	71 (27.5%)	22 (19.3%)
$2 \leq  \log_2FC  < 3$	29 (11.2%)	9 (7.9%)
$3 \leq  \log_2FC $	2 (0.8%)	2 (1.8%)

A total of 372 transcripts were differentially expressed, 258 had lower, 114 had higher expression levels in  $\Delta dgcO$  *Pcc* compared to WT.

into the anaerobic chamber compared to the WT strain; after 30 min, the two strains grow identically again (Figure 4A). When moved from anaerobic to aerobic conditions, the WT strain exhibited a higher final OD<sub>600</sub> than the  $\Delta dgcO$  strain (Figure 4B).

### 3.4. DgcO affects transcript levels of flagellar genes and interacts with chemotaxis proteins

Transcriptomic results revealed the decrease in transcript levels of the *fliJ*, *fliL*, *fliM*, *fliP* flagellar structural and motor proteins (avg.  $\sim -1.16$

$\log_2FC$ ), and the *fliZ* ( $\sim -1.90 \log_2FC$ ) and *flhC* ( $\sim -1.02 \log_2FC$ ), *flhD* ( $\sim -1.68 \log_2FC$ ) transcriptional regulators in *Pcc*  $\Delta dgcO$ , as compared to WT (Figure 5A). The transcription factors *flhC* and *flhD* form the *flhDC* complex *in vivo*; this complex is known to positively regulate motility in *P. carotovorum* (Chatterjee et al., 2009). Based on these findings, and previously published interaction data suggesting that DgcO interacts with the chemotaxis proteins TarH, CheA, and CheW, Ni-NTA was used to pulldown His-tagged DgcO and interacting proteins (Burns et al., 2017). The pull-down was performed with cleared lysate, so it is possible that DgcO makes additional interactions with membrane proteins that were not identified in these experiments. Results indicate that DgcO interacts with the chemotaxis protein CheY, which upon phosphorylation serves as a switch signal to the flagellar motor, suggesting that DgcO localizes with the motility machinery complex. High confidence hits (as assessed by #PSM) for peptides eluting in the same fraction as DgcO can be found in Table 4; it is important to note that many peptides are found as low-quality hits, resulting in a large number of potential interacting partners. Full list of peptide hits from the pulldown can be found in Supplementary Data Sheet 3.

### 3.5. DgcO alters metal homeostasis

From the RNAseq data, the ferric uptake regulator transcription factor, *fur* ( $\sim -1.36 \log_2FC$ ), and various metal transporters were found to be differentially expressed in aerobically

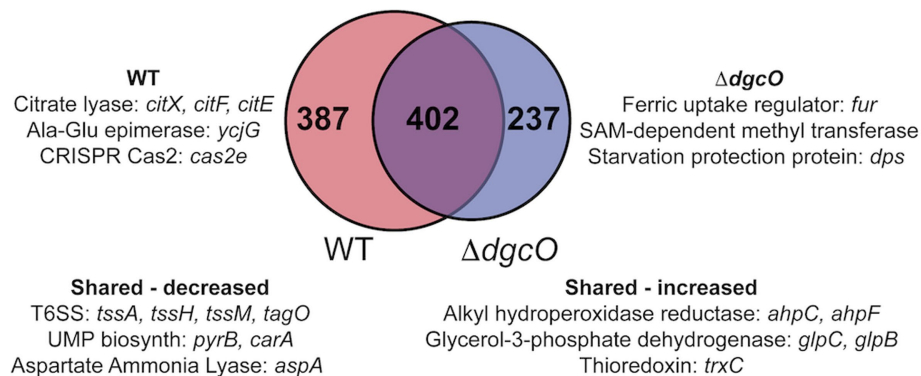


FIGURE 2

Venn diagram depicting overlap between transcripts that are differentially expressed in aerobic vs. anaerobic WT and aerobic vs. anaerobic  $\Delta dgcO$ . Shared and unique genes with the largest differential expression are listed.

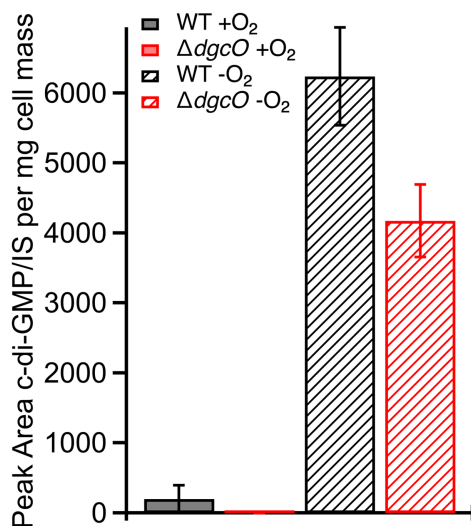


FIGURE 3

Levels of c-di-GMP quantified from WT and  $\Delta dgcO$  *Pcc* grown aerobically or anaerobically. No significant difference between c-di-GMP levels was found between the two strains grown aerobically, but when grown anaerobically, both strains show a significant increase compared to aerobic growths. Additionally, the  $\Delta dgcO$  strain has a significantly lower concentration compared to WT when grown anaerobically.

grown  $\Delta dgcO$ , as compared to WT (Troxell and Hassan, 2013). Genes annotated as responsible for transporting iron (avg.  $\sim 1.25$  Log<sub>2</sub>FC) and the *copA* copper (I) transporter ( $\sim 1.03$  Log<sub>2</sub>FC) have increased transcript levels in *Pcc*  $\Delta dgcO$ , while the zinc ABC transporter substrate-binding protein *znuA* ( $\sim -1.05$  Log<sub>2</sub>FC) and the magnesium starvation response regulator *mgtS* ( $\sim -1.11$  Log<sub>2</sub>FC) had decreased transcript levels in *Pcc*  $\Delta dgcO$  (Figure 5B). To determine if the differential expression of genes for metal transporters, metal homeostasis, and transition metal-containing cofactor synthesis affected cellular metal levels, ICP-MS was used to quantify cellular metal content in WT and  $\Delta dgcO$  *Pcc* grown in minimal media (to allow for complete control of metal concentrations). Out of the 10 metals tested, six showed

statistically significant differences in levels, with WT containing higher levels of each metal (Figure 6). The six metals with differences in their concentrations, Co, Cu, Mo, Mg, Mn, and Zn all play an important role in the cell, often serving as cofactors for enzymatic activity. Interestingly, iron levels were not found to be different between the two strains, suggesting possible compensatory effects to the decreased *fur* transcript levels. No difference was found when cells were grown in M9 in the presence of various concentrations of FeCl<sub>3</sub> (Supplementary Figure S1C). Similarly, when cells were grown in M9 media supplemented with Cu, Mn, and Zn at concentrations seen in ICP-MS, no difference was found between WT and  $\Delta dgcO$  *Pcc* (Supplementary Figure S1D).

