frontiers | Research Topics

# Trends in applications and improved production of biologically active metabolites using microbial fermentations

**Edited by**
Niranjan Koirala, Ameer Khusro and Sailesh Malla

**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 978-2-83250-533-5
DOI 10.3389/978-2-83250-533-5

## 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

# Trends in applications and improved production of biologically active metabolites using microbial fermentations

**Topic editors**

Niranjan Koirala — Gandaki Province Academy of Science and Technology, Nepal

Ameer Khusro — Loyola College, Chennai, India

Sailesh Malla — Chr. Hansen, Denmark

**Citation**

Koirala, N., Khusro, A., Malla, S., eds. (2023). *Trends in applications and improved production of biologically active metabolites using microbial fermentations*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83250-533-5

# Table of contents

frontiers | Frontiers in Microbiology

TYPE Editorial
PUBLISHED 22 November 2022
DOI 10.3389/fmicb.2022.1065888

Check for updates

OPEN ACCESS

EDITED AND REVIEWED BY
William James Hickey,
University of Wisconsin-Madison,
United States

*CORRESPONDENCE
Niranjan Koirala
koirala.biochem@gmail.com

SPECIALTY SECTION
This article was submitted to Microbiotechnology, a section of the journal Frontiers in Microbiology

RECEIVED 10 October 2022
ACCEPTED 09 November 2022
PUBLISHED 22 November 2022

CITATION
Koirala N, Khusro A and Malla S (2022) Editorial: Trends in applications and improved production of biologically active metabolites using microbial fermentations.
*Front. Microbiol.* 13:1065888.
doi: 10.3389/fmicb.2022.1065888

COPYRIGHT
© 2022 Koirala, Khusro and Malla. 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.

# Editorial: Trends in applications and improved production of biologically active metabolites using microbial fermentations

Niranjan Koirala[1]*, Ameer Khusro[2] and Sailesh Malla[3]

[1]Gandaki Province Academy of Science and Technology, Pokhara, Nepal, [2]Department of Biotechnology, Centre for Research and Development, Hindustan College of Arts and Science, Chennai, India, [3]Department of Strain Development, Discovery R&D, Chr. Hansen, Hørsholm, Denmark

KEYWORDS

**fermentation, microorganisms, genome mining, metabolomics, biologically active metabolites, chemotherapeutics**

Editorial on the Research Topic
Trends in applications and improved production of biologically active metabolites using microbial fermentations

Microorganisms are considered factories of metabolites. Fermentation is one of the oldest processes of biotechnology which represents the production of bioactive metabolites from distinct groups of microorganisms (Koirala et al., 2014; Khusro and Aarti, 2022). A plethora of microbes from soil, marine, fermented foods, and fecal and other environmental sources produces diversified classes of bioactive metabolites *via* fermentation process.

Amino acids, nucleotides, and fermentation end products (ethanol and organic acids) are the primary metabolites which help in the growth and metabolism of microorganisms. On the other hand, secondary metabolites viz. enzymes, proteins, bacteriocins, antibiotics, pigments, growth hormones, etc. are generally synthesized during the stationary phase of microbial growth and exhibit disparate applications in nutrition, medicine, agriculture, food industries, environment, and livestock sectors (Koirala et al., 2017; Singh et al., 2017; Yang et al.). The production or secretion of these metabolites from microbes can be enhanced by optimizing independent variables *via* solid state or submerged fermentation process using cost-effective substrates (Martău et al., 2021).

Although a plethora of reports has been published in the past revealing the pivotal roles of microorganisms as bioactive metabolites producers, our understanding on the mechanism of production, its enhancement strategies, and its applications in diversiform sectors is still limited. The Research Topic entitled "*Trends in applications and improved production of biologically active metabolites using microbial fermentations*" is a significant attempt to bring together renowned scientists and researchers worldwide and spotlight the novel research activities being carried out in the theme of "biologically

active metabolites from microbes." This Research Topic comprises 11 articles covering crucial aspects of microbial fermentation and its associated metabolites' applications.

To facilitate the use of woody plant as a natural biomass resource for addressing the shortage of feed for ruminants in the tropics, Du et al. used PacBio SMRT sequencing method to explore the microbial co-occurrence network and silage fermentation of Gliricidia and Leucaena prepared with Napier grass and corn stover. The findings suggested that a woody plant can be mixed with corn stover to make high-quality silage, which can alleviate the shortage of feed and promote local animal production. In another interesting investigation, Al-Askar et al. focussed on the bioprocessing of biomass residuals (peanut plant residual) into a beneficial substance (citric acid) using *Trichoderma longibrachiatum*. The crude citric acid that was obtained showed inhibitory potential against three toxinogenic fungi (*A. flavus*, *A. ochraceus*, and *F. oxysporum*) too. The study demonstrated the utility of T. longibrachiatum of an endophytic fungus to produce citric acid through the fermentation of peanut plant residual biomass. Native grass is widely utilized for grazing and haymaking, and is the prime source of forage in pastoral areas. Seasons affect the quality and productivity of native grass. Ensiling is a traditional method to preserve forage nutrients in the harvesting season. However, there is limited information available related to the microbial community and fermentation products during the ensiling process of native grass with additive treatments. Therefore, to improve the usability of native grass resources as feed, the effects of lactic acid bacteria and molasses additions on the microbial population, fermentation quality, and nutritional quality of native grass during silage were investigated by Li et al.. Outcomes of this study suggested that the supplementation of lactic acid bacteria and molasses improved the relative abundance of lactobacilli of native grass silage and enhanced the fermentation quality.

Daptomycin (a cyclic lipopeptide antibiotic) shows antibacterial activity against antibiotic-resistant Gram-positive bacteria. It is produced by *Streptomyces roseosporus via* non-ribosomal peptide synthetases. Lyu et al. successfully utilized multi-level metabolic engineering strategies in *S. roseosporus* to reconstruct high-quality daptomycin- overproducing strain L2797-VHb, including precursor engineering, regulatory pathway reconstruction, byproduct engineering, multicopy biosynthetic gene cluster, and fermentation process engineering.

The quality of cigar tobacco leaves is affected by the microbiota. Zheng et al. improved the quality of cigar tobacco leaves by fermenting it with *Acinetobacter* sp. 1H8 and *Acinetobacter indicus* 3B2. The inoculation of these two bacterial strains completely changed the original bacterial community. The study indicated the improvement of fermentation product quality by regulating microbial community, and gain insight into the microbial ecosystem.

Similarly Feng et al. studied the effect of exogenous electrons on electroactive *Escherichia Coli*, Shi et al. studied the catotenoid synthesis in *Phaffia rhodozyma*, Malla et al. used functional metagenomis approach for the identification of a novel efficient L-Lysine exporter. The production and characterization of lanthomicins by promoter engineering in *Streptomyces chattanoogensis* L10 was performed by Liu et al. Finally She et al. studied the yield improvements of Albofungins, and alpha amylase production using fermentation technology was researched by Geissler et al..

In summary, the articles published in the Research Topic "*Trends in applications and improved production of biologically active metabolites using microbial fermentations*" provides great insights on the novel approaches implemented (statistical optimization, metagenomics, metabolic engineering, whole genome sequencing, proteomics, and transcriptomic) not only in the quality improvement of various biomasses but also the enhanced production of biologically active metabolites from different microbes *via* fermentation process.

## Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

## Acknowledgments

We thank the authors of the papers published in this Research Topic for their valuable contributions, all the reviewers for their rigorous reviews and editors for editing the articles until making the final decision.

## 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.

# References

Khusro, A., and Aarti, C. (2022). Metabolic heterogeneity and techno-functional attributes of fermented foods-associated coagulase-negative staphylococci. *Food Microbiol.* 105, 104028. doi: 10.1016/j.fm.2022.104028

Koirala, N., Pandey, R. P., Thang, D. V., Jung, H. J., and Sohng, J. K. (2014). Glycosylation and subsequent malonylation of isoflavonoids in *E. coli*: Strain development, production and insights into future metabolic perspectives. *J. Ind. Microbiol. Biotechnol.* 41, 1647–1658. doi: 10.1007/s10295-014-1504-6

Koirala, N., Pandey, R. P., Thuan, N. H., Ghimire, G. P., Jung, H. J., Oh, T. J., et al. (2017). Metabolic engineering of *Escherichia coli* for the production of isoflavonoid-4?-*O*-methoxides and their biological activities. *Biotechnol. Appl. Biochem.* 66, 484–493. doi: 10.1002/bab.1452

Martău, G. A., Unger, P., Schneider, R., Venus, J., Vodnar, D. C., and López-Gómez, J. P. (2021). Integration of solid state and submerged fermentations for the valorization of organic municipal solid waste. *J. Fungi* (Basel). 7, 766. doi: 10.3390/jof7090766

Singh, R., Kumar, M., Mittal, A., and Mehta, P. K. (2017). Microbial metabolites in nutrition, healthcare and agriculture. *Biotech.* 3, 15. doi: 10.1007/s13205-016-0586-4

frontiers in Microbiology

ORIGINAL RESEARCH
published: 10 March 2022
doi: 10.3389/fmicb.2022.772417

# A Novel Endophytic *Trichoderma longibrachiatum* WKA55 With Biologically Active Metabolites for Promoting Germination and Reducing Mycotoxinogenic Fungi of Peanut

*Abdulaziz A. Al-Askar[1]*, Ehsan M. Rashad[2], Zeiad Moussa[3], Khalid M. Ghoneem[2], Ashraf A. Mostafa[1], Fatimah O. Al-Otibi[1], Amr Abker Arishi[4] and WesamEldin I. A. Saber[3]**

[1]*Department of Botany and Microbiology, Faculty of Science, King Saud University, Riyadh, Saudi Arabia,* [2]*Department of Seed Pathology Research, Plant Pathology Research Institute, Agricultural Research Center, Giza, Egypt,* [3]*Microbial Activity Unit, Department of Microbiology, Soils, Water and Environment Research Institute, Agricultural Research Center, Giza, Egypt,* [4]*School of Molecular Sciences, The University of Western Australia, Perth, WA, Australia*

Plant residuals comprise the natural habitat of the plant pathogen; therefore, attention is currently focusing on biological-based bioprocessing of biomass residuals into benefit substances. The current study focused on the biodegradation of peanut plant residual (PNR) into citric acid (CA) through a mathematical modeling strategy. Novel endophytic *Trichoderma longibrachiatum* WKA55 (GenBank accession number: MZ014020.1), having lytic (cellulase, protease, and polygalacturonase) activity, and tricalcium phosphate (TCP) solubilization ability were isolated from peanut seeds and used during the fermentation process. As reported by HPLC, the maximum CA (5505.1 μg/g PNR) was obtained after 9 days in the presence of 15.49 mg TCP, and 15.68 mg glucose. GC–MS analysis showed other bioactive metabolites in the filtrate of the fermented PNR. Practically, the crude product (40%) fully inhibited (100%) the growth and spore germination of three mycotoxinogenic fungi. On peanuts, it improved the seed germination (91%), seedling features, and vigor index (70.45%) with a reduction of abnormal seedlings (9.33%). The current study presents the fundamentals for large-scale production in the industry for the sustainable development of PNR biomass as a natural source of bioactive metabolites, and safe consumption of lignocellulosic-proteinaceous biomass, as well. *T. longibrachiatum* WKA55 was also introduced as a novel CA producer specified on PNR. Application of the resulting metabolite is encouraged on a large scale.

**Keywords: citric acid, hydrolytic enzymes, biocontrol, biomass, fermentation, response surface methodology, peanut, bioactive metabolites**

OPEN ACCESS

***Edited by:***
*Niranjan Koirala, Gandaki Province Academy of Science and Technology, Nepal*

***Reviewed by:***
*Sadia Sultan, Universiti Teknologi MARA Puncak Alam, Malaysia*
*Surendra Sarsaiya, Zunyi Medical University, China*
*Monica Butnariu, Banat University of Agricultural Sciences and Veterinary Medicine, Romania*

****Correspondence:***
*Abdulaziz A. Al-Askar*
*aalaskara@ksu.edu.sa*
*WesamEldin I. A. Saber*
*wesameldin.saber@arc.sci.eg*

***Specialty section:***
*This article was submitted to Microbiotechnology, a section of the journal Frontiers in Microbiology*

***Received:*** *08 September 2021*
***Accepted:*** *21 February 2022*
***Published:*** *10 March 2022*

***Citation:***
*Al-Askar AA, Rashad EM, Moussa Z, Ghoneem KM, Mostafa AA, Al-Otibi FO, Arishi AA and Saber WIA (2022) A Novel Endophytic Trichoderma longibrachiatum WKA55 With Biologically Active Metabolites for Promoting Germination and Reducing Mycotoxinogenic Fungi of Peanut. Front. Microbiol. 13:772417. doi: 10.3389/fmicb.2022.772417*

## INTRODUCTION

The agricultural proteinaceous-lignocellulosic residues (PNR) of the above-ground part of groundnut or peanut (*Arachis hypogaea* L.) represent an environmental issue if they are not wisely managed. Moreover, they can be biologically converted into valuable biomolecules. On the dry matter (93.39%) base, ash composed 5.33%, and the organic matter; 94.67%, the latter has a high level of protein (8.08%), and the remaining mostly composed of cellulose, hemicellulose, lignin, and pectin, and thus, it been used as a ruminant feedstuff (Alebel et al., 2019). The degradation of cellulose, hemicellulose, lignin, pectin, and protein polymers requires the catalytic action of cellulases pectinase and proteases enzymes to yield various corresponding monomers, such as mono-sugars and D-galacturonic acid as carbon and energy source, whereas the proteolytic activity is referred to the protease enzymes that yield various amino acids, which perform a critical role in the growth and multiplication of the microorganisms (Jacob et al., 2008; Elsayed et al., 2021; Al-Askar et al., 2021a).

Experiments in recent decades have produced a large quantity of evidence on the prospective use of helpful endophytic microorganisms to manage diseases. Endophytic microbiome (endosymbionts) is defined as the diverse array of microbial communities that live and grow intra-and/or intercellularly in the plant's tissues, at least part of its life cycle, without causing over symptoms on the host plants. Endophytes may benefit hosts through a variety of mechanisms, including molecules that increase their ability to compete for space, nutrients, and/or ecological niches, as well as the synthesis of antimicrobial substances, phytohormones, and peptides that may keep vegetables and plant organs healthy. Fortunately, all these substances have no negative impact neither on consumers nor on the environment (Shehata et al., 2016; López et al., 2018).

Of the seed-borne endophytes, the fungal seed microbiome is an important variable branch that may be deeply impacted by local conditions and host genotypes (Klaedtke et al., 2016). Although fungal seed-borne endophytes comprise a significant part of the seed microbiome, to date, this fungal microbiome has not been fully explored. Superior to other fungi, endophytic *Trichoderma* spp. produce a novel array of metabolites that exhibit different bioactivities with diverse applications in pharmaceutical, industrial, plant growth regulation, and many others (De Silva et al., 2019; Sridharan et al., 2020, 2021). Therefore, this study is a trial to shed some light and share some information about such a limited-explored area.

Despite being an inconvenient task, finding out a suitable microorganism having the complete and complementary catalytic action to manage the biodegradation of PNR is of great importance. Proper bioprocessing of PNR is a very beneficial and talented approach and potentially offers a low-cost raw resource to various valuable bioproducts. Alternative to the conventional substrates, proteinaceous-lignocellulosic materials such as PNR are abundant, attractive, and usually low-priced as well as, provide the required nutrients to the microorganisms, and are suitable substrates for their growth and activity (Kumari et al., 2008; Max et al., 2010).

Most of the previous work on plant biomass used two-phase bioconversion, i.e., the hydrolysis of complex materials to fermentable monomers, followed by microbial fermentation of monomers to a target product(s). The biodegradation is usually catalyzed by hydrolytic enzymes and the fermentation process is run by a proper microbe (Saber et al., 2015; Al-Askar et al., 2021b).

Solid-state fermentation (SSF) was applied during the bioconversion process. What is nominated as the best choice of fermentation of plant biomass is the low wastewater output, minimum production cost, fewer energy requirements, high rate of efficiencies, easier aeration, and simple fermentation medium (Show et al., 2015). These advantages could be maximized when using proper experimental design, of which central composite design (CCD) of response surface methodology (RSM) is a precise and easily applicable statistical approach for building a mathematical empirical model for maximization of a target product, in contrast to conventional procedures; this design is used to identify the importance and interactions of the input factors (independent variables) and select conditions that optimize the output response (variable) using a limited number of experiments (runs) even when little is known about the bioprocess conditions (Saber et al., 2015; Asadi and Zilouei, 2017).

The importance of the current study arises from the point of view that much remains to be studied for developing a commercially feasible process for the utilization of alternative disposal of PNR. Also, of the points to be improved is to find out alternative efficient fungus in bioconversion of PNR directly into the target biopolymer. The current work was set up to merge the two-phase bioconversion of PNR in one single step, in which the resulting monomers from the biodegradation process can share directly to the formation of the target molecules, e.g., citric acid (CA) (2-hydroxy-propane-1,2,3-tricarboxylic acid).

Herein, a novel endophytic isolate of *Trichoderma* sp. was reported as a biologically active fungus capable to produce CA, as a value-added product by the bioprocessing of PNR biomass under a one-phase conversion process, applying SSF technique and CCD. The fungal exudate was evaluated as a suppressor to the associated mycotoxinogenic and pathogenic fungi that settled in biomass and peanut seeds as well.

## MATERIALS AND METHODS

Unless otherwise specified, all chemicals used through the enzymatic and biochemical analysis are of a high grade obtained from Sigma-Aldrich Company (Egypt brach).

### Isolation of Endophytic Fungi of Peanut

Fifteen-seed samples of peanut plants were collected from different locations in Al Nubaria district, and AL-Behaira and Sohage governorates. The surveyed area lays between latitudes of 26°49′N and 31°11′N, and longitudes of 29°68′E and 31°73′E, Egypt,

**Abbreviations:** PNR, Peanut plant residual; CA, Citric acids; TCP, Tricalcium phosphate; FPase, Filter-paperase; CMCase, Carboxymethyl cellulase; PGase, Polygalacturonases; VOCs, Volatile organic compounds.

during June 2020. The samples were gathered from an area of 50 m× 50 m around each sampling site in a random zigzag pattern. Full mature peanut pods were gathered in cotton bags, labeled in the field, and stored at 4°C until the isolation process.

The discovery of endophytic seed-borne fungi was performed, utilizing the agar plate technique recommended by ISTA (2007). A total of 200 healthy-looking seeds of each sample were surface sterilization, applying ethanol (70% for 60 s), then by sodium hypochlorite (2% for 90 s), and again in ethanol (100% for 30 s); finally, the seeds were washed using sterilized water for five times and dried using sterilized filter papers under laminar air flow chamber. For each plate, 10 were plated on a potato dextrose agar (PDA) plate (Difco, United States). To prevent any bacterial growth, L-chloramphenicol and streptomycin sulfate (5 mg/L each) were added to the medium. The seed-containing dishes were reared at 20 ± 2°C under cool white fluorescent light with alternating cycles of 12 h light/darkness.

Along 7 days, plates were inspected under stereoscopic and compound microscopes to pick the emerged fungi. Hyphal-tip and/or single-spore isolation procedures were used to gain pure cultures of endophytic fungi. The percentage of occurrence of each isolated endophyte fungal species was calculated using equation (10) follows:

$$\text{Mean percentage of a fungus} = \frac{\left(\Sigma \text{ fungus percentage in examined samples}\right)}{\text{Total number of examined samples}} \quad (1)$$

All endophyte fungi were maintained on slants of PDA for further investigation. Fungi were identified according to their cultural features, fungal morphology, and microscopic characteristics (Ellis, 1971, 1976; Domsch et al., 1980; Nelson et al., 1983; Leslie and Summerell, 2008; Samson et al., 2014).

## Screening the Lytic Activity of the Endophytic Fungi

The isolated endophytic fungi were first descriptively screened for cellulolytic activity by growing on plates of agar medium containing 0.5% carboxymethyl cellulose (CMC) at 25°C and examined daily up to 5 days. Active cellulolytic isolates were determined by the formation of a clear zone after flooding with 0.2% Congo red for 15 min, then de-stained by washing twice with 1 M NaCl for 15 min (Meddeb-Mouelhi et al., 2014). Following the previous primary screening, seven isolates were selected, and quantitatively screened for cellulase, and protease activities, as well as phosphate solubilization ability by growing on a fermentation medium based on PNR.

### Preparation of PNR

The whole PNR was collected from the regions reported above, dried at 50°C overnight, and blinded in an electric grinder to get fine residue (1 mm in length) to serve as solid sustenance and substrate for during fermentation production. PNR was a natural growth substrate to simulate the natural growth situations of the fungi, and therefore, was not subjected to any other pretreatment.

### Fermentation Procedure

The SSF technique was used for quantitatively screening the capability of the selected seven isolates to degrade PNR into a simpler form, as well as the solubilization of complex tricalcium phosphate (TCP). Unless otherwise specified, 1 g of PNR was mixed thoroughly with 20 mg TCP, and 5 ml tap water, pH 5, in 250 ml conical flasks, the PNR-based fermentation medium was autoclaved for 15 min at 121°C. After sterilization, 1 ml of spore suspension, containing 108 spores/ml, from each fungus was used to inoculate the fermentation media. During the incubation period, the moisture content was retained at approximately 65%. Unless otherwise specified, the fermentation lasted for 7 days at 28°C. By the end of the incubation period, 10 ml of 0.01% Tween 80 was integrated with the fermented matter and kept on a rotary shaker (150 rpm for 30 min), then filtered through Whatman No. 1 filter paper before being centrifuged at 5,000 rpm for 20 min. The residue of PNR after SSF was oven-dried (50°C) to constant weight for estimation of reduction in PNR weight after fermentation. The post-culture filtrate (PNR hydrolysate) was biochemically evaluated. The time course profile of PNR bioconversion into organic acids was determined.

### Biochemical Analysis

Cellulases in the post-culture filtrate was carried out following the protocols of Elsayed et al. (2021) with slight modification, in which the activities of filter-paperase (FPase), carboxymethyl cellulase (CMCase), and hemicellulose-degrading xylanase on 1% of microcrystalline cellulose, carboxymethyl cellulose, and oat-spelt xylan were assayed, respectively. All substrates were individually suspended in citrate buffer (0.05 M, pH 4.8), the reaction blend (1 ml of the filtrate and 1 ml appropriate substrate-buffer solution) was incubated at 50°C for 60, 30, and 15 min, respectively. The resulting free reducing sugars by the enzymatic action was detected using the 3,5-dinitrosalicylic acid method (Miller, 1959), and the developed color was measured using UV-VIS spectrophotometer (Jenway 7,315) at wave length $A_{575}$ nm. Enzyme unit (U) is defined as the amount of enzyme required to release one μmol $min^{-1}$ of glucose (FPase or CMCase) or xylose (xylanase) per g PNR/min under the assay conditions.

Polygalacturonases (PGase) was assayed by detecting the reducing monomers released from 1% polygalacturonic acid in 0.1 M sodium acetate buffer (pH 5.2) after 30 min incubation at 40°C (Al-Askar et al., 2021a), using the 3,5-dinitrosalicylic acid method as described above with D-galacturonic acid monohydrate as the standard. One PGase unit was described as that amount of enzyme that yields 1 μmole reducing ends per g PNR/min under the assay conditions.

The proteolytic activity (protease) was quantitatively assayed in the crude extract using 2.5 ml casein (1%) as a substrate incubated with 1 ml crude enzyme, then incubation at 37°C for 10 min, and 2.5 ml TCAA was added. One milliliter of the supernatant was used for detection of the resulted-free amino acids spectrophotometrically at $A_{660}$ nm (Negm et al., 2021). One protease U was termed as the amount of the enzyme that releases 1 μg of tyrosine equivalent per g PNR/min under the assay settings.

The soluble phosphorus was determined in an aliquot of the sample (5 ml) was pipetted into a 25 ml volumetric flask; then, $NH_4$-molybdate solution (5 ml) was slowly added then shaken gently. To which, 1 ml sodium sulfite (20%) was added. The mixture was shaken and 1 ml hydroquinone (0.5%) was added and then was shaken. Total phosphorus was determined spectrophotometrically $A_{660}$ nm after 15 min (Jackson, 2005).

The total organic acids were also determined (Montgomery et al., 1962). A combination of 0.5 ml of the sample, 1.5 ml of ethylene glycol, and 0.2 ml of diluted sulfuric acid were mixed, followed by heating in a boiling water bath for 3 min, then immediately cooling in cold water. To the mixture, 0.5 ml of hydroxyl ammonium sulfate solution and 2.0 ml of 4.5 N sodium hydroxide were added. Then, 10 ml of acidic ferric chloride was added. The optical density was measured at $A_{500}$ nm within 1 h of color development.

The post-culture filtrate was analyzed for glucose content utilizing the kit of glucose oxidase (Spainreact Co., Spain). The pH of the post-culture filtrate was measured by a glass electrode pH-meter (CP-501, Elmetron). The dry weight of the residual fermented PNR was measured by drying at 50°C until a constant weight was attained.

## The Matrix of CCD Used for Maximization of CA Biosynthesis

For bioprocessing the maximization of CA biosynthesis by *Trichoderma* sp. WKA55, the SSF of PNR-based medium was applied, using the experimental matrix of CCD. In which one-gram PNR was mixed with two independent nutritional variables (glucose and TCP) at five concentrations each. According to the design matrix of CCD, five coded levels from each variable were allocated at the center (0), low (−1), high (1), and two axial levels (alpha, $\alpha \pm 1.414$), the corresponding uncoded (actual) points were 10, 20, 15, 7.929, and 22.071 mg TCP per g PNR of both glucose and TCP. A total of 11 combinations of the medium (run) were prepared, autoclaved, and inoculated with the spore suspension (108/ml per g PNR) of the fungus. After incubation (28°C), CA was quantified by high-performance liquid chromatography (HPLC). The next second-order polynomial quadratic model (Equation 2) is used for theoretical valuation of the optimum level of each variable that maximizes CA biosynthesis:

$$Y = \beta_0 + \sum \beta_i X_i + \sum \beta_{ij} X_i X_j + \beta_{ii} X_i^2 \quad (2)$$

Where, Y, the predicted CA; $\beta_0$, the model coefficient; $\beta_i$, the linear coefficients; $\beta_{ij}$, the interaction coefficients; $\beta_{ii}$, quadratic coefficients; and $X_i$, and $X_j$, the independent variables of glucose and TCP, respectively.

## Determination of CA by HPLC

Methanol (20 ml) was used for extraction of CA from the sample at 40°C, then filtering and concentrating under reduced pressure until dryness, followed by redissolving in acidified water (pH 2) with HCl, then evaporated at 40°C under dryness, and reliquefied again in 1 ml $H_2SO_4$ (0.01 N). For the CA examination, 20 µl was injected in HPLC (Agilent 1,200 Infinity Series, United Kingdom), with a C18 column at 30°C. Elution was carried out isocratically with $H_2SO_4$ (0.01 N), as the mobile phase, at a flow rate of 0.1 ml/min, for 120 min. The UV detection was accomplished at 214 nm. CA was quantified by the absorbance verified in the chromatograms relative to the CA standard (Silva et al., 2002).

## Evaluation of Crude CA

### *In vitro* Assessment of Crude CA on Mycotoxinogenic Fungi

The toxigenic fungi: *Aspergillus flavus* PPRI3, *Aspergillus ochraceus* PPRI5, and *Fusarium oxysporum* PPRI10 were kindly provided by Seed Pathology Research Department, Plant Pathology Research Institute, Agricultural Research Center (ID: 60019332), Giza, Egypt. These fungal strains are toxigenic and are reported by the provider to produce aflatoxin, ochratoxin, and fusaric acid, respectively. CA was produced under the optimized fermentation conditions of PNR with TCP and glucose after incubation at 28°C for 9 days as the optimum fermentation period. The resulting crude CA was sterilized by microfiltration using membrane filters with 0.45 µ pore size.

The minimum antifungal concentration (MIC) of the crude CA of *Trichoderma* sp. WKA55 was evaluated against the three toxigenic fungi. One hundred milliliters Erlenmeyer flasks containing 20 ml sterilized Czapek-Dox broth medium (Sigma-Aldrich) with various concentrations of CA (0%, 10%, 20%, 30%, and 40%, v/v) were inoculated with 0.2 ml of $10^6$ conidia/ml of the tested fungal spore suspensions prepared from 6-day-old cultures, then incubated at 28°C for 6 days. At the end of the incubation period, the mycelial mats were harvested, washed with distilled water, and oven-dried at 50°C to constant weight for measuring the growth of tested fungi. The reduction in the growth of mycotoxinogenic fungi was calculated according to the following Equation (3):

$$Growth\ reduction\,(\%) = C - T / C \times 100 \quad (3)$$

Where $C$ is the average mycelial dry weight of each pathogen in the control and $T$ is the average dry weight of each pathogen in the crude CA treatment.

The anti-activity of the crude CA on germination of the fungal conidia was assayed in test tubes containing 10 ml sterilized Czapek-Dox broth medium supplemented with the same previous concentrations of *CA*. All tubes were incubated individually with 0.2 ml of $10^6$ conidia/ml of the spore suspensions, then incubated at 28°C for 12 h. With the aid of a light microscope, the percentage of spore germination was then counted using a hemocytometer slide (Elad, 1992).

### CA Against Seed-Borne Fungi and Seed Germination

The influence of CA on the occurrence of seed-borne pathogenic fungi, germination, and vigor index of peanut seeds (Giza 5),

a common susceptible greenhouse cultivar to Fusarium wilt in Egypt, was investigated. Three seed lots, with the lowest germination (~70%), having high infection by the previously tested; *A. flavus*, *A. ochraceus*, and *F. oxysporum* phytopathogens were selected.

The seeds of each sample were individually soaked in the bio-prepared CA for 30 min, at the best concentration obtained from the preceding MIC test, and the seeds were dried for 2 h under a sterile air stream (Singh et al., 2016). Seeds that were soaked in sterile distilled water were served as control. The seeds were incubated (20 ± 2°C) and relative humidity of 50% for 14 days with rotating cycles of light and darkness every 12-h, in a germination moist chamber. The germination test was conducted on 400 seeds using the standard moist blotter technique (ISTA, 2007). After 7 days of incubation, the fungal community that developed on the seeds was screened with the use of a stereoscopic microscope and confirmed with the optical microscope. The results were expressed as the percentage per pathogen detected.

Germinated seeds were counted at 7 and 14 days and results were expressed as germination percentage. Seed with a radical length of five millimeters or more was counted as germinated. For vigor evaluation, the root length of 100 seeds was measured at 7 and 14 days after planting. Seedling's vigor index (SVI) was calculated with the average root length and germination percentage of each seed sample. The vigor index was assessed utilizing the following Equation (4; Abdul-Baki and Anderson, 1973).

$$SVI = Germination\ percentage(\%) \times average\ root\ length(cm) \quad (4)$$

Other disinfected seeds were treated with a recommended fungicide [Pink-S (5-methylisoxazol-3-ol) at 30% SL] for 15 min and used as a positive control. A set of untreated seeds was served as a negative control. The germination (%) and the measurement of the root lengths (cm), fresh and dry weight (g) were estimated on the 14th day.

## Gas Chromatography-Mass Spectrophotometry

The investigation of GC-MS of the cell-free fungal filtrate was carried out, using Agilent 6,890 gas chromatography prepared with an Agilent mass spectrometric detector, with a direct capillary interface and fused silica capillary column PAS-5 ms (30 mm × 0.25 µm film thickness) (GC-MS-QP 2010 Shimadzu, Japan). The carrier gas (helium) was used at approximately 1 ml/min, pulsed split less mode. The solvent delay (3 min) and the injection size (1.0 ml) were adjusted. The mass spectrophotometric detector was functioned in electron impact ionization mode in ion energy of 70 eV, scanning from *m/z* 50 to 500. The temperatures ion source and quadrupole were adjusted at 230°C and 150°C, respectively. The electron multiplier voltage was kept 1,250 v above autotune. The instrument was manually tuned by perfluorotributylamine. The temperature program of GC was started at 60°C for 2 min then elevated to 280°C at a rate of 5°C/min, and 10 min hold at 280°C, the injector temperature was set at 280°C. Separated peaks were recognized using NIST08s.LIB mass spectral database (Sparkman et al., 2011).

## Fungal Identification on Molecular Base

The selected fungus was molecularly identified. Frozen-dried mycelium (20 mg) was crushed with Kontes pestles in a 1.5 ml tube with 500 µl of Dellaporta buffer (100 mM Tris pH 8. 50 mM ethylenediamine-tetraacetate EDTA, 500 mM NaCl, and 10 mM beta-mercaptoethanol) in liquid nitrogen; then, 33 µl of sodium dodecyl sulfate (20%, w/v) was added, the mixture was vortexed and incubated (65°C) for 10 min. Potassium acetate (160 µl of 5 M) was added and vortexed. Then, spun for 10 min at 10,000 rpm in a microcentrifuge tube. The supernatant (450 µl) was mixed with 450 µl phenol, chloroform, and isoamyl-alcohol (25:24:1) and vortexed for 5 min. Then, spun for 5 min at 10,000 rpm. Four hundred microliter of the upper phase was moved to a clean microcentrifuge tube and mixed with 0.5 volumes of isopropanol, vortex, and spun (10 min at 14,000 rpm). The supernatant was detached, and the precipitated nucleic acid was washed with 70% ethanol, spun (5 min at 10,000 rpm), and finally resuspended in 100 µl of $ddH_2O$.

Amplification was carried out *via* PCR based on 18 s rRNA primers and sequencing of the internal transcribed spacer (ITS) region (White et al., 1990). The forward IT5 primer (5′-GGAAGTAAAAGTCGTAACAAGG-3′) and the reverse ITS4 primer (5′-TCCTCCGCTTATTGATATGC-3′) were utilized to magnify the entire ITS region. PCR was done in a reaction, containing 1 µl of the fungal DNA extract (40 ng of total DNA), 2 mM $MgCl_2$, 2.5 of 10x PCR buffer, 1.5 µl of 10 µM of each primer, 2.5 µl of 10 mM dNTPs, and 0.3 µl of 5 U Taq DNA polymerase and the reaction was completed to 25 µl with Nuclease-free water. PCR was conducted in the ESCO Swift Maxi Thermal Cycler with initial denaturation at 95°C for 2 min, followed by 35 cycles of 95°C for 30 s, 52°C for 30 s, and 72°C for 30 s, and the final cycle is a polymerization that was made at 72°C for 10 min. PCR products were purified using QIAquick® PCR Purification Kit (Cat. No. 28106) as the manufacturing procedures. The purified PCR products were sequenced (Macrogene Company, Seoul, Korea) in both directions using ITS5 and ITS4 primer pairs. The sequence (613 bp) was aligned and deposited into the GenBank database[1] to get similarities with the related fungal sequence (nr/nt) database.

A phylogenetic tree was built based on the ITS region sequence comparisons of length polymorphism of the amplified PCR, and sequences from the database using blast tree and aligned using aligned sequences nucleotide BLAST. The phylogeny was tested using the bootstrap method with 1,000 replications and generated based on the Jukes-Cantor model. The software package: MEGA ver. 10 (Newman et al., 2016) was used for multi-alignments and phylogenetic analysis based on the neighbor-joining method. The obtained sequence was deposited in GenBank, and the accession number of the fungal strain was received.

[1]http://www.ncbi.nlm.nih.gov/BLAST

## Statistical Design and Analysis

All data were the mean of at least triplicates. For optimization of CA production, Design-Expert software (version 12, Stat-Ease, Inc. Minneapolis, MN, United States) was used for constructing the CCD matrix and ANOVA of the data. One-way randomized blocks design was used to construct the experimental scheme of both the germination test and fungal load test of peanuts. CoStat software version 6.4 (CoHort Software, United States) was used to perform the one-way ANOVA and verify the difference between the averages based on Tukey's HSD test. The other data were introduced as mean ± SD. All the statistical hypothesis tests were made at a probability ($P$) value of $\leq 0.05$.

# RESULTS

## Prevalence of Endophytic Fungal in Peanut Seeds

During the present investigation, endophytic fungi were isolated from peanut seeds. Most isolates were recorded the first 5 days of incubation. Twenty-six endophytic fungal species belonging to 11 genera were isolated from collected peanut seed samples (**Figure 1**). Most species belong to Ascomycetes and Deuteromycetes. Of these, the most frequently occurring genera were Fusarium, *Aspergillus*, and *Penicillium*, being 1.20%–10.8%, 2.5%–11.0%, and 13.20%, respectively. *Fusarium* (six species) and Aspergillus (seven species) were the most common genera, representing 23.08% and 26.92% of total fungi, respectively. In this respect, *F. oxysporum* (17.4%) and *F. subglutinans* (11.0%) were the dominant species recovered among all *Fusarium* species, followed by *Fusarium verticillioides* (6.30%), whereas *F. nygamai*, *F. solani*, and *F. incarnatum* were the lowest dominant species. *Penicillium* was the second most prevalent genus of the recovered endophytic fungi. *Aspergillus* spp. ranked third in the frequency in peanut, comprising 26.92% of total fungi. The genus was represented by seven species of which the most dominant species were *Aspergillus terreus*, *A. flavus*, and *Aspergillus sulphureus*, being 10.8%, 9.2%, and 7.0%, respectively, whereas *A. niger*, *A. ochraceus*, *Aspergillus oryzae*, and *Aspergillus nidulans* were the least dominant species. Meanwhile, *Trichoderma* spp. was recorded at moderate percentages (8.60%). Other genera, including *Chaetomium*, *Curvularia*, *Drechslera*, *Macrophomina*, *Mucor*, and *Nigrospora*, were recorded as the lowest endophytes in peanut seeds.

## Cellulolytic Profile of the Endophytic Fungi

To perform the bioconversion of PNR into a valuable biomolecule, two sequential screening tests were performed, the first was a descriptive screening experiment, which was conducted to explore the cellulolytic activity of all the isolated endophytic fungi. For the selection of the endophytic fungus that can degrade PNR, the study was initiated by testing the isolated endophytic fungi for their ability to clarify the medium of microcrystalline cellulose agar plates. The ability of a fungus to form a clear zone around the growing colony was considered as a response to the cellulose-degrading ability of the fungus. *Trichoderma* spp. WKA55 exhibited the highest cellulolytic zone around the fungal growth on the agar plate (**Figure 2**). Accordingly, seven colonies with cellulase activity were obtained. The other isolates were eliminated since they failed to either grow or solubilize cellulose on agar plates. Subsequently, the selected fungi were subjected to further evaluation procedures.