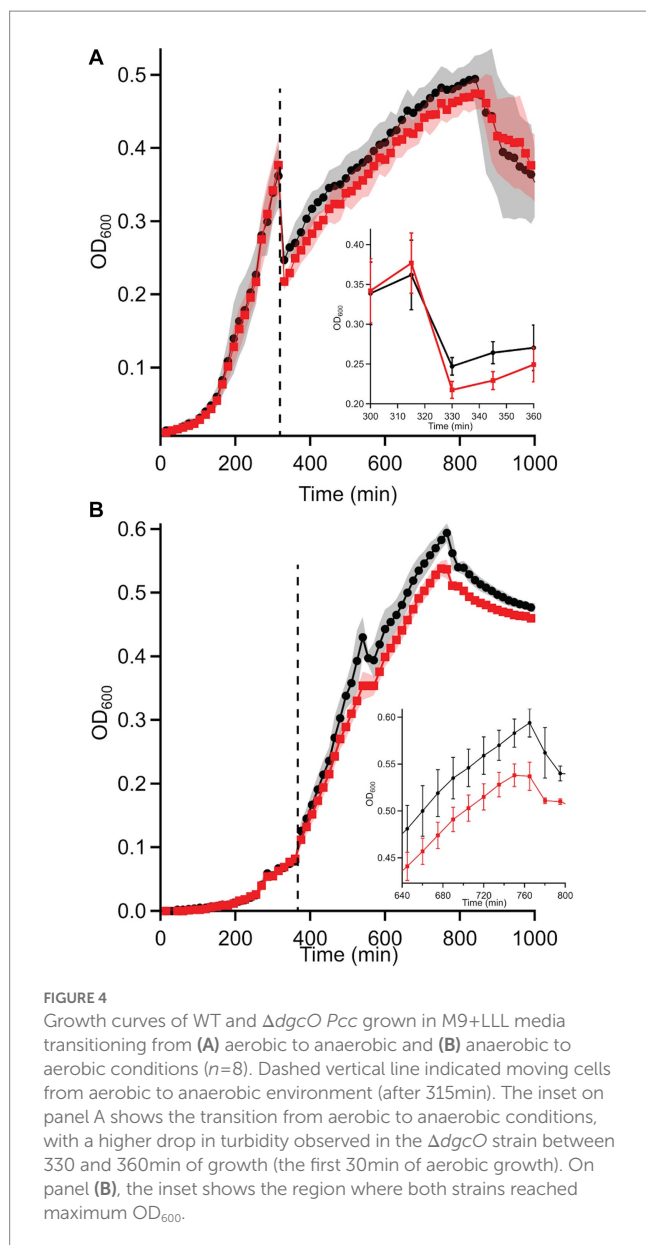
Growth assays performed in LB supplemented with Cu, Mn, and Zn also showed no difference in copper tolerance between the strains (Figure 7B), while neither strain could tolerate the high levels of zinc. With manganese, no difference was seen at 3 mM and 12 mM concentrations, but at 9 mM, the  $\Delta dgcO$  strain showed an increased time (approx. 45 min) to mid-log, but reached the same stationary phase OD<sub>600</sub> as WT (Figure 7A).

### 3.6. PCWDE expression is reduced in $\Delta dgcO$

It was previously demonstrated that the deletion of *dgcO* reduces PCWDE production by  $\sim 15\%$  under aerobic conditions, but the regulation of this activity was unknown (Burns et al., 2017). Transcriptomic data shows that pectate lyase (avg.  $\sim -1.29$  Log<sub>2</sub>FC), polysaccharide lyase ( $\sim -1.04$  Log<sub>2</sub>FC), and cellulase ( $\sim -1.19$  Log<sub>2</sub>FC) transcript levels show a decrease in  $\Delta dgcO$  *Pcc*, providing evidence that *PccDgcO* controls PCWDE expression.

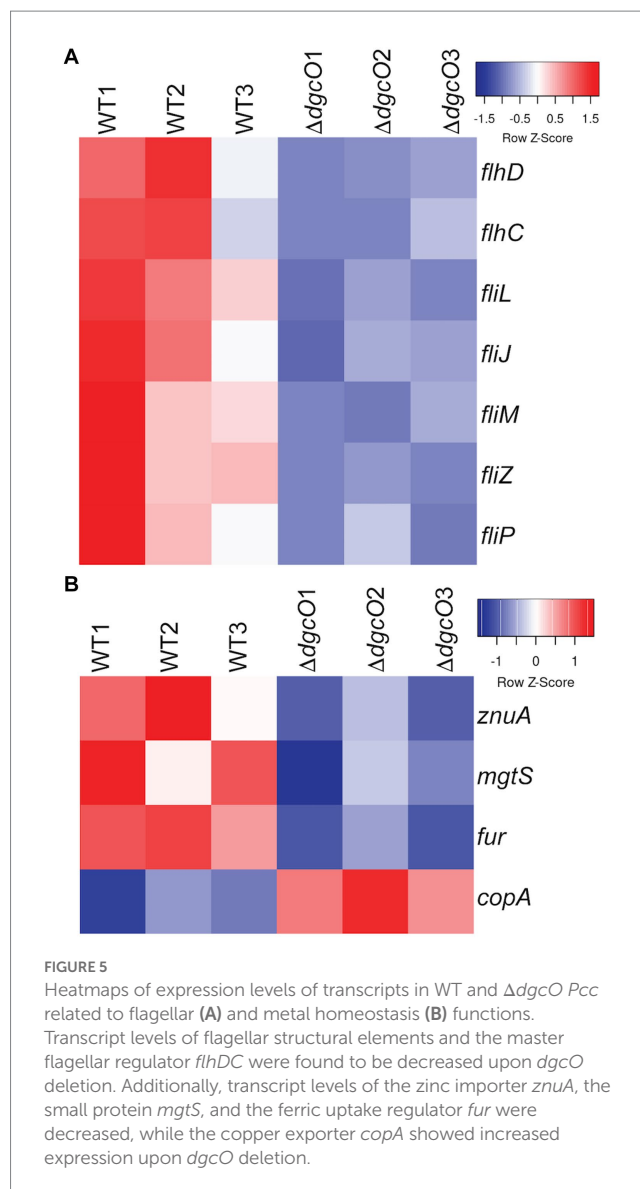
### 3.7. DgcO plays a role in regulating T3SS and T6SS expression

The expression of the type 3 secretion system (T3SS) and T6SS was shown to be regulated by c-di-GMP in *Pseudomonas aeruginosa*, with T3SS being expressed at low, and T6SS being expressed at high c-di-GMP levels, respectively (Dadashi et al., 2021). The exact



mechanism by which the expression of these systems is regulated by c-di-GMP is not fully understood and previously has not been studied in *Pcc*. Upon *dgcO* deletion (and thus the removal of c-di-GMP synthesized by DgcO) differential expression of genes involved in T3SS and T6SS formation and activity in *Pcc* follows the pattern described above. A key T3SS ATPase, *sctN*, shows increased expression in  $\Delta dgcO$  ( $\sim 1.04 \text{ Log}_2\text{FC}$ ). In *Chlamydia trachomatis*, when active, SctN, plays a role in unfolding the chaperone-effector complex, a step necessary for effector secretion and host colonization (Grishin et al., 2018). Besides *sctN*, the decrease in transcript levels of the *flhDC* regulatory complex also suggests transcriptional control of T3SS, as *flhDC* is a known regulator of T3SS expression in *P. aeruginosa* (Soscia et al., 2007). Another possible T3SS regulator is the sigma factor RpoE, which plays a role in response to cellular stress, and has been found to have a decrease of  $\sim -1.07 \text{ Log}_2\text{FC}$  in its transcript levels (Flores-Kim and Darwin, 2012).

Deletion of *Pcc* DgcO results in the decrease in expression of *tssB*, *tssC*, *tssE*, *tssG*, and *tssJ* (avg.  $\sim -1.19 \text{ Log}_2\text{FC}$ ); these genes all encode



for T6SS structural elements. Additionally, a total of 9 transcripts annotated as Hcp family effectors also showed an average decrease in transcript levels of  $\sim -1.88 \text{ Log}_2\text{FC}$ . These results suggest c-di-GMP-dependent control of both the assembly of T6SS and the synthesis of effectors used in interspecies competition.

### 3.8. Regulation of cell-cell communication by DgcO

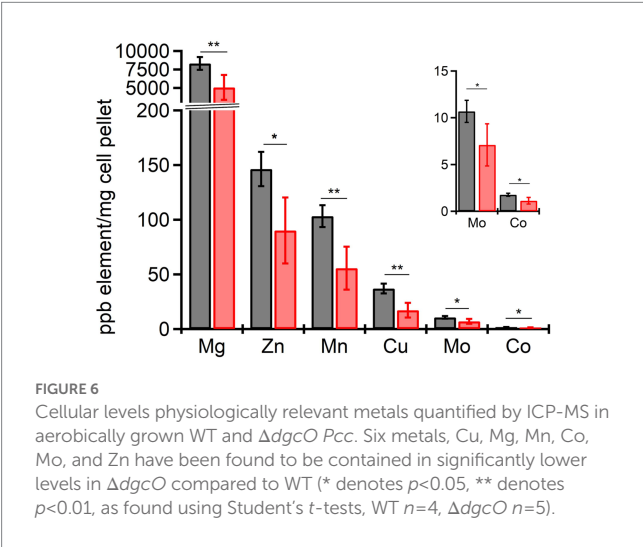
Besides T6SS, systems involved in interactions with other microbes appear to be regulated by DgcO. The quorum sensing master regulator, *expR*, shows a transcript abundance decrease  $\sim -1.38 \text{ Log}_2\text{FC}$ , along with additional LuxR family transcriptional regulators (avg.  $\sim -1.48 \text{ Log}_2\text{FC}$ ). These results, taken together with previously quantified AHL levels, suggest that the deletion of *dgcO* causes no defect in AHL synthesis, but it is possible that AHL gets exported from the cell through pumps, although no differential expression of known AHL-binding pumps was identified in the transcriptomic data (Burns et al., 2017).



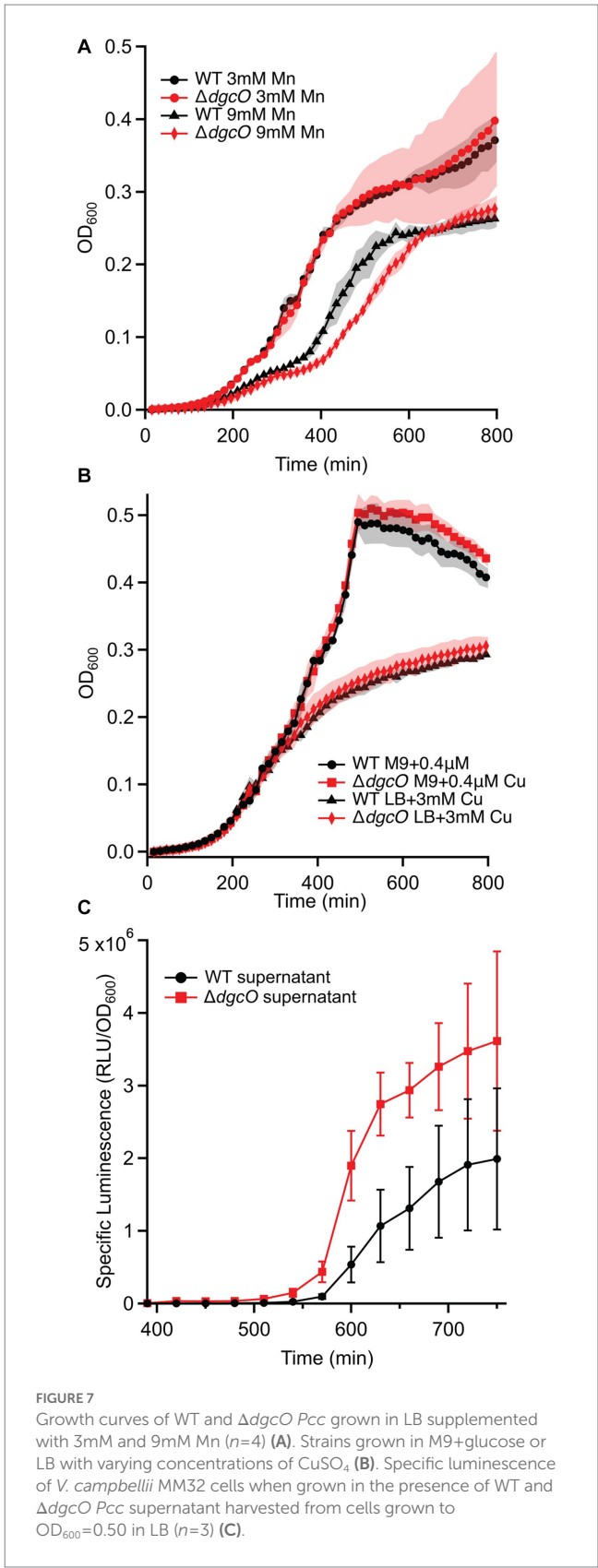
TABLE 4 Proteins with the highest quality hits (as assessed by #PSM) interacting with DgcO as identified in Ni-NTA pulldown.