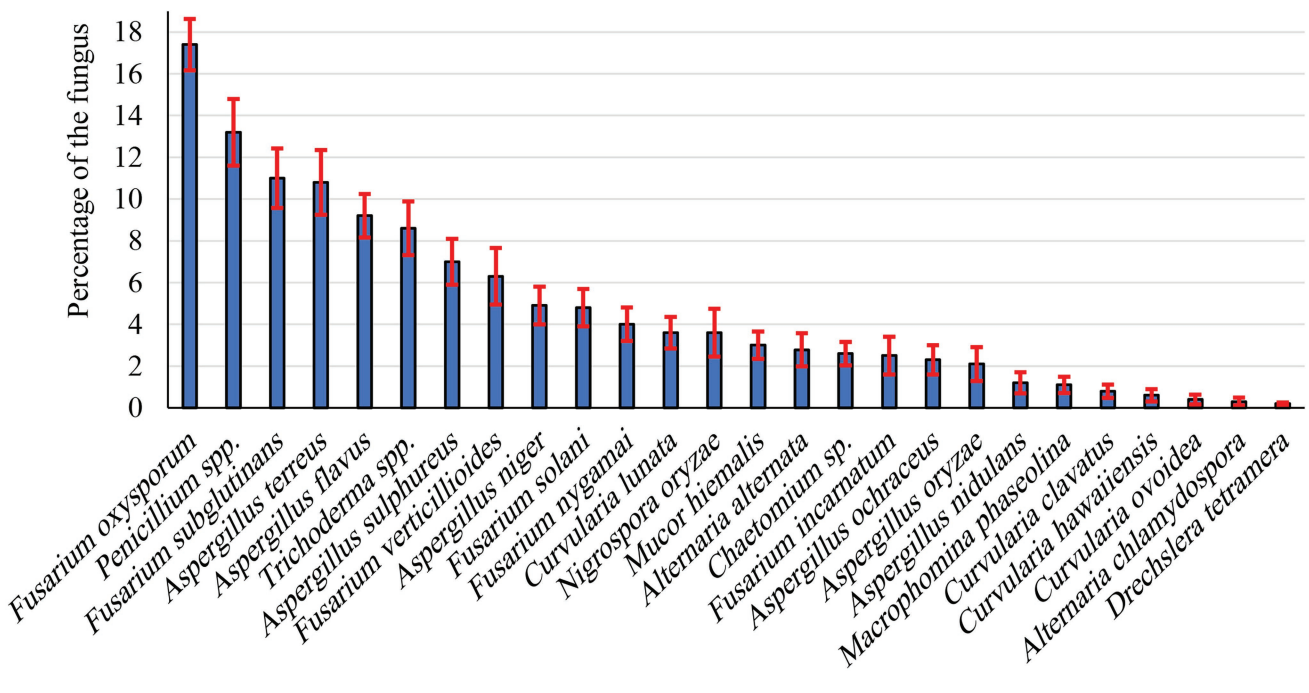



**FIGURE 1** | Prevalence of endophytic fungal species isolated from peanut seeds.

During the second screening test (**Figure 2**), the seven fungal isolates selected based on the aforementioned screening study were undergone to quantitative screening of FPase activity and the release of glucose on the PNR-based medium, to compare the PNR degrading ability of various fungi. Also, of the fungal solubilization of complex TCP was evaluated, such a process is usually coupled with the synthesis of diverse organic acids. All fungi possess FPase activity but differed greatly in the degree of cellulolytic activity, *Trichoderma* sp. WKA55 (6.914 U), *Mucor* sp. WKA29 (6.554 U), and *Aspergillus* sp. WKA69 (5.872 U) recorded the highest activity. Despite Trichoderma sp. WKA55 was the most active cellulolytic fungus in PNR degradation on both descriptive and quantitative tests, the net accumulation of liberated glucose (1.06 mmol/g PNR) was lower compared with the other fungi.

## Time *vis* Biodegradation of PNR by *Trichoderma* sp. WKA55

The profile of PNR biodegradation against 15 days period by *Trichoderma* sp. WKA55 was explored (**Table 1**). It looks that the time course sketch of FPase, CMCase, and xylanase during SSF time was similar and these enzymes are synthesized under similar circumstances. The enzymatic system (FPase, CMCase, and xylanase) associated with and involved in the biodegradation process raised continuously from the binging of fermentation up to the 11th day, then decreased markedly with the progress of the incubation period. Data of protease and PGase showed also a reasonable peak of proteolytic (2.746 U) and pectinolytic activity (0.431 U) after 11 days of fermentation. The activity started from the beginning and lasted until the end of incubation (15 days). Moreover, the filtrate of the fermentation PNR biomass showed positive hydrolysis and liberation of soluble P, which reached its maximum (27.46 μg/g PNR) after 11 days of incubation. The data show also that the maximum biosynthesis of the total organic acids (227.12 mmol/g PNR) by *Trichoderma* sp. WKA55 was reached after 9 days of PNR fermentation, this was associated with an obvious reduction in both post-culture filtrate pH and residual PNR biomass by 3.01% and 37.81%, respectively.

## CCD for Maximization of CA Biosynthesis

The content of organic acids was differentiated in the post-culture filtrate using HPLC, CA was found to be the major

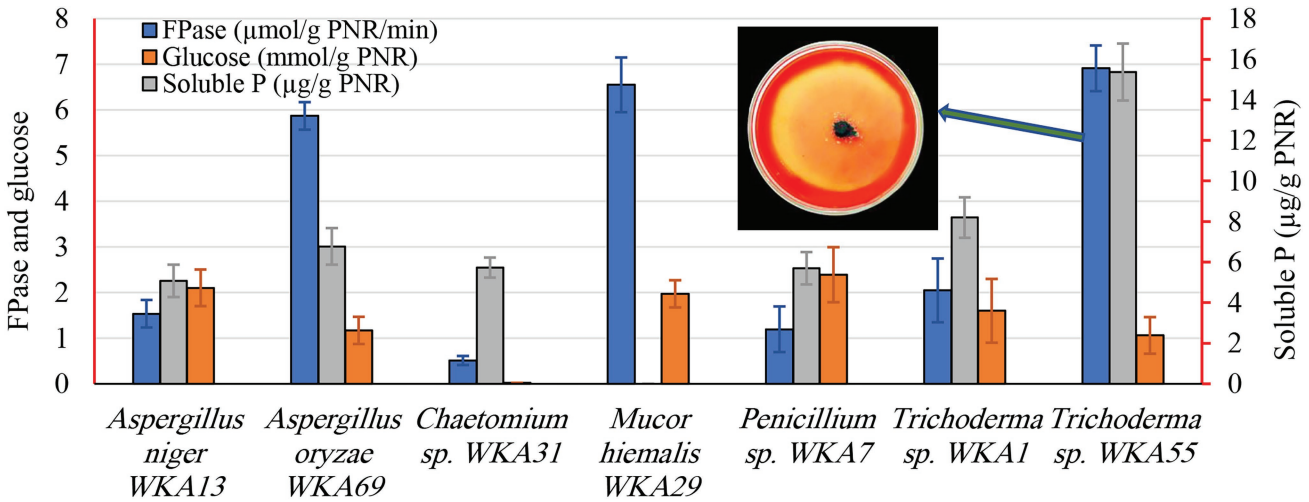



**FIGURE 2 |** Screening of the endophytic fungi for the cellulase, glucose, and the solubilization of tricalcium phosphate (TCP) after growing on a PNR-based fermentation medium. The figure shows the descriptive assay on the agar plate for detection of cellulolytic activity of *Trichoderma* sp. WKA55.

**TABLE 1 |** Time course profile of PNR biodegradation for organic acid production in association with cellulolytic activity and TCP solubilization by *Trichoderma* sp. WKA55 grown on a PNR-based fermentation medium.

| Time, day | Lytic activity (U/min) | | | | | Reduction, % | | Soluble P (μg/g PNR) | OA (mmol/g PNR) |
|---|---|---|---|---|---|---|---|---|---|
| | FPase | CMCase | Xylanase | Protease | PGase | pH | Dry weight | | |
| 1 | 1.42 ± 0.49 | 0.46 ± 0.03 | 0.20 ± 0.08 | 0.12 ± 0.03 | 0.10 ± 0.01 | 0.30 ± 0.02 | 1.52 ± 0.31 | 1.22 ± 0.50 | 12.0 ± 1.7 |
| 3 | 4.36 ± 0.87 | 1.10 ± 0.07 | 2.63 ± 0.89 | 1.36 ± 0.04 | 0.21 ± 0.03 | 1.71 ± 0.27 | 13.00 ± 0.40 | 13.57 ± 2.64 | 75.7 ± 2.6 |
| 5 | 6.12 ± 0.38 | 2.46 ± 0.31 | 13.07 ± 2.36 | 1.44 ± 0.05 | 0.23 ± 0.09 | 2.47 ± 0.47 | 21.18 ± 0.51 | 14.44 ± 3.41 | 94.2 ± 2.5 |
| 7 | 6.91 ± 0.20 | 3.93 ± 0.16 | 14.96 ± 3.44 | 1.61 ± 0.13 | 0.25 ± 0.05 | 3.15 ± 0.99 | 29.03 ± 0.53 | 16.06 ± 3.11 | 170.5 ± 2.7 |
| 9 | 8.07 ± 0.51 | 4.96 ± 0.32 | 42.90 ± 5.59 | 2.25 ± 0.03 | 0.35 ± 0.04 | 3.01 ± 0.20 | 37.81 ± 1.33 | 22.51 ± 3.27 | 227.1 ± 5.3 |
| 11 | 9.07 ± 0.43 | 6.36 ± 0.51 | 60.46 ± 5.91 | 2.75 ± 0.21 | 0.43 ± 0.05 | 3.41 ± 0.43 | 37.96 ± 2.08 | 27.46 ± 4.01 | 216.8 ± 4.5 |
| 13 | 7.48 ± 0.43 | 4.09 ± 0.30 | 23.89 ± 2.76 | 2.54 ± 0.23 | 0.40 ± 0.06 | 3.02 ± 0.48 | 38.26 ± 2.38 | 25.39 ± 3.55 | 202.6 ± 4.1 |
| 15 | 6.91 ± 0.20 | 4.05 ± 0.35 | 13.16 ± 1.46 | 2.44 ± 0.15 | 0.38 ± 0.08 | 2.97 ± 0.52 | 37.62 ± 1.59 | 24.78 ± 1.69 | 198.8 ± 4.8 |

one, signifying about 80% of the detected organic acids. Therefore, the next optimization trial was carried out on CA production on a PNR-based medium under SSF by *Trichoderma* sp. WKA55. The SSF medium of PNR was supplemented with glucose and TCP to support the biosynthesis of *CA*. The incubation lasted for 9 days based on the preceding results. The production of CA was maximized, depending on the full CCD matrix of RSM (**Table 2**).

The produced CA in the 11 experimental runs was determined by HPLC, runs 9, 10, and 11 represent the best combination of medium composition for CA production (average = 5457.4 μg CA/g PNR). These were the replicates of three center points of the experimental design for the tested variables (15 mg/g PNR from glucose and TCP). On the other hand, the fermentation conditions of run 1 excreted the lowest amount of CA (885.5 μg CA/g PNR). The highest value of CA situated nearby the center points of the investigated variables reflected the accuracy and fitness of the selected concentrations of each variable.

## ANOVA of CA Production by CCD

Data recovered from the CCD matrix were statistically evaluated for the effect of various concentrations of glucose and TCP as well as their interaction on CA production (**Table 3**). The ANOVA indicates that the overall model design, all individuals, and quadratics are significant model terms, the only exception was the interaction between both tested variables.

Various model evaluation statistics were calculated, of which the lack-of-fit was found to be not significant (*F*-value = 14.8 and value of $p = 0.0640$). The statistical analysis revealed, also, elevated rates of the coefficient of determination

**TABLE 2 |** Values of CA levels detected by HPLC based on the central composite design (CCD) matrix of the studied variables (glucose and TCP) on the PNR-based medium fermented by *Trichoderma* sp. WKA55.

| Run | Input variables | | | | Output variable |
|---|---|---|---|---|---|
| | Coded unit | | Actual unit (mg/g PNR) | | Citric acid (μg/g PNR ± SD) |
| | Glucose | TCP | Glucose | TCP | |
| 1 | −1 | −1 | 10 | 10 | 885.5 ± 3.5 |
| 2 | −1 | 1 | 10 | 20 | 966.9 ± 6.4 |
| 3 | 1 | −1 | 20 | 10 | 1478.4 ± 28.0 |
| 4 | 1 | 1 | 20 | 20 | 2887.5 ± 14.4 |
| 5 | 0 | −1.414 | 15 | 7.929 | 1322.2 ± 23.1 |
| 6 | 0 | 1.414 | 15 | 22.071 | 2159.3 ± 20.8 |
| 7 | −1.414 | 0 | 7.929 | 15 | 1587.3 ± 11.6 |
| 8 | 1.414 | 0 | 22.071 | 15 | 2413.4 ± 8.1 |
| 9* | 0 | 0 | 15 | 15 | 5555.7 ± 12.1 |
| 10* | 0 | 0 | 15 | 15 | 5357.1 ± 11.7 |
| 11* | 0 | 0 | 15 | 15 | 5459.3 ± 11.6 |

**The middle concentrations.*

**TABLE 3 |** ANOVA of CCD matrix of CA production on the PNR-based medium based supplemented with TCP and glucose by *Trichoderma longibrachiatum* WKA55.

| Source | | Sum of squares | Degree of freedom | Mean square | *F* | *p* |
|---|---|---|---|---|---|---|
| Model | | 33693393.6 | 5 | 6738678.7 | 73.7 | 0.0001 (Significant) |
| Individual | TCP | 894010.6 | 1 | 894010.6 | 9.8 | 0.0261 (Significant) |
| | Glucose | 1694439.7 | 1 | 1694439.7 | 18.5 | 0.0077 (Significant) |
| Interaction (TCP × Glucose) | | 440696.8 | 1 | 440696.8 | 4.8 | 0.0796 (Not significant) |
| Quadratic | $(TCP)^2$ | 21194198.8 | 1 | 21194198.8 | 231.7 | <0.0001 (Significant) |
| | $(Glucose)^2$ | 18449301.4 | 1 | 18449301.4 | 201.7 | <0.0001 (Significant) |
| Residual | | 457363.7 | 5 | 91472.7 | | |
| lack-of-fit | | 437637.1 | 3 | 145879.0 | 14.8 | 0.0640 (Not significant) |
| Pure error | | 19726.6 | 2 | 9863.3 | | |
| Corrected total | | 34150757.3 | 10 | | | |
| Model evaluation statistics | | | | | | |
| Determination coefficient ($R^2$) | | | | 0.9866 | | |
| Adjusted-$R^2$ | | | | 0.9732 | | |
| Predicted-$R^2$ | | | | 0.9076 | | |
| Adequate precision, % | | | | 19.463 | | |
| The predicted residual sum of squares | | | | 164.20 | | |

$R^2$ (0.9866) and adjusted-$R^2$ (0.9732), similarly, predicted $R^2$ recorded a high value, being 0.9076. The adequate precision ratio was high and measured to be 19.463. The predicted residual sum of squares (PRESS) recorded a reasonably small value of 164.20.

## Contour Plot and Model Evaluation

The surface plot of CA was constructed in the form of a contour graph to explore the association between the two tested input variables (glucose and TCP) and the CA response (**Figure 3A**). As is shown, the continual increase in CA biosynthesis was noted when both concentrations of glucose and TCP amplified, reaching the maximum of CA at the center points (middle concentrations) of both variables, then declined with the higher concentrations afterward.

For model evaluation, the validity checking tests were performed to verify the cogency of the attained polynomial model and detect a value, or group of values, that are not easily predicted by the model. The graph of the Box-Cox plot (**Figure 3B**) declares that the current lambda ($\lambda$) was equal to 1.0, the best $\lambda$ was 1.81, whereas the CI was between low (0.70) and high (2.78). In consent with the Box-Cox plot, the plot of actual values versus the estimated (predicted) response values of CA was generated (**Figure 3C**). As could be seen all data points scatter symmetrically and fragmented evenly along the 45-degree line.

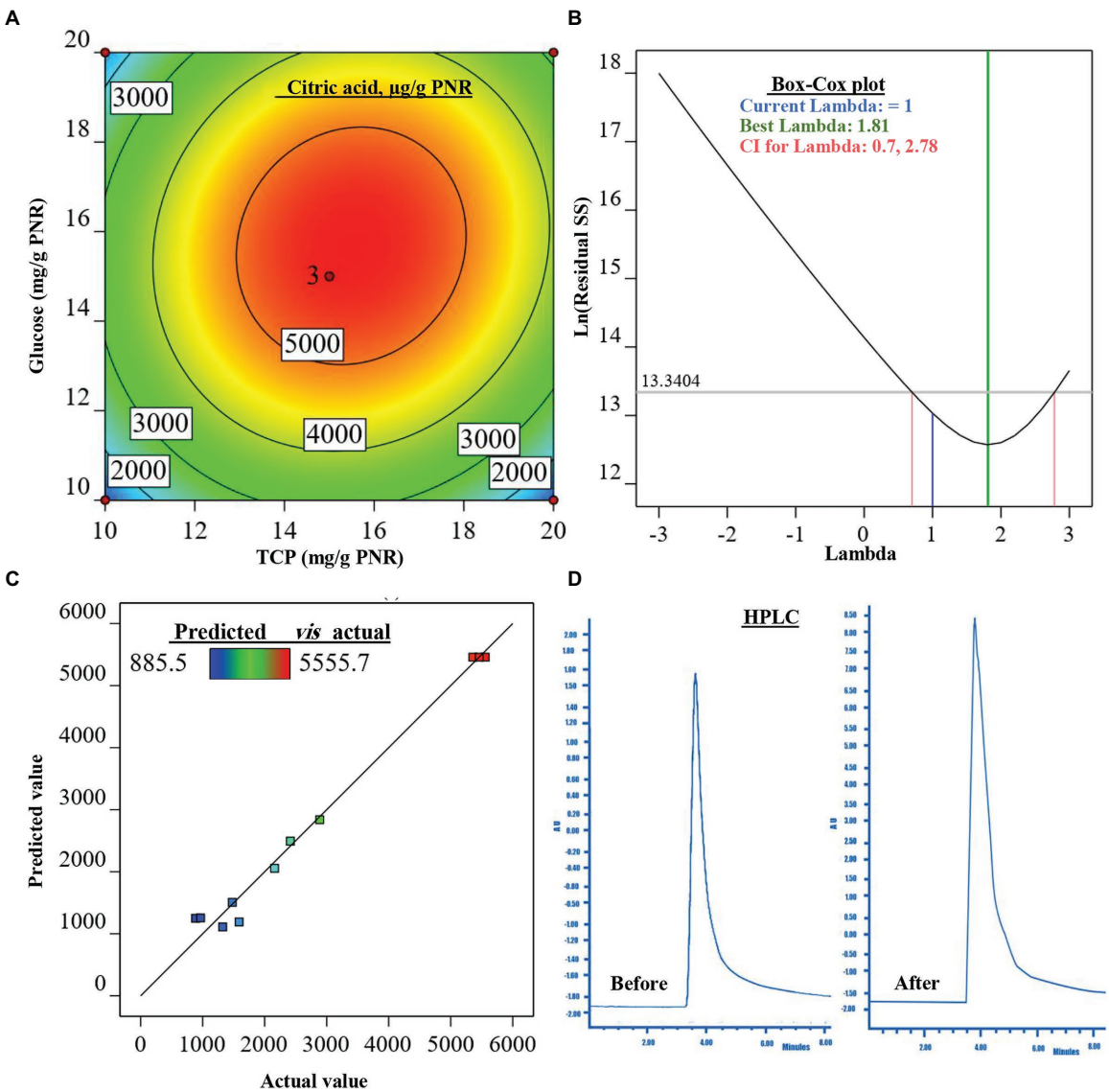



**FIGURE 3 |** The figure shows the contour plot of citric acid (CA; µg/g PNR) biosynthesis by *Trichoderma* sp. WKA55 as a response to the combination of various concentrations of glucose and TCP on a PNR-based medium **(A)**, the plot of Box-Cox for power transforms **(B)**, predicted versus actual data of CA **(C)**, and **(D)** the diagram of the detected CA by HPLC, before and after the optimization process.

### Modeling of CA Production

The preceding statistical tools approved and confirmed a satisfactory adjustment of the quadratic model to the investigational data. This model was used to estimate the optimum level of each tested variable (glucose and TCP), as well as the estimated amount of *CA*. The equation that describes the theoretical relationship between the two tested variables that maximize CA production in actual units was generated in terms of the second-order polynomial model to be:

$$\begin{aligned} \mathrm{CA}\left(\mu g / \mathrm{g\ PNR}\right) = & -27642.09 + 2061.89 \\ & \times \mathrm{glucose} + 2192.47 \times \mathrm{TCP} + 13.28 \times \mathrm{glucose} \\ & \times \mathrm{TCP} - 72.30 \times \mathrm{Glucose}^2 - 77.49 \times \mathrm{TCP}^2 \end{aligned} \quad (5)$$

The mathematical solution of the previous equation led to the theoretical maximum value of CA = 5505.1 μg/g PNR, this could be achieved when fermenting 1 g of PNR with 15.683 mg glucose and 15.490 mg TCP. To confirm the aptness of this mathematical hypothesis, these theoretical concentrations were laboratory validated at the same fermentation conditions. The actually-resulted yield of CA at the laboratory was found to be 5515.9 μg/g PNR. The diagram (**Figure 3D**) of the HPLC analysis of CA before (898.1 μg/g PNR) and after optimization was introduced, showing a 6.14-fold increase in CA biosynthesis.

## Evaluation of Crude CA

### CA Against Mycotoxinogenic Fungi

The biological potential of CA was evaluated against the growth and spore germination of three toxinogenic fungi (*A. flavus*, *A. ochraceus*, and *F. oxysporum*). Data (**Figure 4**) of the MIC show the inhibitory impact on the growth of the tested fungi as the action of CA, even at the lowest tested level (10%) as compared to the controls (0%). Generally, the inhibitory effect was raised with higher concentrations. In this respect, CA at 20% reduced the growth of *A. flavus* down to 31.09%, *A. ochraceus* to 73.51%, and *F. oxysporum* to 39.46%. By increasing the CA concentrations up to 30%, *A. flavus* growth was reduced to 43.67%, and a complete reduction of mycelial growth was observed against *A. ochraceus* and *F. oxysporum*. Spore germination was found to have the same reduction pattern, in which a significant reduction in spore germination of the tested fungi due to CA was noticed, even in the lowest concentrations (10%). Higher concentrations up to 30%, strongly reduced the spore germination of *A. flavus* (95.22%) and inhibited the spore germination of *A. ochraceus* and *F. oxysporum* as compared to the control's treatments. At 40%, the fungicidal activity appeared against growth and spore germination of all tested fungi.

### Seed-Borne Fungi Associated With Peanut *vis* CA

The effect of CA on the frequency of seed-borne fungi on peanuts is evaluated (**Table 4**). Thirteen fungal species belonging to five genera were recovered from the tested samples. *A. niger*, *A. flavus*, *Penicillium* spp., and *Rhizopus stolonifer* were the most abundant, being 28.7%, 21.94%, 20.71%, and 19.22%, respectively. *Fusarium oxysporum* was the most dominant species among all *Fusarium* spp. (8.94%), followed by *F. solani* (4.39%), whereas *F. incarnatum* and *F. verticillioides* were the least dominant (2.79% and 1.48%, respectively). The data declared the diverse effect of CA on the frequency of seed-borne pathogenic fungi of peanut. In this respect, treated seeds by each of CA or the fungicide led to a significant decrease of *A. flavus*, *F. incarnatum*, and *R. stolonifera*. On the other side, both CA and chemical treatments led to the complete absence of *A. ochraceus*, *A. terres*, *F. oxysporum*, and *F. solani*. On the

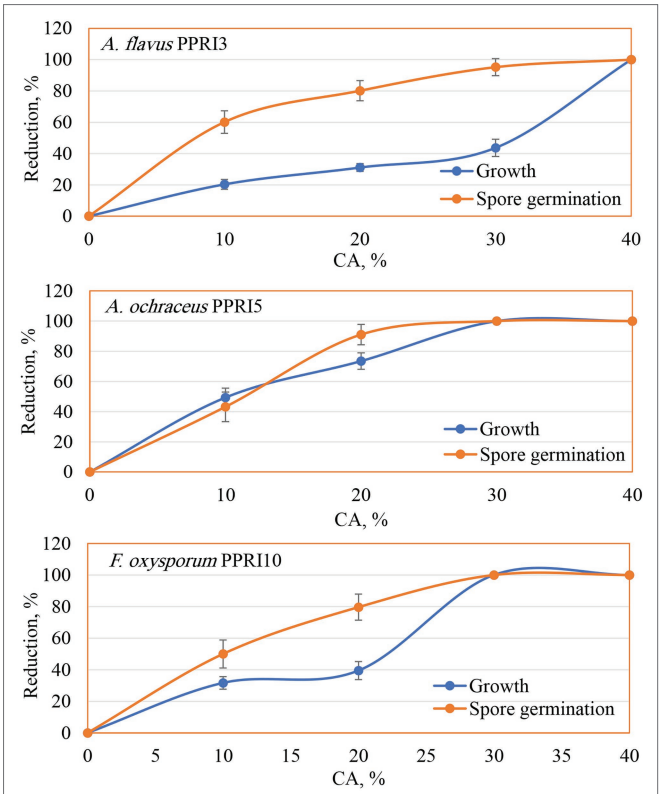



**FIGURE 4 |** Reduction of growth and spore germination of toxinogenic fungi as affected by various crude CA concentrations.

**TABLE 4 |** Mean initial values (%) of the incidence of fungi in peanut seeds as a response to CA treatment.

| Fungus | Control | CA | Fungicide |
|---|---|---|---|
| *Aspergillus flavus* | 21.94[a] | 5.00[a,b] | 1.670[b] |
| *Aspergillus niger* | 28.70[a] | 10.72[b] | 5.470[c] |
| *Aspergillus ochraceus* | 2.340[a] | 0.00[b] | 0.00[b] |
| *Aspergillus tamarii* | 2.280[a] | 1.67[a] | 0.00[a] |
| *Aspergillus terres* | 6.670[a] | 0.00[b] | 0.00[b] |
| *Fusarium incarnatum* | 2.790[a] | 0.72[b] | 0.17[b] |
| *Fusarium oxysporum* | 8.940[a] | 0.00[b] | 0.00[b] |
| *Fusarium solani* | 4.390[a] | 0.00[b] | 0.00[b] |
| *Fusarium verticillioides* | 1.480[a] | 0.00[b] | 0.00[b] |
| *Penicillium* spp. | 20.71[a] | 7.64[a] | 2.570[a] |
| *Phoma* sp. | 1.110[a] | 0.00[a] | 0.00[a] |
| *Rhizopus stolonifer* | 19.22[a] | 4.67[b] | 0.830[b] |

*Different letter(s) within a raw indicates significant differences.*

contrary, the frequency of *A. tamarii* and *Penicillium* spp. was not affected by any of the tested CA or chemical treatments.

### Crude CA *Trichoderma* sp. WKA55 CA *vis* Germination

The influence of crude CA on the germination of peanut seeds and seedling parameters was also evaluated (**Figures 5, 6A–C**). The CA was effective in improving the germination percentage (91%) with no significant difference with the chemical fungicide (85%) as compared to the control treatment (70.52%). Compared with the untreated control, seed treatment with each CA and chemical fungicide showed a significant decrease in the percentages of abnormal seedlings. The vigor index was also affected by the application of CA, which increased the vigor

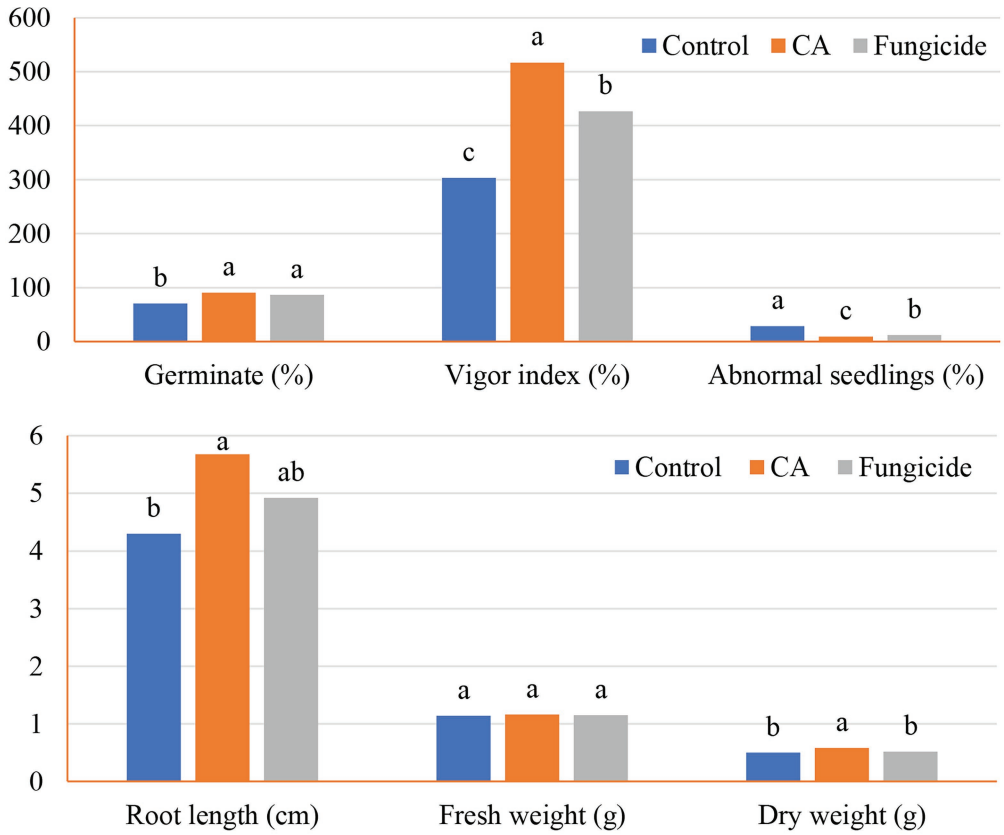



**FIGURE 5 |** Germination and seedlings characteristics of the tested peanut seeds in accordance with CA treatment.


**FIGURE 6 |** Response of the germinated peanut seeds to various treatments after 14 days. Where, CA treatment **(A)**, fungicide treatment **(B)**, and controls **(C)**, showing the various promoting responses.

index in comparison with the control and showed a comparable trend as in the other parameters. Seeds treated with CA showed a 70.45% vigor index compared with the control of peanut seedlings. Chemical seed treatment came in the second rank, recording 40.61%. Regarding seedling features, the positive effect of the tested CA treatment extended to the root length and dry weight characters of the peanut seedlings, being 36.28% and 16%, respectively, in comparison with the control. On the other hand, no significant variation between CA, chemical fungicide, and control in their effect on fresh weight character.

## GC-MS Analysis of *Trichoderma* sp. WKA55 Crude Extract

In order to explore the other secondary metabolites in the fungal filtrate produced by the endophytic *Trichoderma* sp. WKA55, the GC-MS was performed (**Table 5**; **Figure 7**). A total of 20 unique compounds were detected based on mass spectral properties with area sum % less than 0.05% and up to 65.04% and retention time 5.2 and 33.16 min. These molecules fell into classes of acids, aldehydes, alcohols, ketones, terpenoids, pyrones (lactones), and furanes. The major components in the endophytic filtrates of

**TABLE 5** | GC-MS analysis of the endophytic *Trichoderma* sp. WKA55 filtrate, sowing the various compounds detected in the PNR hydrolysate.

| Peak | RT | Name | RI | Formula | Area | Area sum % | MW |
|---|---|---|---|---|---|---|---|
| 1 | 5.195 | Methyl alcohol | 0 | $CH_4O$ | 8,584 | 0.477 | 32 |
| 2 | 5.427 | Propanoic acid, 2-oxo-, ethyl ester | 822 | $C_5H_8O_3$ | 20,162 | 1.120 | 116 |
| 3 | 10.607 | Bicyclo[3.1.0]hexane, 4-methylene-1-(1-methylethyl)- | 897 | $C_{10}H_{16}$ | 29,735 | 1.651 | 136 |
| 4 | 12.259 | 1,3,8-p-Menthatriene | 1,029 | $C_{10}H_{14}$ | 154,791 | 8.596 | 134 |
| 5 | 13.507 | 1,4-Cyclohexadiene, 1-methyl-4-(1-methylethyl)- | 998 | $C_{10}H_{16}$ | 39,873 | 2.214 | 136 |
| 6 | 17.599 | Sorbic acid vinyl ester | 990 | $C_8H_{10}O_2$ | 17,422 | 0.968 | 138 |
| 7 | 17.758 | Cyclohexanol, 5-methyl-2-(1-methylethyl)-, (1.alpha.,2.beta.,5.alpha.)- | 1,164 | $C_{10}H_{20}O$ | 30,883 | 1.715 | 156 |
| 8 | 17.964 | 3-Cyclohexen-1-ol, 4-methyl-1-(1-methylethyl)- | 1,137 | $C_{10}H_{18}O$ | 23,848 | 1.324 | 154 |
| 9 | * | Benzeneacetic acid, 2-methoxy-, methyl ester | 1,349 | $C_{10}H_{12}O_3$ | * | * | 180 |
| 10 | * | Estragole | 1,172 | $C_{10}H_{12}O$ | * | * | 148 |
| 11 | * | Isoflurane | 252 | $C_3H_2ClF_5O$ | * | * | 184 |
| 12 | 20.247 | Benzaldehyde, 4-(1-methylethyl)- | 1,230 | $C_{10}H_{12}O$ | 186,796 | 10.373 | 148 |
| 13 | 20.365 | 2-Cyclohexen-1-one, 2-methyl-5-(1-methylethenyl)-, (R)- | 1,190 | $C_{10}H_{14}O$ | 82,803 | 4.598 | 150 |
| 14 | 21.308 | 1-Triazene, 3,3-dimethyl-1-phenyl- | 1,112 | $C_8H_{11}N_3$ | 4,511 | 0.251 | 149 |
| 15 | 21.963 | Phenol, 2-methyl-5-(1-methylethyl)- | 1,262 | $C_{10}H_{14}O$ | 1,171,140 | 65.038 | 150 |
| 16 | * | Phenol, 2-ethyl-4,5-dimethyl- | 1,340 | $C_{10}H_{14}O$ | * | * | 150 |
| 17 | 26.401 | 1,3,6,10-Dodecatetraene, 3,7,11-trimethyl-, (Z,E)- | 1,458 | $C_{15}H_{24}$ | 14,573 | 0.809 | 204 |
| 18 | 31.587 | Caryophyllene oxide | 1,507 | $C_{15}H_{24}O$ | 50,159 | 2.786 | 220 |
| 19 | 33.156 | 1-Naphthalenol, 1,2,3,4,4a,7,8,8a-octahydro-1,6-dimethyl-4-(1-methylethyl)-, | 1,580 | $C_{15}H_{26}O$ | 23,912 | 1.328 | 222 |
| 20 | * | Pentanal, 2,4-dimethyl- | 777 | $C_7H_{14}O$ | * | * | 114 |

*RT, retention time; *, the compound detected at less than 0.05%; RI, retention index; and MW, molecular weight.*

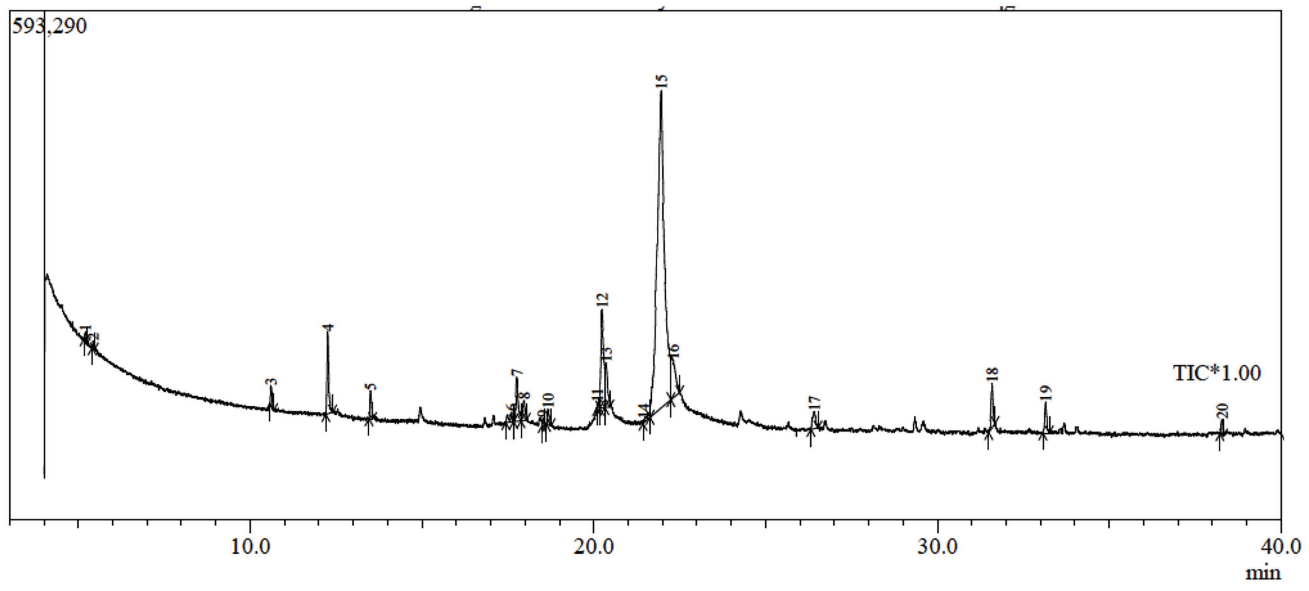



**FIGURE 7** | GC-MS chromatogram of fermented peanut plant residual (PNR) by the endophytic *Trichoderma* sp. WKA55, displaying the peaks of various components.

WKA55 strain were phenol, 2-methyl-5-(1-methylethyl)- (65.04%), followed by benzaldehyde, 4-(1-methylethyl)- (10.37%), 1,3,8-p-menthatriene (8.60%) and 2-cyclohexen-1-one, 2-methyl-5-(1-methylethenyl)-, and (R)- (4.60%). Also, caryophyllene oxide (2.79%), 1,4-cyclohexadiene, 1-methyl-4-(1-methylethyl)- (2.21%), cyclohexanol, 5-methyl-2-(1-methylethyl)-, (1.alpha.,2.beta.,5.al) (1.72%), bicyclo[3.1.0]hexane, 4-methylene-1-(1-methylethyl)- (1.65%), 1-naphthalenol, 1,2,3,4,4a,7,8,8a-octahydro-1,6-dimethyl-4-(1-methylethyl)-(1.33%), 3-cyclohexen-1-ol, 4-methyl-1-(1-methylethyl)-(1.32%), and propanoic acid, 2-oxo-, ethyl ester (1.12%) were found in moderate levels, while sorbic acid vinyl ester (0.98%), 1,3,6,10-dodecatetraene, 3,7,11-trimethyl-, (Z,E)- (0.81%), methyl alcohol (0.48%), and 1-triazene, 3,3-dimethyl-1-phenyl- (0.25%) were found to be the minor metabolites.