Protein	Function
TsaA	Peroxidase
LuxS	S-ribosylhomocysteine lyase
Mdh	Malate dehydrogenase (NAD)
Tsf	Translation elongation factor Ts (EF-Ts)
OsmY	Periplasmic protein
SdhA	Succinate dehydrogenase subunit A
FusA	Translation elongation factor 2 (EF-2/EF-G)
NuoG	NADH dehydrogenase subunit G
KatG	Catalase-peroxidase
CheY	Chemotaxis regulatory protein CheY
ArnA	Bifunctional UDP-glucuronic acid decarboxylase/UDP-4-amino-4-deoxy-L-arabinose formyltransferase
SlyD	FKBP-type peptidyl prolyl cis-trans isomerase/Apo-metallochaperone SlyD
Crp	cAMP-regulatory protein
Fur	Ferric uptake regulator

It is important to note that ArnA, SlyD, Crp, and Fur are known common contaminants in Ni-NTA pulldowns (Bolanos-Garcia and Davies, 2006; Chen et al., 2017). Arabinose utilization was tested but yielded no significant differences between the strains (Supplementary Figure S1E).



In addition to AHL-mediated cell–cell communication, DgcO appears to have an effect on interspecies communication through localized signaling, as Ni-NTA pulldown showed that LuxS, a major component of the AI-2 synthesis pathway, associates with DgcO (Pöllumaa et al., 2012; Joshi et al., 2016). A *Vibrio* reporter assay using the *V. campbellii* MM32 AI-2 reporter strain showed that when the reporters were grown in the presence of supernatant harvested from WT and  $\Delta dgcO$  *Pcc* cells grown to  $OD_{600}=0.50$  in LB, specific luminescence of the reporters was higher when they were grown in the presence of supernatant from  $\Delta dgcO$  *Pcc* (Figure 7C). Interestingly, when sterile-filtered supernatant was used from *Pcc* WT or  $\Delta dgcO$  grown media to



mimic a plant host (M9+LLL), *V. campbellii* and *V. fischeri* reporters failed to grow, suggesting that lettuce extract induced *Pcc* excretion of inhibitory factors in a *dgcO*-independent manner.

### 3.9. DgcO affects translation-related genes

Transcripts encoding for a total of 6 ribosomal proteins (*rpmC*, *rpmI*, *rpmJ*, *rpsJ*, *rpsM*, and *rmsQ*) were found to have lower levels in *Pcc*  $\Delta$ *dgcO* (avg.  $\sim -1.15$  Log<sub>2</sub>FC), while the ribosome hibernation promoting factor *hpf* has an increase of  $\sim 1.12$  Log<sub>2</sub>FC in transcript levels, as compared to WT (Ueta et al., 2008; Polikanov et al., 2012). Additionally, Ni-NTA pulldown identified the translation elongation factors FusA and Tsf to be interacting with *Pcc* DgcO, suggesting that DgcO plays a role in regulating translation.

### 3.10. Regulation of anaerobic processes

As mentioned, *Pcc* DgcO serves as an O<sub>2</sub> sensor and diguanylate cyclase activity is greater aerobically. Based on transcriptomic data however, deletion of *Pcc* DgcO also modulates anaerobic processes, namely anaerobic citrate and glycerol metabolism, even under aerobic growth conditions. Genes encoding for citrate lyase and various associated proteins (*citCDEF* and *citX*) were had an average increase of  $\sim 1.80$  Log<sub>2</sub>FC in their transcript levels, while two subunits of the anaerobic glycerol-3-phosphate dehydrogenase (*glpBC*) had an average increase of  $\sim 2.63$  Log<sub>2</sub>FC in  $\Delta$ *dgcO* *Pcc*, suggesting a possible c-di-GMP repression of these processes aerobically. However, when WT and  $\Delta$ *dgcO* *Pcc* cells were grown aerobically with 2% or 4% glycerol as a sole carbon source, no significant difference was found between the two strains (Supplementary Figure S1F).

### 3.11. Putative retron identified in the *Pcc* genome

Based on RNA mapping and expression levels, a putative noncoding RNA has been identified upstream of a genomically encoded reverse transcriptase. This, along with the presence of an ATPase and a HNH effector immediately downstream of the reverse transcriptase, suggests that a possible retron is present in *Pcc* (Millman et al., 2020). Furthermore, expression of these three genes is decreased ( $\sim -1.80$  Log<sub>2</sub>FC) in the  $\Delta$ *dgcO* strain, suggesting that DgcO might play a role in the regulation of this putative retron, which are believed to be at least partially controlled by cyclic nucleotides (Maxwell, 2021; Tal et al., 2021). However, so far c-di-GMP-dependent regulation of retrons has not been identified. While the architecture matches those previously reported (Millman et al., 2020), further validation is needed to confirm the presence and identity of the retron.

## 4. Discussion

Transcriptomic studies comparing aerobic and anaerobic growth of *Pcc* WT and  $\Delta$ *dgcO* found that while the WT strain had more differentially expressed genes (789 vs. 639), a large portion (402 genes) of these genes were differentially expressed in both strains. When compared to differential expression data from *P. atrosepticum* grown under aerobic or anaerobic conditions, a number of functions show similar regulation between the two conditions, such as genes related to the TCA cycle (e.g., *sdhCD*), stress response (e.g., *trxC*, *ahpC*), or transporters (e.g., *copA*, PTS sugar transporters) (Babujee et al., 2012). Interestingly, a large number of genes (142) found to be differentially

expressed between the same strains aerobically vs. anaerobically was also found to be differentially expressed between WT and  $\Delta$ *dgcO* aerobically. These genes include the *citXDEF* operon and genes involved in T6SS expression. These results, along with the changes in growth exhibited by  $\Delta$ *dgcO* *Pcc* when transitioning between aerobic and anaerobic conditions in M9 + LLL media (Figures 4A,B), suggest that DgcO plays an important regulatory role in *Pcc* transitioning between aerobic and anaerobic conditions.

When comparing WT and  $\Delta$ *dgcO* *Pcc*, analysis identified transcript differential expression upon *dgcO* deletion only when cells are grown aerobically, which is supported by previous studies that DgcO diguanylate cyclase activity is highest in the presence of O<sub>2</sub> (Burns et al., 2014). The results also identified changes in transcript levels relating to previously identified phenotypes, such as PCWDE production, AHL synthesis, biofilm formation, and motility (Burns et al., 2017). Possible explanations include control through riboswitches, direct binding of c-di-GMP (e.g., the cellulose biosynthesis protein BcsE, which contains a GIL domain), or through currently unidentified transcription factors (Weinberg et al., 2007; Xin et al., 2014). The widely studied c-di-GMP binding PilZ domain is not found in *Pcc*, suggesting the presence of additional, currently unidentified c-di-GMP domains (Ryjenkov et al., 2006).

LC-MS/MS quantification of c-di-GMP showed no significant difference in concentration between aerobically grown WT and  $\Delta$ *dgcO*, further supporting the hypothesis that aerobic DgcO-dependent c-di-GMP signaling is primarily regulated through local concentrations rather than changes in the concentration of the cellular c-di-GMP pool. Interestingly, anaerobically, the  $\Delta$ *dgcO* strain was found to have a significantly lower c-di-GMP concentration compared to WT, potentially loss of localized signaling that alters activity of other c-di-GMP metabolic enzymes, as previous work has demonstrated that DgcO is most active when oxygen is bound (Burns et al., 2014).

Transcriptomic data shows evidence of *dgcO* affecting the transcript levels of numerous transcription factors, such as *fur*, *flhDC*, and *19xpr* (Pöllumaa et al., 2012; Tanui et al., 2017), and genes known to be part of the transcription factors regulons also are differentially expressed, suggesting that *dgcO* may regulate cellular function by regulating transcription factors expression.

The decreased cellular metal content in  $\Delta$ *dgcO* (Figure 6), and changes in Mn tolerance (Figure 7) are likely due to differential expression of *fur* and various metal pumps (Supplementary Table S1). In addition, the changes in metal concentration possibly affect previously observed decreased *Pcc*  $\Delta$ *dgcO* biofilm formation, as Mg, Cu, Mn, and Zn are known to induce and stabilize biofilms in *Pectobacterium brasiliense*, and in *Pcc* str. PC1 (Manjurul Haque et al., 2012; Burns et al., 2017; Haque et al., 2017). While a few proteins related to metal regulation were identified as interacting with DgcO, the hit probabilities were low and the proteins were only identified in a subset of the biological replicates, suggesting that metal levels are modulate by DgcO primarily by regulating transcription (Table 5). Decrease in metal content and related growth phenotypes of the  $\Delta$ *dgcO* strain provide the first evidence of a GCS protein regulating cellular metal levels, although this putative regulation requires further investigation.

Interactions of *Pcc*DgcO protein with the chemotaxis proteins CheW and CheY suggest regulation of these proteins through specific localized c-di-GMP signaling. Interestingly, in general, decreased c-di-GMP levels correlate with increased motility, but  $\Delta$ *dgcO* was shown to have reduced motility compared to WT *Pcc* (Burns et al., 2017; Jenal et al., 2017). Additional chemotaxis and motility related proteins were identified in the pull-down experiments, as were a number of

TABLE 5 Comparison of identified *PccDgcO* interacting proteins and transcripts from key cellular pathways.

	Interacting proteins	Transcripts
Quorum sensing	<b>LuxS</b>	LuxR family txn regulator (3)
Motility	<b>FliC</b> , FliD, CheY, CheZ	<i>fliM</i> , <i>fliJ</i> , <i>fljP</i> , <i>fliL</i> , <i>flhC</i> , <i>flhD</i> , <i>fliZ</i> , <i>fliM</i> , methyl accepting chemotaxis protein (4)
ROS detoxification	<b>TsaA</b> , <b>Tpx</b> , <b>AhpF</b> , KatG, SodA	<i>ahpF</i>
Metal regulation	AfuA, Bfr, CysG1, Zur	<i>fur</i> , <i>ftnA</i> , <i>copA</i> , <i>znuA</i> , Fe permease (3), Fe transporter, Fe transporter/permease
Glycerol utilization	<b>GapA</b> , <b>GcvP</b> , <b>GpmA</b> , <b>GlpD</b> , Pkg	<i>glpB</i> , <i>glpC</i> , <i>ugpC</i> , glycerate kinase, <i>gapA</i>

Interacting proteins with the highest quality hits (as assessed by #PSM) are highlighted in bold.

differentially expressed transcripts (Table 5). These data suggest that DgcO might possibly regulate motility through stabilizing the Che chemotaxis scaffold, producing c-di-GMP that binds to chemotaxis machinery, and modulating expression of motility and chemotaxis genes. Overall, while these results only show low and moderate level changes in transcript levels, a model can be hypothesized by which DgcO primarily controls function through specific local c-di-GMP signaling, without a significant contribution to the overall cellular c-di-GMP pool. While this model requires further validation, it provides new insights into the cellular effects of oxygen-dependent c-di-GMP regulation in this important phytopathogen.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: NCBI GEO - GSE214075.