## Molecular Identification of *Trichoderma* sp. WKA55

Based on the preceding trials, the isolated fungal strain (*Trichoderma* sp. WKA55) was molecularly identified based on 18s rRNA, as a perfect identification tool. From BLAST results, strain WKA55 displayed high similarity, up to 99.65%, with the previously identified *T. longibrachiatum* strains (MT218356.1, MT102396.1, and MN511324.1) on the GenBank. The constructed phylogenetic tree of *Trichoderma* sp. WKA55 was depicted (**Figure 8**), which comes in line with the previous morphological identification. The GenBank accession number of the present fungal strain (*T. longibrachiatum* WKA55) was MZ014020.1.

## DISCUSSION

In the current study, a survey of fungal endophytes in of the peanut seeds was managed. The survey yielded 26 endophytic species belonging to 11 genera, of which *Trichoderma* spp., an endophyte bioagent was recorded at moderate percentages (8.60%).

Most of the previous endophytic fungi are isolated from parts other than seeds, and little is known about the endosymbionts of seeds. Most of the endophytic isolates belong to Ascomycota and Deuteromycetes. Previously, 72 species belonging to 21 genera of endophytic fungi were isolated from the roots of three leguminous plants (peanut, alfalfa, and broad bean). In this work, *Fusarium*, *Penicillium*, and *Aspergillus* were the commonly isolated genera, of which *F. solani*, *F. subglutinans*, *F. oxysporum*, *P. duclauxii*, *P. funiculosum*, *A. tubingensis,* and *A. flavus* were the most prevalent (El-Maghraby et al., 2013). Similarly, 17 and 16 species belonging to 11 genera were isolated from common bean and cowpea seeds, respectively

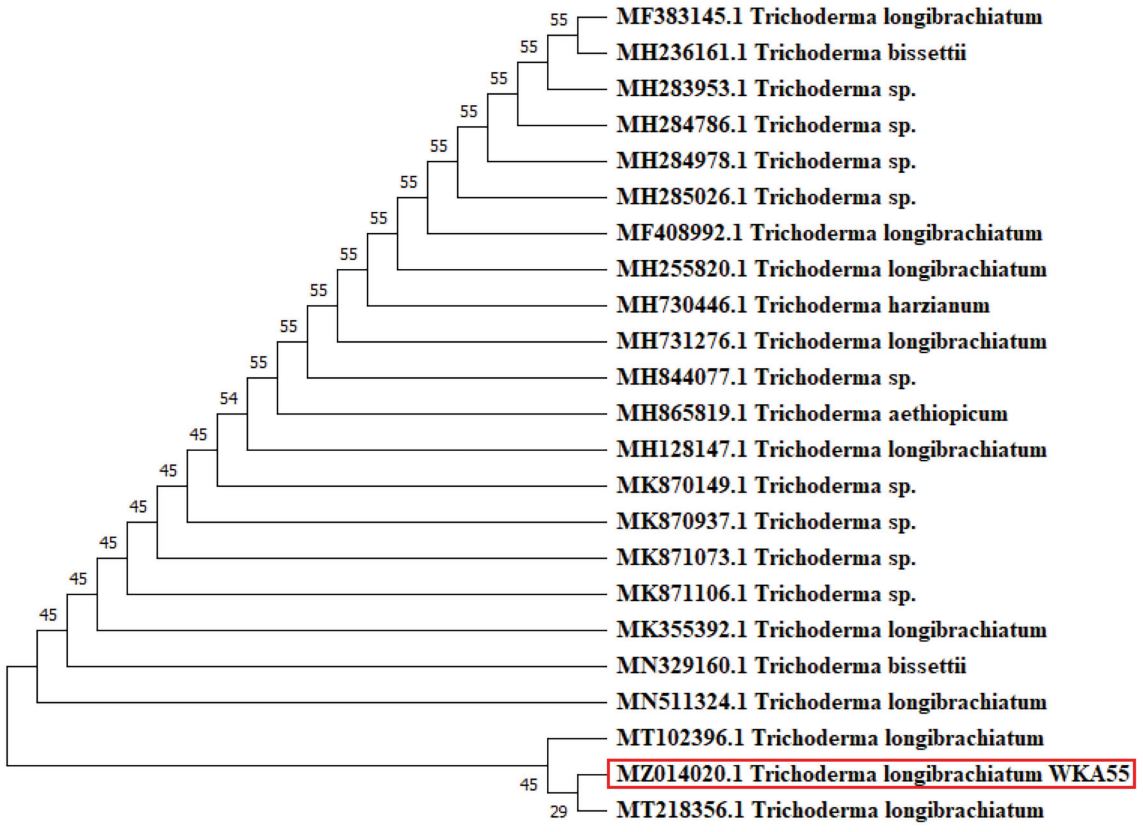



**FIGURE 8 |** Molecular phylogenetic tree of the sequence of 18s rRNA, showing the position of *Trichoderma* sp. strain WKA55 (surrounded with the red rectangle) within the closely related sequences in GenBank.

(El-Samawaty et al., 2014). However, these studies did not attain fungal strains with significant features. On the other side, isolation of diverse fungal genera from peanut seeds in the current investigation may be back to the nutrient composition of such pulses, which resemble a nutritious substrate for fungal growth and spore establishment (El-Samawaty et al., 2014). The high incidence of some pathogenic fungal genera in peanut seeds may be back to the aggressivity of these pathogens in the field, and later under improper storage conditions play a life-threatening role in humans and cause several damages in pulses as well (Santos et al., 2016; El-Zefzaf, 2020).

The current study reported endophytic fungi with lytic activity on plant biomass residuals. The biodegradation of such residuals comprises the first step in the bioconversion process. Earlier, two *Aspergillus* spp., two *Penicillium* spp., and one *Trichoderma* sp. were isolated from rice straw and showed high cellulase activity (Saber et al., 2010). Fungal ability to decompose the crystalline cellulosic biomass depends upon the presence of complementary cellulases profiles (endoglucanase, exo-glucanase, and β-glucosidase) in sufficient amounts (Al-Askar et al., 2021b). The fungal cellulolytic activity could be considered indicators of the degradation of PNR by the tested fungi, leading to the release of pentoses and hexoses monomers as the main units, representing the cornerstone for the formation of various organic molecules, including organic acids.

Seven selected endophytic fungi were further investigated for cellulosic activity and complex phosphate solubilization ability. The released glucose as a function of cellulase activity in the present work is lower than stated earlier, suggesting the speedy and/or direct biotransformation of the liberated glucose into biomolecules such as value-added organic acids. This type of bioprocessing that accompanied by low free glucose in the fermentation medium could be considered the initial indicator for the direction of production of valuable compounds. FPase is the overall cellulolytic activity and plays a critical role in the bioconversion of lignocellulosic into fermentable sugars (Saber et al., 2015; Elsayed et al., 2021). Accordingly, this study could be considered a pioneer regarding the screening of the lytic enzymes for the selection of endophytic fungus. What is more, these fungi could be a new source of enzymes with wide biotechnological applications.

The maximum liberation of free P was achieved by *Trichoderma* sp. WKA55 (15.37 μg/g PNR), confirming the competence of such isolate as a potent solubilizer of complex phosphate. Several filamentous fungi, mainly black Aspergilli, some species of *Penicillium*, and rarely *Trichoderma* spp. were reported as complex phosphate solubilizers, the solubilization of complex phosphate and release soluble P is usually linked with the generation of organic acids, reduction of medium pH, and/or production of various organic molecules (Max et al., 2010; Show et al., 2015; Elsayed et al., 2021).

Both cellulases (FPase and CMCase) of *Trichoderma* sp. WKA55 showed obvious activity during the 15 days of fermentation. This, in turn, indicates that both enzymes are induced and could be also constitutive, that is to say, it could be found associated with the fungus, even in the absence of their inducer substrate, also indicate that *Trichoderma* sp. WKA55 has a complete cellulolytic system that can completely degrade the majority of PNR cellulose. In this connection, hydrolysis of PNR cellulose requires (i) endoglucanase (1,4-β-D-glucan-4-glucanohydrolases, EC 3.2.1.4) that randomly cleaves the internal β-1,4-glucosidic bonds, (ii) exoglucanases (1,4-β-D-glucan glucanohydrolases, EC 3.2.1.74 and 1,4-β-D-glucan cellobiohydrolases, EC 3.2.1.91), which act on the reducing or nonreducing terminals of cellulose chains, liberating glucose and/or cellobiose, exoglucanases can also catalyze the peeling of the microcrystalline structure on cellulose chains, and (iii) β-glucosidase (β-glucoside glucohydrolases, EC 3.2.1.21) that releases glucose monomers from cellodextrins and cellobiose (Saber et al., 2015; Elsayed et al., 2021).

Marked xylanase activity along the 15 days incubation period was noticed by *Trichoderma* sp. WKA55, suggesting the ability of such fungus to catalyze xylan of hemicellulose part of PNR into xylose monomers. The continual secretion of such xylanase along the fermentation period is an indication of the induced and constitutive nature of the enzyme. The complete hydrolysis of xylan begins with endoxylanase (1, 4-β-D-xylan xylanohydrolase; EC 3.2.1.8), which cleaves the internal glycosidic bonds, accumulating xylooligomers of β-D-xylopyranosyl that may inhibit the endoxylanase, then xylosidases (1,4-β-D-xylan xylohydrolase; EC 3.2.1.37), hydrolyzes the xylooligomers, thus removing the cause of inhibition and releasing D-xylose from the xylooligomers; finally, acetylxylan esterase (EC 3.1.1.6) eliminates the O-acetyl groups from positions 2 and/or 3 on the β-D-xylopyranosyl residues (Abdelwahed et al., 2011; Elsayed et al., 2021).

Several kinds of proteases (such as peptidyl/peptide hydrolases, EC 3.4.21-24, and 99) can degrade nitrogenous complexes, and hydrolysis the peptide bonds in a protein fragment (Negm et al., 2021). For a long time, the function of proteases was solely dedicated to protein digestion as a feeding source for the microorganism. Next, it becomes clear that the catalytic action of protease can have a wide range of very specific functions, besides their critical role in the pathogenesis. Six types of proteases were classified based on the catalytic activity, i.e., cysteine-proteases, threonine-proteases, glumatic-proteases, aspartic-proteases, metalloproteases, and serine-proteases, the latter three are the most abundant protease types (Barrett et al., 2012).

Pectin is another complex polysaccharide (composed of α-1, 4-linked D-galacturonic acid) located mainly in the middle lamella, representing the intercellular cement, and helping to bind cells together. Also, in the cell wall pectin forms an amorphous gel, filling the spaces between the cellulose microfibrils (Jacob et al., 2008). Polygalacturonases (EC 3.2.1.15) degrade pectic substances and split the pectin chain by adding a water molecule and breaking the linkage between two galacturonan molecules. Pectin lyases (EC 4.2.2.2) are another enzyme that split the chain by removing a molecule of water from the linkage, releasing unsaturated double-bond products. The synergistic action of both enzymes breaks the pectin chain and releases a single unit of galacturonan (Sathiyaraj et al., 2011), leading to weakening

of the cell walls and tissue maceration (liquefaction) thus development of many diseases, particularly those characterized by soft-rotting of tissues. This enzyme system is involved in many phytopathogens that invade the plant tissue (Rahman and Punja, 2005; Al-Askar et al., 2021a).

PNR did not receive any pretreatment before the fermentation, suggesting that *Trichoderma* sp. WKA55 potentially has complete lytic activities (cellulolytic, xylanolytic, proteolytic, and pectinolytic), which enabled the fungus to penetrate and degrade the complex structure of PNR. Reduction in residual PNR biomass after fermentation could be considered a shred of evidence for the transformation of the fermented PNR into other organic molecules. Another, the scenario of pH during fermentation is very important, especially for CA production, since spore germination requires pH higher than 5; then, the absorption of ammonia in the medium by germinating spores causes the release of protons, this, is in turn, lowers the pH and improves the generation of CA, and the low pH value during the fermentation process inhibits the production of unwanted organic acids (Max et al., 2010).

It was better to manage the minimization of the bioconversion process into one direst step by supplementing the fermentation medium with a complex phosphate, e.g., tricalcium phosphate, as a sole phosphorus source. This urged the microorganism to subtilize complex phosphate for getting the free phosphorus required for various physiological activities in microbial cells, such as the formation of nucleic acids and the generation of ATP. This directed microbial solubilization of complex phosphate is usually convoyed by and attributed to the formation of organic acids, which are also stated as the end-molecules of microbial decomposition of cellulose (Kumari et al., 2008; Al-Askar et al., 2021b). That is why cellulose-degrading enzymes and complex phosphate are prerequested in the organic acid production medium.

Three microbial mechanisms were reported for solubilization of complex phosphate; (1) the catalytic action of phytase and/or acid phosphatase, (2) the secretion of organic acids; mainly citric, oxalic, and succinic acids, and/or (3) the drop in pH (Saber et al., 2010; Elsayed et al., 2021). It has been stated that the gradual reduction in pH during biofermentation led to the formation of organic acids; therefore, the presence of complex phosphate as the only source of P in SSF medium induced the fungus toward the assimilation, solubilization, and utilization of complex phosphate, by the secretion of enzymes (phytase and/or acid phosphatase) and/or the production of organic acids (Kumari et al., 2008; Max et al., 2010; Elsayed et al., 2021), the latter is in charge for lowering the pH.

The hydrolytic enzymes reached their maximum on the 11th day of fermentation, but the production of organic acids was maximized on the 9th day so, 9 day was considered as the chosen biofermentation period. However, this period is relatively accepted in relation to the expected valuable products, and further, is comparable with 28 days fermentation for biodegradation of wheat straw and cellulase production under SSF (Dinis et al., 2009) and 4 weeks on wheat and rice straws (Saber et al., 2010).

However, plant biomass residuals, such as PNR, are a suitable raw material for fermentation only if the microorganism has the enzymatic system required for fermentation and catalyze effectively at low pH values (Max et al., 2010), this is what exactly happened by the present *Trichoderma* sp. WKA55. Nevertheless, the previous tests could conclude two approaches: one of them is an indirect, represented by the cellulolytic activity, reduction of residual PNR biomass, and solubilization of TCP by the tested fungal strain; the other direct approach is the biosynthesis of organic acids.

The current endophytic fungus (*Trichoderma* sp. WKA55) might be a novel source of hydrolytic enzymes of broad biotechnological uses. Obtaining fungal strain that has a complementary degradation system of plant biomass residuals and, at the same time, solubilizing the complex phosphate could be managed genetically, but isolation from the native environment may be better for adaptation and succession when incubated with a medium containing plant biomass residual. Because of *Trichoderma* sp. WKA55 has met these hypotheses, it was chosen for identification based on the morphological structures of vegetative mycelia and light microscope investigation, in addition to the molecular identification.

Perfectly, the present candidate fungus was proved to have the complete and complementary catalytic action to manage the transformation of PNR into other molecules, i.e., organic acids. A primary test using HPLC declares CA as the major detected organic acid in the fungal filtrate. Therefore, an optimization trial using CCD was performed, taking into account the maximization of CA biosynthesis using PNR as a fermentation substrate.

The CCD of RSM procedure generates the maximum amount of knowledge from the minimum number of experiments. This is an important issue, considering the cost of HPLC analysis. In contrast to the one-variable-at-a-time method, RSM can find the true optimum conditions, and interactions of the bioprocess would have been made available. What is more, all the experimental trials are carried out simultaneously; thus, the results are obtained rapidly (Saber et al., 2015).

Once applying the CCD, it was found to support the fermentation medium with a low level of glucose to support the biosynthesis of *CA*. As it is already known that glucose has a stimulatory action on CA production (Max et al., 2010). Besides reducing the overall process cost, the application of RSM in biofermentation technology maximizes the target yields. Herein, PNR was applied as one support-substrate phase, i.e., solid medium support for SSF, and at the same time, carbon and nitrogen sources for CA production. This is expected to be friendly for both coasts and the environment. The highest value of CA has located nearby the design center points of the examined variables, returning the precision and fitness of the selected concentrations of both variables.

The significance of model terms through ANOVA indicates the importance of the parameters in the experimental design for CA production. Model evaluation statistics were calculated to assess the model, the lack-of-fit was calculated to give information about the suitability of the chosen variables,

consequently, the appropriateness of the obtained data to create a new model.

The regression model exhibits a significant lack-of-fit when it fails to designate the functional relationship between the experimental input factors and the response output variable(s). Lack-of-fit can occur if individuals, interactions, and/or quadratic terms from the model are not included in the final model. It can also occur if several unusually large residuals resulted when fitting the model. Fortunately, the present lack-of-fit of the model was not significant; therefore, the model is fitted and reflects a good parameter.

$R^2$ is a diagnostic tool, that ranges from 0 to 1, that measures the exactness of the model, the higher value, the greater the association between the variables and target response (CA). High values of the three kinds of $R^2$ in the current work indicate a high relationship between the experimental parameters (glucose and TCP) and the response (CA). Similarly, high predicted $R^2$ indicates how the high accuracy of the model to forecast the new observations, the higher the predicted $R^2$ value, the greater the predictive capability of the model. Accordingly, the current model could help in the prediction and estimation of CA from any given concentration from glucose and TCP with 90.76% accuracy.

The adequate precision that measures signal-to-noise ratio was found to be 19.463; however, a value greater than 4 is desired. The high ratio of the present model indicates an adequate signal. Thus, the models can be used to predict the values of responses and can be used to predict CA values along with the design space; this is a critical requirement for a model together with the non-significant lack-of-fit, as the fitting of the model is required.

PRESS is another statistical parameter that is usually used for evaluating the model and calculation of the predicted $R^2$. The lesser the PRESS value, the superior the model's forecasting capability. Both predicted $R^2$ and PRESS minimize the overfitting of the model (misleading regression coefficients) by preventing random error or noise.

The contour plot showed the maximum of CA at the middle concentrations of both glucose and TCP, then declined with the higher concentrations; this, in turn, indicates the accuracy of the tested variables and their concentrations as well.

Regarding the validity checking tests, the Box-Cox plot for power transforms displayed the best value of $\lambda$ value between the points of low and high C.I., indicating that the current model predicts and reflects efficiently the real system and thus, there was no need for data transformation. Plotting the actual vis, the predicted values showed that the data points divided evenly along the 45-degree line, indicating that all values can be easily predicted by the model. Such a scattering pattern around a 45-degree diagonal line approves an acceptable correlation between the predictive and experimental values of *CA*. This is the ideal distribution of the data points, which fulfill the optimum fitness of the model. This, in turn, indicates that all run points could be precisely predicted by the model.

Laboratory validation of the fermentation conditions, based on the CCD modeling process, revealed that the variance between the theoretical (5505.10) and experimental (5,515 ± 21) values of CA is small enough (10.11 μg/g PNR) to fortify the fitness of the model. However, such medium constitution is characterized by simplicity and economy, in relation to the high valued-added CA produced under SSF of PNR by *Trichoderma* sp. WKA55.

During this study, the resulting CA in the hydrolysate of SSF by the cellulolytic, proteolytic, and TCP-solubilizing *Trichoderma* sp. WKA55 was quantified using HPLC. Because of speed, simplicity, and accuracy, the HPLC technique is an appealing procedure for such detection. CA could be produced through SSF and submerged fermentation. Here, the SSF technique was applied to get the advantages of simplicity, low energy requirements, high volumetric productivity, ease of aeration, and simulation of the natural habitat of most fungi. CA was reported to be delivered by several fungi, involving *Aspergillus* spp. and *Penicillium* spp. (Max et al., 2010; Saber et al., 2015), but little is known about its production by the current endophytic strain; *Trichoderma* sp. WKA55, representing a new source for CA production. Hydrolysis, biochemical metabolism, and microbial assimilation are suggested mechanisms for the formation of organic acids in the fermented matter. Likewise, all organic acids, CA has its biosynthesis pathway, it was found to be promoted as the major acidic metabolite of *Penicillium bilaii* under nitrogen-limited conditions, in some cases CA share the same pathway of oxalic acid formation, and both are at one enzymatic step from the main metabolism of D-glucose and D-fructose pathway (Max et al., 2010; Show et al., 2015; Al-Askar et al., 2021b). By the end of optimization trials, a total yield of 5,515 ± 21 μg CA/g PNR was attained, representing a new pioneer fungal candidate for CA biosynthesis. The current yield is higher than (2.007 μg/g) these obtained recently on maize stover (Al-Askar et al., 2021b).

The biological potential of CA against some important toxicogenic fungi showed a marked reduction in growth and spore germination against the toxinogenic fungi to prevent or at least reduce their activities. Because of their direct hazards and devastating influence on human health, diseases produced by mycotoxigenic fungus have received special attention. Their mycotoxins are widespread, and they are produced under a wide range of environmental circumstances (Bennett and Klich, 2003; Marin et al., 2013). Individuals of the genus *Fusarium* and *Aspergillus* were reported to produce groups of mycotoxins, contamination of agricultural products by such toxigenic and pathogenic fungi plays a crucial role in terms of economy, hygiene, and health. Among mycotoxins, aflatoxin, fumonisin, and ochratoxin are the more toxic to organisms, causing an assortment of dangerous impacts including teratogenicity, hepatotoxicity, and mutagenicity (Marin et al., 2013). Mycotoxins secreted by such fungi devastate the yields during production, storage, processing, and even in the markets, leading to losing their nutritious value (Jimoh and Kolapo, 2008).

The utilization of fabricated fungicides for the elimination of toxinogenic fungi tends to the development of various known perils. Alternative bio-fabricated ones have been proposed. In this regard, the mycotoxin content was reduced after treatment with the fungal filtrate, and more, the antimicrobial property of CA against a vast array of

microorganisms was demonstrated. For instance, in an *in vitro* study, great inhibition was noticed on the linear growth, dry weight, and sporulation of the soil-borne pathogens; *F. solani* FP2 and *F. oxysporum* FP4 by the action of CA (Abdel-Monaim and Ismail, 2010). Similarly, a significant growth reduction of wheat seed-borne pathogens (*F. culmorum*, *F. moniliforme*, and *F. graminearum*) grown on media amended with different concentrations of CA (Seadh and El-Metwally, 2015). In the two investigations, the reduction in fungal growth was linked to the increase in CA levels.

The antifungal action of CA could be ascribed to several pathways such as the decline of internal pH of the microbial cell by ionization of undissociated acid molecules and disturbance of system transportation by changing the permeability of cell membrane and/or reduction of proton motive power (Shokri, 2011). In addition, some antifungal metabolites in the culture filtrate may synergistically have contributed to the antifungal nature of *Trichoderma* sp. WKA55 filtrate. Moreover, the evolution of the endophytic fungi to elaborate many novel metabolites may be back to their multiple interactions inside the host plant, i.e., with the plant tissue, or with the endemic pathogens, or even among each other (Schulz et al., 2015). The antifungal role of the lytic activity of *Trichoderma* species was reported against the growth of *Rhizoctonia solani* pathogen. The isolate can produce pectinase and chitinase and solubilize phosphorus (Sallam et al., 2021).

As a practical applicable approach, peanut seed treatment by the filtrate of *Trichoderma* sp. WKA55 led to a marked improvement in reducing the fungal load on the seed. Such data encourage the application of the current *Trichoderma* filtrate as pretreatment to minimize the fungal load of peanut seeds, especially the toxicogenic ones. However, the current data have another advantage from the nutritional point of view, since the mycotoxins contamination caused by mycotoxinogenic fungi is a great concern in nuts production worldwide.

Several studies stated that metabolites produced by *Trichoderma* spp. had a positive promoted plant growth without negative impact on the plants. In the present study, there was a positive effect of crude CA on seed germination and seedling features that may be back to the high capacity of *Trichoderma* species to secret plant-growth promotors that act as auxin-like molecules, e.g., gibberellic acid and indole acetic acid, and the secretion of antibiotic-like substances as a biocontrol agent, which possibly act as auxin-like molecules (Vinale et al., 2014; Saber et al., 2017; Singh et al., 2018; Al-Askar et al., 2021b). Recently, a pronounced increase of seed germination, root and shoot lengths, plant dry matter, and vigor index of durum wheat seeds coated with three different strains of *T. harzianum* and the *Trichoderma*-based commercial product; Trianum-T22 were reported (Kthiri et al., 2020). These results were also confirmed on tomato seeds by *T. asperellum* CA, which has the potential to enhance germination percentage, root length, and vigor index parameters up to 91%, 8.85%, and 49.40%, respectively (Al-Askar et al., 2021b). Another critical role of *T. asperellum* was stated to induce $H^+$-ATPase, which is an important enzyme responsible for cell growth and plasma membrane elongation during the growth of maize seedlings (López-Coria et al., 2016).

In the present study, secondary metabolites other than CA in the fungal metabolite were explored, through GC-MS, for a possible role of antibiosis and pathogenic activity, a total of 20 compounds from the culture filtrate of the endophytic *T. longibrachiatum* WKA55 strain were determined. Among microorganisms, the secondary metabolites of the genus *Trichoderma* have a wide biological role in agricultural, industrial, and medical aspects. With minimal nutritional requirements, *Trichoderma* species secrete a plethora of metabolites such as terpenes, gliotoxin, pyrones, gliovirin, and peptaibols (Vinale et al., 214; Khan et al., 2020). These secondary metabolites have antimicrobial features, acting as potential biocontrol agents against phytopathogenic antifungal activities (Vizcaino et al., 2005). *Trichoderma longibrachiatum* is one of the recent studied fungal biocontrol agents, to date, several strains have been reported to generate a diversity of new volatile and nonvolatile metabolites, including peptides, polyketides, and terpenes, which led to the appreciation of *T. longibrachiatum* as a potential source of unique antibiotics and as talented biocontrol agents against phytopathogenic microorganisms and nematodes (Sridharan et al., 2020, 2021). These metabolites were found to delay the colonization of the pathogens; for instance, the antifungal secondary metabolites of *T. longibrachiatum* were reported against *Candida albicans* and *Pyricularia oryzae* (Xuan et al., 2014). Further, the secondary metabolites were found to induce disease resistance and regulate plant growth (Khan et al., 2020).

Among active constituents detected in the current fungal filtrate, phenol, 2-methyl-5-(1-methylethyl)-, or carvacrol was found to be the major unique volatile compounds in the tested extract (65.04%), which was, recently, demonstrated as among the main active ingredients of *T. harzianum* filtrate with significant antifungal activity against the phytopathogenic *Rhizoctonia solani* (Yousef et al., 2018). Similarly, phenol 2,4, bis (1,1-dimethyl ethyl) was reported as a main secondary metabolite in the filtrates of *T. harzianum* and *T. asperellum* that prompted the germination and the growth of the tomato seedling (Yuef et al., 2018).

The present fungal strain showed the presence of benzaldehyde, 4-(1-methylethyl)-, caryophyllene oxide, and propanoic acids in the filtrate. These compounds showed antifungal properties that inhibit the mycelial growth of different phytopathogens (Lee, 2015; Yassin et al., 2021). Benzaldehyde, 4-(1-methylethyl)- was reported as a volatile organic compound in *T. atroviride* extract and exhibited very strong antifungal activity at low concentrations (Lee, 2015). Several positive biological benefits of caryophyllene oxide, secreted by *T. viride*, which reduced the growth of *Fusarium solani*, *Rhizoctonia solani*, and *Sclerotium rolfsii* pathogens (Awad et al., 2018). Also, caryophyllene oxide, propanoic acid, and 1 pentanol were reported among active constituents of *T. viride* filtrate, against the growth of *F. verticillioides* and *F. proliferatum* maize fungal pathogens (Yassin et al., 2021).

The antifungal activity of *T. longibrachiatum* metabolites was recently explained against *Sclerotium rolfsii* and

*Macrophomina phaseolina*. During the direct *T. longibrachiatum*-pathogen interactions, an alteration and reduction of the mycelial growth and inhibited the production of sclerotia of the pathogens were noticed, which was associated with secretion of several antifungal compounds including longifolene, 1-butanol 2-methyl, cedrene, caryophyllene, and cuprenene, which are involved in the biosynthetic pathways of the sesquiterpenoid and alkane, and the degradation pathway of trimethylamine. What is more, 1-pentanol, 1-hexanol, myristonyl pantothenate, bisabolol, d-Alanine, and diethyl trisulphide draws attention as a plant-growth promoting and unique antimicrobial compounds (Sridharan et al., 2020, 2021).

GC-Mass analysis of the endophytic fungus studied *T. longibrachiatum*_strain WKA55 (MZ014020.1) revealed the presence of several volatile organic compounds (VOCs) belonging to a large class of compounds comprising hydrocarbons, alcohols, ketones, aldehydes, esters, acids, and terpenes, as well as caryophyllene derivative. In this respect, epi-β-caryophyllene was detected among the bioactive VOCs, which have been reported to possess properties of antifungal, and antioxidants activity, as well as promoting seedling growth and chlorophyll content (Lee et al., 2016; Rajani et al., 2021). *Trichoderma* species have been reported as excellent plant-growth-promoting fungi, which can improve plant health by creating a favorable environment and production of a large number of secondary metabolites such as harzianic acid, harzianolde, abscisic acid, auxin-related compounds indole-3-acetic acid, indol-3-acelaldehyde, and indol-3 ethanol (Zin and Badaluddin, 2020). Some others produce valuable VOCs with different mechanisms involved in biological control, inducing defense responses and promoting plant growth (Phoka et al., 2020; Wonglom et al., 2020; Ruangwong et al., 2021). However, the approach of biological actions of the current fungal metabolites are unique and numerous, yet a small portion is known. Collectively, the potential uses of these metabolites include may include the volatile-mediated inhibition of pathogen growth and augmented plant systemic resistance. The antifungal activity cannot be attributed to CA or the main metabolite [phenol, 2-methyl-5-(1-methylethyl)-] only but also to the occurrence of other bioactive constituents, which may attribute synergistically to exert an antifungal power.

Finally, the current endophytic fungal strain (*Trichoderma* sp. WKA55) was identified using the molecular procedure as a perfect means for identification. Because of high sensitivity and specificity, molecular identification is widely used for the rapid identification of filamentous fungi at various taxonomic levels. The technique is set up for the comparison of the sequence coding for 18s rRNA gene after PCR amplification, through the ITS, which fragment size is uniform in numerous groups of fungi, making nucleotide sequencing of ITS fractions prerequisite for revealing interspecific, and in some cases, intraspecific variation (Al-Askar et al., 2021b). *Trichoderma longibrachiatum* belongs to the clade longibrachiatum, sordariomycetes, order: hypocreales, and family hypocreaceae (Hewedy et al., 2020).

Summing up, a novel endophytic isolate was reported as a biologically active fungus capable to produce CA, through a safe and economic disposal procedure of PNR biomass. The basic concepts for the bioconversion process of PNR into CA, on a simple and low-cost medium using a new endophytic *T. longibrachiatum* WKA55, were illustrated. The unique complementary hydrolytic system of *T. longibrachiatum* WKA55 enables efficient bioconversion of plant biomass without any previous pretreatments, suggesting a novel source of broad biotechnological enzymes as well. The crude fungal exudate was able to suppress the pathogenic and mycotoxin-producing fungi associated with peanut seeds; as well as improve the germination and growth of peanut seedlings. Therefore, large-scale production is very encouraged especially with incremental production of PNR. The current concept may be applied for the breakdown of the infection cycle of the various pathogens that probably exist. It is worthy to note that the seasonal climatic variation may affect the stability of the percentage of the compositional constituents of PNR, which may reflect on the production stability of fungal metabolites. However, this drawback is very limited, and variation is ranging within a narrow range. Nevertheless, the suggested point for further studies could be that the residual fermented PNR matter could be used as a soil amendment or as an animal feed supplement, representing integrated management of such biomass residue. To gain a better understanding of the influence of crude CA produced *via T. longibranchiatum* WKA55 in restricting mycotoxgenic pathogens growth and stimulation the seedlings growth, other work recommended under field conditions to extensively evaluate the obtained compounds on the large scale.

## 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 in the article/supplementary material.

## AUTHOR CONTRIBUTIONS

ER, ZM, KG, and WS contributed to the conception and design of the study. AA, ER, ZM, AM, and FA-O organized the database. AA-A, ER, ZM, AM, and WS performed the statistical analysis. AA-A, ER, ZM, KG, and FA-O wrote the first draft of the manuscript. AA-A, ER, ZM, KG, FA-O, AA, and WS wrote the discussion section. AA and WS substantially contributed to the conception of the work and interpretation of data and critically revised the intellectual content of the manuscript. All authors contributed to manuscript revision, read, and approved the final version.

## FUNDING

The authors extend their appreciation to the researcher supporting project number (RSP2022R505), King Saud University, Riyadh, Saudi Arabia for funding this work.

## REFERENCES


Abdel-Monaim, M. F., and Ismail, M. E. (2010). The use of antioxidants to control root rot and wilt diseases of pepper. *Not. Sci. Biol.* 2, 46–55. doi: 10.15835/nsb.2.2.3699

Abdelwahed, N. A. M., El-Naggar, N. E., and Saber, W. I. A. (2011). Factors and correlations controlling cellulase-free xylanase production by *Streptomyces halstedii* NRRL B-1238 in submerged culture. *Aust. J. Basic Appl. Sci.* 5, 45–53. doi: 10.22587/ajbas

Abdul-Baki, A., and Anderson, J. D. (1973). Vigour determination of soyabean seed by multiple criteria. *Crop Sci.* 13, 630–633. doi: 10.2135/cropsci1973.0011183X001300060013x

Al-Askar, A. A., Rashad, E. M., Ghoneem, K. M., Mostafa, A. A., Al-Otibi, F. O., and Saber, W. I. A. (2021a). Discovering *Penicillium polinicum* with high-lytic capacity on *Helianthus tuberosus* tubers: oil-based preservation for mold management. *Plan. Theory* 10:413. doi: 10.3390/plants10020413

Al-Askar, A. A., Saber, W. I. A., Ghoneem, K. M., Hafez, E. E., and Ibrahim, A. A. (2021b). Crude citric acid of *Trichoderma asperellum*: tomato growth promotor and suppressor of *Fusarium oxysporum* f. sp. *lycopersici*. *Plan. Theory* 10:222. doi: 10.3390/plants10020222

Alebel, M. G., Urge, M., Assefa, G., Worku, B., and Abebe, A. (2019). The effect of using either soybean or groundnut straw as part of basal diet on body weight gain, and carcass characteristics of Gumuz sheep. *Int. J. Livest. Prod.* 10, 70–76. doi: 10.5897/IJLP2018.0549

Asadi, N., and Zilouei, H. (2017). Optimization of organosolv pretreatment of rice straw for enhanced biohydrogen production using *Enterobacter aerogenes*. *Bioresour. Technol.* 227, 335–344. doi: 10.1016/j.biortech.2016.12.073

Awad, N. E., Kassem, H. A., Hamed, M. A., El-Feky, A. M., Elnaggar, M. A. A., Mahmoud, K., et al. (2018). Isolation and characterization of the bioactive metabolites from the soil derived fungus *Trichoderma viride*. *Mycology* 9, 70–80. doi: 10.1080/21501203.2017.1423126

Barrett, A. J., Woessner, J. F., and Rawlings, N. D. (eds.) (2012). *Handbook of Proteolytic Enzymes. 2nd Edn. Vol. 1*. London, UK: Elsevier Academic Press.

Bennett, J. W., and Klich, M. (2003). Mycotoxins. *Clin. Microbiol. Rev.* 16, 497–516. doi: 10.1128/CMR.16.3.497-516.2003

De Silva, N. I., Brooks, S., Lumyong, S., and Hyde, K. D. (2019). Use of endophytes as biocontrol agents. *Fungal Biol. Rev.* 33, 133–148. doi: 10.1016/j.fbr.2018.10.001

Dinis, M. J., Bezerra, R. M., Nunes, F., Dias, A. A., Guedes, C. V., Ferreira, L. M., et al. (2009). Modification of wheat straw lignin by solid state fermentation with white-rot fungi. *Bioresour. Technol.* 100, 4829–4835. doi: 10.1016/j.biortech.2009.04.036

Domsch, K. H., Gams, W., and Anderson, T. H. (1980). *Compendium of Soil Fungi. Vol. 1*. London, UK: Academic Press.

Elad, Y. (1992). The use of antioxidants (free radical scavengers) to control gray mould (*Botrytis cinerea*) and white mould (*Sclerotinia sclerotiorum*) in various crops. *Plant Pathol.* 41, 417–426. doi: 10.1111/j.1365-3059.1992.tb02436.x

Ellis, M. B. (1971). *Dematiaceous Hyphomycetes. 1st Edn.* Commonwealth Mycological Institute, Kew, Surrey, UK.

Ellis, M. B. (1976). *More Dematiaceous Hyphomycetes*. Commonwealth Mycological Institute, Kew, Surrey, England, 507 pp.

El-Maghraby, O. M. O., Soltan, S. M., Mohammed, R. M., and Mohammed, M. M. (2013). Endophytic fungi of three leguminous plant roots in Egypt. *J. Basic Appl. Mycol.* 4, 59–68.

El-Samawaty, A. M. A., Moslem, M. A., Sayed, S. R. M., and Yassin, M. A. (2014). Fungal endophytes survey of some legume seeds. *J. Pure Appl. Microbiol.* 8, 153–160.

Elsayed, M. S., Eldadamony, N. M., Alrdahe, S. S. T., and Saber, W. I. A. (2021). Definitive screening design and artificial neural network for modeling a rapid biodegradation of date palm fronds by a new *Trichoderma* sp. PWN6 into citric acid. *Molecules* 26:5048. doi: 10.3390/molecules26165048

El-Zefzaf, H. M. (2020). Fungi associated with peanut seeds and their extracellular enzyme activities under the influence of physical factors. *Asian J. Appl. Sci.* 13, 144–151. doi: 10.3923/ajaps.2020.144.151

Hewedy, O. A., Abdel Lateif, K. S., Seleiman, M. F., Shami, A., Albarakaty, F. M., and El-Meihy, R. M. (2020). Phylogenetic diversity of *Trichoderma* strains and their antagonistic potential against soil-borne pathogens under stress conditions. *Biology* 9:189. doi: 10.3390/biology9080189

ISTA (2007). International Seed Testing Association. International Rules for Seed Testing; International Seed Testing Association. ISTA: Bassersdorf, Switzerland, 2007.

Jackson, M. L. (2005). *Soil Chemical Analysis: Advanced Course*. University of Wisconsin-Madison Libraries, USA: Parallel Press, 930.

Jacob, N., Asha Poorna, C., and Prema, P. (2008). Purification and partial characterization of polygalacturonase from *Streptomyces lydicus*. *Bioresour. Technol.* 99, 6697–6701. doi: 10.1016/j.biortech.2007.10.002

Jimoh, K. O., and Kolapo, A. L. (2008). Mycoflora and aflatoxin production in market samples of some selected Nigerian foodstuffs. *Res. J. Microbiol.* 3, 169–174. doi: 10.3923/jm.2008.169.174

Khan, R. A. A., Najeeb, S., Hussain, S., Xie, B., and Li, Y. (2020). Bioactive secondary metabolites from *Trichoderma* spp. against phytopathogenic fungi. *Microorganisms* 8:817. doi: 10.3390/microorganisms8060817

Klaedtke, S., Jacques, M. A., Raggi, L., Préveaux, A., Bonneau, S., Negri, V., et al. (2016). Terroir is a key driver of seed-associated microbial assemblages. *Environ. Microbiol.* 18, 1792–1804. doi: 10.1111/1462-2920.12977

Kthiri, Z., Jabeur, M. B., Machraoui, M., Gargouri, S., Hiba, K., and Hamada, W. (2020). Coating seeds with *Trichoderma* strains promotes plant growth and enhance the systemic resistance against *Fusarium* crown rot in durum wheat. *Egypt. J. Biol. Pest Control.* 30:139. doi: 10.1186/s41938-020-00338-6

Kumari, A., Kapoor, K. K., Kundu, B. S., and Kumari, M. R. (2008). Identification of organic acids produced during rice straw decomposition and their role in rock phosphate solubilization. *Plant Soil Environ.* 54, 72–77. doi: 10.17221/2783-PSE

Lee, S. Y. J. (2015). Analysis of Volatile Organic Compounds Emitted by Filamentous Fungi and Volatile Mediated Plant Growth. doctoral dissertation. Rutgers University-Graduate School-New Brunswick.

Lee, S., Yap, M., Behringer, G., Hung, R., and Bennett, J. W. (2016). Volatile organic compounds emitted by *Trichoderma* species mediate plant growth. *Fungal Biol. Biotechnol.* 3:7. doi: 10.1186/s40694-016-0025-7

Leslie, J. F., and Summerell, B. A. (2008). *The Fusarium Laboratory Manual*. New York, NY, USA: John Wiley & Sons.

López, S. M. Y., Pastorino, G. N., Franco, M. E. E., Medina, R., Lucentini, C. G., Saparrat, M. C. N., et al. (2018). Microbial endophytes that live within the seeds of two tomato hybrids cultivated in Argentina. *Agronomy* 8:136. doi: 10.3390/agronomy8080136

López-Coria, M., Hernández-Mendoza, J. L., and Sánchez-Nieto, S. (2016). *Trichoderma asperellum* induces maize seedling growth by activating the plasma membrane H+-ATPase. *Mol. Plant-Microbe Interact.* 29, 797–806. doi: 10.1094/MPMI-07-16-0138-R

Marin, S., Ramos, A. J., Cano-Sancho, G., and Sanchis, V. (2013). Mycotoxins: occurrence, toxicology, and exposure assessment. *Food Chem. Toxicol.* 60, 218–237. doi: 10.1016/j.fct.2013.07.047

Max, B., Salgado, J. M., Rodríguez, N., Cortés, S., Converti, A., and Domínguez, J. M. (2010). Biotechnological production of citric acid. *Braz. J. Microbiol.* 41, 862–875. doi: 10.1590/S1517-83822010000400005

Meddeb-Mouelhi, F., Moisan, J. K., and Beauregard, M. (2014). A comparison of plate assay methods for detecting extracellular cellulase and xylanase activity. *Enzym. Microb. Technol.* 66, 16–19. doi: 10.1016/j.enzmictec.2014.07.004

Miller, G. L. (1959). Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal. Chem.* 31, 426–428. doi: 10.1021/ac60147a030

Montgomery, H. A. C., Dymock, J. F., and Thom, N. S. (1962). The rapid colorimetric determination of organic acids and their salts in sewage-sludge liquor. *Analyst* 87, 949–955. doi: 10.1039/an9628700949

Negm, S., El-Metwally, M. M., Saber, W. I. A., Abo-Neima, S., Moustafa, M., and El-Kott, A. (2021). Nanoparticles induce the biosynthesis and activity of the new possible therapeutic proteinase source, *Talaromyces purpureogenus* KJ584844. *Biocell* 45, 119–127. doi: 10.32604/biocell.2021.012011

Nelson, P. E., Toussoun, T. A., and Marasas, W. F. O. (1983). *Fusarium Species, An Illustrated Manual for Identification*. The Pennsylvania State University Press, USA, 193 pp.

Newman, L., Duffus, A., and Lee, C. (2016). Using the free program MEGA to build phylogenetic trees from molecular data. *Am. Biol. Teach.* 78, 608–612. doi: 10.1525/abt.2016.78.7.608

Phoka, N., Suwannarach, N., Lumyong, S., Ito, S.-I., Matsui, K., Arikit, S., et al. (2020). Role of volatiles from the endophytic fungus *Trichoderma asperelloides* PSU-P1 in biocontrol potential and in promoting the plant growth of Arabidopsis thaliana. *J. Fungi* 6:341. doi: 10.3390/jof6040341

Rahman, M., and Punja, Z. K. (2005). Factors influencing development of root rot on ginseng caused by *Cylindrocarpon destructans*. *Phytopathology* 95, 1381–1390. doi: 10.1094/PHYTO-95-1381

Rajani, P., Rajasekaran, C., Vasanthakumari, M. M., Olsson, S. B., Ravikanth, G., and Shaanker, R. U. (2021). Inhibition of plant pathogenic fungi by endophytic *Trichoderma* spp. through mycoparasitism and volatile organic compounds. *Microbiol. Res.* 242:126595. doi: 10.1016/j.micres.2020.126595

Ruangwong, O.-U., Wonglom, P., Suwannarach, N., Kumla, J., Thaochan, N., Chomnunti, P., et al. (2021). Volatile organic compound from *Trichoderma asperelloides* TSU1: impact on plant pathogenic fungi. *J. Fungi* 7:187. doi: 10.3390/jof7030187

Saber, W. I. A., El-Naggar, N. E., and AbdAl-Aziz, S. A. (2010). Bioconversion of lignocellulosic wastes into organic acids by cellulolytic rock phosphate-solubilizing fungal isolates grown under solid-state fermentation conditions. *Res. J. Microbiol.* 5, 1–20. doi: 10.3923/jm.2010.1.20

Saber, W. I. A., El-Naggar, N. E., El-Hersh, M. S., and El-khateeb, A. Y. (2015). An innovative synergism between *Aspergillus oryzae* and *Azotobacter chroococcum* for bioconversion of cellulosic biomass into organic acids under restricted nutritional conditions using multi-response surface optimization. *Biotechnology* 14, 47–57. doi: 10.3923/biotech.2015.47.57

Saber, W. I. A., Ghoneem, K. M., Rashad, Y. M., and Al-Askar, A. A. (2017). *Trichoderma harzianum* WKY1: an indole acetic acid producer for growth improvement and anthracnose disease control in sorghum. *Biocontrol Sci. Tech.* 27, 654–676. doi: 10.1080/09583157.2017.1321733

Sallam, N., Ali, E. F., Seleim, M. A. A., and Bagy, H. M. M. K. (2021). Endophytic fungi associated with soybean plants and their antagonistic activity against *Rhizoctonia solan. Egypt. J. Biol. Pest Control.* 31:54. doi: 10.1186/s41938-021-00402-9

Samson, R. A., Visagie, C. M., Houbraken, J., Hong, S. B., Hubka, V., Klaassen, C. H. W., et al. (2014). Phylogeny, identification and nomenclature of the genus *Aspergillus*. *Stud. Mycol.* 78, 141–173. doi: 10.1016/j.simyco.2014.07.004

Santos, F., Medina, P. F., Lourenção, A. L., Parisi, J. J. D., and Godoy, I. J. (2016). Damage caused by fungi and insects to stored peanut seeds before processing. *Bragantia* 75, 184–192. doi: 10.1590/1678-4499.182

Sathiyaraj, G., Srinivasan, S., Kim, H. B., Subramaniyam, S., Lee, O. R., Kim, Y. J., et al. (2011). Screening and optimization of pectin lyase and polygalacturonase activity from ginseng pathogen *Cylindrocarpon destructans*. *Braz. J. Microbiol.* 42, 794–806. doi: 10.1590/S1517-83822011000200048

Schulz, B., Hass, S., Junker, C., Andrée, N., and Schobert, M. (2015). Fungal endophytes are involved in multiple balanced antagonisms. *Curr. Sci.* 109, 39–45.

Seadh, S. E., and El-Metwally, M. A. (2015). Influence of antioxidants on wheat productivity, quality and seed-borne fungi management under NPK fertilization levels. *Asian J. Crop Sci.* 7, 87–112. doi: 10.3923/ajcs.2015.87.112

Shehata, H. R., Lyons, E. M., Jordan, K. S., and Raizada, M. N. (2016). Bacterial endophytes from wild and ancient maize are able to suppress the fungal pathogen *Sclerotinia homoeocarpa*. *J. Appl. Microbiol.* 120, 756–769. doi: 10.1111/jam.13050

Shokri, H. (2011). Evaluation of inhibitory effects of citric and tartaric acids and their combination on the growth of *Trichophyton mentagrophytes*, *Aspergillus fumigatus*, *Candida albicans* and *Malassezia furfur*. *Comp. Clin. Pathol.* 20, 543–545. doi: 10.1007/s00580-011-1195-6

Show, P. L., Oladele, K. O., Siew, Q. Y., Aziz Zakry, F. A., Lan, J. C. W., and Ling, T. C. (2015). Overview of citric acid production from *Aspergillus niger*. *Front. Life Sci.* 8, 271–283. doi: 10.1080/21553769.2015.1033653

Silva, B. M., Andrade, P. B., Mendes, G. C., Seabra, R. M., and Ferreira, M. A. (2002). Study of the organic acids composition of quince (*Cydonia oblonga* Miller) fruit and jam. *J. Agric. Food Chem.* 50, 2313–2317. doi: 10.1021/jf011286+

Singh, A., Shukla, N., Kabadwal, B. C., Tewari, A. K., and Kumar, J. (2018). Review on plant-*Trichoderma*-pathogen interaction. *Int. J. Curr. Microbiol. App. Sci.* 7, 2382–2397. doi: 10.20546/ijcmas.2018.702.291

Singh, V., Upadhyay, R. S., Sarma, B. K., and Singh, H. B. (2016). *Trichoderma asperellum* spore dose depended modulation of plant growth in vegetable crops. *Microbiol. Res.* 193, 74–86. doi: 10.1016/j.micres.2016.09.002

Sparkman, O. D., Penton, Z., and Kitson, F. G. (2011). *Gas Chromatography and Mass Spectrometry: A Practical Guide Ed. 2.* London, UK: Elsevier Academic Press.

Sridharan, A. P., Sugitha, T., Karthikeyan, G., Nakkeeran, S., and Sivakumar, U. (2021). Metabolites of *Trichoderma longibrachiatum* EF5 inhibits soil borne pathogen, *Macrophomina phaseolina* by triggering amino sugar metabolism. *Microb. Pathog.* 150:104714. doi: 10.1016/j.micpath.2020.104714

Sridharan, A. P., Thankappan, S., Karthikeyan, G., and Uthandi, S. (2020). Comprehensive profiling of the VOCs of *Trichoderma longibrachiatum* EF5 while interacting with *Sclerotium rolfsii* and *Macrophomina phaseolina*. *Microbiol. Res.* 236:126436. doi: 10.1016/j.micres.2020.126436

Vinale, F., Sivasithamparam, K., Ghisalberti, E. L., Woo, S. L., Nigro, M., Marra, R., et al. (2014). *Trichoderma* secondary metabolites active on plants and fungal pathogens. *Open Mycol. J.* 8, 127–139. doi: 10.2174/1874437001408010127

Vizcaino, J. A., Sanz, L., Cardoza, R. E., Monte, E., and Gutierrez, S. (2005). Detection of putative peptide synthetase genes in *Trichoderma* species: application of this method to the cloning of a gene from *T. harzianum* CECT 2413. *FEMS Microbiol. Lett.* 244, 139–148. doi: 10.1016/j.femsle.2005.01.036

White, T. J., Bruns, T., and Lee, S. T. J. (1990). "Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics" in *PCR Protocols: A Guide to Methods and Applications*. eds. M. A. Innis, D. H. Gelfand, J. J. Sninsky and T. J. White (New York, NY, USA: Academic Press), 315–322.

Wonglom, P., Ito, S., and Sunpapao, A. (2020). Volatile organic compounds emitted from endophytic fungus *Trichoderma asperellum* T1 mediate antifungal activity, defense response and promote plant growth in lettuce (*Lactuca sativa*). *Fungal Ecol.* 43:100867. doi: 10.1016/j.funeco.2019.100867

Xuan, Q. C., Huang, R., Chen, Y. W., Miao, C. P., Ma, K. X., Wang, T., et al. (2014). Cyclonerol derivatives from *Trichoderma longibrachiatum* YM311505. *Nat. Prod. Commun.* 9, 313–314. doi: 10.1177/1934578X1400900307

Yassin, M. T., Mostafa, A. A., Al-Askar, A. A., Sayed, S. R. M., and Rady, A. M. (2021). Antagonistic activity of *Trichoderma harzianum* and *Trichoderma viride* strains against some fusarial pathogens causing stalk rot disease of maize, *in vitro*. *J. King Saud Univ. Sci.* 33:101363. doi: 10.1016/j.jksus.2021.101363

Yousef, S. A. M., Salem, S. H., and EL-Sharkawy, H. H. A. (2018). Effect of some chemical inducers on antagonistic potential of *Trichoderma harzianum* against *Rhizoctonia solani* and it's metabolites production. *J. Plant Prot. Pathol.* 9, 497–506. doi: 10.21608/jppp.2018.43741

Yuef, M. H., Ariel, T. J., Raúl, R., Alberto, L. J., Benigno, E., and Eduardo, O. (2018). Identification and evaluation of secondary metabolites by gas chromatography-mass spectrometry (GC-MS) in native strains of *Trichoderma* species. *Afr. J. Biotechnol.* 17, 1162–1171. doi: 10.5897/AJB2018.16546

Zin, N. A., and Badaluddin, N. A. (2020). Biological functions of *Trichoderma* spp. for agriculture applications. *Ann. Agric. Sci.* 65, 168–178. doi: 10.1016/j.aoas.2020.09.003

**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.

*Copyright © 2022 Al-Askar, Rashad, Moussa, Ghoneem, Mostafa, Al-Otibi, Arishi and Saber. 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.*

frontiers in Microbiology

ORIGINAL RESEARCH
published: 14 March 2022
doi: 10.3389/fmicb.2022.756209

# Microbial Co-occurrence Network and Fermentation Information of Natural Woody-Plant Silage Prepared With Grass and Crop By-Product in Southern Africa

*Zhumei Du[1,2], Seishi Yamasaki[1], Tetsuji Oya[1], Damiao Nguluve[3], Denise Euridse[3], Benedito Tinga[3], Felicidade Macome[3] and Yimin Cai[1]**

[1] Japan International Research Center for Agricultural Sciences, Tsukuba, Japan, [2] College of Grassland Science and Technology, China Agricultural University, Beijing, China, [3] Agricultural Research Institute of Mozambique, Matola, Mozambique

OPEN ACCESS

**Edited by:**
Sailesh Malla,
Chr. Hansen, Denmark

**Reviewed by:**
Robin Anderson,
Agricultural Research Service, United States Department of Agriculture, United States
Yantyati Widyastuti,
Indonesian Institute of Sciences, Indonesia
Sureewan Sittijunda,
Mahidol University, Thailand

***Correspondence:**
Yimin Cai
cai@affrc.go.jp

**Specialty section:**
This article was submitted to Microbiotechnology, a section of the journal Frontiers in Microbiology

**Received:** 10 August 2021
**Accepted:** 27 January 2022
**Published:** 14 March 2022

**Citation:**
Du Z, Yamasaki S, Oya T, Nguluve D, Euridse D, Tinga B, Macome F and Cai Y (2022) Microbial Co-occurrence Network and Fermentation Information of Natural Woody-Plant Silage Prepared With Grass and Crop By-Product in Southern Africa. Front. Microbiol. 13:756209. doi: 10.3389/fmicb.2022.756209

To facilitate the use of woody plant (WP) as a natural biomass resource to address the shortage of feed for ruminants in the tropics, we use PacBio SMRT sequencing to explore the microbial co-occurrence network and silage fermentation of gliricidia and leucaena prepared with Napier grass (NG) and corn stover (CS) in Southern Africa. Based on dry matter, the crude protein contents of WP are as high as 25%. Compared with NG, the addition of CS speed up the dynamic succession of microorganisms in the silage fermentation process from Gram-negative bacteria to Gram-positive bacteria, and promoted *Lactiplantibacillus plantarum* to become the dominant community and enhanced the metabolic pathways of lactic acid and citric acid, thus improved the fermentation flavour and quality of WP silage. WP can be mixed with CS to make high-quality silage, which can alleviate the shortage of feed and promote local animal production.

**Keywords: feed shortage, microbial co-occurrence network, natural biomass resource, silage fermentation, woody plant**

## INTRODUCTION

To cope with insufficient feed caused by the rapid development of animal husbandry in the tropics, new and nutrient-rich locally available feed resources including natural woody plant (WP) have been investigated (Zhang et al., 2019). In addition to their contributions to a sustainable agricultural and animal ecosystem, tropical legume WP are rich sources of crude protein (CP) and minerals; they also promote biomass and animal production (Luscher et al., 2014).

Gliricidia [*Gliricidia sepium* (Jacq.) Kunth ex Walp.] is native to tropical, arid forests in Mexico and Central America. Beyond its region of origin, it is also grown in many tropical and subtropical

**Abbreviations:** WP, woody plant; PM, paper mulberry; LBC, lactic acid buffer capacity; WSC, water-soluble carbohydrate; DM, dry matter; FM, fresh matter; OM, organic matter; CP, crude protein; EE, ether extract; NDF, neutral detergent fibre; ADF, acid detergent fibre; ADL, acid detergent lignin; $NH_3$-N, ammonia nitrogen; LAB, lactic acid bacteria; OTU, operational taxonomic unit; cfu, colony-forming unit; NG, Napier grass; CS, corn stover.

regions, including the Caribbean, northern parts of South America, central Africa, portions of India, and Southeast Asia (Oliveira et al., 2018). Leucaena [*Leucaena leucocephala* (Lam.) de Wit] is a small, fast-growing mimosa tree native to southern Mexico and northern Central America (Belize and Guatemala), which has been naturalised throughout the tropics (Rengsirikul et al., 2011). The fresh branches and leaves of these WP are used as natural feed for grazing cattle because of their abundant nutrients, low lignin content, and good palatability (Speedy and Pugliese, 1992). Both types of WP have high biomass production capacity; the yield of their leaves can be as high as 20,000 kg-dry weight/ha/year. Under favourable climatic conditions, the yield can reach twice this level (Rajvanshi, 2019). Generally, the WP harvest time is concentrated in the rainy season, which is not conducive to hay preparation; however, silage fermentation is considered an ideal storage method. WP, like legume grass, have a high protein content, but they also have high moisture and low water-soluble carbohydrate (WSC) contents. Therefore, WP must be mixed with local gramineous forage or crop by-products to prepare silage.

Napier grass (NG, *Pennisetum purpureum* Schumach.) is the major feed resource in cattle production systems in Africa (Du et al., 2020). NG is widely planted in Africa because of its high biomass yield and adaptations for survival under a wide range of soil types, fertility levels, and weather conditions (Negawo et al., 2017). In contrast, corn (*Zea mays* L.) stover (CS) is the main crop by-product in Africa; dry stover is widely used for ruminant feed, but it is most often discarded in the field to be burned and used as fertiliser (Cai et al., 2020a). NG and CS are used to adjust the moisture content and increase the fermentation substrate level to optimise WP silage fermentation.

PacBio Single-Molecule Real-Time (SMRT) sequencing technology can cover the full read length of DNA fragments and obtain long sequence reads, providing microbial diversity information at the species level. This technology has been used to study the microbial community and fermentation mechanism in silage. However, there is limited information on how to improve the fermentation quality of WP silage in combination with local forage resources in Africa. To develop a high-quality WP preparation technology to alleviate the shortage of feed for ruminants and evaluate silage-related microbial communities, we used SMRT sequencing to assess the microbial community and co-occurrence network related to silage fermentation of WP prepared with NG and CS in Southern Africa.

# MATERIALS AND METHODS

## Woody Plant and Silage Fermentation

The experiment was conducted at an experimental farm of the Agricultural Research Institute of Mozambique (IIAM, Matola, Mozambique) on February 26, 2019. NG was harvested from the first cutting at the early flowering stage. Corn was cultivated by a local farm in the same area, and fresh CS was collected after the harvest of corn cobs were harvested during the experiment. Gliricidia and leucaena grow naturally in hilly areas in the region (Matola, Mozambique). Young branches and leaves of both WP at the juvenile stage were obtained from first cuttings in different hilly areas with three replicates.

After harvest, gliricidia, leucaena, NG, and CS were immediately chopped into lengths of approximately 1–2 cm using a chopper (130DX; ARS Co., Ltd., Osaka, Japan). The chopped materials were homogenised according to the experimental design and the silages were prepared with gliricidia, leucaena, NG, CS, and their mixture. The mixing proportions of WP with NG or CS are 10, 25, and 50% based on a fresh matter (FM) basis. These homogenised materials were divided into two fractions. The first fraction was collected as fresh samples and placed into sterilised bags that were kept in an ice box. These were immediately transported to a laboratory to determine their lactic acid buffering capacity (LBC), chemical and protein compositions, energy, macro-minerals, and microbial community. The second fraction was used for making silage. As shown in **Figure 1**, the silages were prepared with three replicates by using polyethylene drum silos (20 L, Ka-Kosher Co., Ltd., Sinaloa, Mexico). Approximately 16 kg of mixture material was packed into the silo, all of the silos were then compacted to exclude air and storage was at ambient temperature (24–37°C). After 60 days of ensiling, silage samples (approximately 500 g) from each replicate from the top, middle, and bottom layers of the drum silos were taken and mixed thoroughly before taking subsamples. The fresh samples in each treatment were divided into three portions. The first is to store the samples (approximately 50 g) in a freezer at −80°C for future analysis of microbial community; the second is to dry the samples (approximately 200 g) for subsequent analysis of dry matter (DM), chemical and protein compositions, CP loss; and the third is to use 10 g samples to prepare the extract liquid to analyse the microbial population and fermentation quality of the silage.

## Microbial Analysis

The microbial population of materials and silages consisted of lactic acid bacteria (LAB), aerobic bacteria, coliform bacteria, yeasts, and moulds, which were measured by the plate counting method described by Cai et al. (1999). Samples (10 g) were blended with 90 mL of sterilised saline solution (8.50 g/L NaCl) and homogenised for 5 min in a Stomacher lab blender (400; Seward, United Kingdom). The resulting suspension was serially diluted from $10^{-1}$ to $10^{-8}$ with saline solution. A 0.05 mL aliquot from each diluted suspension was spread on agar plates. LAB were counted on Lactobacilli de Man, Rogosa, and Sharpe (MRS) agar medium (Difco Laboratories, Detroit, MI, United States) in an anaerobic box (TEHER Hard Anaerobox, ANX-1; Hirosawa Ltd., Tokyo, Japan). The LAB were identified by the Gram-positive and catalase-negative rods or cocci that produced lactic acid. Aerobic bacteria were grown on nutrient agar medium (Nissui-Seiyaku Co., Ltd., Tokyo, Japan) under aerobic conditions. Coliform bacteria were counted on blue light broth agar medium (Nissui-Seiyaku Co., Ltd., Tokyo, Japan), and their colonies were distinguished from other bacteria by the blue-colour colonies. All bacterial agar plates were incubated at 30°C for 2–3 days. Yeasts and moulds were counted on potato dextrose agar medium (Nissui-Seiyaku Co., Ltd., Tokyo, Japan) with sterilised tartaric acid solution (pH 3.5) at 30°C for 3–5 days


**FIGURE 1 |** Silage preparation of woody plant.

of incubation. Yeasts and moulds were distinguished from other bacteria based on colony appearance and cell morphology. All the microbial colonies were reported as viable microbial numbers in colony-forming unit (cfu)/g of FM.

For SMRT sequencing analysis, the triplicate samples (10 g) were mixed with 90 mL of sterilised 0.85% NaCl solution and shaken at the speed of 250 rpm in a 4°C refrigerator for 45 min. The liquid mixture was filtered through a four-layer cheesecloth pre-autoclaved, and then the filtrate was centrifuged at 10,000 rmp for 10 min at 4°C to obtain microbial precipitate. Then the precipitated sample was used for DNA extraction by a DNA kit (D5625-01, Omega, Norcross, GA, United States) as described by Du et al. (2021a). The quality of the extracted DNA was monitored on 1% agarose gel electrophoresis and spectrophotometry (optical density at 260/280 nm ratio). All the DNA samples were stored at −20°C for future analysis.

The forward primer 27F and reverse primer 1492R were used to amplify the full-length 16S rRNA gene by PCR for SMRT sequencing (Du et al., 2021b). The PCR amplicons were purified using Agencourt AMPure XP Beads (Beckman Coulter, Indianapolis, IN, United States), and quantified by Qubit dsDNA HS Assay Kit and Qubit 3.0 fluorometer (Invitrogen, Thermo Fisher Scientific, Waltham, MA, United States). After the individual quantification step, amplicons were pooled in equal amounts. SMRTbell libraries were prepared from the amplified DNA by SMRTbell Express Template Prep Kit 2.0 according to the manufacturer's instructions (PacBio, Menlo Park, CA, United States). Purified SMRTbell libraries from the pooled and barcoded samples were sequenced on a single PacBio Sequel II 8M cell using the Sequel II Sequencing kit 2.0.

Single-Molecule Real-Time sequencing was performed on a PacBio RS II instrument (Pacific Biosciences, Menlo Park, CA, United States) using P6-C4 chemistry (Mosher et al., 2013). Raw data were processed using the protocol RS_Readsofinsert.1 in SMRT Portal version 2.7 software (PacBio) (Du et al., 2021a). Low-quality sequences were removed using the Quantitative Insights Into Microbial Ecology (QIIME) package (version 1.7) (Caporaso et al., 2010a). Using 100% clustering of sequence identity, the extracted high-quality circular consensus sequence (CCS) were aligned to obtain representative sequences using Python nearest alignment space termination (PyNAST) and clustering and classification inference with U-statistics (UCLUST) analysis (Caporaso et al., 2010b; Edgar, 2010). Unique sequences were classified into operational taxonomic unit (OTU) based on a 99% threshold identity using the UCLUST algorithm (Lozupone and Knight, 2005). Potential chimeric sequences in the representative set of OTU were removed using the Chimera Slayer tool (Haas et al., 2011). The SILVA database version 132 was implemented to classify different OTU and annotate the taxonomic information for each OTU representative sequence based on Bergey's taxonomy at the genus, family, order, class, and phylum levels, according to classification at an 80% minimum bootstrap threshold (Quast et al., 2013). OTU that occurred only once or twice were discarded. In order to describe the shared and unique microorganisms in all samples following the OTU clustering analyses, Venn diagrams were produced using open-source software package (version 1.2) of R statistical tools (Shade and Handselman, 2012). Since mixture silages showed similar bacterial community, we used gliricidia or leucaena + NG or CS (90 + 10) mixed silages as a representative

for Venn diagram analysis. The relative abundances of different bacterial communities at the species level were also analysed for Windows statistical software package. Hierarchical cluster and heat map analyses were performed using R package pvclust (version 3.0.2) (Suzuki and Shimodaira, 2006; Mansfeldt et al., 2014). A microbial network analysis was drawn with Python language tool (Langfelder and Horvath, 2008).

The metabolic potential of the microbial community and the composition of functional genes were assigned to functional annotations of sequenced metagenomic sequences through 16S rRNA marker gene, and were postulated based on the Kyoto Encyclopedia of Genes and Genomes (KEGG). KEGG is utilised for bioinformatics research and education, including data analysis in genomics, metagenomics, metabolomics and other omics studies, modelling and simulation in systems biology, and translational research in drug development. The functional profiles and differences among different groups were analysed with phylogenetic investigation of communities by reconstruction of unobserved states (PICRUSt2) (Langille et al., 2013). Due to the same reason as the Venn diagram analysis, that is, two types of woody silage have similar fermentation characteristics, we only use gliricidia material and silage, and gliricidia silage and gliricidia + CS (90 + 10) mixed silage for metabolic pathway analysis.

## Chemical Analysis

The pre-ensiled materials and the silage were dried in an oven for 48 h at 65°C until a constant mass was attained. After drying, the samples were ground using a high speed vibrating sample mill (T1-200; for use with two containers with a working capacity of 50 mL; CMT Japan Co., Ltd., Yokohama, Japan). According to the methods of the AOAC International (2000), samples were analysed for DM (method 930.15), ash (method 923.03), CP (method 990.03), and ether extract (EE, method 920.39). The organic matter (OM) content was calculated as the weight lost after ashing. Neutral detergent fibre (NDF), acid detergent fibre (ADF), and acid detergent lignin (ADL) were determined as described by Van Soest et al. (1991). The NDF and ADF were expressed exclusive of residual ash. Heat stable amylase and sodium sulphite were used for the NDF procedure. ADL was analysed by solubilisation of cellulose with sulphuric acid. LBC was determined by titrating with NaOH from pH 4.0 to 6.0 (mmol/kg DM) after first reducing the pH to below 4.0 using HCl, as described by Muck et al. (1991). The WSC including glucose, sucrose, and fructose were determined by high performance liquid chromatography (HPLC, LC-2000 plus; JASCO Co., Tokyo, Japan) as described by Cai (2004). The analytical conditions were as follows: column, SC 1011 (8.0 mm × 30 cm, Shoko, Tokyo, Japan); oven temperature, 80°C; mobile phase, water; flow velocity, 1.0 mL/min; and detector, Jasco RI-1530.

Binding protein, effective protein, and neutral detergent insoluble protein (NDIP, indicates the CP contents of NDF residue) were analysed by the method of Cai et al. (2020b). CP loss (%) = (CP content of fresh material–CP content of silage)/CP content of fresh material × 100%. Gross energy (GE) was determined using an automatic oxygen bomb calorimeter (CA-4PJ; Shimadzu, Kyoto, Japan) (Gao et al., 2019). Digestible energy (DE) and metabolisable energy (ME) concentrations of the test ingredients were calculated using a difference procedure (Baker and Stein, 2009; Li et al., 2018). Herein, CP, EE, NDF, ADF, and ash are expressed as a percentage of DM, and the units for GE, DE, and ME are MJ/kg of DM. The macro-minerals contents of samples, including calcium, phosphorous, magnesium, and potassium, were measured using a wet-digestion method, and then analysed with an atomic absorption spectrophotometer (PerkinElmer, LAMBDA 1050, Yokohama, Japan) as described by Pequerul et al. (1993).

## Analysis of Silage Fermentation

The terminal fermentation products of the silages were analysed using the method of cold-water extracts, as described by Cai (2004). The remaining wet silage sample (10 g) was homogenised in 90 mL of sterilised distilled water and kept in a refrigerator at 4°C for 24 h. Thereafter, the extract samples were filtered through quantitative ashless filter paper (circle size: 5A, 110 mm; Advantec Co., Ltd., Tokyo, Japan). The filtrate was used to determine pH, ammonia nitrogen ($NH_3$-N), and organic acid (lactic acid, acetic acid, propionic acid, and butyric acid) contents. The pH was measured using a glass electrode pH meter (D-71; Horiba Co., Ltd., Kyoto, Japan). The $NH_3$-N contents of silages were determined by steam distillation of the filtrates, as described by Cai (2004) using Kjeltec auto distillation (2200; Foss Tecator, Höganäs, Sweden). The silage filtrates were shaken with cation exchange resin (Amberlite, IR 120B H AG; Organo Corporation, Tokyo, Japan) and centrifuged at 6,500 × *g* and 4°C for 5 min. The supernatants were passed through a 0.45 mm filter under pressure, and the filtrates were then injected into an HPLC system (LC-2000 plus; JASCO Co., Tokyo, Japan) to determine organic acid contents in accordance with the methods described by Cai (2004). The HPLC system was equipped with a Shodex RSpak KC-811 column (8.0 mm × 30 cm; Showa Denko K. K., Tokyo, Japan) at an oven temperature of 60°C. The detector was a Jasco UV-2070 used at 450 nm with an eluent of 3 m*M* $HClO_4$ and reagent of 0.2 m*M* bromothymol blue + 8 m*M* $Na_2HPO_4$ + 2 m*M* NaOH. The flow rate was 1.0 mL/min.

## Statistical Analysis

ANOVA was performed using the general linear model (GLM) procedure of Statistical Package for the Social Sciences (SPSS Version 19.0, SPSS Inc., Chicago, IL, United States) to examine the differences between samples, and significance was declared at $P < 0.05$. The LBC, microbial population, chemical and protein compositions, energy, and macro-minerals contents of samples were subjected to one-way ANOVA. Tukey's honest significant difference (HSD) test was employed for different sample means (Steel and Torrie, 1980).

The hierarchical cluster and heat map analyses showed the correlation analyses of the bacterial community with lactic acid, LAB, pH, and $NH_3$-N at species level, respectively. LAB, organic acid, pH, and $NH_3$-N information are displayed horizontally, respectively, and the bacterial community information is displayed vertically. The corresponding value of the middle heat

map is the Spearman correlation coefficient r, which ranges between −1 and 1, $r < 0$ indicates a negative correlation (blue), $r > 0$ indicates a positive correlation (red), and "*," "**," and "***" represent $P < 0.05$, $P < 0.01$, and $P < 0.001$, respectively. The network analysis showed the correlation networks among microorganisms at species levels. The circle represents the microorganism species, the circle size represents the average abundance of the species, the line represents the correlation between the two species, the thickness of the line represents the strength of the correlation, and the colour of the line: orange represents positive correlation, green means negative correlation.

## RESULTS

The pH, LBC, microbial population, chemical and protein compositions, energy, and macro-mineral values of WP, NG, and CS before ensiling are shown in **Table 1**. The WSC content was lower ($P < 0.001$) and the LBC was higher ($P < 0.001$) in WP than in both forages. The WSC content was significantly lower ($P < 0.001$) in NG than in CS, but LBC showed the opposite pattern ($P < 0.01$). The LAB count was higher ($P < 0.001$) in CS than in WP and NG. Aerobic bacteria were predominant in all samples, and their counts were higher (all $P < 0.05$) than the counts of other microorganisms. The yeast count was lower ($P < 0.05$) in WP than in forages, whereas coliform bacteria showed the opposite pattern. The mould counts of all samples were similar at 3 lg cfu/g of FM. The DM contents in WP, NG, and CS ranged from 21.18 to 38.94%. The CP contents in gliricidia and leucaena were >25.91% of the contents in DM; the CP contents were higher ($P < 0.001$) in gliricidia and leucaena than in NG and CS. The EE and ADL contents were higher ($P < 0.001$) in WP than in NG and CS, but the NDF and ADF contents showed the opposite pattern. The effective protein, GE, DE, ME, calcium, magnesium, and potassium contents were higher (all $P < 0.001$) in gliricidia and leucaena than in NG and CS, whereas the binding protein and NDIP contents showed the opposite pattern. The phosphorous contents were similar in all samples.

The chemical and protein compositions and CP loss values of WP, NG, CS, and their mixture silages are shown in **Table 2**. For the WP mixture silages, as the proportion of NG or CS increased, the OM, CP, EE, and ADL contents decreased; conversely, the DM, NDF, and ADF contents and CP loss increased. The WP, NG, and their mixture silages had higher (all $P < 0.05$) CP losses than did the CS and its mixture silages.

The microbial population and fermentation quality of WP, NG, CS, and their mixture silages are shown in **Table 3**. Moulds were below the detectable level (<2 lg cfu/g of FM). LAB (7 lg cfu/g of FM) was predominant in CS and its mixture silages prepared with WP; LAB values were higher (all $P < 0.05$) in these samples than in other silages. Aerobic bacteria and coliform bacteria were present in WP, NG, and their mixture silages at 4–7 lg cfu/g of FM, but they were below detectable levels in CS and its mixture silages. Compared with NG, CS and its mixture silages had better fermentation quality with a lower ($P < 0.001$) pH and lower (both $P < 0.001$) acetic acid and $NH_3$-N contents, as well as a higher ($P < 0.05$) lactic acid content. Propionic acid and butyric acid were produced in WP, NG, and their mixture silages, but they were below detectable levels in CS and its mixture silages.

**TABLE 1 |** pH, LBC, microbial population, chemical and protein compositions, energy, and macro-mineral of material.

| Items | Gliricidia | Leucaena | NG | CS | *P*-value |
|---|---|---|---|---|---|
| pH | 6.38*b* | 6.31*b* | 6.47*a* | 6.60*a* | 0.02 |
| LBC (mEq/kg of DM) | 577.16*a* | 508.18*a* | 332.10*b* | 266.62*c* | <0.001 |
| **Microbial population (lg cfu/g of FM)** | | | | | |
| Lactic acid bacteria | 4.04*b* | 4.02*b* | 4.00*b* | 5.14*a* | <0.001 |
| Aerobic bacteria | 8.28*a* | 8.10*a* | 7.33*b* | 7.43*b* | 0.03 |
| Coliform bacteria | 5.73*b* | 5.66*b* | 3.88*c* | 6.11*a* | 0.04 |
| Yeast | 4.51*b* | 4.44*b* | 5.35*a* | 5.45*a* | 0.04 |
| Mould | 3.07 | 3.13 | 3.22 | 3.37 | 0.06 |
| **Chemical composition** | | | | | |
| DM (%) | 24.92*b* | 21.18*b* | 38.94*a* | 36.88*a* | <0.001 |
| OM (% of DM) | 90.30*d* | 92.62*b* | 91.19*c* | 93.26*a* | <0.001 |
| CP (% of DM) | 25.91*b* | 26.31*a* | 7.99*c* | 7.27*c* | <0.001 |
| EE (% of DM) | 4.02*a* | 3.41*b* | 2.69*c* | 1.85*d* | <0.001 |
| NDF (% of DM) | 52.10*c* | 60.62*b* | 66.74*a* | 65.05*a* | <0.001 |
| ADF (% of DM) | 34.52*d* | 37.49*c* | 45.92*a* | 40.53*b* | <0.001 |
| ADL (% of DM) | 11.09*b* | 13.40*a* | 5.03*c* | 4.53*d* | <0.001 |
| WSC (% of DM) | 4.62*c* | 4.97*c* | 5.53*b* | 10.48*a* | <0.001 |
| **Protein composition (% of CP)** | | | | | |
| Binding protein | 13.64*c* | 16.60*b* | 23.98*a* | 22.64*a* | <0.001 |
| NDIP | 37.40*b* | 34.69*c* | 42.20*a* | 42.76*a* | <0.001 |
| Effective protein | 76.77*a* | 74.95*b* | 69.01*c* | 68.81*c* | <0.001 |
| **Energy (MJ/kg of DM)** | | | | | |
| GE | 20.07*a* | 20.94*a* | 14.03*c* | 15.44*b* | <0.001 |
| DE | 12.55*a* | 11.62*b* | 10.01*c* | 10.73*c* | <0.001 |
| ME | 9.47*a* | 9.53*a* | 8.07*b* | 8.30*b* | <0.001 |
| **Macro-mineral (g/kg of DM)** | | | | | |
| Calcium | 1.57*a* | 1.26*b* | 0.84*c* | 0.53*c* | <0.001 |
| Phosphorous | 0.27 | 0.23 | 0.42 | 0.19 | 0.06 |
| Magnesium | 0.59*a* | 0.42*b* | 0.32*c* | 0.20*c* | <0.001 |
| Potassium | 2.64*a* | 2.56*b* | 2.42*c* | 1.69*d* | <0.001 |

*a–d: Data are means of three samples, means in the same row followed by different letters differ ($P < 0.05$).*

*NG, Napier grass; CS, corn stover; LBC, lactic acid buffering capacity; DM, dry matter; cfu, colony-forming unit; FM, fresh matter; OM, organic matter; CP, crude protein; EE, ether extract; NDF, neutral detergent fibre; ADF, acid detergent fibre; ADL, acid detergent lignin; WSC, water-soluble carbohydrate; NDIP, neutral detergent insoluble protein; GE, gross energy; DE, digestible energy; ME, metabolisable energy.*

A Venn diagram of the OTU at 97% sequence identity in WP prepared with NG and CS before and after ensiling is shown in **Figure 2**. The dominant microbiome of the gliricidia, NG, and CS materials (**Figure 2A**) contained 138 shared OTU, as well as 21, 48, and 100 unique OTU, respectively. The leucaena, NG, and CS materials (**Figure 2B**) contained 152 shared OTU and 25, 34, and 81 unique OTU, respectively. The WP silages shared 12 and 11 OTU with the gliricidia (**Figure 2C**) and Leucaena (**Figure 2D**) silages prepared with NG and CS, respectively. As seen in **Figure 2C**, the unique OTU ranged from 2 to 93 in CS, NG, gliricidia, and their mixture silages. As seen in **Figure 2D**, the unique OTU ranged from 2 to 50 in CS, NG, leucaena, and their mixture silages.