Author contributions

FF, NM, XL, and EW performed experiments and analyzed data. FF, NM, and EW wrote and edited manuscript. All authors contributed to the article and approved the submitted version.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Babujee, L., Apodaca, J., Balakrishnan, V., Liss, P., Kiley, P. J., Charkowski, A. O., et al. (2012). Evolution of the metabolic and regulatory networks associated with oxygen availability in two phytopathogenic enterobacteria. *BMC Genomics* 13:110. doi: 10.1186/1471-2164-13-110

Bolanos-Garcia, V. M., and Davies, O. R. (2006). Structural analysis and classification of native proteins from *E. coli* commonly co-purified by immobilised metal affinity chromatography. *Biochim. Biophys. Acta Gen. Subj.* 1760, 1304–1313. doi: 10.1016/j.bbagen.2006.03.027

Bowden, S. D., Eyres, A., Chung, J. C. S., Monson, R. E., Thompson, A., Salmond, G. P. C., et al. (2013). Virulence in *Pectobacterium atrosepticum* is regulated by a coincidence circuit involving quorum sensing and the stress alarmone, (p)ppGpp. *Mol. Microbiol.* 90, 457–471. doi: 10.1111/mmi.12369

Funding

This work was supported by NSF CHE1352040 (EW), NSF CHE2003350 (EW), and Frasci Foundation Grant 824-H17 (EW).

Acknowledgments

The authors thank members of the Weinert lab for assistance and helpful suggestions. The authors also thank Tim Miyashiro for providing the *V. fischeri* strains used in the signaling molecule experiments. The authors would like to acknowledge SeqCenter, LLC for running whole genome and RNA sequencing, the Penn State Laboratory for Isotopes and Metals in the Environment for running ICP-MS, and Creative Biogene for constructing the  $\Delta$ dgcO deletion strain. The authors would also like to acknowledge the Huck Metabolomics Core Facility for use of the Thermo TSQ Quantis Plus system and Sergei Koshkin for developing the LC/MS method and collecting the data shown in Figure 3.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2023.1134742/full#supplementary-material>

Boyd, C. D., and O’Toole, G. A. (2012). Second messenger regulation of biofilm formation: breakthroughs in understanding c-di-GMP effector systems. *Annu. Rev. Cell Dev. Biol.* 28, 439–462. doi: 10.1146/annurev-cellbio-101011-155705

Burns, J. L., Deer, D. D., and Weinert, E. E. (2014). Oligomeric state affects oxygen dissociation and diguanylate cyclase activity of globin coupled sensors. *Mol. Biosyst.* 10, 2823–2826. doi: 10.1039/C4MB00366G

Burns, J. L., Jariwala, P. B., Rivera, S., Fontaine, B. M., Briggs, L., and Weinert, E. E. (2017). Oxygen-dependent globin coupled sensor signaling modulates motility and virulence of the plant pathogen *Pectobacterium carotovorum*. *ACS Chem. Biol.* 12, 2070–2077. doi: 10.1021/acschembio.7b00380

Charkowski, A. O. (2006). “The soft rot *Erwinia*,” in *Plant-Associated Bacteria* ed. S. S. Gnanamanickam (Dordrecht: Springer Netherlands), 423–505.

Chatterjee, A., Cui, Y., and Chatterjee, A. K. (2009). RsmC of *Erwinia carotovora* subsp. *carotovora* negatively controls motility, extracellular protein production, and



virulence by binding FlhD and modulating transcriptional activity of the master regulator, FlhDC. *J Bacteriol* 191, 4582–4593. doi: 10.1128/JB.00154-09

Chen, X., Nomani, A., Patel, N., and Hatefi, A. (2017). Production of low-expressing recombinant cationic biopolymers with high purity. *Protein Expr. Purif.* 134, 11–17. doi: 10.1016/j.pep.2017.03.012

Colton, D. M., Stabb, E. V., and Hagen, S. J. (2015). Modeling analysis of signal sensitivity and specificity by *Vibrio fischeri* LuxR variants. *PLoS One* 10, 1–21. doi: 10.1371/journal.pone.0126474

Dadashi, M., Chen, L., Nasimian, A., Ghavami, S., and Duan, K. (2021). Putative rna ligase rtcB affects the switch between t6ss and t3ss in *Pseudomonas aeruginosa*. *Int. J. Mol. Sci.* 22:12561. doi: 10.3390/ijms222212561

Doğan, T., Macdougall, A., Saidi, R., Poggioli, D., Bateman, A., O'Donovan, C., et al. (2016). UniProt-DAAC: domain architecture alignment and classification, a new method for automatic functional annotation in UniProtKB. *Bioinformatics* 32, 2264–2271. doi: 10.1093/bioinformatics/btw114

Duggal, Y., Fontaine, B. M., Dailey, D. M., Ning, G., and Weinert, E. E. (2020). RNase I modulates *Escherichia coli* motility, metabolism, and resistance. *ACS Chem. Biol.* 15, 1996–2004. doi: 10.1021/acschembio.0c00390

Edwards, J. C., Johnson, M. S., and Taylor, B. L. (2006). Differentiation between electron transport sensing and proton motive force sensing by the Aer and Trs receptors for aerotaxis. *Mol. Microbiol.* 62, 823–837. doi: 10.1111/j.1365-2958.2006.05411.x

Ernst, C. M., Braxton, J. R., Rodriguez-Osorio, C. A., Zagieboylo, A. P., Li, L., Pironti, A., et al. (2020). Adaptive evolution of virulence and persistence in carbapenem-resistant *Klebsiella pneumoniae*. *Nat. Med.* 26, 705–711. doi: 10.1038/s41591-020-0825-4

Fan, J., Ma, L., Zhao, C., Yan, J., Che, S., Zhou, Z., et al. (2020). Transcriptome of *Pectobacterium carotovorum* subsp. *carotovorum* PccS1 infected in calla plants *in vivo* highlights a spatiotemporal expression pattern of genes related to virulence, adaptation, and host response. *Mol. Plant Pathol.* 21, 871–891. doi: 10.1111/mpp.12936

Flores-Kim, J., and Darwin, A. J. (2012). Links between type III secretion and extracytoplasmic stress responses in *Yersinia*. *Front. Cell. Infect. Microbiol.* 2:125. doi: 10.3389/fcimb.2012.00125

Fontaine, B. M., Martin, K. S., Garcia-Rodriguez, J. M., Jung, C., Briggs, L., Southwell, J. E., et al. (2018). RNase I regulates *Escherichia coli* 2',3'-cyclic nucleotide monophosphate levels and biofilm formation. *Biochem. J.* 475, 1491–1506. doi: 10.1042/BCJ20170906

Goedhart, J., and Luijsterburg, M. S. (2020). VolcanoR is a web app for creating, exploring, labeling and sharing volcano plots. *Sci. Rep.* 10:20560. doi: 10.1038/s41598-020-76603-3