**TABLE 2 |** Chemical and protein compositions, and CP loss of WP, NG, CS, and their mixture silages.

| Items | DM (%) | Chemical composition (% of DM) | | | | | | CP loss (%) |
|---|---|---|---|---|---|---|---|---|
| | | OM | CP | EE | NDF | ADF | ADL | |
| Gliricidia | 21.34d | 90.89e | 23.20a | 3.67a | 50.02e | 33.91g | 10.24b | 19.70a |
| Leucaena | 20.31d | 91.36c | 24.82a | 3.32b | 57.36b | 34.99f | 13.39a | 19.32a |
| NG | 38.78c | 90.11f | 6.52f | 1.77h | 64.01a | 44.70a | 4.72i | 18.40a |
| CS | 37.25c | 91.40c | 6.47f | 1.78h | 63.58a | 38.89e | 3.87j | 5.36b |
| **Gliricidia + NG** | | | | | | | | |
| 90 + 10 | 32.54c | 91.10d | 18.65c | 3.14c | 53.08c | 40.22c | 8.06e | 19.04a |
| 75 + 25 | 47.94b | 90.97e | 15.26d | 2.46f | 55.28c | 43.07b | 6.55g | 19.42a |
| 50 + 50 | 56.93a | 90.75e | 13.48e | 2.09g | 59.42b | 44.78a | 5.64h | 19.65a |
| **Gliricidia + CS** | | | | | | | | |
| 90 + 10 | 31.65c | 92.98a | 20.89b | 3.36b | 51.81d | 37.93ef | 8.27de | 6.09b |
| 75 + 25 | 43.79b | 92.56a | 18.59c | 3.02d | 53.76c | 39.23d | 6.40g | 6.28b |
| 50 + 50 | 52.71a | 91.87b | 15.14d | 2.82de | 58.58b | 40.04c | 5.28g | 6.60b |
| **Leucaena + NG** | | | | | | | | |
| 90 + 10 | 34.81c | 91.44c | 17.27c | 2.69e | 53.68c | 41.48c | 9.35c | 18.82a |
| 75 + 25 | 45.21b | 91.64b | 14.54d | 2.23f | 55.21c | 43.47b | 7.32f | 19.40a |
| 50 + 50 | 55.87a | 91.97b | 12.86e | 1.95g | 56.13b | 44.66a | 6.10gh | 19.61a |
| **Leucaena + CS** | | | | | | | | |
| 90 + 10 | 36.67c | 92.91a | 21.53b | 3.19c | 52.84d | 38.97e | 8.95d | 6.12b |
| 75 + 25 | 45.66b | 92.87a | 19.03c | 3.00d | 53.94c | 39.95cd | 6.58g | 6.76b |
| 50 + 50 | 53.97a | 92.85a | 14.86d | 2.92de | 54.61bc | 40.53c | 5.16h | 6.84b |
| SEM | 0.14 | 0.20 | 1.46 | 0.17 | 0.86 | 0.91 | 0.22 | 0.39 |
| *P*-value | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 |

*a–j: Data are means of three samples, means in the same column followed by different letters differ ($P < 0.05$).*
*90 + 10, 75 + 25, 50 + 50, indicated the mixing ratio (%) of silage based on fresh matter; CP, crude protein; WP, woody plant; NG, Napier grass; CS, corn stover; DM, dry matter; OM, organic matter; EE, ether extract; NDF, neutral detergent fibre; ADF, acid detergent fibre; ADL, acid detergent lignin; NDIP, neutral detergent insoluble protein; SEM, standard error of the mean.*

Relative bacterial abundances at the species level in the WP, NG, and CS materials, and their mixture silages are shown in **Figure 3**. The dominant species in the WP material was *Pantoea agglomerans*. The dominant species in the NG material were *Microbacterium trichothecenolyticum* and *P. agglomerans*; the dominant species in CS material were *M. trichothecenolyticum*, *Streptococcus sanguinis*, and *Methylobacterium adhaesivum*. *Lactiplantibacillus plantarum* was present at a low level in all materials. After ensiling, *L. plantarum* was the predominant species in all silages, and its counts were higher (both $P < 0.05$) in NG and CS silages than in WP silage; CS silage had the highest number. Compared with the WP silages, the addition of NG and CS significantly increased (both $P < 0.05$) the proportion of *L. plantarum* in the mixture silages, with the highest relative abundances in the WP and CS mixture silages.

Correlation heatmap and hierarchical cluster analysis of the bacterial community at the species level and terminal fermentation products are shown in **Figure 4**. Lactic acid was positively correlated with *Lactobacillus* and *Weissella* species, such as *L. plantarum*, *Levilactobacillus brevis*, *Limosilactobacillus fermentum*, and *Weissella paramesenteroides*. LAB was positively correlated with *L. plantarum*, but negatively correlated with *P. agglomerans*. pH showed the opposite patterns. $NH_3$-N content was negatively correlated with *L. plantarum* and positively correlated with *Akkermansia muciniphila*.

The correlation networks among all microorganisms at the species level are shown in **Figure 5**. The most important species appeared to be *L. plantarum*, which was positively correlated with *M. trichothecenolyticum* and negatively correlated with both *Kosakonia cowanii* and *Capnocytophaga granulosa*. The second most important species was *P. agglomerans*, which was positively correlated with *Porphyromonas pasteri* and negatively correlated with *M. adhaesivum*.

The KEGG pathways enriched in gliricidia material and silage are shown in **Figure 6A**; the KEGG pathways enriched in gliricidia silage and gliricidia + CS mixed silage are shown in **Figure 6B**. **Figure 6A** shows that the carbohydrate metabolism, amino acid metabolism, and energy metabolism pathways were the predominant metabolic categories. Among them, the largest proportion of the metabolism pathway involved carbohydrate metabolism. The proportion of the carbohydrate metabolism pathway was higher ($P < 0.01$), whereas the proportions of amino acid and energy metabolism pathways were lower (both $P < 0.01$) in gliricidia silage than in gliricidia material. **Figure 6B** shows that the biosynthesis of secondary metabolites, biosynthesis of antibiotics, and microbial metabolism in diverse environments were the predominant metabolic categories in gliricidia silage and mixture silages. Compared with the gliricidia silage, the mixture silage had lower ($P < 0.01$) proportions of the above three metabolic pathways and a higher ($P < 0.01$) proportion of the tricarboxylic acid (TCA) cycle.

**TABLE 3 |** Microbial population and fermentation quality of WP, NG, CS, and their mixture silages.

| Items | Microbial population (lg cfu/g of FM) | | | | | pH | Organic acid (% of FM) | | | | $NH_3$-N (% of DM) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | LAB | Aerobic bacteria | Coliform bacteria | Yeast | Mould | | Lactic acid | Acetic acid | Propionic acid | Butyric acid | |
| Gliricidia | 6.52e | 5.39c | 5.87b | ND | ND | 5.40b | 0.28d | 0.30c | 0.04b | 0.28c | 0.42c |
| Leucaena | 6.54e | 6.91ab | 4.98c | 5.32a | ND | 5.83a | 0.30d | 0.24d | 0.05b | 0.35b | 0.35d |
| NG | 6.54e | 7.35a | 6.06a | 5.49a | ND | 4.87e | 0.45c | 0.53a | 0.12a | 0.46a | 0.58a |
| CS | 7.58b | ND | ND | 5.31a | ND | 4.09f | 1.44a | 0.34c | ND | ND | 0.26e |
| **Gliricidia + NG** | | | | | | | | | | | |
| 90 + 10 | 6.83d | 7.33a | 5.44b | 5.21a | ND | 5.04d | 0.41c | 0.35c | 0.03b | 0.43a | 0.55a |
| 75 + 25 | 6.68d | 6.64b | 6.21a | 4.31b | ND | 5.22d | 0.30d | 0.45b | 0.06b | 0.35b | 0.50b |
| 50 + 50 | 6.60e | 6.36b | 6.75a | 4.22b | ND | 5.35c | 0.35d | 0.50a | 0.10a | 0.30b | 0.44c |
| **Gliricidia + CS** | | | | | | | | | | | |
| 90 + 10 | 7.90a | ND | ND | 5.75a | ND | 3.72h | 1.45a | 0.34c | ND | ND | 0.41c |
| 75 + 25 | 7.68b | ND | ND | 4.79b | ND | 3.97g | 1.10b | 0.37b | ND | ND | 0.35d |
| 50 + 50 | 7.59b | ND | ND | ND | ND | 4.01f | 0.72c | 0.39b | ND | ND | 0.29e |
| **Leucaena + NG** | | | | | | | | | | | |
| 90 + 10 | 6.97d | 6.33b | 4.84c | ND | ND | 5.48b | 0.44c | 0.29d | 0.05b | 0.46a | 0.51b |
| 75 + 25 | 6.58e | 5.00d | 5.41b | ND | ND | 5.69b | 0.37d | 0.42b | 0.07a | 0.41a | 0.45c |
| 50 + 50 | 6.05f | 5.00d | 5.49b | ND | ND | 5.83a | 0.31d | 0.50a | 0.11a | 0.38b | 0.37d |
| **Leucaena + CS** | | | | | | | | | | | |
| 90 + 10 | 7.82a | ND | ND | 4.80b | ND | 4.01f | 1.45a | 0.25d | ND | ND | 0.27e |
| 75 + 25 | 7.57b | ND | ND | ND | ND | 4.38ef | 0.83b | 0.27d | ND | ND | 0.32d |
| 50 + 50 | 7.16c | ND | ND | ND | ND | 4.47ef | 0.54c | 0.32c | ND | ND | 0.46c |
| SEM | 0.30 | 0.36 | 0.54 | 0.29 | NS | 0.15 | 0.06 | 0.05 | 0.02 | 0.03 | 0.02 |
| *P*-value | <0.001 | <0.001 | <0.001 | <0.001 | NS | <0.001 | <0.001 | 0.02 | <0.001 | <0.001 | <0.001 |

*a–h: Data are means of three samples, means in the same column followed by different letters differ ($P < 0.05$).*
*90 + 10, 75 + 25, 50 + 50, indicated the mixing ratio (%) of silage based on fresh matter; WP, woody plant; cfu, colony-forming unit; FM, fresh matter; LAB, lactic acid bacteria; $NH_3$-N, ammonia nitrogen; NG, Napier grass; CS, corn stover; DM, dry matter; SEM, standard error of the mean; ND, not detected.*

# DISCUSSION

## Microbial Population and Chemical Composition of Material

The desire to reduce feed cost and ensure ruminant production system profitability has prompted the development of WP resources as an alternative source of protein for sustainable animal production. WP are recommended as ruminant feed to improve forage quality, reduce the need to purchase concentrate, and decrease the cost of feeding animals. Generally, animals prefer forage with low fibre content and high NDF digestibility, as well as a high CP content (Egan, 1980), because this type of forage is associated with high animal intake and efficient assimilation of energy, minerals, and vitamins. In this study, as shown in **Table 1**, the WP material was rich in CP (>23.91% of DM) and protein (74.95% of CP); however, it had low fibre, binding protein, and NDIP contents. In addition, the WP material also had high contents of energy and macro-minerals. Therefore, the WP material contains high levels of digestible feed ingredients for use in livestock feed as a source of protein. It is difficult to produce good-quality silage using WP because of their high moisture content (Du et al., 2021b). In addition, the WSC, LAB, and LBC contents of materials affect silage fermentation (Cai et al., 2020a). The moisture content of WP material was >76%, possibly leading to poor fermentation. When NG and CS were added to WP, the moisture content was <70%, within the range (60–70%) suitable for fermentation (McDonald et al., 1991; Du et al., 2021b). A suitable moisture content inhibits the growth of harmful microorganisms such as *Clostridium* species; it also prevents butyric acid fermentation and protein decomposition (Cai et al., 2020b; Du et al., 2021a). For silage fermentation, other important factors are WSC > 5% of DM, LAB > 5 lg cfu/g of FM, and low LBC (Cai et al., 1999). In this study, the LAB count and WSC content in WP material were below these levels. This indicates that a small number of LAB is associated with poor silage fermentation, thus reducing fermentation quality. In addition, WP material has a high LBC value similar to the value of legume alfalfa, which hampers silage production (Xu et al., 2019). During ensiling, these cations neutralise the organic acids formed by silage fermentation, thus preventing pH reduction. Therefore, preparation of high-quality silage by direct ensiling of WP is problematic. NG and CS are locally available feed resources with good fermentation characteristics, such as low LBC count, high WSC, and optimal DM content. The addition of NG and CS can overcome some of the difficulties in WP silage fermentation.

## Fermentation Characteristics of Silage

Woody plant can usually be harvested multiple times during the rainy season. If it is prepared and stored in the manner used for hay, some nutrients will be lost during drying and storage

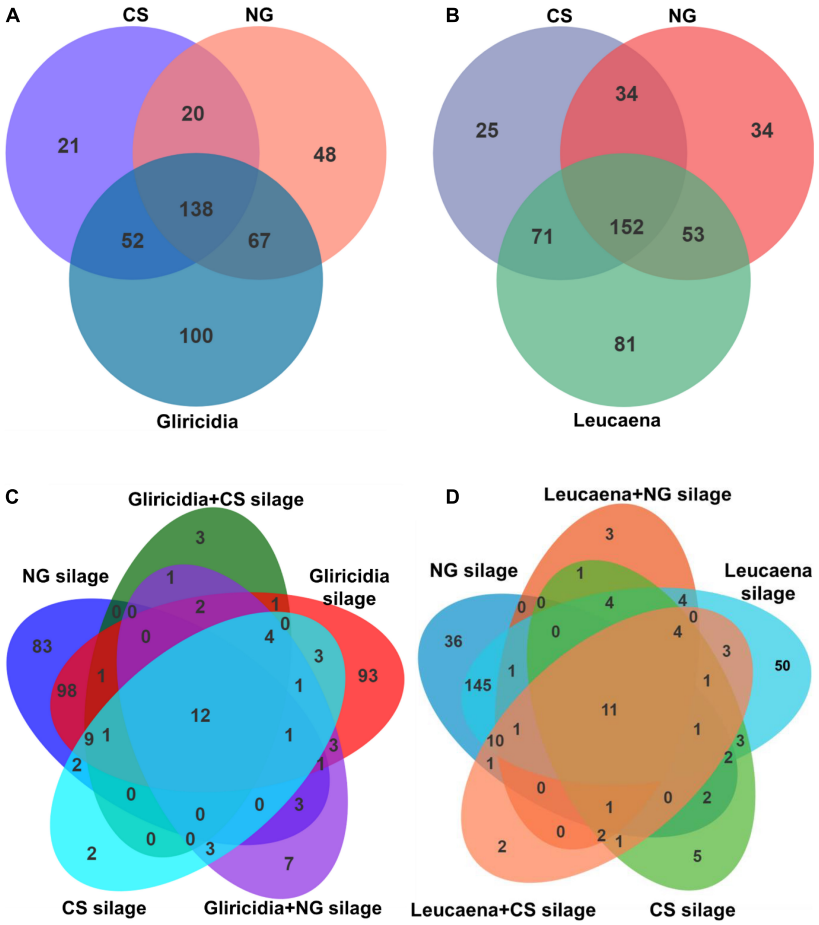



**FIGURE 2 |** Venn diagram depicting unique or shared bacterial OTU (97% sequence identity) in gliricidia, leucaena, NG, CS materials and their mixture silages. **(A)** The number of OTU shared by gliricidia, NG, and CS before ensiling. **(B)** The number of OTU shared by leucaena, NG, and CS before ensiling. **(C)** The number of OTU shared by gliricidia, NG, and CS and their mixture silages. **(D)** The number of OTU shared by leucaena, NG, and CS and their mixture silages. OTU, operational taxonomic unit; NG, Napier grass; CS, corn stover. The mixing ratio of gliricidia or leucaena were 90% of the mixture silage based on a fresh matter.

(Zhang et al., 2019). After drying, the palatability and nutritional value of the leaves for livestock are greatly reduced (Du et al., 2021b). Silage can preserve the nutrients in fresh forage and ensure the supply of animal feed throughout the year (Yan et al., 2019). LAB have important roles in the preparation of silage, converting WSC into lactic acid, lowering the pH value, and inhibiting the growth of harmful bacteria; thus, they produce high-quality silage (Cai et al., 2020a). In this study (**Tables 2**, **3**), particularly in the gliricidia or leucaena + CS (90 + 10) mixed silages, the LAB population significantly increased after 60 days of ensiling, inhibiting the growth of harmful bacteria. Adjustment of the moisture content and enhancement of fermentation substrate levels are methods that improve fermentation quality. Therefore, when using WP to prepare silage, the recommended grass or crop stover proportion is approximately 10%; this should be adjusted to 60–70% depending on moisture content. After harvest, fresh CS is generally rich in LAB and WSC, and the addition of CS can improve the fermentation quality of WP mixture silage. However, NG itself does not have good fermentation characteristics because of its low WSC content and low LAB count. The NG silage and WP + NG mixed silage did not show improved fermentation quality. Therefore, CS is recommended for the preparation of mixture silage with WP in Africa.

## Microbial Community of Material and Silage

Silage fermentation is a dynamic process of microbial community succession and shifts in metabolite levels. Using traditional culture methods, only some microorganisms in silage can be isolated and identified; this hampers the analysis of fermentation mechanisms. Next-generation sequencing (NGS) technology has been used to evaluate silage microbial community, abundance, and diversity. It is a high-throughput approach that can detect a large number of microbial taxa. However, microorganisms can be identified only to the genus level, and the complete microbial community cannot be profiled based on partial 16S rRNA gene sequences (Bokulich and Mills, 2012; Mayo et al., 2014). The SMRT sequencing technology used in this study enables accurate evaluation of the microbial community and diversity

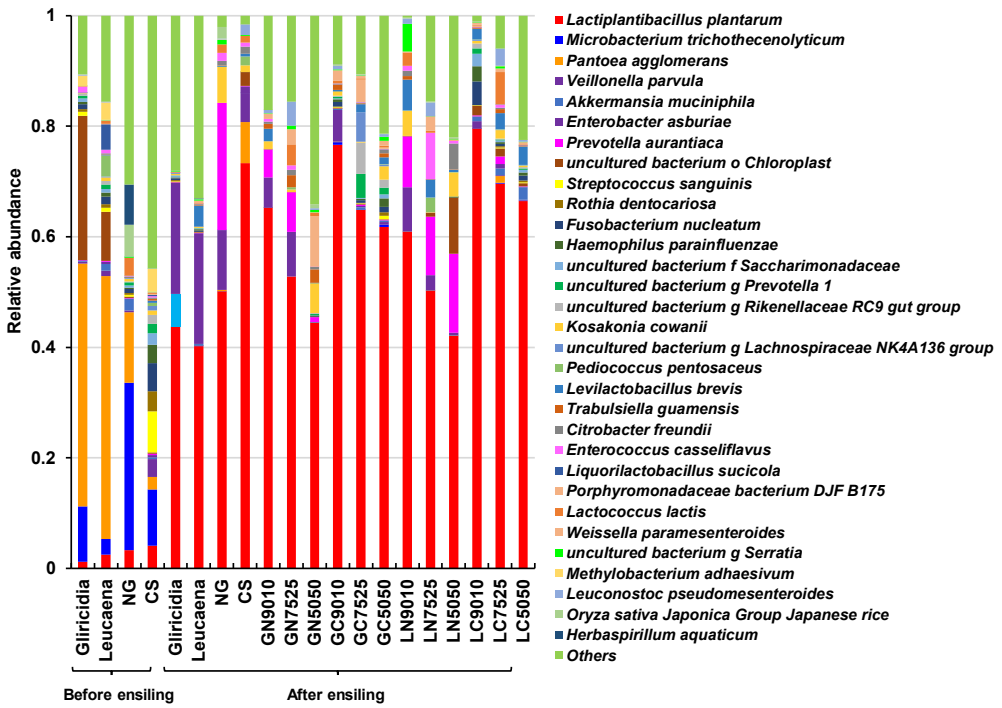



**FIGURE 3 |** The relative abundances of the bacterial communities at the species levels in the gliricidia, Leucaena, NG, and CS materials, and their mixture silages. NG, Napier grass; CS, corn stover; GN, gliricidia + NG; GC, gliricidia + CS; LN, leucaena + NG; LC, leucaena + CS; 9010, 90 + 10%; 8020, 80 + 20%; 7030, 70 + 30%.

within silage (Xu et al., 2019). As shown in **Figure 1**, compared with the NG and CS materials, the greater OTU numbers in gliricidia and leucaena may reflect epiphytic microorganisms on the surface, because of the high moisture content and rich nutrition of fresh branches and leaves. After ensiling, natural LAB create an acidic environment to inhibit core microorganisms, reducing their abundance and forming a unique microbial community in the silages.

Microbial community structure and function affect silage fermentation (McEniry et al., 2010; Guan et al., 2018; Cai et al., 2020b). As shown in **Figure 2**, the Gram-negative bacterium *P. agglomerans* was the major species in WP materials. This plant pathogen is often present in forage and is considered harmful in silage (Du et al., 2021a). The dominant species in the NG and CS materials was *M. trichothecenolyticum*, a Gram-positive and aerobic or weakly anaerobic non-spore-producing and acid-fast species. In contrast, the abundance of *L. plantarum* was low in all materials, indicating that WP were similar to forage and crop by-products; WP typically have a low abundance of LAB but a high abundance of undesirable bacteria before ensiling (Tian et al., 2020; Du et al., 2021b). Therefore, measures to control the microbial community are needed to improve silage fermentation.

Silage fermentation is conducted in an anaerobic, closed, solid-fermentation system (Cai et al., 1998). A higher relative abundance of LAB is frequently observed in this system, and the different microbial community structure have differential effects on the silage fermentation (Queiroz et al., 2013; Cai et al., 2020b). In this study (**Figure 3**), when WP and CS were mixed to prepare WP silage, the LAB can quickly become dominant bacteria and complete the succession process from Gram-negative bacteria to Gram-positive bacteria, indicating that mixture silage can effectively improve the fermentation quality by improving the microbial community. However, there was not great difference between the microbial communities in the NG silage and WP + NG mixed silages. A reasonable explanation for this phenomenon is that, compared with NG, CS contains more lactic acid bacteria, such as *L. plantarum*, which have stronger fermentation ability and acid tolerance. In the process of mixture silage fermentation, these epiphytic LAB existing on CS can quickly respond to the double pressure of anaerobic and acidic silage environment, carry out lactic acid fermentation and improve the fermentation quality of silage (Wang et al., 2019; Du et al., 2021a).

## Interaction Between Bacterial Community and Fermentation Product

Microorganisms affect silage fermentation through a series of metabolites. For example, *Lactobacillus* species use WSC to produce lactic acid, while *Enterobacter* species ferment lactic acid into products such as acetic acid (McDonald et al., 1991), thereby influencing silage fermentation quality and aroma. In addition, these metabolites affect the microbial community structure. For example, lactic acid and acetic acid lower the pH and inhibit the growth of aerobic bacteria and moulds, thus affecting the aerobic stability of silage. Accordingly, silage fermentation is

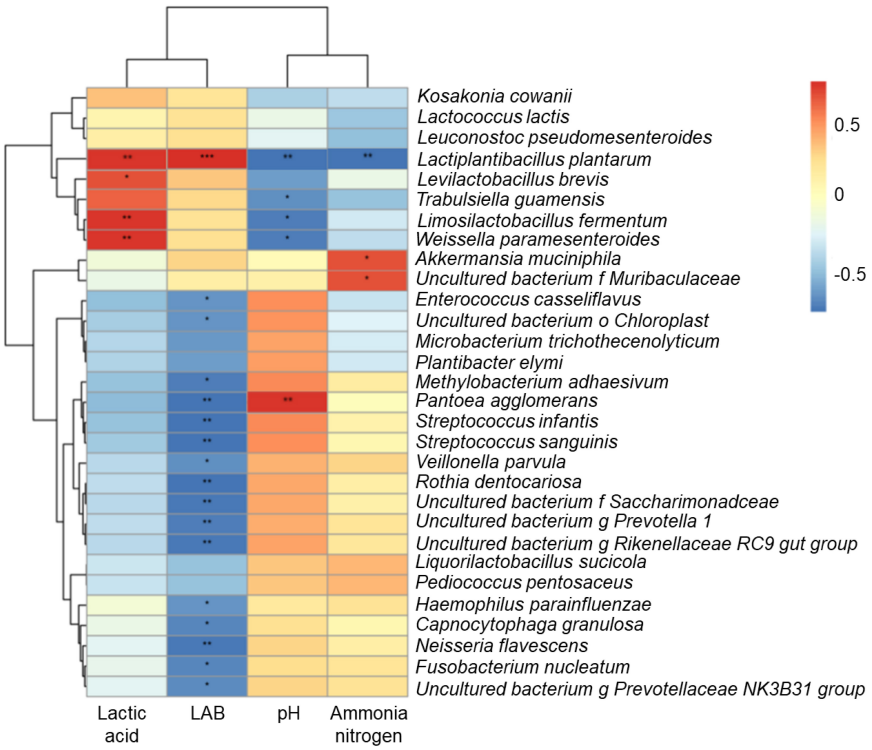



**FIGURE 4 |** Correlation heatmap and hierarchical cluster analysis between bacterial community and terminal fermentation products at species level. LAB, lactic acid bacteria; $NH_3$-N, ammonia nitrogen. $^*P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$.

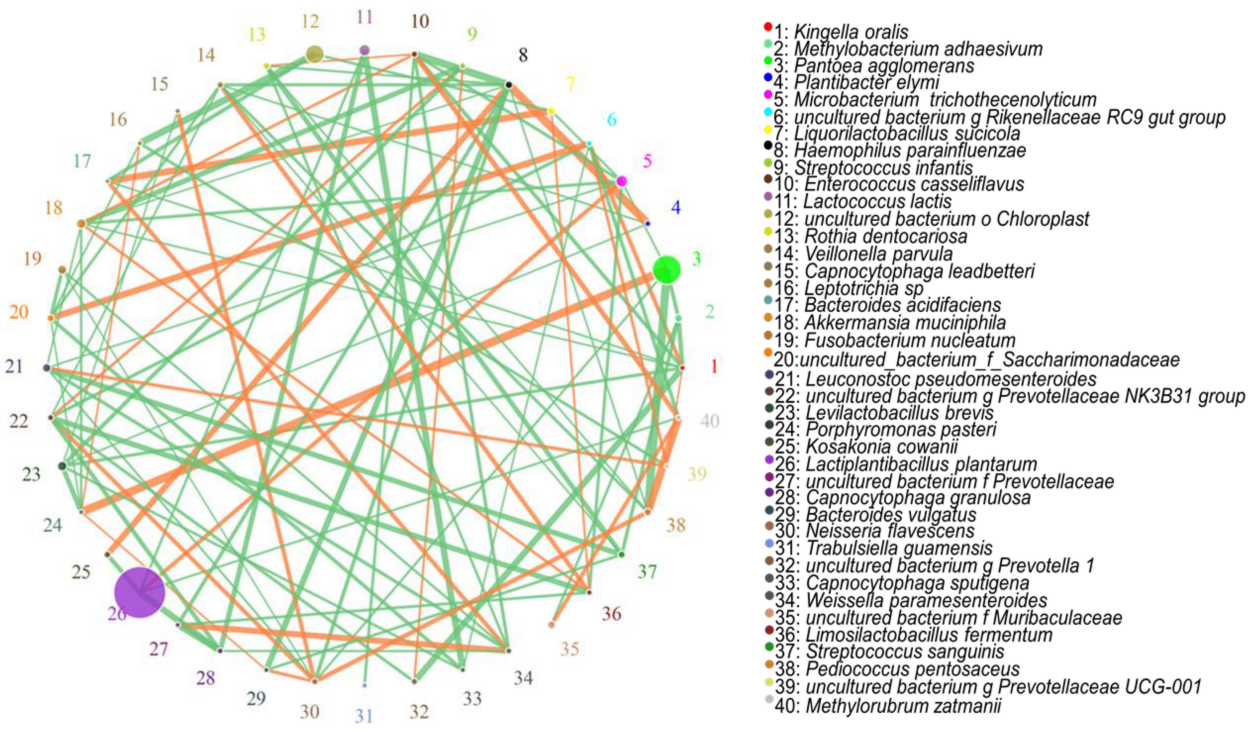



**FIGURE 5 |** Correlation networks among all the microorganisms at the species level.

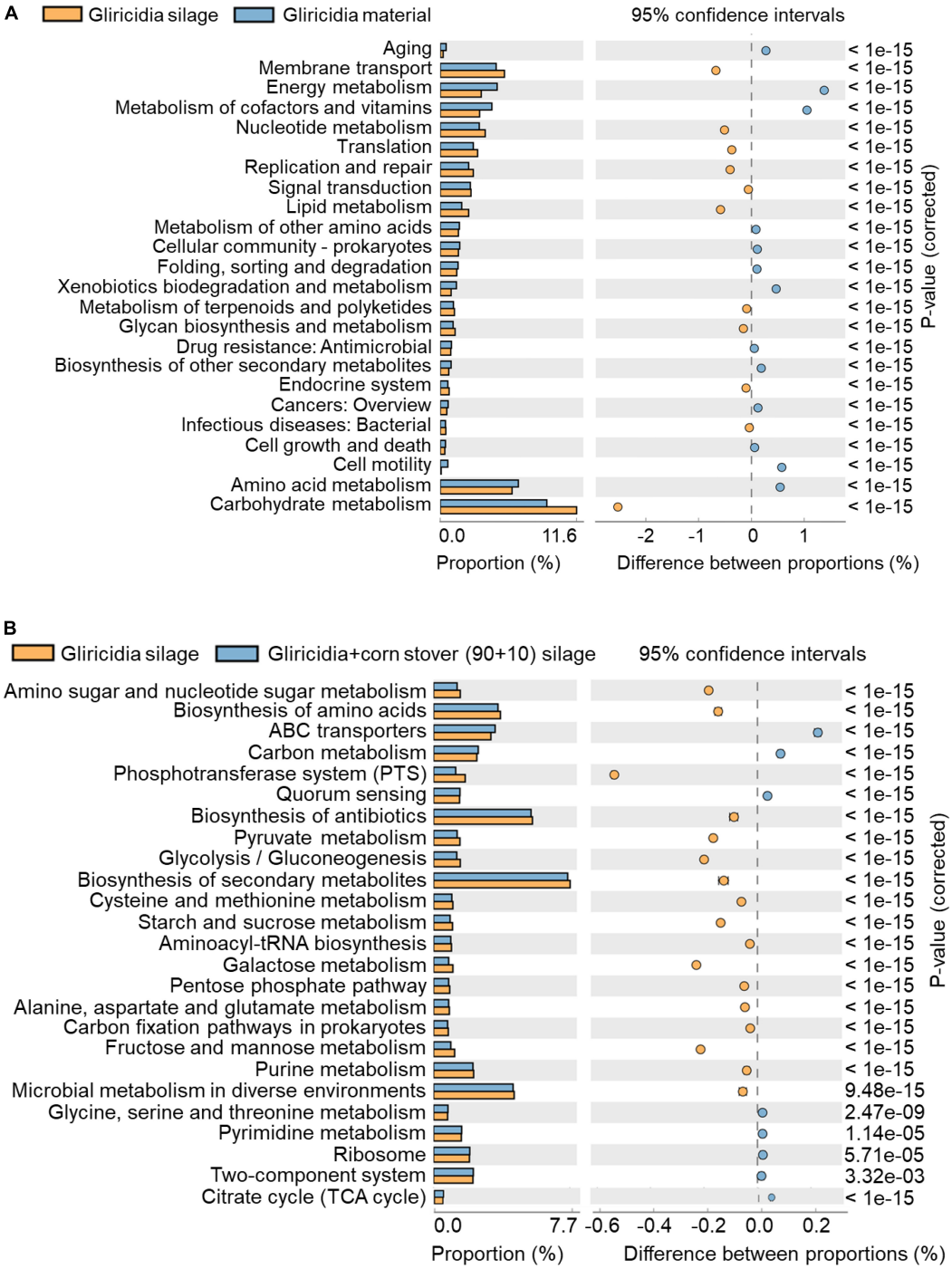



**FIGURE 6 |** KEGG metabolic pathways in gliricidia material and silage **(A)**, and gliricidia silage and gliricidia + CS (90 + 10) mixed silage **(B)**. KEGG, Kyoto Encyclopedia of Genes and Genomes.

modulated by microbial community structure and metabolites, which interact during fermentation.

In this study, the production of lactic acid was attributed to *Lactobacillus* species. Lactobacilli are homofermentative type LAB that inhibit the activities of harmful microorganisms by rapid acidification in the later stage of silage fermentation (Cai, 2004; Mu et al., 2020). In the **Figure 4**, *L. plantarum, L. brevis, L. fermentum, and W. paramesenteroides* grew well and produced lactic acid in silage, inhibiting the growth of *P. agglomerans* by reducing the pH. The *A. muciniphila* competed for nutrients with *L. plantarum* during ensiling, while degrading proteins to produce $NH_3$-N. However, silage fermentation by

*L. plantarum* inhibited the growth of these bacteria and the production of $NH_3$-N.

In the **Figure 5**, during silage fermentation, *L. plantarum* was positively correlated with *M. trichothecenolyticum* and negatively correlated with both *K. cowanii* and *C. granulosa*. The nature of this relationship is unclear; however, *M. trichothecenolyticum* is dependent on respiratory metabolism that produces acid from glucose and other sugars during the ensiling process. Therefore, *L. plantarum* inhibits the growth of other Gram-negative bacteria; it may promote the growth of *M. trichothecenolyticum*. The abundances of *K. cowanii* and *C. granulosa* were negatively correlated with the abundance of *L. plantarum*. These bacteria are often distributed on the surface of forage and compete with LAB for nutrients in the early stage of silage fermentation. As silage fermentation progresses, they are inhibited by the lactic acid from LAB. The *P. agglomerans* and *P. pasteri* are frequently present in silage fermentation; they produce n-butyric acid, acetic acid, propionic acid, isobutyric acid, isovaleric acid, and phenylacetic acid, thus reducing fermentation quality. Therefore, the two bacteria are positively correlated with each other, but both are negatively correlated with *L. plantarum*.

### Kyoto Encyclopedia of Genes and Genomes Metabolism Pathway Analysis

In the KEGG analysis, carbohydrate metabolism was the most important metabolic category in gliricidia materials and silage. Pyruvate metabolism, glycolysis, and butyrate metabolism were predominant components of the carbohydrate metabolic pathways, consistent with the main silage metabolism pathway reported by Du et al. (2021b). In this study (**Figure 6**), silage had a higher proportion of carbohydrate metabolism pathways than did the corresponding material, reflecting the community dynamics and metabolic activities of the major LAB. Amino acid metabolism can be oxidised into carbon dioxide and water through the tricarboxylic acid cycle, releasing energy (Halperin and Jungas, 1983). Therefore, fresh materials have higher rates of amino acid metabolism and energy metabolism pathways than do silages.

The roles of secondary metabolites and antibiotics in silage are unclear. Secondary metabolites and antibiotics may be produced by microorganisms—including bacteria, fungi, and actinomycetes—during silage fermentation, which interfere with the development of other cells. LAB grew more vigorously in mixture silage than in gliricidia silage, inhibiting the growth and metabolic activities of other microorganisms while suppressing the biosynthesis of secondary metabolites and antibiotics. However, the proportion of citric acid cycle was higher in the gliricidia + CS mixed silage than in gliricidia silage. In the citric acid cycle, the condensation reaction between oxaloacetate and the acetyl group of acetyl-CoA produces citric acid. Citric acid is used as a food acidifier, flavouring agent, and chelating agent. Citric acid-mediated adjustment of pH value can improve the performance of antioxidants, inhibit the activity of enzymes, and extend the shelf life of food (Ke et al., 2017). The WP and CS mixture silage has an increased citric acid content, which will improve its flavour and quality, facilitating long-term storage.

## CONCLUSION

To improve fermentation quality and develop a preparation method for nutrient-rich WP, we used SMRT sequencing technology to study the microbial community and fermentation characteristics of WP silage following the addition of NG and CS in Southern Africa. WP have high CP, energy, and mineral contents; they offer ideal feed for ruminants. However, the preparation of good-quality WP silage is hampered by the low WSC content, low LAB count, and high LBC in WP. In PM silage prepared with NG and CS, the dominant microbial community shifted from Gram-negative bacteria to LAB, promoting lactic acid fermentation. *L. plantarum* was predominant, which improved silage flavour and quality. CS has good silage characteristics, helping to overcome the influence of unfavourable factors in silage fermentation. Addition of CS improves the microbial community structure and metabolic pathways of silage; thus, the combination of CS and WP is suitable for high-quality silage. WP offer a potential source of high-protein feed for ruminants, which can alleviate the feed shortage in Southern Africa.

## 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

ZD and YC conceived and designed the study. ZD, SY, and TO carried out the experiments. ZD and DN carried out the data analysis. ZD, YC, DE, BT, and FM wrote and revised the manuscript. All authors read and approved the manuscript.