Gorshkov, V., Kwenda, S., Petrova, O., Osipova, E., Gogolev, Y., and Moleleki, L. N. (2017). Global gene expression analysis of cross-protected phenotype of *Pectobacterium atrosepticum*. *PLoS One* 12:e0169536. doi: 10.1371/journal.pone.0169536

Grishin, A. V., Luyksaar, S. I., Kapotina, L. N., Kirsanov, D. D., Zayakin, E. S., Karyagina, A. S., et al. (2018). Identification of chlamydial T3SS inhibitors through virtual screening against T3SS ATPase. *Chem. Biol. Drug Des.* 91, 717–727. doi: 10.1111/cbdd.13130

Haque, M. M., Oliver, M. M. H., Nahar, K., Alam, M. Z., Hirata, H., and Tsuyumu, S. (2017). CytR homolog of *Pectobacterium carotovorum* subsp. *carotovorum* controls air-liquid biofilm formation by regulating multiple genes involved in cellulose production, c-di-GMP signaling, motility, and type III secretion system in response to nutritional and e. *Front. Microbiol.* 8, 1–17. doi: 10.3389/fmicb.2017.00972

Hengge, R. (2010). Cyclic-di-GMP reaches out into the bacterial RNA world. *Sci. Signal.* 3, 1–4. doi: 10.1126/scisignal.3149pe44

Hengge, R. (2021). High-specificity local and global c-di-GMP signaling. *Trends Microbiol.* 29, 993–1003. doi: 10.1016/j.tim.2021.02.003

Hengge, R., Galperin, M. Y., Ghigo, J. M., Gomelsky, M., Green, J., Hughes, K. T., et al. (2016). Systematic nomenclature for GGDEF and EAL domain-containing cyclic di-GMP turnover proteins of *Escherichia coli*. *J. Bacteriol.* 198, 7–11. doi: 10.1128/JB.00424-15

Illumina. (2019). Illumina bcl2fastq: 2.20.0.445 ed.

Jenal, U., Reinders, A., and Lori, C. (2017). Cyclic di-GMP: second messenger extraordinaire. *Nat. Rev. Microbiol.* 15, 271–284. doi: 10.1038/nrmicro.2016.190

Joshi, J. R., Burdman, S., Lipsky, A., Yariv, S., and Yedidia, I. (2016). Plant phenolic acids affect the virulence of *Pectobacterium atrosepticum* and *P. carotovorum* ssp. *brasiliense* via quorum sensing regulation. *Mol. Plant Pathol.* 17, 487–500. doi: 10.1111/mpp.12295

Kim, H. S., Ma, B., Perna, N. T., and Charkowski, A. O. (2009). Phylogeny and virulence of naturally occurring type III secretion system-deficient *Pectobacterium* strains. *Appl. Environ. Microbiol.* 75, 4539–4549. doi: 10.1128/AEM.01336-08

Kim, D., Paggi, J. M., Park, C., Bennett, C., and Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* 37, 907–915. doi: 10.1038/s41587-019-0201-4

Koiv, V., Andresen, L., Broberg, M., Frolova, J., Somervuo, P., Auvinen, P., et al. (2013). Lack of RsmA-mediated control results in constant hypervirulence, cell elongation, and

hyperflagellation in *Pectobacterium wasabiae*. *PLoS One* 8:e54248. doi: 10.1371/journal.pone.0054248

Liao, Y., Smyth, G. K., and Shi, W. (2014). Feature counts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatic.* 30, 923–930.

Liu, Y., Helmann, T. C., Stodghill, P., and Filiatrault, M. J. (2020). Complete genome sequence of the necrotrophic plant-pathogenic bacterium *Pectobacterium carotovorum* WPP14. *Plant Pathol. Cornell University* 105, 196–198. doi: 10.1094/PDIS-05-20-1059-A

Luo, P., He, X., Liu, Q., and Hu, C. (2015). Developing universal genetic tools for rapid and efficient deletion mutation in *Vibrio* species based on suicide T-vectors carrying a novel Counterselectable marker, vmi480. *PLoS One* 10:e0144465. doi: 10.1371/journal.pone.0144465

Manjurul Haque, M., Hirata, H., and Tsuyumu, S. (2012). Role of PhoP-PhoQ two-component system in pellicle formation, virulence and survival in harsh environments of *Dickeya dadantii* 3937. *J. Gen. Plant Pathol.* 78, 176–189. doi: 10.1007/s10327-012-0372-z

Marquez-Villavicencio, M. D. P., Weber, B., Witherell, R. A., Willis, D. K., and Charkowski, A. O. (2011). The 3-Hydroxy-2-butanone pathway is required for *Pectobacterium carotovorum* pathogenesis. *PLoS One* 6:e22974. doi: 10.1371/journal.pone.0022974

Maxwell, K. L. (2021). Cyclic pyrimidines jump on the anti-phage bandwagon. *Cells* 184, 5691–5693. doi: 10.1016/j.cell.2021.10.012

Miller, M. B., and Bassler, B. L. (2001). Quorum sensing in bacteria. *Annu. Rev. Microbiol.* 55, 165–199. doi: 10.1146/annurev.micro.55.1.165

Miller, S. T., Xavier, K. B., Campagna, S. R., Taga, M. E., Semmelhack, M. F., Bassler, B. L., et al. (2004). *Salmonella typhimurium* recognizes a chemically distinct form of the bacterial quorum-sensing signal AI-2. *Mol. Cell* 15, 677–687. doi: 10.1016/j.molcel.2004.07.020

Millman, A., Bernheim, A., Stokar-Avihail, A., Fedorenko, T., Voichek, M., Leavitt, A., et al. (2020). Bacterial Retrons function in anti-phage defense. *Cells* 183, 1551–1561.e12. doi: 10.1016/j.cell.2020.09.065

Mole, B., Habibi, S., Dangl, J. L., and Grant, S. R. (2010). Gluconate metabolism is required for virulence of the soft-rot pathogen *Pectobacterium carotovorum*. *Mol. Plant-Microbe Interact.* 23, 1335–1344. doi: 10.1094/MPMI-03-10-0067

Polikanov, Y. S., Blaha, G. M., and Steitz, T. A. (2012). How hibernation factors RMF, HPF, and YfiA turn off protein synthesis. *Science* 336, 915–918. doi: 10.1126/science.1218538

Pöllumaa, L., Alamäe, T., and Mäe, A. (2012). Quorum sensing and expression of virulence in *pectobacteria*. *Sensors* 12, 3327–3349. doi: 10.3390/s120303327