## FUNDING

This work was supported by the project "Development of Sustainable Technologies to Increase Agricultural Productivity and Improve Food Security in Africa (Food Security in Africa)," Japan International Research Center for Agricultural Sciences (JIRCAS), Japan, and Instituto de Investigação Agrária de Moçambique (IIAM), Mozambique.

## ACKNOWLEDGMENTS

We would like to thank Japan International Research Center for Agricultural Sciences (JIRCAS, Japan) and Instituto de Investigação Agrária de Moçambique (IIAM, Mozambique). Plant samples imported from Mozambique to Japan were permitted by the Minister of Agriculture, Forestry and Fisheries of Japan under Act on Domestic Animal Infectious Diseases Control.

## REFERENCES


AOAC International (2000). *Official Methods of Analysis*, 17th Edn. Arlington, VA: Association of Official Analytical Chemists.

Baker, K. M., and Stein, H. H. (2009). Amino acid digestibility and concentration of digestible and metabolizable energy in soybean meal produced from conventional, high-protein, or low-oligosaccharide varieties of soybeans and fed to growing pigs. *J. Anim. Sci.* 87, 2282–2290. doi: 10.2527/jas.2008-1414

Bokulich, N. A., and Mills, D. A. (2012). Next-generation approaches to the microbial ecology of food fermentations. *BMB Rep.* 45, 377–389. doi: 10.5483/bmbrep.2012.45.7.148

Cai, Y. (2004). "Analysis method for silage," in *Field and Laboratory Methods for Grassland Science*, ed. Japanese Society of Grassland Science (Tokyo: Tosho Printing Co. Ltd), 279–282.

Cai, Y., Benno, Y., Ogawa, M., and Kumai, S. (1999). Effect of applying lactic acid bacteria isolated from forage crops on fermentation characteristics and aerobic deterioration of silage. *J. Dairy Sci.* 82, 520–526.

Cai, Y., Benno, Y., Ogawa, M., Ohmomo, S., Kumai, S., and Nakase, T. (1998). Influence of Lactobacillus spp. from an inoculant and of *Weissella* and *Leuconostoc* spp. from forage crops on silage fermentation. *Appl. Environ. Microbiol.* 64, 2982–2987. doi: 10.1128/AEM.64.8.2982-2987.1998

Cai, Y., Du, Z., Jethro, B. D., Nignan, M., and Yamasaki, S. (2020a). Analysis of main factors affecting silage fermentation of sorghum prepared with whole crop and stover in semiarid West Africa. *Afr. J. Range For. Sci.* 38, 169–178. doi: 10.2989/10220119.2020.1794959

Cai, Y., Du, Z., Yamasaki, S., Nguluve, D., Tinga, B., Macome, F., et al. (2020b). Community of natural lactic acid bacteria and silage fermentation of corn stover and sugarcane tops in Africa. *Asian Australas. J. Anim. Sci.* 33, 1252–1264. doi: 10.5713/ajas.19.0348

Caporaso, J. G., Bittinger, K., Bushman, F. D., DeSantis, T. Z., Andersen, G. L., and Knight, R. (2010a). PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* 26, 266–267. doi: 10.1093/bioinformatics/btp636

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010b). QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336. doi: 10.1038/nmeth.f.303

Du, Z., Sun, L., Chen, C., Lin, J., Yang, F., and Cai, Y. (2021a). Exploring microbial community structure and metabolic gene clusters during silage fermentation of paper mulberry, a high-protein woody plant. *Anim. Feed Sci. Technol.* 275:114766. doi: 10.1016/j.anifeedsci.2020.114766

Du, Z., Sun, L., Lin, Y., Yang, F., and Cai, Y. (2021b). The use of PacBio SMRT technology to explore the microbial network and fermentation characteristics of woody silage prepared with exogenous carbohydrate additives. *J. Appl. Microbiol.* 131, 2193–2211. doi: 10.1111/jam.15124

Du, Z., Yamasaki, S., Oya, T., Nguluve, D., Tinga, B., Macome, F., et al. (2020). Ensiling characteristics of total mixed ration prepared with local feed resources in Mozambique and their effects on nutrition value and milk production in Jersey dairy cattle. *J. Anim. Sci.* 91, e13370. doi: 10.1111/asj.13370

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, 2460–2461. doi: 10.1093/bioinformatics/btq461

Egan, A. (1980). Host animal-rumen relationships. *Proc. Nutr. Soc.* 39, 79–87. doi: 10.1079/pns19800011

Gao, J., Wang, P., Zhou, C., Li, P., Tang, H., Zhang, J., et al. (2019). Chemical composition and in vitro digestibility of corn stover during field exposure and the fermentation characteristics of silage prepared with microbial additives. *Asian Australas. J. Anim. Sci.* 32, 1854–1863. doi: 10.5713/ajas.18.0886

Guan, H., Yan, Y., Li, X., Li, X., Shuai, Y., Feng, G., et al. (2018). Microbial communities and natural fermentation of corn silages prepared with farm bunker-silo in Southwest China. *Bioresour. Technol.* 265, 282–290. doi: 10.1016/j.biortech.2018.06.018

Haas, B. J., Gevers, D., Earl, A. M., Feldgarden, M., Ward, D. V., Giannoukos, G., et al. (2011). Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res.* 21, 494–504. doi: 10.1101/gr.112730.110

Halperin, M. L., and Jungas, R. L. (1983). Metabolic production and renal disposal of hydrogen ions. *Kidney Int.* 24, 709–713. doi: 10.1038/ki.1983.217

Ke, W., Ding, W., Xu, D., Ding, L., Zhang, P., Li, F., et al. (2017). Effects of addition of malic or citric acids on fermentation quality and chemical characteristics of alfalfa silage. *J. Dairy Sci.* 100, 8958–8966. doi: 10.3168/jds.2017-12875

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559. doi: 10.1186/1471-2105-9-559

Langille, M. G. I., Zaneveld, J., Caporaso, J. G., McDonald, D., Knights, D., Reyes, J. A., et al. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat. Biotechnol.* 31, 814–821. doi: 10.1038/nbt.2676

Li, E., Zhao, J., Liu, L., and Zhang, S. (2018). Digestible energy and metabolizable energy contents of konjac flour residues and ramie in growing pigs. *Anim. Nutr.* 4, 228–233. doi: 10.1016/j.aninu.2018.01.001

Lozupone, C., and Knight, R. (2005). UniFrac: a new phylogenetic method for comparing microbial communities. *Appl. Environ. Microbiol.* 71, 8228–8235. doi: 10.1128/AEM.71.12.8228-8235.2005

Luscher, A., Mueller-Harvey, I., Soussana, J. F., Rees, R. M., and Peyraud, J. L. (2014). A review: potential of legume-based grassland–livestock systems in Europe. *Grass Forage Sci.* 69, 206–228. doi: 10.1111/gfs.12124

Mansfeldt, C. B., Rowe, A. R., Heavner, G. L. W., Zinder, S. H., and Richardson, R. E. (2014). Meta-analyses of *Dehalococcoides mccartyi* Strain 195 transcriptomic profiles identify a respiration rate-related gene expression transition point and interoperon recruitment of a key oxidoreductase subunit. *Appl. Environ. Microb.* 80, 6062–6072. doi: 10.1128/AEM.02130-14

Mayo, B., Rachid, C., Alegria, A., Leite, A., Peixoto, R., and Delgado, S. (2014). Impact of next generation sequencing techniques in food microbiology. *Curr. Genomics* 15, 293–309. doi: 10.2174/1389202915666140616233211

McDonald, P., Henderson, A. R., and Heron, S. J. E. (1991). *The Biochemistry of Silage*, 2nd Edn. Marlow: Chalcombe Publications.

McEniry, J., O'Kiely, P., Clipson, N. J. W., Forristal, P. D., and Doyle, E. M. (2010). Assessing the impact of various ensilage factors on the fermentation of grass silage using conventional culture and bacterial community analysis techniques. *J. Appl. Microbiol.* 108, 1584–1593. doi: 10.1111/j.1365-2672.2009.04557.x

Mosher, J. J., Bernberg, E. L., Shevchenko, O., Kan, J., and Kaplan, L. A. (2013). Efficacy of a 3rd generation high-throughput sequencing platform for analyses of 16S rRNA genes from environmental samples. *J. Microbiol. Methods* 95, 175–181. doi: 10.1016/j.mimet.2013.08.009

Mu, L., Xie, Z., Hu, L., Chen, G., and Zhang, Z. (2020). Cellulase interacts with Lactobacillus plantarum to affect chemical composition, bacterial communities, and aerobic stability in mixed silage of high-moisture amaranth and rice straw. *Bioresour. Technol.* 315:123772. doi: 10.1016/j.biortech.2020.123772

Muck, R. E., O'Kiely, P., and Wilson, R. K. (1991). Buffering capacities in permanent grasses. *Irish J. Agr. Res.* 30, 129–141.

Negawo, A. T., Teshome, A., Kumar, A., Hanson, J., and Jones, C. S. (2017). Opportunities for Napier grass (Pennisetum purpureum) improvement using molecular genetics. *Agronomy* 7:28. doi: 10.3390/agronomy7020028

Oliveira, A. P. D., Bagaldo, A. R., Loures, D. R. S., Bezerra, L. R., Moraes, S. A., Yamamoto, S. M., et al. (2018). Effect of ensiling gliricidia with cassava on silage quality, growth performance, digestibility, ingestive behavior and carcass traits in lambs. *Anim. Feed Sci. Technol.* 241, 198–209. doi: 10.1016/j.anifeedsci.2018.05.004

Pequerul, A., Pérez, C., Madero, P., Val, J., and Monge, E. (1993). "A rapid wet digestion method for plant analysis," in *Optimization of Plant Nutrition-Developments in Plant and Soil Sciences*, eds M. A. C. Fragoso, M. L. Van Beusichem, and A. Houwers (Dordrecht: Springer), 3–6.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. doi: 10.1093/nar/gks1219

Queiroz, O. C. M., Arriola, K. G., Daniel, J. L. P., and Adesogan, A. T. (2013). Effects of 8 chemical and bacterial additives on the quality of corn silage. *J. Dairy Sci.* 96, 5836–5843. doi: 10.3168/jds.2013-6691

Rajvanshi, A. K. (2019). *Subabul Reloaded: How One Tree Could be the Answer to India's Fodder, Fuel Needs*. Bengaluru: The Better India.

Rengsirikul, K., Kanjanakuha, A., Ishii, Y., Kangvansaichol, K., Sripichitt, P., Punsuvon, V., et al. (2011). Potential forage and biomass production of newly introduced varieties of leucaena (*Leucaena leucocephala* (Lam.) de Wit.) in Thailand. *Grass Sci.* 57, 94–100. doi: 10.1111/j.1744-697X.2011.00213.x

Shade, A., and Handselman, J. (2012). Beyond the Venn diagram: the hunt for a core microbiome. *Environ. Microbiol.* 14, 4–12. doi: 10.1111/j.1462-2920.2011.02585.x

Speedy, A., and Pugliese, P. L. (1992). "Fodder trees and shrubs in range and farming systems in dry tropical africa," in *Legume Trees and Other Fodder Trees*

*as Protein Sources for Livestock*, eds A. W. Speedy and P.-L. Pugliese (Rome: Food and Agriculture Organization of the United Nations), 43–59.

Steel, R. G., and Torrie, J. H. (1980). *Principles and Procedures of Statistics: A Biometrical Approach*. New York, NY: McGraw Hill Company.

Suzuki, R., and Shimodaira, H. (2006). Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* 22, 1540–1542. doi: 10.1093/bioinformatics/btl117

Tian, H., Chen, Y., Zhu, N., Guo, Y., Deng, M., Liu, G., et al. (2020). Effect of *Broussonetia papyrifera* silage on the serum indicators, hindgut parameters and fecal bacterial community of Holstein heifers. *AMB Express* 10:197. doi: 10.1186/s13568-020-01135-y

Van Soest, P. J., Robertson, J. B., and Lewis, B. A. (1991). Methods for dietary fiber, neutral detergent fiber, and nonstarch polysaccharides in relation to animal nutrition. *J. Dairy Sci.* 74, 3583–3597. doi: 10.3168/jds.S0022-0302(91)78551-2

Wang, C., He, L., Xing, Y., Zhou, W., Yang, F., Chen, X., et al. (2019). Fermentation quality and microbial community of alfalfa and stylo silage mixed with *Moringa oleifera* leaves. *Bioresour. Technol.* 284, 240–247. doi: 10.1016/j.biortech.2019.03.129

Xu, D., Ding, W., Ke, W., Li, F., Zhang, P., and Guo, X. (2019). Modulation of metabolome and bacterial community in whole crop corn silage by inoculating homofermentative *Lactobacillus plantarum* and heterofermentative *Lactobacillus buchneri*. *Front. Microbiol.* 9:3299. doi: 10.3389/fmicb.2018.03299

Yan, Y., Li, X., Guan, H., Huang, L., Ma, X., Peng, Y., et al. (2019). Microbial community and fermentation characteristic of Italian ryegrass silage prepared with corn stover and lactic acid bacteria. *Bioresour. Technol.* 279, 166–173. doi: 10.1016/j.biortech.2019.01.107

Zhang, Y., Li, D., Wang, X., Lin, Y., Zhang, Q., Chen, X., et al. (2019). Fermentation dynamics and diversity of bacterial community in four typical woody forages. *Ann. Microbiol.* 69, 233–240. doi: 10.1007/s13213-018-1398-z

**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.

*Copyright © 2022 Du, Yamasaki, Oya, Nguluve, Euridse, Tinga, Macome and Cai. 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.*

frontiers in Microbiology

ORIGINAL RESEARCH
published: 16 March 2022
doi: 10.3389/fmicb.2022.830121

# Effects of Lactic Acid Bacteria and Molasses Additives on Dynamic Fermentation Quality and Microbial Community of Native Grass Silage

*Yuyu Li[1,2], Shuai Du[3], Lin Sun[4], Qiming Cheng[5], Junfeng Hao[1,2], Qiang Lu[1,2], Gentu Ge[1,2], ZhiJun Wang[1,2] and Yushan Jia[1,2]**

[1] Key Laboratory of Forage Cultivation, Processing and High Efficient Utilization of Ministry of Agriculture and Rural Affairs, Inner Mongolia Agricultural University, Hohhot, China, [2] Key Laboratory of Grassland Resources of Ministry of Education, Inner Mongolia Agricultural University, Hohhot, China, [3] National Engineering Laboratory of Biological Feed Safety and Pollution Prevention and Control, Key Laboratory of Animal Nutrition and Feed Science of Zhejiang Province, Institute of Feed Science, Zhejiang University, Hangzhou, China, [4] Inner Mongolia Academy of Agricultural & Animal Husbandry Sciences, Hohhot, China, [5] College of Animal Science, Guizhou University, Guiyang, China

OPEN ACCESS

**Edited by:**
Niranjan Koirala,
Gandaki Province Academy of Science and Technology, Pokhara, Nepal

**Reviewed by:**
Xianjun Yuan,
Nanjing Agricultural University, China
Robin Anderson,
Agricultural Research Service (USDA), United States

***Correspondence:**
Yushan Jia
jys_nm@sina.com

**Specialty section:**
This article was submitted to Microbiotechnology, a section of the journal Frontiers in Microbiology

**Received:** 06 December 2021
**Accepted:** 26 January 2022
**Published:** 16 March 2022

**Citation:**
Li Y, Du S, Sun L, Cheng Q, Hao J, Lu Q, Ge G, Wang Z and Jia Y (2022) Effects of Lactic Acid Bacteria and Molasses Additives on Dynamic Fermentation Quality and Microbial Community of Native Grass Silage. Front. Microbiol. 13:830121. doi: 10.3389/fmicb.2022.830121

Ensiling native grass is an effective method to protect the nutritional quality of forage and alleviate feed shortages in the cold winter of the Inner Mongolian Plateau. To improve the usability of native grass resources as feed in China, the effects of lactic acid bacteria and molasses additions on the microbial population, fermentation quality, and nutritional quality of native grass during silage were investigated. Treatments were a control treatment with no additive (CK), lactic acid bacteria (L), molasses (M), and lactic acid bacteria in combination with molasses (L+M), all of which were stored at ambient temperature (17–28°C) for 7, 14, 30, and 60 days. The results showed that all additives improved nutritional value and fermentation quality with low pH and ammonia nitrogen ($NH_3$–N) and high crude protein (CP) and water soluble carbohydrate (WSC) than control silage over the ensiling period. Compared with L or M silage, the L+M silage combination improved fermentability, as evidenced by higher LA content and a faster pH drop during the first 7 days of ensiling. With prolonged ensiling time, the combined addition of L and M could increase the count of desirable *Lactobacillus*, decrease microbial diversity, and inhibit the growth of undesirable microorganism, such as Clostridia, *Escherichia*, and *Enterobacter* abundance compared with silage treated with CK, L. or M. Application of L together with M could further improve the silage quality of native grass by altering bacterial community structure. In summary, the addition of lactic acid bacteria and molasses increased the relative abundance of *Lactobacillus* of native grass silage and improved fermentation quality.

**Keywords: native grass, lactic acid bacteria, molasses, microorganisms, silage**

## INTRODUCTION

Native grass is an herb that grows on native grasslands, which are widely scattered across the Mongolian Plateau; large needlegrass (*Stipa grandis* P. Smirn) and Chinese leymus [*Leymus chinensis* (Trin.) Tzvel] dominate typical steppe communities (Zhang et al., 2016), are mostly utilized for grazing and haymaking, and are the main source of forage in pastoral areas (Du S. et al., 2020). Seasonal changes affect the quality and productivity of native grass (Kang et al., 2007). From June to September, native grass grows well, and animals usually get enough nutrition and gain weight during this time (Long et al., 1999). However, feed deficiency during the long cold season (November to June) is one of the major issues confronting the Mongolian Plateau's traditional animal grazing system (Zhou et al., 2017; Ding et al., 2020), which is due to both the limited quantity and quality of native grass during this period. Farmers usually begin storing grass in mid-August. During harvest and storage, there is a significant loss of dry matter (DM) and crude protein (CP) (Scarbrough et al., 2002). Therefore, efficient native grass preservation is regarded as a crucial farming strategy for alleviating feed shortages in the Mongolian Plateau's traditional year-round grazing system.

Ensiling has been considered as a traditional method to preserve forage nutrients in the harvesting season because of various advantages, such as easy to operate, prolonging supply time, and suitability for a wide range of feedstock (Ding et al., 2020), and can effectively address animal feed shortages during the winter months on the Mongolian Plateau (Zhang et al., 2016). Ensiling mainly depends on lactic acid bacteria to transfer soluble carbohydrates to lactic acid, resulting in a low pH environment and inhibiting the growth of aerobic bacteria, making the feed preserved (Lee et al., 2018) and improving palatability (Driehuis, 2013). Ensiling has been shown in previous research to improve fermentation quality while also preserving the nutrients of natural grass, resulting in a high feeding value (Khota et al., 2016; Du Z. M. et al., 2020). However, due to the low water content, low water soluble carbohydrate, low lactic acid bacteria attachment, and high buffer energy of natural forage, direct silage is difficult to succeed (Zhang et al., 2016; Hou et al., 2017). To solve this problem, Cai et al. (2020) evaluated the effects of additives on native grass silage and found that the use of L and cellulase can improve the fermentation quality of native grass silage. L is the most common type of silage additive, capable of ensuring extensive fermentation and efficient substrate utilization in ensiled materials. It has been reported that adding L to rice straw silage improved its quality (Zhao et al., 2019). M is a sugar industry byproduct that is high in soluble carbohydrate and serves as a fermentable substrate. M not only improved fermentation quality but also significantly altered the microbial community of cassava foliage silage (Li et al., 2021). However, there is limited information related to the microbial community and fermentation products during the ensiling process of native grass with additive treatments. Therefore, this study was conducted to increase understanding to inform the regulation of fermentation of native grass on the Mongolian Plateau. The purpose of this study was to assess the effects of adding L, M, and their combination on the fermentation quality and microbial community dynamics of native grass silage.

## MATERIALS AND METHODS

### Study Sites and Silage Preparation

Native grass was harvested at the heading stage from a typical grassland (45°58′E, 118°57′N) on the Inner Mongolian Plateau, People's Republic of China, on July 15, 2020. The dominant species harvested were large needlegrass (*Stipa grandis* P. Smirn), Chinese Leymus [*Leymus chinensis* (Trin.) Tzvel.], and Dahurian Bushclover [*Lespedeza davurica* (Laxm.) Schindl.]. Harvested native grass was chopped to lengths of 2–3 cm. Lactic acid bacteria (L) inoculant and molasses (M) were used as additives for ensiling the native grass. L (*Lactobacillus plantarum*, Snow Brand Seed Co. Ltd., Sapporo, Japan) was provided by the Institute of Animal Husbandry and Grassland of Japan. M was obtained from Baihui Biological Technology Co. Ltd. (Chifeng, Inner Mongolia, China). The L inoculant was applied at a level of $10^6$ colony-forming units (cfu) per gram of fresh material (FM), and molasses was applied at 30 g/kg of FM. The chopped native grass was mixed and divided into equal portions for four treatments: no additive control (CK), L treatment (L), M treatment (M), and L + M treatment (L + M). All the additives were mixed homogenously with native grass. In detail, each treated batch was divided into four replicates (one for backup) of 200 g each, which were packed into polyethylene bags and vacuum sealed. A total of 64 bags (4 treatments × 4 ensiling days × 4 repeats) were prepared and kept at ambient temperature (17–28°C). After 7, 14, 30, and 60 days of ensiling, fermentation characteristics, chemical compositions, microorganism counts, and the microbial community were analyzed. Initial fresh forage samples were taken before ensiling.

### Analysis of Chemical Composition, Organic Acid, and Microbial Population

For chemical analysis, each sample set had three parallel determinations. Dry matter (DM) contents were determined by oven drying at 65°C for 48 h (Zhang et al., 2016). The crude protein (CP) content was computed by multiplying TN content by 6.25 (Sun et al., 2021). The neutral detergent fiber (NDF) and acid detergent fiber (ADF) contents of fresh material were determined by the methods of Van Soest et al. (1991). The WSC was determined using the method of Thomas (1977).

Each polyethylene bag was opened on a clean bench. About 10 g of fresh forage or silage was blended with 90 ml of sterilized water, and the extract was serially diluted to quantify the bacterial group in a sterile solution. The numbers of LAB, aerobic bacteria, coliform bacteria, yeast, and molds were measured through the method of Sun et al. (2021).

Last, the liquid extract was filtered through four layers of cheesecloth and filtered paper. The prepared filtrate was used to analyze the pH values, organic acids, and ammonia nitrogen ($NH_3$–N) of the sample. The lactic acid (LA), acetic acid

(AA), propionic acid (PA), and butyric acid (BA) content of silage was analyzed by high-performance liquid chromatography according to the method of Cheng et al. (2021). The nitrogen ($NH_3$–N) content was determined according to the method of Broderick and Kang (1980).

## DNA Extraction, Polymerase Chain Reaction Amplification, and Sequencing

The E.Z.N.A.® sample DNA kit was used to isolate bacterial community genomic DNA from native grass and silage samples. The reagent, which was developed to extract DNA from trace amounts of sample, has been demonstrated to be effective for preparing DNA from most bacteria. The blank was made of non-nuclear water. Total DNA was eluted in 50 L of elution buffer and stored at −80°C until it was measured in the PCR (LC-Bio Technology Co. Ltd., Hangzhou, Zhejiang Province, China).

Subsequently, extracted DNA samples were subjected to PCR amplification on bacteria 16S rDNA (V3 and V4 regions). For bacterial amplification, the following primers were used: 341F and 805R. The PCR reactions were conducted using the following program: 3 min of denaturation at 95°C, 27 cycles of 30 s at 95°C, 30 s for annealing at 55°C, and 45 s for elongation at 72°C, and a final extension at 72°C for 10 min. PCR reactions were performed in triplicate 20-μl mixture containing 4 μl of 5× FastPfu buffer, 2 μl of 2.5 mM dNTPs, 0.8 μl of each primer (5 μM), 0.4 μl of FastPfu Polymerase, and 10 ng of template DNA. Agarose gel electrophoresis [2% (w/w)] was used to confirm the PCR products. To eliminate the possibility of false-positive PCR results as a negative control, ultrapure water was used throughout the DNA extraction process instead of a sample solution. AMPure XT beads (Beckman Coulter Genomics, Danvers, MA, United States) were used to purify the PCR products, and Qubit was used to quantify them (Invitrogen, Carlsbad, CA, United States).

The PCR products were sequenced on an Illumina platform (Guangzhou Gene Denovo Co. Ltd., Guangzhou, China) after purification and quantification. Fqtrim was used to perform quality filtering on the raw reads under specific filtering conditions in order to obtain high-quality clean tags (v0.94). Vsearch software (v2.3.4) was used to filter chimeric sequences. QIIME2 was used to calculate alpha and beta diversity, with the same number of sequences extracted randomly by reducing the number of sequences to the minimum of some samples, and relative abundance (X bacteria count/total count) was used in bacteria taxonomy. The QIIME2 process was used to examine alpha and beta diversity. Blast was used to perform species annotation sequence alignment, and the alignment databases were SILVA and NT-16S.

## Statistical Analyses

The data was subjected using two-way analysis of variance with the fixed effects of additives, ensiling time, and additives × ensiling time using the general linear model (GLM) procedure of SAS 9.3 software (SAS Institute, Inc., Cary, NC, United States). The effect was considered significant when $p < 0.05$. The data of high-throughput sequencing were analyzed on the free online platform www.omicstudio.cn.

# RESULTS

## Chemical Parameters and Microbial Compositions of Silage Materials

The chemical parameters and microbial compositions by native grass before ensiling are presented in **Table 1**. The DM, WSC, CP, NDF, and ADF were 37.72, 4.31, 10.83, 62.98, and 36.58%, respectively. Microbial compositions in the native grass for L, aerobic bacteria, coliform bacteria, and yeasts were 4.26, 8.75, 5.82, and 6.35 $\log_{10}$ cfu/g FM, respectively. Fresh native grass molds were below the detectable level.

## Effect of Additives and Ensiling Days on Chemical Parameters of Native Grass Silage

Effects of additives and ensiling days on chemical parameters of native grass silages are shown in **Table 2**. The additive treatments significantly altered the contents of DM, CP, and WSC, but had no effects on NDF and ADF contents. L + M silage had higher WSC and CP contents and lower ADF and DNF contents compared with L and M silage after 60 days of silage. Ensiling days had significant ($p < 0.05$) effects on the chemical composition (except DM) of silage. All additives, except for the CK treatment, significantly decreased ($p < 0.05$) NDF content of native grass silage. In addition, the L + M silage had significantly ($p < 0.05$) higher CP contents than other silage at the end of ensiling. The WSC contents in M and M + L silages were significantly ($p < 0.05$) lower than the CK silage. The NDF and ADF content exhibited a continuous downtrend during the whole ensiling process. The interaction of additive treatments and ensiling days had a significant impact on silage WSC ($p < 0.05$), but did not affect the silage DM, CP, NDF, or ADF contents ($p > 0.05$).

## Effect of Additives and Ensiling Days on Fermentation Quality of Native Grass Silage

**Table 3** illustrates the dynamics of the fermentation quality of native grass silage during ensiling. These parameters were significantly affected by additive treatments, ensiling days, and their interaction ($P < 0.05$). Compared with the control, all additives significantly ($p < 0.05$) increased lactic acid (LA) concentration, and decreased acetic acid (AA), propionic acid (PA), and $NH_3$–N concentration in silage. Compared with the other treatments, L + M addition further increased LA concentration and decreased the contents of AA, PA, and $NH_3$–N; butyric acid (BA) was not detected in the native grass silage. The highest LA concentration was recorded in L + M silage, which had a value of 29.94% on the 60th day of ensiling ($p < 0.05$). The pH values of the additive silages were significantly decreased ($p < 0.05$) compared with the control during 14 days of ensiling, and L + M silage always maintained a lower pH value during ensiling. Moreover, $NH_3$–N content exhibited a continuous increasing trend during the whole ensiling process, and the content of $NH_3$–N in the L + M treatment group remained low after 14 days of ensiling.

**TABLE 1 |** Chemical and microbial compositions of fresh native grass.

| | Items | Sample | *SEM* |
|---|---|---|---|
| Chemical composition | DM (%FM) | 37.72 | 2.38 |
| | WSC (%DM) | 4.31 | 0.03 |
| | CP (%DM) | 10.83 | 0.43 |
| | NDF (%DM) | 62.98 | 3.03 |
| | ADF (%DM) | 36.58 | 1.07 |
| Microbial counts | L (log cfu/g FM) | 4.26 | 0.11 |
| | Aerobic bacteria (log cfu/g FM) | 8.75 | 0.03 |
| | Coliform bacteria (log cfu/g FM) | 5.82 | 0.15 |
| | Yeast (log cfu/g FM) | 6.35 | 0.29 |
| | Molds (log cfu/g FM) | ND | ND |

*FM, fresh material; DM, dry matter; CP, crude protein; NDF, neutral detergent fiber; ADF, acid detergent fiber; WSC, water-soluble carbohydrates; L, lactic acid bacteria; ND, not detected, SEM, standard error of the mean.*

## Effect of Additives and Ensiling Days on Microorganism Counts of Native Grass Silage

**Table 4** shows the effects of additives and ensiling days on microbial counts of native grass silages. Overall, additive treatments, ensiling days, and their interaction had a significant impact on microbial counts ($p < 0.05$). Additive treatment silages had lower ($p < 0.05$) coliform bacteria content and higher L and yeast numbers compared with the control silage after 60 days of ensiling. Compared with other treatments, the L and yeast contents in L + M-treated silage were significantly ($p < 0.05$) higher, while the coliform bacteria contents were lower than that of other silages at the end of ensiling. The microbial count of silages exhibited a continuous decreasing trend during the whole ensiling process. Mold counts were not detected in any treatment after 14 days of ensiling.

## Microbial Community of Native Grass Silage

The bacterial diversity of native grass silage was discovered using high-throughput analyses targeting variable regions 3 and 4 of 16S rDNA. The valid sequences of all 51 triplicate samples added up to 1,294,499 after unqualified sequences were removed. The average Good's coverage for all of the samples was greater than 99%, indicating that the sequencing depth was sufficient for reliable bacterial community analysis (**Table 5**). The interaction of additive treatments and ensiling days had a significant impact on Simpson and Shannon ($p < 0.05$), but did not affect Reads, OUT, and Chao1 ($p > 0.05$). According to the OTU and Chao 1 indexes, the bacterial community's richness decreased with increased ensiling time in the CK silage. The Shannon and Simpson indexes were remarkably lower in the M and L + M-treated silages than in the CK silage at days 7 and 14 of ensiling ($p < 0.05$). In addition, the Chao 1 index and Simpson index

**TABLE 2 |** Effect of additives and ensiling days on chemical compositions of native grass silages.

| Items | Treatment | Ensiling days | | | | Significance | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 7 | 14 | 30 | 60 | SEM | *T* | *D* | *T * D* |
| DM (%FM) | CK | 35.97aA | 35.79bA | 34.72bAB | 34.63cB | 0.13 | ** | NS | NS |
| | L | 36.37aA | 35.92bAB | 35.33bB | 35.19abB | | | | |
| | M | 36.16aA | 35.47bAB | 35.06bB | 35.02bcB | | | | |
| | L + M | 37.19aA | 36.62aA | 36.17aA | 36.01aA | | | | |
| CP (%DM) | CK | 9.71aA | 8.97bAB | 8.49bBC | 7.87bC | 0.11 | ** | ** | NS |
| | L | 10.07aA | 9.67aA | 9.34aAB | 8.57bB | | | | |
| | M | 10.26aA | 9.91aA | 9.17abB | 8.58bB | | | | |
| | L + M | 10.48aA | 9.86aB | 9.65aB | 9.08aC | | | | |
| NDF (%DM) | CK | 61.11aA | 58.03aA | 57.63aA | 57.45aA | 0.42 | NS | ** | NS |
| | L | 59.43aA | 58.66aAB | 57.73aAB | 56.46aB | | | | |
| | M | 59.77aA | 58.87aA | 56.11aB | 55.52aB | | | | |
| | L + M | 57.70aA | 55.92aB | 54.45aB | 53.92aC | | | | |
| ADF (%DM) | CK | 33.79aA | 31.34aA | 31.90aA | 31.04aA | 0.33 | NS | * | NS |
| | L | 32.58aA | 32.32aA | 30.58aA | 30.47aA | | | | |
| | M | 32.72aA | 31.69aA | 29.99aA | 29.08aA | | | | |
| | L + M | 31.01aA | 30.37aA | 29.53aA | 28.26aB | | | | |
| WSC (%DM) | CK | 3.78bA | 3.04cB | 2.87bBC | 2.54cC | 0.08 | ** | ** | * |
| | L | 3.95abA | 3.17bcB | 3.10bB | 2.82bcB | | | | |
| | M | 4.26abA | 3.72abB | 3.40aBC | 3.20abC | | | | |
| | L + M | 4.34aA | 3.98aB | 3.63aC | 3.50aC | | | | |

*FM, fresh material; DM, dry matter; CP, crude protein; NDF, neutral detergent fiber; ADF, acid detergent fiber; WSC, water-soluble carbohydrates; LAB, lactic acid bacteria; ND, not detected; CK, no additive control; L, lactic acid bacteria; M, molasses; L + M, lactic acid bacteria + molasses; SEM, standard error of the mean; T, treatments; D, ensiling days; T * D, interaction between treatments and ensiling days. *Significant at 0.05. **Significant at 0.01. Means in the same column (a–c) or row (A–C) with different superscript letters differ significantly from each other ($p < 0.05$).*

**TABLE 3 |** Effect of additives and ensiling days on fermentation quality of native grass silages.

| Items | Treatment | Ensiling days | | | | Significance | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 7 | 14 | 30 | 60 | SEM | *T* | *D* | *T * D* |
| pH value | CK | 4.66aA | 4.63aA | 4.48aB | 4.32aC | 0.04 | ** | ** | ** |
| | L | 4.54aA | 4.14bB | 4.05bB | 4.03cB | | | | |
| | M | 4.63aA | 4.50aAB | 4.53aAB | 4.25bB | | | | |
| | L + M | 4.18bA | 4.01bB | 3.95bBC | 3.91dC | | | | |
| LA (%DM) | CK | 9.17bD | 11.05cC | 14.76cB | 17.47cA | 0.88 | ** | ** | ** |
| | L | 13.08aD | 13.34bC | 19.49bB | 26.61bA | | | | |
| | M | 12.3aD | 13.82bC | 17.3bB | 23.74bA | | | | |
| | L + M | 14.64aD | 21.76aC | 26.00aB | 29.94aA | | | | |
| AA (%DM) | CK | 2.26aD | 5.38aC | 10.86B | 14.75aA | 0.56 | ** | ** | ** |
| | L | 1.16bD | 3.18bC | 5.39cB | 7.61cA | | | | |
| | M | 2.67aD | 3.28bC | 6.65bB | 11.49bA | | | | |
| | L + M | 1.09bD | 2.34cC | 3.51dB | 5.92dA | | | | |
| PA (%DM) | CK | 0.97aD | 1.79aC | 2.91aB | 4.89aA | 0.17 | ** | ** | ** |
| | L | 0.56bD | 1.45bC | 2.52bB | 2.96bA | | | | |
| | M | 0.88aC | 1.58abBC | 1.96cB | 3.57bA | | | | |
| | L + M | 0.3cD | 1.42cC | 1.91cB | 2.65bA | | | | |
| $NH_3$–N (%TN) | CK | 1.39aB | 2.01aB | 2.87aB | 6.05aA | 0.18 | * | ** | ** |
| | L | 1.22aC | 1.44aC | 2.19aB | 3.33bA | | | | |
| | M | 1.05aC | 1.52aBC | 2.03aB | 2.69bA | | | | |
| | L + M | 1.12aC | 1.42aBC | 1.81aB | 2.3bA | | | | |

*LA, lactic acid; AA, acetic acid; PA, propionic acid; $NH_3$–N, ammonia nitrogen; TN, total nitrogen; ND, not detected; SEM; ND, no detected; CK, no additive control; L, lactic acid bacteria; M, molasses; L + M, lactic acid bacteria + molasses; SEM, standard error of the mean; T, treatments; D, ensiling days; T * D, interaction between treatments and ensiling days. *Significant at 0.05. **Significant at 0.01. Means in the same column (a–d) or row (A–D) with different superscript letters differ significantly from each other ($p < 0.05$).*

**TABLE 4 |** Effect of additives and ensiling days on microorganism counts of native grass silages.

| Items | Treatment | Ensiling days | | | | Significance | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 7 | 14 | 30 | 60 | *SEM* | *T* | *D* | *T * D* |
| L (log cfu/g FM) | CK | 8.74cA | 8.07aB | 6.50bC | 5.74cD | 0.16 | ** | ** | ** |
| | L | 8.99abA | 8.16aB | 6.81abC | 6.56bD | | | | |
| | M | 8.89bcA | 8.06aB | 6.45bC | 6.45bC | | | | |
| | L + M | 9.09aA | 8.29aB | 7.00aC | 6.77aC | | | | |
| Yeast (log cfu/g FM) | CK | 6.37dA | 5.08bB | 3.35cC | 3.38cC | 0.28 | ** | ** | ** |
| | L | 8.84bA | 7.06aB | 3.51cD | 3.91bC | | | | |
| | M | 6.92cA | 7.54aA | 3.84bB | 3.50bcB | | | | |
| | L + M | 9.35aA | 6.99aB | 5.91aC | 5.62aD | | | | |
| Aerobic bacteria (log cfu/g FM) | CK | 6.38aA | 5.68aA | 3.67aB | 3.63aB | 0.15 | ** | ** | * |
| | L | 4.73abA | 3.45bB | 3.23bB | 3.16bB | | | | |
| | M | 4.53bA | 3.49bB | 3.53aB | 3.30bB | | | | |
| | L + M | 3.66bA | 3.38bB | 3.17bB | 2.62cC | | | | |
| Coliform bacteria (log cfu/g FM) | CK | 5.76a | ND | ND | ND | 0.29 | ** | ** | ** |
| | L | 3.62b | ND | ND | ND | | | | |
| | M | 3.95b | ND | ND | ND | | | | |
| | L + M | 2.85c | ND | ND | ND | | | | |

*CK, no additive control; L, lactic acid bacteria; M, molasses; L + M, lactic acid bacteria + molasses; ND, no detected; SEM, standard error of the mean; T, treatments; D, ensiling days; T * D, interaction between treatments and ensiling days. *Significant at 0.05. **Significant at 0.01. Means in the same column (a–d) or row (A–D) with different superscript letters differ significantly from each other ($p < 0.05$).*

showed trends similar to the trend in OTUs. Compared with other treatment groups, L + M addition resulted in the decrease in the richness of the bacterial community without dose effect in silage at day 60 of ensiling ($p > 0.05$). The lowest Shannon index of bacterial diversity was observed in the L + M-treated silage (i.e., 3.01).

**TABLE 5 |** Effect of additives and ensiling days on bacterial alpha diversity of native grass silages.

| Items | Treatment | Ensiling days | | | | Significance | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 7 | 14 | 30 | 60 | *SEM* | *T* | *D* | *T * D* |
| Reads | CK | 72825aA | 77336aA | 75397aA | 77922aA | 407.39 | NS | NS | NS |
| | L | 75263aA | 75315aA | 77059aA | 78374aA | | | | |
| | M | 75654aA | 78011aA | 77415aA | 75825aA | | | | |
| | L + M | 76901aA | 75853aA | 77805aA | 78222aA | | | | |
| OUT | CK | 254aA | 181aA | 189aA | 117aA | 12.53 | NS | NS | NS |
| | L | 117aA | 135aA | 135aA | 199aA | | | | |
| | M | 88aA | 136aA | 133aA | 134aA | | | | |
| | L + M | 118aA | 127aA | 235aA | 115aA | | | | |
| Chao1 | CK | 271.9aA | 195.0aA | 193.2aA | 119.0aA | 13.71 | NS | NS | NS |
| | L | 119.0aA | 142.0aA | 141.3aA | 204.9aA | | | | |
| | M | 91.2aA | 138.4aA | 136.2aA | 136.1aA | | | | |
| | L + M | 120.3aA | 134.2aA | 249.3aA | 118.6aA | | | | |
| Simpson | CK | 4.05aA | 4.34aA | 4.46aA | 2.94aB | 0.09 | ** | ** | * |
| | L | 3.30abA | 3.72abA | 3.57aA | 3.72aA | | | | |
| | M | 2.77bA | 3.36bA | 3.55aA | 3.66aA | | | | |
| | L + M | 2.96bA | 3.28bA | 3.63aA | 3.01aA | | | | |
| Shannon | CK | 0.86aA | 0.88aA | 0.90aA | 0.67aA | 0.01 | NS | * | * |
| | L | 0.79abA | 0.84aA | 0.82aA | 0.83aA | | | | |
| | M | 0.77bA | 0.82aA | 0.84aA | 0.87aA | | | | |
| | L + M | 0.77bA | 0.83aA | 0.84aA | 0.76aA | | | | |
| Coverage | CK | 0.9988aA | 0.9995aA | 0.9997aA | 0.9998aA | 0.0001 | NS | NS | NS |
| | L | 0.9998aA | 0.9997aA | 0.9997aA | 0.9995aA | | | | |
| | M | 0.9998aA | 0.9997aA | 0.9997aA | 0.9998aA | | | | |
| | L + M | 0.9997aA | 0.9996aA | 0.9991aA | 0.9997aA | | | | |

*CK, no additive control; L, lactic acid bacteria; M, molasses; L + M, lactic acid bacteria + molasses; ND, no detected; SEM, standard error of the mean; T, treatments; D, ensiling days; T * D, interaction between treatments and ensiling days. *Significant at 0.05. **Significant at 0.01. Means in the same column (a,b) or row (A,B) with different superscript letters differ significantly from each other ($p < 0.05$).*

The PCoA analysis (**Figure 1**) showed the differences in bacterial communities between fresh samples and different additive treatments. The bacterial communities of fresh samples and different additive treatments were significantly separated, indicating that the microbial communities of fresh samples and different additive treatments changed during silage. A clear difference in the additive treatments in contrast with the control indicates that the additive treatments and control silage bacterial communities were different on day 60 of ensiling.

The dynamics of bacterial communities in native grass silage at the phylum level are shown in **Figure 2A**. *Proteobacteria* (89.82%) and *Firmicutes* (9.87%) dominated the majority of the total phylum sequence in the fresh sample. During the process of ensiling, *Firmicutes* and *Proteobacteria* dominated in all groups, but the community composition was affected by the ensiling treatments. Compared with the CK treatment, *Firmicutes* increased and *Proteobacteria* decreased in the L and L + M-treated groups. However, the relative abundance of *Proteobacteria* in M treatments was higher than that of the other treatments. In addition, the increase in *Firmicutes* abundance was greater with added L + M, and the relative abundance of *Firmicutes* (at 60 days) was higher.

**Figure 2B** illustrates the dynamics of bacterial populations in native grass silage at the genus level. The main epiphytic genera of fresh native grass were *Pantoea* and *Pseudomonas*, followed by *Erwinia*. Their abundance decreased greatly during the ensiling process, especially in the treated silage. *Lactobacillus* was the most abundant genus in silage with *L. plantarum* added (L and L + M; >54%) after 30 days of ensiling. The addition of L and L + M decreased the abundance of *Enterobacter* and *Escherichia* compared with silage with added M alone. In the absence of additives, after 60 days of ensiling, the dominant species were *Pantoea* and *Lactobacillus*. Under the influence of the fermentation promoter, there was a slight difference from the CK treatment. The main microbes were *Lactobacillus* (60.96%) and *Enterobacter* (11.26%) in the control group, whereas the main genera were *Lactobacillus* (58.29%), *Enterobacter* (7.72%), *Escherichia* (4.00%), and *Pantoea* (4.17%) in the L group. In the M group, the main genera were *Lactobacillus* (23.86%), *Enterobacter* (35.31%), *Escherichia* (20.12%), and *Pantoea* (1.74%), and the main genera were *Lactobacillus* (67.44%), *Escherichia* (4.85%), and *Enterobacter* (3.29%) in the L + M combination group.

To further reveal the effect of additives on the bacterial community in natural forage silage, one-way analysis of variance bar plots of the genus level among native grass silage groups after 60 days of ensiling are shown in **Figure 3**. In this study, *Pantoea* was the dominant genus in the fresh native grass (FM) (26.48%). Furthermore, *Clostridium_sensu_stricto*_12

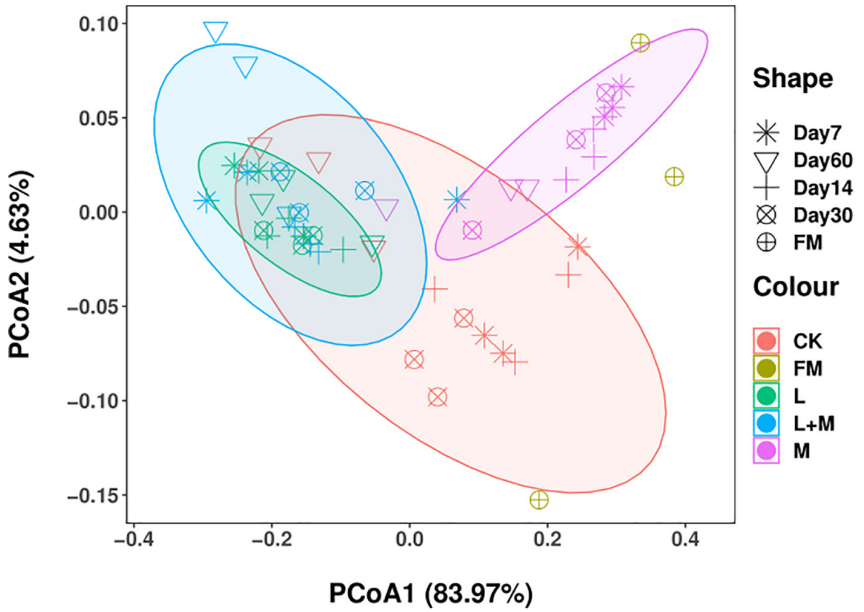



**FIGURE 1 |** The community dissimilarities in different additives treatments and fermentation time, calculated via weighted UniFrac distances, with coordinates calculated using principal coordinates analysis (PCoA). FM, fresh material; CK, no additive control; L, lactic acid bacteria; M, molasses; L + M, lactic acid bacteria + molasses.

was the subdominant family in FM, following *Pantoea*. The additive treatments significantly changed the relative abundance of *Lactobacillus*, *Enterobacter*, *Escherichia*, *Pantoea*, *Pseudomonas*, *Clostridium_sensu_stricto*_12, *Erwinia*, *Klebsiella*, and *Lactococcus* after 60 days of native grass silage. Lactobacillus was the most abundant genus in the L + M-treated silage (L + M; >65%). Molasses enhanced the growth of *Enterobacter* and *Escherichia*.

The effects of metabolic pathway changes in bacterial communities were quantified by predicting the functional characteristics of bacterial communities. Therefore, **Figure 4** showed the metabolic pathways of silage treated with different additives. In the present study, higher relative abundance of carbohydrate metabolism was observed in the L-inoculated silage during the early phase of ensiling (7 days). With increased fermentation time, the relative abundance of carbohydrate metabolism gradually increased. Amino acid metabolism was lowest in the M-inoculated silage from 7 days until the end of ensiling. Metabolism of other amino acids showed the same pattern. The relative abundance of energy metabolism in L and L + M silages were higher than that of the CK and M treatments during the entire ensiling period, and nucleotide metabolism showed the same trend. After 60 days of fermentation of silage treated with additives, the metabolism of cofactors and vitamins were increased.

## DISCUSSION

The amount of epiphytic L to raw materials is the key factor for the success of silage (Cai et al., 1999). The amount of epiphytic L is more than $10^5$ cfu $g^{-1}$ of FM; it is easier for silage to be successful and preserved (Cai et al., 1999). In addition, WSC is a significant factor for silage fermentation; the content of WSC in raw material is more than 5% DM, which can ensure the success of silage fermentation (Moselhy et al., 2015; Ni et al., 2017). The content of WSC and the number of epiphytic L influence the rate of increase in LA and drop in pH during the early stages of ensiling, which are critical for successful silage (Wang S. et al., 2019). In our study, the counts of L on native grass were $<10^5$ cfu $g^{-1}$ of FM (**Table 1**), and the WSC content of FM was 4.31, which indicated difficulty in good preservation of native grass through ensiling without any additives. Thus, it was necessary to add exogenous L and/or M to enhance lactic acid fermentation during ensiling.

Silage treated with L + M had higher content of DM than that of other silages and could be attributed to molasses providing additional substrates, promoting sufficient lactic acid production, and resulting in pH reduction and DM loss prevention (Muck et al., 2018); this finding is similar to that of Chen et al. (2017). The addition treatment silages had lower NDF and ADF content than the control silages, which might be due to the additives encouraging microorganism proliferation and enhancing microbial respiration and fermentation of the fiber component. This finding is in line with those of Baytok et al. (2014) and Li et al. (2021). During the ensiling process of native grass, the content of CP significantly decreased. This was mostly due to the fact that some microorganisms involved in silage fermentation degrade protein (Du Z. M. et al., 2020). Sufficient content of WSC could provide more substance for lactic acid fermentation. The study of Ge et al. (2018) found that inoculating with L would ensure efficient conversion and utilization of WSC during ensiling, followed by a considerable drop in pH and

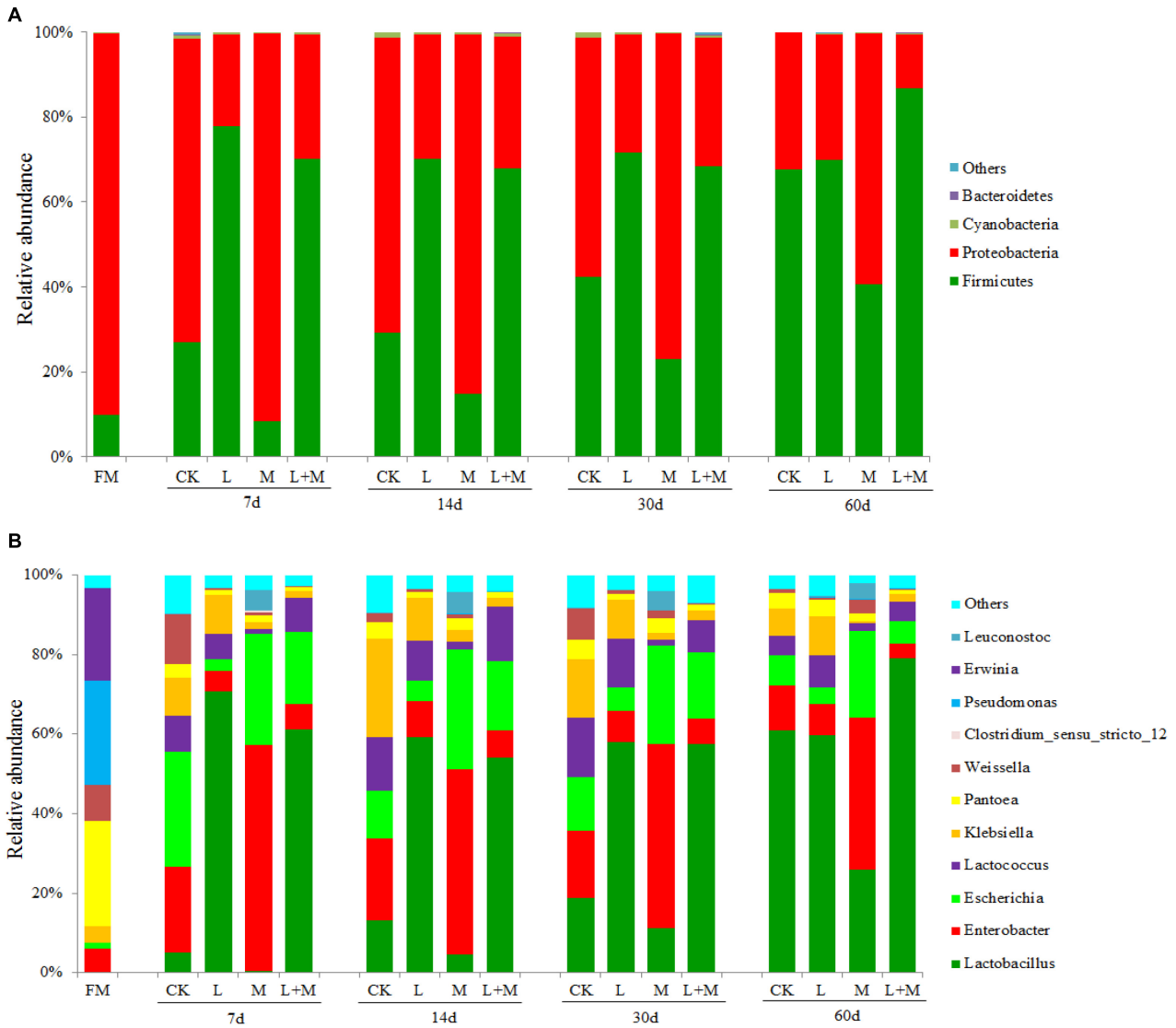



**FIGURE 2 |** The bacterial abundance at phylum and genus level in native grass silage. FM, fresh material; CK, no additive control; L, lactic acid bacteria; M, molasses; L + M, lactic acid bacteria + molasses. The bacterial communities are shown at the phylum level **(A)** and the genus level **(B)**.

reduction of nutrient loss, because L transformed the WSC into lactic acid during the early stage of fermentation. Therefore, the addition of M and L prior to ensiling seems to be necessary for high-quality native grass silage.

Ensiling is a complex bacterial fermentation process that causes a buildup of organic acid and a drop in pH (Ding et al., 2020). Silage pH is a key indication for determining fermentation quality, with a pH of 4.2 or below indicating well-fermented silage (Mcdonald et al., 1991). All of the treatments in our research had lower pH values than the CK treatment for the whole ensiling time, and a combination of L and M had the highest reduction in pH among the treatments. In this paper, it was found that the decrease in pH value occurred mainly in the first 7 or 14 days of silage, and then decreased with the increase in silage time, which was following the report of Ni et al. (2017). After 60 days of ensiling, the pH values of the L + M treatments declined to below 4.00, which could be due to the direct increase in fermentable substrate that promoted sufficient LA production and preserved silage due to the low pH, and which may also explain the decreased concentration of PA. In contrast, the pH value of the CK treatment remained above 4.20. The above results further showed that the addition of L and M prior to ensiling could stimulate the decrease in pH. In addition, L addition showed lower pH values than with M addition alone, which may be because the WSC were decomposed to a certain extent by L in the early ensiling stage, which, in turn, might accelerate the domination of L (Wang et al., 2021). Our present study showed that additive treatments decreased pH, AA, PA, and $NH_3$–N contents, as a result of the increased LA contents in silage. The L and L + M silages had higher

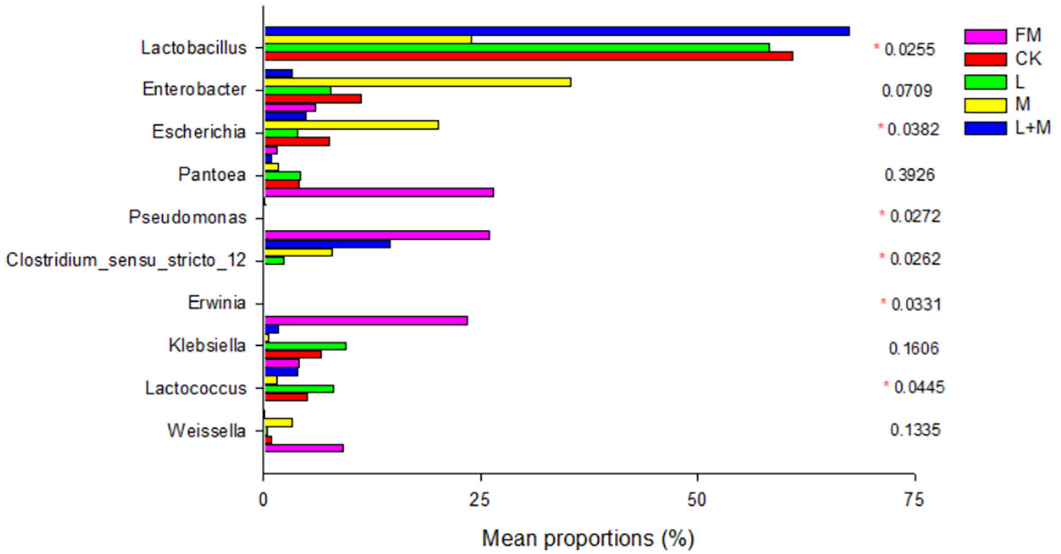



**FIGURE 3 |** One-way analysis of variance bar plots of the genus level (10 most abundant genera) among native grass silage groups after 60 days of silage. *$p < 0.05$. FM, fresh material; CK, no additive control; L, lactic acid bacteria; M, molasses; L + M, lactic acid bacteria + molasses.

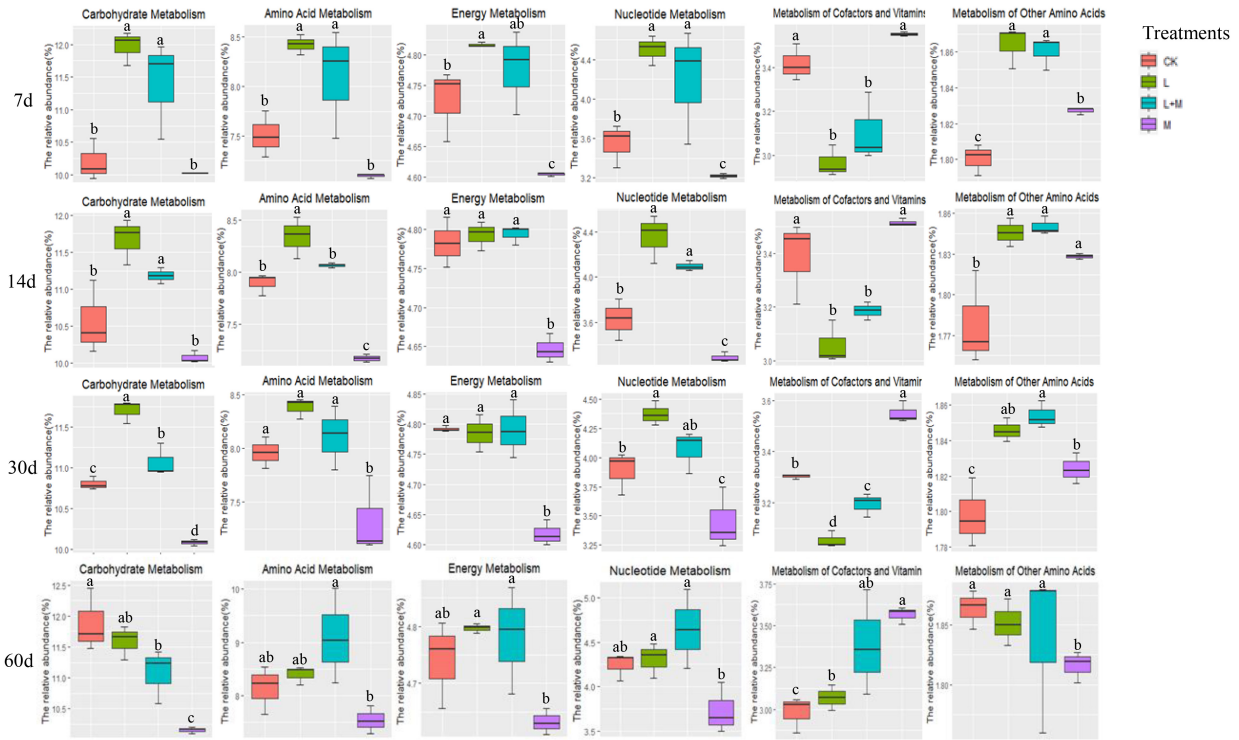



**FIGURE 4 |** Statistics of additives and ensiling days on the abundance of native grass silage microorganism KEGG pathway. Summary of significant functional shifts predicted using Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt). For each KEGG pathway, the second level of the predicted functional shift is shown with respect to the fermentation processes and additive treatments. a–d indicates significant differences between different additive treatments in the same silage period at $p < 0.05$. CK, no additive control; L, lactic acid bacteria; M, molasses; L + M, lactic acid bacteria + molasses.

LA contents than other silage treatments. This may be because both additives contained L, which could initiate an accumulation of LA, thereby resulting in a decrease in the final pH of silage (Martinez-Fernandez et al., 2010). This is in line with studies on typical woody forage (Wang et al., 2021), fodder ramie (Ning et al., 2012), and soybean (Ni et al., 2017). $NH_3$–N production

is related to CP degradation in all silages and reveals the extent of proteolysis in silage (Wilkinson, 2005). Compared with the CK silage, adding L and M alone or in combination significantly decreased $NH_3$–N content in silage, which might be because the addition of additives reduces the pH value, and the acidic environment reduced undesirable fermentation and proteolysis. After 60 days of ensiling, the decline in $NH_3$–N content in L and L + M might be because some LAB can induce nitrification, which transformed ammonia nitrogen to nitrate nitrogen (Si et al., 2019). Similarly, Hou et al. (2017) found that additive treatments improved the fermentation quality of native grass silage. The abovementioned results indicated that L and M addition in the ensiling process could promote fermentation quality, and that the combination treatment enhanced the fermentation more efficiently.

Before ensiling, L, aerobic bacteria, coliform bacteria, molds, and yeasts are often found in Inner Mongolian native grasses. Aerobic bacteria are the dominating microorganisms during the initial stage of ensiling, resulting in some fermentation loss (Ge et al., 2018). In the present study, L counts of all silages during ensiling were $>10^5$ cfu $g^{-1}$ of FM. Compared with the control group, the content of lactic acid bacteria in the additive treatment group increased by 12.4–17.9%, and the growth of non-important bacteria was inhibited. The physiological growth of L and an adequate supply of M as a substrate could explain an increase in L count (So et al., 2020). Control silage had no satisfactory microorganisms (e.g., aerobic bacteria and yeasts), while the addition treatments dramatically reduced the counts of aerobic bacteria and molds. This was most likely due to the additive treatments' ability to create enough lactic acid to lower pH and limit the growth of other hazardous bacteria during silage fermentation (Henderson, 1993). Ni et al. (2017) discovered that a combination of L and 2% M might prevent the growth of clostridia and enterobacter in soybean silage. Overall, these findings demonstrated that additives can lower the number of unwanted microorganisms in native grass silage, with the L + M treatment having the greatest impact.

Bacteria in fresh native grass and silage samples were sequenced using amplicon sequencing. The recovered reads retrieved from each sample ranged from 69,317 to 78,375. All of the samples had coverage values of approximately 0.99. This demonstrates that the sequencing breadth was rather broad, and the microbial high-throughput data were sufficient to define the bacterial microbial community's features (Yang et al., 2019). It is generally recognized that these specific OTUs may have contributed to differences in silage quality (Mu et al., 2021). The OTUs, Shannon index, and Chao1 value in L + M were lower after 60 days of ensiling than in CK, L, and M. Presumably, the combined addition of L and M was more likely to the native grass fermentation, thereby inhibiting the growth of other microorganisms, because the additive treatments decreased the pH, inhibited harmful microorganisms, and promoted the growth of L species. As a consequence, the silage of the combined addition of L and M in combination had the lowest microbial diversity and richness.

The quality of silage is determined by the outcome of the competition between L and spoilage microorganisms, as well as the competition and collaboration between L (Bai et al., 2021). The PCoA plot revealed a distinct separation of bacteria in FM and additive-treated silages, indicating that ensiling rebuilds the microbial community. These results are in line with the work of Zeng et al. (2020), who also found that the bacterial communities in FM and silages were distinct. In the present study, in the early fermentation stage (7 days), the bacterial communities in the L and L + M silages were clearly separated from the other groups. This might be due to L quickly initiating lactic acid fermentation, lowering the pH, and as such impacting bacterial community succession. In addition, compared with the control treatment, the PCoA of additive treated silages was also separate, which showed that additives significantly influenced the microbial community. The variation in bacterial community might explain the difference in silage quality (Ni et al., 2017; Dong et al., 2019; Li et al., 2021). Therefore, based on the results of α diversity and beta diversity, L and M treatments could affect the microbial diversity and community structure of native grass silage.

This study found that *Proteobacteria* was the most abundant phylum in FM, accounting for more than 85%, which is consistent with previous studies that *Proteobacteria* was the dominant phylum in fresh native grass (You et al., 2021). In the present study, *Firmicutes* and *Proteobacteria* were prevalent in all treatments in our research, and the community composition altered with ensiling time. Compared with the CK group, *Firmicutes* increased, while *Proteobacteria* declined in the L and L + M groups. For maize stover and red clover silage, Xu et al. (2017) and Dong et al. (2019) found comparable findings, which might be explained by the higher microbial populations of L (You et al., 2021). *Proteobacteria* have a major role in polysaccharide utilization, organic matter degradation, and carbon cycling (Ma et al., 2018). The higher relative abundance of *Proteobacteria* in M compared with other treatments (**Figure 2A**) may have been due to increased fermentation substrates, which can be rapidly hydrolyzed and utilized of M by *Proteobacteria*, as was also observed by Mu et al. (2021).

In the present study, *Pantoea* was found to be the most common facultative aerobic genus in FM. *Pantoea* has been observed in fresh native grass (You et al., 2021), alfalfa (Sa et al., 2021), and soybean (Ni et al., 2017). *Pantoea* abundance decreased after 30 days of fermentation, which might be attributable to their great sensitivity to pH decline (Ogunade et al., 2018). *Lactobacillus* is a well-known microbe that decreases pH during the ensilaging process by producing lactic acid, which impacts the quality of the fermentation (Li et al., 2021). *Lactobacillus*, *Weissella*, and *Leuconostoc* are the genera with the most bacteria involved in lactic acid fermentation in silage (Yang et al., 2019), Similarly, *Lactobacillus* was the predominant genus in native grass silage treated with L + M. Ni et al. (2017) found a similar outcome when they ensiled soybean infected with L + M, and *Lactobacillus* became the dominating genus after silage fermentation was completed. *Enterobacteria* are facultative anaerobes (i.e., can live in anaerobic and acidic environments) and can metabolize WSCs and LA to produce AA, PA, and other fermentation products (Li et al., 2019; Wang Y. et al., 2019). Combined addition of M and L also

contributed to the growth of *Lactobacillus*, and *Lactobacillus* abundance reached nearly 70% of the total population. It is known that lactic acid-producing cocci (*Weissella*, *Leuconostocs*, and *Lactococcus*) initiate lactic fermentation in the early ensiling process, whereas lactic acid rod (*Lactobacillus*) play a key role in pH reduction at the later stage (Cai et al., 1998). The high abundance of *Lactobacillus* in the L + M combined addition treatment could explain their relatively good fermentation quality compared with the other treatments (Ni et al., 2017; Li et al., 2021). Interestingly, in this study, the abundance of *Enterobacter* and *Escherichia* increased in M silage compared with other silages. This may be attributed to the separate addition of molasses providing a rich availability of WSC, which directly improves the competitiveness of *Enterobacter* and *Escherichia* in the silage, and thus, L could not quickly become a dominant flora. A similar trend was also observed by Mu et al. (2021). Combined addition of M and L treatments could increase the abundance of *Lactobacillus*, decrease the abundance of the *Pseudomonas*, and improve fermentation quality of native grass silage. Furthermore, their combination had a positive synergistic effect on silage fermentation and the microbial community.

Silage fermentation is a very complex biological process involving a variety of microorganisms that produces a variety of metabolites during ensiling by degrading substrates or transforming metabolites via sophisticated metabolic pathways (Guan et al., 2018). Carbohydrate, amino acid, energy and cofactor metabolism, and vitamin metabolism were found to be metabolic pathways linked to silage fermentation in previous investigations. As a result, we chose these metabolic pathways for statistical analysis, including nucleotide metabolism (Bai et al., 2021; Xu et al., 2021). In L and L + M-treated silage, carbohydrate metabolism was higher during the whole ensiling process. This might be because the addition of L accelerated the decomposition of WSC in the silage. Similarly, Bai et al. (2021) also discovered that the expression of the carbohydrate metabolism pathway was connected to the relative abundance of total L in the bacterial community. The abundance of amino acid metabolism, nucleotide metabolism, and cofactor and vitamin metabolism gradually increased with fermentation, which contradicted the findings of Sa et al. (2021) who observed that amino acid metabolism, nucleotide metabolism, and metabolism of cofactors and vitamins were decreased in the inoculated silages during the fermentation stage. This might be because the raw materials attached to L and WSC content were low, while the exogenous additives contributed to the increase in this metabolism.

## CONCLUSION

The present study illustrated that additives could improve the silage quality of native grass to different degrees, and that native grass silage treated with combined addition of L and M had better fermentation quality than other treatments. The use of additives prior to ensiling could reduce undesirable microorganisms and improve the nutritional value of forage native grass silage, with L + M having the best effects. In summary, the results confirmed that lactic acid bacteria and molasses exerted a beneficial synergistic effect on silage fermentation, which effectively improved silage quality, enhanced the relative abundance of *Lactobacillus*, and inhibited the growth of undesirable microorganisms on native grass.

## 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: https://www.ncbi.nlm.nih.gov/, PRJNA783886.

## AUTHOR CONTRIBUTIONS

YL methodology, visualization, validation, data curation, and wrote the original draft. SD and LS interpreted the data and edited the language. QC wrote, reviewed, and edited the manuscript. GG and ZW conceptualization, acquisition of funding, and writing—reviewing and editing of the manuscript. QL and JH software. YJ conceptualization and funding acquisition. All authors have read and agreed to the published version of the manuscript.

## FUNDING

This work was supported by the Key Laboratory of Forage Cultivation, Processing and High Efficient Utilization of Ministry of Agriculture and Rural Affairs, the Key Laboratory of Grassland Resources of Ministry of Education, the program for the National Natural Science Foundation of China (31760710), National Key Research and Development Program (2017YFD0502103), and Technology Project of Inner Mongolia (2020GG0032), China.

## REFERENCES

Bai, J., Ding, Z. T., Ke, W. C., Xu, D. M., Wang, M. S., Huang, W. K., et al. (2021). Different lactic acid bacteria and their combinations regulated the fermentation process of ensiled alfalfa: ensiling characteristics, dynamics of bacterial community and their functional shifts. *Microb. Biotechnol.* 14, 1171–1182. doi: 10.1111/1751-7915.13785

Baytok, E., Aksu, T., Karsli, M. A., and Muruz, H. (2014). The effects of formic acid, molasses and inoculant as silage additives on corn silage composition and ruminal fermentation characteristics in sheep. *Turkish J. Vet. Anim. Sci.* 29, 469–474. doi: 10.1021/am506877b

Broderick, G. A., and Kang, J. H. (1980). Automated simultaneous determination of ammonia and total amino acids in ruminal fluid and in vitro media. *J. Dairy Sci.* 63, 64–75. doi: 10.3168/jds.S0022-0302(80)82888-8

Cai, Y. M., Benno, Y., Ogawa, M., and Kumai, S. (1999). Effect of applying lactic acid bacteria isolated from forage crops on fermentation characteristics and aerobic deterioration of silage. *J. Dairy Sci.* 82, 520–526. doi: 10.3168/jds.S0022-0302(99)75263-X

Cai, Y. M., Benno, Y., Ogawa, M., Ohmomo, S., and Nakase, T. (1998). Influence of Lactobacillus spp. from an Inoculant and of *Weissella* and *Leuconostoc* spp. from forage crops on silage fermentation. *Appl. Environ. Microbiol.* 64, 2982–2987. doi: 10.1128/aem.64.8.2982-2987.1998

Cai, Y. M., Du, Z. M., Yamasaki, S., Nguluve, D., Tinga, B., Macome, F., et al. (2020). Influence of microbial additive on microbial populations, ensiling characteristics, and spoilage loss of delayed sealing silage of Napier grass. *Asian Australas J. Anim. Sci.* 33, 1103–1112. doi: 10.5713/ajas.19.0471

Chen, X., Li, W., Gao, C., Zhang, X., Weng, B., and Cai, Y. M. (2017). Silage preparation and fermentation quality of kudzu, sugarcane top and their mixture treated with lactic acid bacteria, molasses and cellulase. *Anim. Sci. J.* 88, 1715–1721. doi: 10.1111/asj.12840

Cheng, Q. M., Li, P., Xiao, B., Yang, F., Li, D., Ge, G. T., et al. (2021). Effects of LAB inoculant and cellulase on the fermentation quality and chemical composition of forage soybean silage prepared with corn stover. *Grassl. Sci.* 67, 83–90. doi: 10.1111/grs.12289

Ding, Z., Bai, J., Xu, D., Li, F., Zhang, Y., and Guo, X. (2020). Microbial community dynamics and natural fermentation profiles of ensiled alpine grass *Elymus nutans* prepared from different regions of the qinghai-tibetan plateau. *Front. Microbiol.* 11:855. doi: 10.3389/fmicb.2020.00855

Dong, Z., Li, J., Chen, L., Wang, S., and Shao, T. (2019). Effects of freeze–thaw event on microbial community dynamics during red clover ensiling. *Front. Microbiol.* 10:1559. doi: 10.3389/fmicb.2019.01559

Driehuis, F. (2013). Silage and the safety and quality of dairy foods: a review. *Agric. Food Sci.* 22, 16–34. doi: 10.23986/afsci.6699

Du, S., You, S. H., Bao, J., Ge, G. T., and Cai, Y. M. (2020). Growth performance, carcass characteristics, and meat quality of Mongolian lambs fed native grass or hay with or without concentrate on the Inner Mongolian Plateau. *Can. J. Anim. Sci.* 100, 470–478. doi: 10.1139/cjas-2019-0126

Du, Z. M., Risu, N., Ge, G. T., Jia, Y. S., and Cai, Y. M. (2020). Dynamic changes and characterization of the protein and carbohydrate fractions of native grass grown in Inner Mongolia during ensiling and the aerobic stage. *Asian Australas J. Anim. Sci.* 33, 556–567. doi: 10.5713/ajas.19.0212

Ge, G. T., Hou, M. L., Liu, T. Y., Jia, Y. S., and Cai, Y. M. (2018). Microbial population, chemical composition and silage fermentation of native grasses growing on the Inner Mongolian Plateau. *Grassl. Sci.* 64, 226–233. doi: 10.1111/grs.12207

Guan, H., Yan, Y. H., Li, X. L., Li, X. M., Shuai, Y., Feng, G. Y., et al. (2018). Microbial communities and natural fermentation of corn silages prepared with farm bunker-silo in Southwest China. *Bioresour. Technol.* 265, 282–290. doi: 10.1016/j.biortech.2018.06.018

Henderson, N. (1993). Silage additives. *Anim. Feed Sci. Technol.* 45, 35–56. doi: 10.1016/0377-8401(93)90070-Z

Hou, M., Ge, G. T., Liu, T. Y., Jia, Y. S., and Cai, Y. M. (2017). Silage preparation and fermentation quality of natural grasses treated with lactic acid bacteria and cellulase in meadow steppe and typical steppe. *Asian Australas J. Anim. Sci.* 30, 788–796. doi: 10.5713/ajas.16.0578

Kang, L., Han, X., Zhang, Z., and Sun, O. J. (2007). Grassland ecosystems in China: review of current knowledge and research advancement. *Philos. Trans. R. Soc. Lond.* 362, 997–1008. doi: 10.1098/rstb.2007.2029

Khota, W., Pholsen, S., Higgs, D., and Cai, Y. M. (2016). Natural lactic acid bacteria population of tropical grasses and their fermentation factor analysis of silage prepared with cellulase and inoculant. *J. Dairy Sci.* 99, 9768–9781. doi: 10.3168/jds.2016-11180

Lee, I. S., Lee, S. Y., Choi, M. K., Kang, C. H., and Kim, J. M. (2018). Evaluation of the fermentation ability of lactic acid bacteria to manufacture highest quality rice straw silage. *Korean J. Crop Sci.* 63, 106–111. doi: 10.7740/kjcs.2018.63.2.106

Li, D. X., Ni, K. K., Zhang, Y. C., Lin, Y. L., and Yang, F. Y. (2019). Fermentation characteristics, chemical composition and microbial community of tropical forage silage under different temperatures. *Asian Australas J. Anim. Sci.* 32, 665–674. doi: 10.5713/ajas.18.0085

Li, M., Zi, X., Zhou, H., Lv, R., Tang, J., and Cai, Y. (2021). Effect of lactic acid bacteria, molasses, and their combination on the fermentation quality and bacterial community of cassava foliage silage. *Anim. Sci. J.* 92, 1–10. doi: 10.1111/asj.13635

Long, R. J., Zhang, D. G., Wang, X., Hu, Z. Z., and Dong, S. K. (1999). Effect of strategic feed supplementation on productive and reproductive performance in yak cows. *Prev. Vet. Med.* 38, 195–206. doi: 10.1016/S0167-5877(98)00125-1

Ma, S. S., Fang, C., Sun, X. X., Han, L. J., He, X. Q., and Huang, G. Q. (2018). Bacterial community succession during pig manure and wheat straw aerobic composting covered with a semi-permeable membrane under slight positive pressure. *Bioresour. Technol.* 259, 221–227. doi: 10.1016/j.biortech.2018.03.054

Martinez-Fernandez, A., Soldado, A., Vicente, F., Martinez, A., and de la Roza-Delgado, B. (2010). Wilting and inoculation of *Lactobacillus buchneri* on intercropped triticale-fava silage: effects on nutritive, fermentative and aerobic stability characteristics. *Agric. Food Sci.* 19, 302–312. doi: 10.2137/145960610794197597

Mcdonald, P., Henderson, A. R., and Heron, S. (1991). *The Biochemistry Of Silage*, 2nd Edn. Marlow: Chalcombe Publications.