Rivera, S., Young, P. G., Hoffer, E. D., Vansuch, G. E., Metzler, C. L., Dunham, C. M., et al. (2018). Structural insights into oxygen-dependent signal transduction within globin coupled sensors. *Inorg. Chem.* 57, 14386–14395. doi: 10.1021/acs.inorgchem.8b02584

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616

Römling, U., Galperin, M. Y., and Gomelsky, M. (2013). Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. *Microbiol. Mol. Biol. Rev.* 77, 1–52. doi: 10.1128/MMBR.00043-12

Ryjenkov, D. A., Simm, R., Römling, U., and Gomelsky, M. (2006). The PilZ domain is a receptor for the second messenger c-di-GMP: the PilZ domain protein YcgR controls motility in enterobacteria. *J. Biol. Chem.* 281, 30310–30314. doi: 10.1074/jbc.C600179200

Soscia, C., Hachani, A., Bernadac, A., Filloux, A., and Bleves, S. (2007). Cross talk between type III secretion and flagellar assembly systems in *Pseudomonas aeruginosa*. *J. Bacteriol.* 189, 3124–3132. doi: 10.1128/JB.01677-06

Sudarsan, N., Lee, E. R., Weinberg, Z., Moy, R. H., Kim, J. N., Link, K. H., et al. (2008). Riboswitches in eubacteria sense the second messenger cyclic Di-GMP. *Science* 321, 411–413. doi: 10.1126/science.1159519

Surette, M. G., Miller, M. B., and Bassler, B. L. (1999). Quorum sensing in *Escherichia coli*, *salmonella typhimurium*, and *Vibrio harveyi*: a new family of genes responsible for autoinducer production. *Proc. Natl. Acad. Sci. U. S. A.* 96, 1639–1644. doi: 10.1073/pnas.96.4.1639

Tal, N., Morehouse, B. R., Millman, A., Stokar-Avihail, A., Avraham, C., Fedorenko, T., et al. (2021). Cyclic CMP and cyclic AMP mediate bacterial immunity against phages. *Cells* 184, 5728–5739.e16. doi: 10.1016/j.cell.2021.09.031

Tanui, C. K., Shyntum, D. Y., Priem, S. L., Theron, J., and Moleleki, L. N. (2017). Influence of the ferric uptake regulator (Fur) protein on pathogenicity in *Pectobacterium carotovorum* subsp. *brasiliense*. *PLoS One* 12, 1–21. doi: 10.1371/journal.pone.0177647

Tarnawski, M., Barends, T. R. M., and Schlichting, I. (2015). Structural analysis of an oxygen-regulated diguanylate cyclase. *Acta Crystallogr. D Biol. Crystallogr.* 71, 2158–2177. doi: 10.1107/S139900471501545X

Taylor, B. L. (1983). How do bacteria find the optimal concentration of oxygen? *Trends Biochem. Sci.* 8, 438–441. doi: 10.1016/0968-0004(83)90030-0

Toth, I. K., and Birch, P. R. J. (2005). Rotting softly and stealthily. *Curr. Opin. Plant Biol.* 8, 424–429. doi: 10.1016/j.pbi.2005.04.001

Troxell, B., and Hassan, H. M. (2013). Transcriptional regulation by ferric uptake regulator (Fur) in pathogenic bacteria. *Front. Cell. Infect. Microbiol.* 4, 1–13. doi: 10.3389/fcimb.2013.00059



- Tuckerman, J. R., Gonzalez, G., Sousa, E. H. S., Wan, X., Saito, J. A., Alam, M., et al. (2009). An oxygen-sensing diguanylate cyclase and phosphodiesterase couple for c-di-GMP control. *Biochemistry* 48, 9764–9774. doi: 10.1021/bi901409g
- Ueta, M., Ohniwa, R. L., Yoshida, H., Maki, Y., Wada, C., and Wada, A. (2008). Role of HPF (hibernation promoting factor) in translational activity in *Escherichia coli*. *J. Biochem.* 143, 425–433. doi: 10.1093/jb/mvm243
- Vinogradov, S. N., Tinajero-Trejo, M., Poole, R. K., and Hoogewijs, D. (2013). Bacterial and archaeal globins - a revised perspective. *Biochimica et Biophysica Acta Prot Proteom* 1834, 1789–1800. doi: 10.1016/j.bbapap.2013.03.021
- Walker, J. A., Rivera, S., and Weinert, E. E. (2017). Mechanism and role of globin coupled sensor Signalling. *Adv. Microb. Physiol.* 71, 133–169. doi: 10.1016/bs.ampbs.2017.05.003
- Wang, C., Pu, T., Lou, W., Wang, Y., Gao, Z., Hu, B., et al. (2018). Hfq, a RNA chaperone, contributes to virulence by regulating plant Cell Wall-degrading enzyme production, type VI secretion system expression, bacterial competition, and suppressing host defense response in *Pectobacterium carotovorum*. *Mol. Plant-Microbe Interact.* 31, 1166–1178. doi: 10.1094/MPMI-12-17-0303-R
- Wavemetrics, I. (2009). Igor pro. 6.10 ed. USA: Lake Oswego, OR.
- Weinberg, Z., Barrick, J. E., Yao, Z., Roth, A., Kim, J. N., Gore, J., et al. (2007). Identification of 22 candidate structured RNAs in bacteria using the CMfinder comparative genomics pipeline. *Nucleic Acids Res.* 35, 4809–4819. doi: 10.1093/nar/gkm487
- Xin, F., Ahmad, I., Blanka, A., Schottkowski, M., Annika, C., Galperin, M. Y., et al. (2014). GIL, a new c-di-GMP-binding protein domain involved in regulation of cellulose synthesis in enterobacteria. *Mol. Microbiol.* 93, 439–452. doi: 10.1111/mmi.12672

# Frontiers in Microbiology

Explores the habitable world and the potential of microbial life

The largest and most cited microbiology journal which advances our understanding of the role microbes play in addressing global challenges such as healthcare, food security, and climate change.

## Discover the latest Research Topics

[See more →](#)

### Frontiers

Avenue du Tribunal-Fédéral 34  
1005 Lausanne, Switzerland  
[frontiersin.org](https://frontiersin.org)

### Contact us

+41 (0)21 510 17 00  
[frontiersin.org/about/contact](https://frontiersin.org/about/contact)