Moselhy, M. A., Borba, J. P., and Borba, A. (2015). Improving the nutritive value, in vitro digestibility and aerobic stability of Hedychium gardnerianum silage through application of additives at ensiling time. *Anim. Feed Sci. Technol.* 206, 8–18. doi: 10.1016/j.anifeedsci.2015.05.001

Mu, L., Xie, Z., Hu, L., Chen, G., and Zhang, Z. (2021). *Lactobacillus plantarum* and molasses alter dynamic chemical composition, microbial community and aerobic stability of mixed (amaranth and rice straw) silage. *J. Sci. Food Agric.* 101, 5225–5235. doi: 10.1002/jsfa.11171

Muck, R. E., Nadeau, E., Mcallister, T. A., Contreras-Govea, F. E., and Kung, L. J. (2018). Silage review: recent advances and future uses of silage additives. *J. Dairy Sci.* 101, 3980–4000. doi: 10.3168/jds.2017-13839

Ni, K., Wang, F., Zhu, B., Yang, J., Zhou, G., Pan, Y., et al. (2017). Effects of lactic acid bacteria and molasses additives on the microbial community and fermentation quality of soybean silage. *Bioresour. Technol.* 238, 706–715. doi: 10.1016/j.biortech.2017.04.055

Ning, T. T., Xu, C. C., Wang, H. L., Hao, W., and Lei, H. (2012). Effects of ensiling fodder ramie (*Boehmeria nivea* L.) with lactic acid bacteria and molasses on fermentation quality. *Adv. Mater. Res.* 524-527, 2167–2171.

Ogunade, I. M., Jiang, Y., Cervantes, A. A. P., Kim, D. H., Oliveira, A. S., Vyas, D., et al. (2018). Bacterial diversity and composition of alfalfa silage as analyzed by Illumina MiSeq sequencing: effects of *Escherichia coli* O157:H7 and silage additives. *J. Dairy Sci.* 101, 2048–2059. doi: 10.3168/jds.2017-12876

Sa, D. W., Lu, Q., Wang, Z., Ge, G. T., Sun, L., and Jia, Y. S. (2021). The potential and effects of saline-alkali alfalfa microbiota under salt stress on the fermentation quality and microbial. *BMC Microbiol.* 21:149. doi: 10.1186/s12866-021-02213-2

Scarbrough, D. A., Coblentz, W. K., Coffey, K. P., Turner, J. E., Davis, G. V., Kellogg, D. W., et al. (2002). Effects of summer management and fall harvest date on ruminal in situ degradation of crude protein in stockpiled bermudagrass. *Anim. Feed Sci. Technol.* 96, 119–133. doi: 10.1016/S0377-8401(01)00349-2

Si, H. Z., Liu, H. L., Li, Z. P., Nan, W. X., Jin, C. N., Sui, Y. T., et al. (2019). Effect of Lactobacillus plantarum and Lactobacillus buchneri addition on fermentation, bacterial community and aerobic stability in lucerne silage. *Anim. Prod. Sci.* 59, 1528–1536.

So, S., Cherdthong, A., and Wanapat, M. (2020). Improving sugarcane bagasse quality as ruminant feed with *Lactobacillus*, cellulase, and molasses. *J. Anim. Sci. Technol.* 62, 648–658. doi: 10.5187/jast.2020.62.5.648

Sun, L., Na, N., Li, X., Li, Z., Wang, C., Wu, X., et al. (2021). Impact of packing density on the bacterial community, fermentation, and in vitro digestibility of whole-crop barley silage. *Agriculture* 11:672. doi: 10.3390/agriculture11070672

Thomas, T. A. (1977). An automated procedure for the determination of soluble carbohydrates in herbage. *J. Sci. Food Agric.* 28, 639–642. doi: 10.1002/jsfa.2740280711

Van Soest, P. J., Robertson, J. B., and Lewis, B. A. (1991). Methods for dietary fiber, neutral detergent fiber, and nonstarch polysaccharides in relation to animal nutrition. *J. Dairy Sci.* 74, 3583–3597. doi: 10.3168/jds.S0022-0302(91)78551-2

Wang, S., Guo, G., Li, J., Chen, L., Dong, Z., and Shao, T. (2019). Improvement of fermentation profile and structural carbohydrate compositions in mixed silages ensiled with fibrolytic enzymes, molasses and Lactobacillus plantarum MTD-1. *Ital. J. Anim. Sci.* 18, 328–335. doi: 10.1080/1828051X.2018.1528899

Wang, X., Liu, H., Xie, Y., Zhang, Y., Lin, Y., Zheng, Y., et al. (2021). Effect of sucrose and lactic acid bacteria additives on fermentation quality, chemical composition and protein fractions of two typical woody forage silages. *Agriculture* 11:256. doi: 10.3390/agriculture11030256

Wang, Y., He, L. W., Xing, Y. Q., Zhou, W., Pian, R. Q., Yang, F. Y., et al. (2019). Bacterial diversity and fermentation quality of *Moringa oleifera* leaves silage prepared with lactic acid bacteria inoculants and stored at different temperatures. *Bioresour. Technol.* 284, 349–358. doi: 10.1016/j.biortech.2019.03.139

Wilkinson, J. (2005). *Analysis and Clinical Assessment of Silage*. Tetbury: Chalcombe Publications.

Xu, D. M., Wang, N., Rinne, M., Ke, W. C., Weinberg, Z. G., Da, M., et al. (2021). The bacterial community and metabolome dynamics and their interactions modulate fermentation process of whole crop corn silage prepared with or without inoculants. *Microb. Biotechnol.* 14, 561–576. doi: 10.1111/1751-7915.13623

Xu, Z. S., He, H. Y., Zhang, S. S., and Kong, J. (2017). Effects of inoculants Lactobacillus brevis and *Lactobacillus parafarraginis* on the fermentation characteristics and microbial communities of corn stover silage. *Sci. Rep.* 7:13614. doi: 10.1038/s41598-017-14052-1

Yang, L., Yuan, X., Li, J., Dong, Z., and Shao, T. (2019). Dynamics of microbial community and fermentation quality during ensiling of sterile and nonsterile alfalfa with or without *Lactobacillus plantarum* inoculant. *Bioresour. Technol.* 275, 280–287. doi: 10.1016/j.biortech.2018.12.067

You, S. H., Du, S., Ge, G. T., Wan, T., and Jia, Y. S. (2021). Microbial community and fermentation characteristics of native grass prepared without or with isolated lactic acid bacteria on the mongolian plateau. *Front. Microbiol.* 12:731770. doi: 10.3389/fmicb.2021.731770

Zeng, T. R., Li, X. L., Guan, H., Yang, W. Y., Liu, W. G., Liu, J., et al. (2020). Dynamic microbial diversity and fermentation quality of the mixed silage of corn and soybean grown in strip intercropping system. *Bioresour. Technol.* 313:123655. doi: 10.1016/j.biortech.2020.123655

Zhang, Q., Wu, B. Y. L., Nishino, N., Wang, X. G., and Yu, Z. (2016). Fermentation and microbial population dynamics during the ensiling of native grass and subsequent exposure to air. *Anim. Sci. J.* 87, 389–397. doi: 10.1111/asj.12427

Zhao, J., Dong, Z., Li, J., Chen, L., Bai, Y., Jia, Y., et al. (2019). Effects of lactic acid bacteria and molasses on fermentation dynamics, structural and nonstructural carbohydrate composition and in vitro ruminal fermentation of rice straw silage. *Asian Australas J. Anim. Sci.* 32, 783–791. doi: 10.5713/ajas.18.0543

Zhou, J., Mi, J., Degen, A., Ding, L., Guo, X., Shang, Z., et al. (2017). Urinary purine derivatives excretion, rumen microbial nitrogen synthesis and the efficiency of utilization of recycled urea in Tibetan and fine-wool sheep. *Anim. Feed Sci. Technol.* 227, 24–31. doi: 10.1016/j.anifeedsci.2017.03.005

**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.

*Copyright © 2022 Li, Du, Sun, Cheng, Hao, Lu, Ge, Wang and Jia. 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.*

frontiers in Microbiology

ORIGINAL RESEARCH
published: 21 March 2022
doi: 10.3389/fmicb.2022.837894

# Whole Genome Sequencing and RNA-seq-Driven Discovery of New Targets That Affect Carotenoid Synthesis in *Phaffia rhodozyma*

*Zhihui Shi[1,2], Xiaoxian He[1,2], Hailiang Zhang[3], Xuena Guo[1], Yanfei Cheng[1], Xuelian Liu[3], Zhaoyue Wang[1]* and Xiuping He[1,2]**

[1] CAS Key Laboratory of Microbial Physiological and Metabolic Engineering, State Key Laboratory of Mycology, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China, [2] College of Life Sciences, University of Chinese Academy of Sciences, Beijing, China, [3] State Key Laboratory of Direct-Fed Microbial Engineering, Beijing DaBeiNong Science and Technology Group Co., Ltd. (DBN), Beijing, China

Carotenoids are unsaturated compounds with terpene groups. Among them, astaxanthin has strong antioxidant properties. It is widely used in aquaculture, food, medicine, and cosmetics with a broad market prospect. *Phaffia rhodozyma* is an important microorganism that synthesizes astaxanthin, but its wild strains have low pigment content, long growth cycle, and low fermentation temperature. Therefore, it is important to research the genetic improvement of the physiological and biochemical properties of *P. rhodozyma*. In this study, the atmospheric and room temperature plasma mutagenesis technology was adopted, through the functional evolution of the carotenoid production performance; then, through the comparative analysis of the genomics and transcriptomics of the wild strain and evolved strain, the key factor *GST1* gene that affects carotenoid synthesis was discovered.

**Keywords:** ***Phaffia rhodozyma*, carotenoids, astaxanthin, genomics, transcriptomics, *GST1***

OPEN ACCESS

***Edited by:***
*Sailesh Malla,*
*Chr. Hansen Holding A/S, Denmark*

***Reviewed by:***
*M. Carmen Limon,*
*University of Seville, Spain*
*Fu-Li Li,*
*Qingdao Institute of Bioenergy and Bioprocess Technology (CAS), China*

***Correspondence:**
*Zhaoyue Wang*
*wangzy@im.ac.cn*
*Xiuping He*
*hexp@im.ac.cn*

***Specialty section:***
*This article was submitted to Microbiotechnology, a section of the journal Frontiers in Microbiology*

***Received:*** *17 December 2021*
***Accepted:*** *07 February 2022*
***Published:*** *21 March 2022*

***Citation:***
*Shi Z, He X, Zhang H, Guo X, Cheng Y, Liu X, Wang Z and He X (2022) Whole Genome Sequencing and RNA-seq-Driven Discovery of New Targets That Affect Carotenoid Synthesis in Phaffia rhodozyma. Front. Microbiol. 13:837894. doi: 10.3389/fmicb.2022.837894*

## INTRODUCTION

Carotenoids are unsaturated organic compounds with terpene groups. At present, carotenoids have been identified with more than 700 different structures, which contain multiple conjugated double bonds (Higuera-Ciapara et al., 2006). The main function of carotenoids is to protect cells from reactive oxygen damage. Among carotenoids, natural astaxanthin has a strong antioxidant ability, which can prevent cell lipid oxidation, improve organism metabolism, and have an anti-aging effect (Ambati et al., 2014; Sandmann et al., 2021). Astaxanthin plays an important role in the pigmentation of salmon and crustaceans in aquaculture. Astaxanthin can also improve the reproduction and growth performance of animals (Sandmann, 2015; Wang et al., 2021).

*Phaffia rhodozyma* (PR) was originally isolated from tree exudates in the remote areas of Japan and the west coast of North America. PR (sexual type, also named *Xanthophyllomyces dendrorhous*) synthesizes astaxanthin as the main carotenoid, which is the most promising and most economical natural source of astaxanthin. β-Carotene (a precursor of astaxanthin) is the second most abundant pigment in PR. The cells can be used directly as feed additives (Johnson and Lewis, 1979; Johnson, 2003).

In PR, the carotenoid synthesis pathway starts with the acetyl-CoA substrate, *via* the mevalonate (MVA) pathway, and by phosphorylation and decarboxylation, eight molecules of isoprene

pyrophosphate (IPP, a common precursor of terpenoids) are synthesized. Two molecules of IPP are condensed to form geranyl pyrophosphate (GPP), which is then to form farnesyl pyrophosphate (FPP), geranyl geranyl pyrophosphate (GGPP), and the colorless carotenoid *cis*-phytoene. Through four-step dehydrogenation reactions and two-step cyclization reactions to synthesize β-carotene, then undergoing an oxidation reaction, astaxanthin is finally synthesized (Verdoes et al., 1999; Sieiro et al., 2003). By overcoming the limitation of a certain pathway, other reactions may become new bottlenecks because their enzymatic activity may not be able to cope with higher flow rates, which is revealed through the accumulation of intermediate products rather than final products. At present, little is known about the regulation of astaxanthin biosynthesis in PR.

Due to the commercial value of astaxanthin, many studies have focused on the mutagenesis and breeding of astaxanthin-producing strains with higher yields. Some key factors for pigment fermentation were searched, and the optimization of fermentative crafts was also adopted to increase the production of astaxanthin in PR strains. For example, oxygen and light were reported to play a synergistic effect (Breitenbach et al., 2011). Increased ventilation and low light intensity lead to a higher production of colored carotenoids and more conversion of phytoene at the expense of ergosterol (Miao et al., 2015; Gutiérrez et al., 2019). These are useful to enhance pigment synthesis; however, less crucial regulatory information based on the strain level will limit the improvement of pigment synthesis to a large extent. Hence, an investigation on the strain genetics and regulation is necessary and important. During the selection of PR strains, the colonies are orange red, and the more astaxanthin is synthesized in the cell, the darker the color. This feature makes it a high-throughput screening method for mutant strains. In this study, the carotenoid synthesis ability of PR was evolved using the atmospheric and room temperature plasma (ARTP) mutagenesis technology, and the key regulatory factors for carotenoid synthesis were searched through the whole genome sequencing (WGS) and RNA-seq of the wild strain and evolved strains. The effect of the key regulators on the physiological characteristics of PR was further investigated in order to improve the synthesis of carotenoids and guide metabolic engineering.

# MATERIALS AND METHODS

## Strains and Culture Conditions

The wild-type PR strain AS2.1557 was obtained from the (China General Microbiological Culture Collection Center, Beijing, China). F94, a strain with higher carotenoid yield, was evolved and screened by ARTP mutagenesis.

The seed medium was a Yeast Extract Peptone Dextrose Medium (YPD) (composed of glucose 20.0 g, yeast extract 10.0 g, and peptone 20.0 g per liter. The fermentation medium was a YPD medium (glucose 20.0 g, yeast extract 5.0 g, and peptone 3.0 g per liter) or a Yeast Malt Agar (YM) medium (malt extract 3.0 g, yeast extract 5.0 g, and peptone 3.0 g per liter).

All experiments were conducted in a shaking flask culture in 250 ml flasks containing a fixed liquid volume of 20 ml. PR and F94 cells were, respectively, transferred from 4° YPD slants to fresh YPD and cultured at 20° for 48 h. Ten percent of the preincubation broth was inoculated to the seed medium for another 24 h, and 10% of the broth from the previous step was further inoculated to YPD for another 24 h to produce a seed culture. Ten percent of the seed culture was then inoculated to a fermentation medium, and the overall fermentation period was 84 h on a rotary shaker with 200 rpm at 20°; all experiments were performed in triplicate. The dry weight of cells was determined by centrifuging 5 ml broth at 12,000 rpm, rinsing with distilled water, and drying at 85° until it attains constant weight (∼15 h).

## Carotenoid and Astaxanthin Measurement

About 5 ml broth was centrifuged at 12,000 rpm for 1 min and washed with distilled water. Cell pellets were mixed with 1.8 ml hydrochloric acid (3 M), fully shaken and soaked for 30 min, and then incubated in a boiling water bath for 8–9 min. After observing that the cells became flocculent, they were immediately cooled down in an ice bath. The cell pellets after distilled water washing were extracted with acetone, then shaken and extracted for 15 min in the dark, until the cells were colorless. They were centrifuged at 8,000 rpm for 5 min, and the supernatant was taken.

Carotenoids were measured at a wavelength of 474 nm using a UV spectrophotometer. The calculation of carotenoid content were as described in the literature (Jiang et al., 2017).

Astaxanthin was analyzed quantitatively by HPLC on an Eclipse Plus C18 column (250 × 4.6 mm; 5 μm; Agilent, Beijing, China), with a temperature of 30°, flow rate of 1.0 ml/min, and wavelength of 478 nm. The mobile phase consisted of 90% methanol and 10% acetonitrile. Astaxanthin was identified based on the retention time in comparison with standard astaxanthin (Aladdin; Shanghai).

## Atmospheric and Room Temperature Plasma Mutagenesis

The helium-based ARTP mutation breeding system (ARTP-IIS) was from Si Qing Yuan Biotechnology Co., Ltd. (Wuxi, China). Ten microliters of approximately $2.5 \times 10^8$ CFU/ml of PR strain were uniformly coated on sterile copper slides, which were then placed in the ARTP chamber for mutagenesis. The distance between the slides and the plasma emitter was adjusted to 2 mm, the output power was 120 W, and the flow rate of the carrier gas (helium) was 10.0 L/min. The mutagenic exposure time was set at 0–80 s. The treated cells were serially diluted to an appropriate concentration and inoculated onto a YPD agar to determine cell survival rates. The lethality rate was calculated as follows: lethality rate = (1 - N1/N0) × 100%, where N0 is the colony number of the control and N1 is the colony number of the mutants.

## General DNA and RNA Manipulations

The general DNA manipulations in *Escherichia coli* or *Saccharomyces cerevisiae* were partly performed according to the standard methods (Adams et al., 1997; Russell and Sambrook, 2001), and the voltage for transformation was changed to 2 kV in PR.

Total RNA extract: The cells of PR and F94 in 600 μl culture liquid were harvested by centrifugation (12,000 rpm, 1 min), frozen immediately in liquid nitrogen, and stored at −80°C. Liquid nitrogen grinding was done to break the cell wall; RNA extraction was performed using the TRIzol Reagent (Invitrogen) according to the manufacturer's instructions. The total RNA concentration and purity were determined by using the NanoDrop instrument.

Plasmid construction: All the plasmids used in this study are described in **Supplementary Table 1**. Plasmid Ycp50-G, *E. coli*-yeast shuttle vector with $Amp^r$ and $G418^r$. For expression of glutathione S-transferase (*GST1*), plasmid YCp-TA-GST1 was constructed with the *GST1* regulated by promoter *TEF1* and terminator *ADH1*. The primers used in the experiment are shown in **Supplementary Table 2**.

Real-time polymerase chain reaction (RT-PCR): RT-PCR analyses were performed with the LightCycler96 software SW 1.1 (Roche, Beijing, China), using RNA samples as a template. Dissociation curves were constructed to test amplification validity. Target genes were obtained from the NCBI[1]. Actin was used as a control gene. Relative gene expression was calculated by the $2^{-\Delta\Delta CT}$ (cycle threshold) method using Sequence Detection software program v1.2.2 (Applied Biosystems, Beijing, China). Each RT-PCR analysis was run in quadruplicate for test consistency. qRT-PCR verification RNA-Seq result analysis is shown in **Supplementary Figure 1**.

## Whole Genome Resequencing

*Phaffia rhodozyma* and F94 strains were selected and their genomic DNA were sent to Nuohe Zhiyuan Biotechnology Co., Ltd. (Beijing, China) for whole genome resequencing. The SAMTOOLS bioinformatics tool was used for comparing the samples with the reference genome Xden1 GCA_001007165.2.

## Transcriptome Analysis (RNA-seq)

*Phaffia rhodozyma* and F94 were, respectively, grown in the YPD or YM medium for 42 h at 20°. The subsequent procedures for RNA sequencing were conducted by Anoroad (Beijing, China). Clean data were *de novo* assembled by the Trinity software program[2]. The genes with fold changes > 1.5-fold were functionally classified using the Munich Information Center for Protein Sequences' FunCat. There were four parallel settings for each sample. Parallelism was demonstrated by PCA analysis (**Supplementary Figure 2**).

# RESULTS AND DISCUSSION

## Atmospheric and Room Temperature Plasma Mutagenesis of *Phaffia rhodozyma* and Performance Analysis of the Mutagenized Strains

The wild strain PR after 24 h of cultivation was collected for the ARTP treatment for about 50 s (the lethality rate was about 85%). From about 10,000 mutants, 500 colonies with darker colors were selected (**Supplementary Figure 3**). By a preliminary comparison of the cell growth and pigment synthesis ability of the mutants, the F94 strain was found fine (**Figure 1**). From the cell pellets' color, it was shown that the evolutionary strain F94 was darker than that of the strain PR (**Figure 2A**). Under two culture conditions to compare the strains' performance, it was found that after being fermented in the YPD medium for 84 h, the carotenoid content and astaxanthin content in the F94 strain, respectively, increased by 24% [225.71–281.90 μg/g dry cell weight, NAD+: Nicotinamide adenine dinucleotide (DCW)] and 27% (48.33–61.26 μg/g DCW) compared with that in the wild strain PR (**Figure 2B**). In the YM medium, the carotenoid content and astaxanthin content of strain F94 increased by 30% (249.49–324.82 μg/g DCW) and 26% (64.71.33–81.43 μg/g DCW), respectively (**Figure 2C**). The YM medium is more conducive to the synthesis of strains of carotenoids, compared with the YPD medium.

Stability was significant for the microbial strain, especially for mutated strains; hence, the F94 strain was cultivated for 15 generations before the later omics analysis. The results of the performance comparison among the multiple single colonies of the 15th generation of F94 showed that the evolutionary strain was stable no matter in cell growth or in the synthesis of carotenoid and astaxanthin (**Figure 2D**).

## Omics Analysis of Mutant F94 and Wild Strain *Phaffia rhodozyma*

*Phaffia rhodozyma* and F94 were selected for the whole genome resequencing, and SAMTOOLS was used for bioinformation analysis using Xden1 GCA_001007165.2 as the reference genome. The same variation information compared with the reference genome was removed from the genome information of PR and F94, and the rest was the mutation information of the F94 strain compared with the original strain PR. The resequencing analysis showed that 28 genes had single-nucleotide polymorphism (SNP) mutations in the F94 strain, 37 genes had small fragment insertion and deletion (SNP_Indel) mutations, 49 fragments had structural variations, and 222 genes had copy number variations. The specific mutation information is in the appendix (**Supplementary Tables 3–6**). It is limited in that there is only protein ID and no gene ID in the reference genome in the current database, so the mutated genes cannot be enriched, and large fragments of structural variation cannot be located on chromosomes.

We further used the RNA-Seq technique to investigate the genomic transcription changes in the mutant strain F94 compared to wild strain PR under different media. The differences of the 4 groups were given after RNA-seq analysis, including the differences of the F94 strain, respectively, under the YM and YPD media (FM_FD), among which, 819 genes were upregulated, while the remaining 485 genes were downregulated, the differences of the PR strain under the YM and YPD medium (PM_PD) with 723 of upregulated genes and 353 of downregulated genes, differences between F94 and PR under the YPD medium (FD_PD), with 153 upregulated and 137

[1] ncbi.nlm.nih.gov/

[2] github.com/trinityrnaseq/trinityrnaseq/wiki

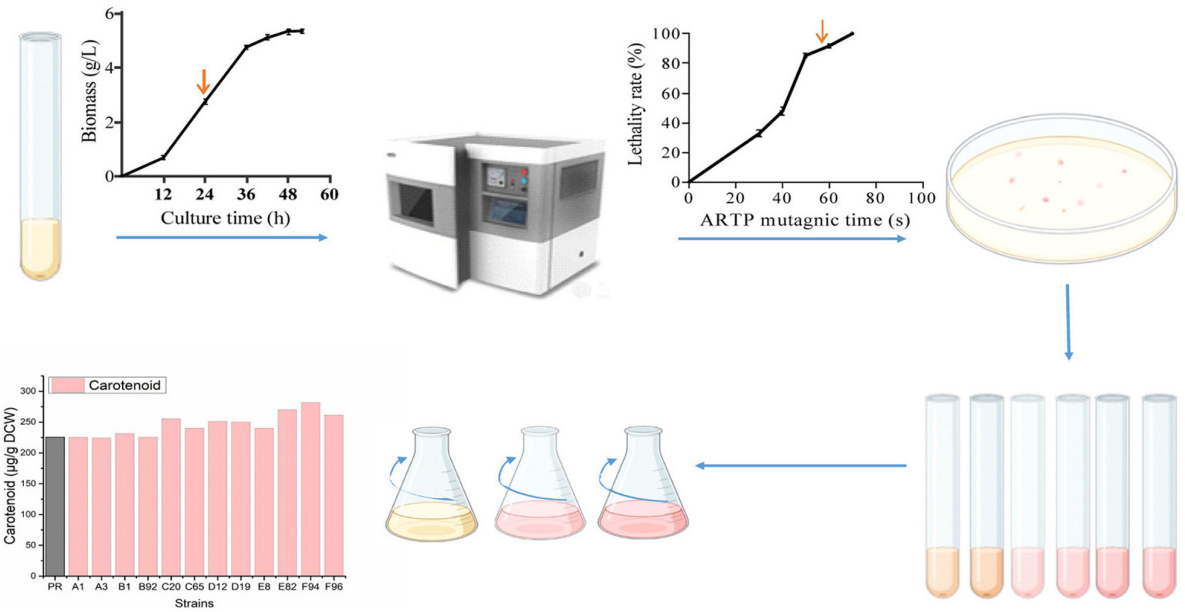



**FIGURE 1 |** Mutagenesis flow chart. Firstly, the growth curve of wild strain PR was measured in Yeast Extract Peptone Dextrose Medium (YPD), indicating that 24 h is the logarithmic phase. Next, it was determined that the lethality rate was 85% when the ARTP mutagenesis treatment time was 50 s. Then, the strain suspension was spread on a solid medium; large and red colonies were selected. There was a preliminary screening at the test tube level, using carotenoids as indicators. Finally, rescreening was performed at the shake flask level to obtain mutagenic strains.

downregulated genes, differences between F94 and PR under the YM medium (FM_PM), among which, 138 genes were upregulated and 226 genes were downregulated. R software is used for the cluster analysis of differentially expressed genes. The clustering results of the differentially expressed genes in each control group showed that the mutant strains had significant differences in transcription levels compared with wild strains, and the transcription levels of the same strain in different media were also significantly different (**Figure 3**). The difference coverage of the FM_PM group is broader than that of FD_PD; hence, the following analysis is mainly based on the differentially expressed genes of the F94 strain and the wild strain PR under the YM condition (FM_PM), and the data from YPD serves as a supplement. The differences in the transcription level caused by the different medium composition is mainly from the PR strain (PM_PD), and the data from F94 serve as a supplement.

## Transcriptional Profiling of F94 Compared to *Phaffia rhodozyma* Strain Fermented in Yeast Malt Agar Medium

The transcription levels of 364 genes in the F94 strain were changed, of which 138 genes were upregulated and 226 genes were downregulated. According to the GO (Gene Ontology) enrichment analysis of differential genes, the number of differential genes in each item was analyzed by a hypergeometric test, and the significantly enriched GO items were found (with $Q < 0.05$ as the threshold, and the GO term satisfying this condition as the significantly enriched GO entries). In the metabolic process category, only alpha-1,4-glucosidase activity, aryl-alcohol dehydrogenase (NAD+) activity, and maltose alpha-glucosidase activity were enriched. The nuclear replication fork, replication fork, nuclear prereplication complex, prereplication complex, MiniChromosome Maintenance (MCM) complex, nuclear chromatin, alpha DNA polymerase:primase complex, and nuclear chromosome part were enriched in the cell components. In the molecular functional categories, numerous entries were shown, including mitotic cell cycle processes, cell cycle processes, DNA replication initiation, cell cycle DNA replication initiation, nuclear cell cycle DNA replication initiation, mitotic DNA replication initiation, deoxyribonucleotide biosynthesis, mitotic cell cycle G1 arrest in response to pheromone, and pheromone response MAPK cascade (**Figure 4A**). Taking $Q < 0.05$ as the standard, the enrichment analysis of each pathway in KEGG (Kyoto Encyclopedia of Genes and Genomes) was carried out by a hypergeometric test, and the significant enrichment of the pathway from differentially expressed genes was found. In the mutagenic strain F94, significant changes have taken place in the following aspects (**Figure 4B**): DNA replication, cell cycle, meiosis, pyrimidine metabolism, and so on. In pyrimidine metabolism, the genes encoding carbamyl phosphate synthase [EC:6.3.5.5], cytidine triphosphate (CTP) synthase [EC:6.3.4.2], ribonucleotide reductase class II [EC:1.17.4.2], and 2′-deoxyuridine 5′-triphosphate (dUTP) pyrophosphatase [EC:3.6.1.23] were downregulated. Downregulated PRPS (ribose-phosphate pyrophosphokinase) catalyzes the formation of PRPP (5-phosphate ribose-1-pyrophosphate) from ribose-5-phosphate and ATP. PRPP is a very important precursor of nucleotide synthesis. In DNA replication, there were ten genes such as

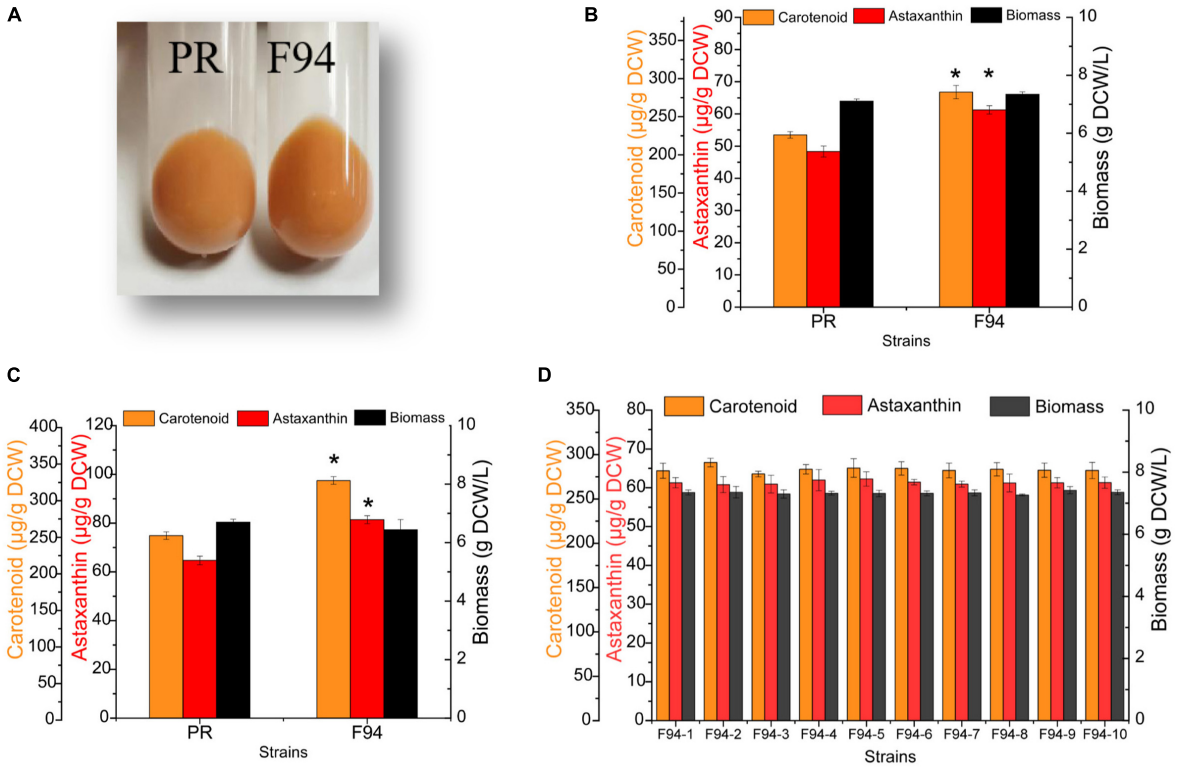



**FIGURE 2 |** Analysis of growth performance, carotenoid synthesis ability, and stability of mutagenized strains. **(A)** Comparison of the color of PR and F94 strains. **(B)** Analysis of growth and pigment content of PR and F94 strains in YPD medium. **(C)** Analysis of growth and pigment content of PR and F94 strains in Yeast Malt Agar (YM) medium. **(D)** Stability analysis of F94 strain. Statistical analysis was performed using Student's *t*-test (*$p < 0.05$).

*MCM2, MCM3, MCM6, MCM7, PCNA*, and *POLA1* that were downregulated. A total of 13 genes in the cell cycle such as *MCM1, MCM2, MCM3, MCM5, MCM6, MCM7, MAD3*, and *BUB1* and 12 genes in mitosis such as *SMC1, SMC3, ESP1*, and *BUB1* were downregulated. The results of GO enrichment and KEGG enrichment indicated that the transcription levels of the mutant strain were changed in multidimension not only in molecular function and metabolic processes but also in the cell cycle, mitosis, and other cellular behaviors due to the change of genomic information after ARTP mutagenesis.

## Changes in Metabolic Flux of Acetyl-CoA in Strain F94

Carotenoids are an unsaturated terpene compound. Acetyl-CoA is the basic precursor for the synthesis of terpenes. Acetyl-CoA is the key molecule of central carbon metabolism and participates in a variety of metabolic processes. During the analysis of the KEGG pathway map, it was found that the metabolic flux of acetyl-CoA in the mutant strain F94 might have changed. As shown in **Figure 5**, the downregulation of genes encoding chorismate mutase [EC:5.4.99.5] and pre-benzoate dehydratase [EC:4.2.1.51] in the phenylalanine, tyrosine, and tryptophan biosynthesis metabolic pathway might cause more phosphoenolpyruvate to flow to pyruvate. The gene encoding pyridoxine kinase [EC:2.7.1.35] in the Vitamin B6 metabolism pathway was downregulated, which reduced carbon loss in the branching pathway and flowed more to pyruvate. Those genes encoding arginine succinate lyase [EC:4.3.2.1] and pyrroline-5-carboxylate reductase [EC:1.5.1.2] in the arginine and proline metabolism pathway and the genes encoding asparagine synthase and argininosuccinate lyase [EC:4.3.2.1] in the alanine, aspartate, and glutamate metabolism pathway were also downregulated, which might reduce the flow of acetyl-CoA to tricarboxylic acid cycle. However, the genes encoding isocitrate lyase [EC 4.1.3.1] and malate synthase [EC:2.3.3.9] in glyoxylic acid and glyoxylate and dicarboxylate metabolism were upregulated. It has reduced the carbon loss of acetyl-CoA and increased the production of α-sandalene in yeast by knocking out the two genes of the glyoxylic acid cycle, citrate synthase, and malate synthase (Chen et al., 2013). The downregulation of the genes encoding ketol-acid reductoisomerase [EC:1.1.1.86] and 2-isopropylmalate synthase [EC:2.3.3.13] in the leucine and isoleucine biosynthesis pathway might cause more acetyl-CoA to flow to the MVA pathway. The gene encoding Delta14-sterol reductase [EC:1.3.1.70] in steroid biosynthesis pathway was downregulated. The steroid biosynthesis pathway competes with the carotenoid biosynthesis precursor FPP. The decrease of steroid biosynthesis might cause more FPP to flow to the carotenoid biosynthesis pathway.

In addition, it was found that the gene encoding the ABC transporter was upregulated in this study. The ABC transporter played an important role in the accumulation of

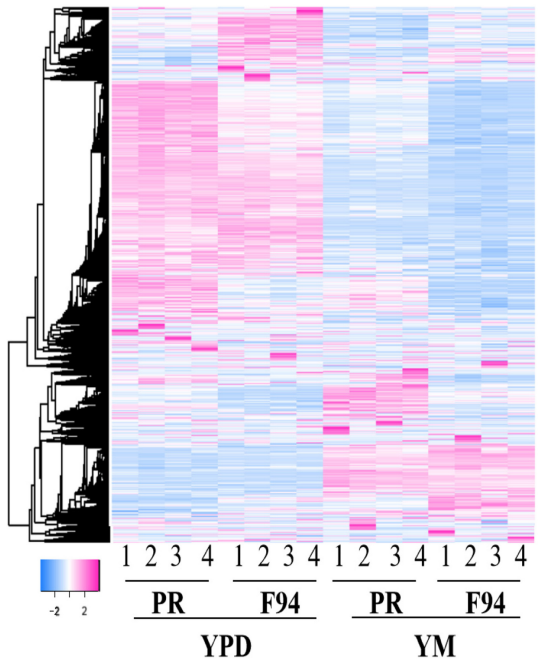



**FIGURE 3 |** Differential gene cluster map. The heatmap of the enrichment of the differentially expressed genes of F94 compared with the wild strain PR and the strains under the Yeast Malt Agar (YM) treatment compared with the YPD. The enrichment illustrates good four repetitions. The legend of the color bar displays the log2 values of the fold change of each replicate of each gene, from –2 to 2.

secondary metabolites and transmembrane transport. Some researchers found that the addition of gibberellin induced the excessive synthesis of astaxanthin. By combining the analysis of transcriptome and metabolome, they also reported that the gene encoding the ABC transporter was upregulated and the proposed strategy of transporter engineering to increase the production of astaxanthin (Liu et al., 2021).

## Transcriptional Profiling of *Phaffia rhodozyma* Strain Fermented in Yeast Malt Agar Compared to That in Yeast Extract Peptone Dextrose Medium

The fermentation in different media showed that the YM medium was more conducive to the synthesis of carotenoids by PR. Besides the analysis of different strains, RNA-seq was also adopted to analyze the transcriptional changes of wild strain PR in the YM medium and YPD medium. The transcription level of 1,076 genes changed, of which 723 genes were upregulated and 353 genes were downregulated in YM.

The results of the GO enrichment of the differential genes of the PR strain under the conditions of the YM medium compared to the YPD medium showed that the metabolic processes were enriched into the oxidation–reduction process, fatty acid metabolic process, oxoacid metabolic process, transmembrane transport, lipid metabolic process, protein refolding, amino acid transmembrane transport, etc. Peroxisome was enriched in the cell components. The molecular functional categories were enriched in oxidoreductase activity, coenzyme binding, cofactor binding, transmembrane transporter activity, vitamin binding, the hydrolase activity on glycosyl bonds, L-ascorbic acid binding, and so on (**Figure 6A**). GO enrichment also showed that different nutrients in the culture medium led to the changes in intracellular metabolic activity, peroxisome, transporter activity, and coenzyme/cofactor binding.

The results of the KEGG pathway enrichment of the differential gene in the PR strain under different media conditions (**Figure 6B**) showed that the PR cells in the YM medium had undergone significant changes in the following aspects: peroxisome, fatty acid metabolism, glycine/serine and threonine metabolism, the PPAR signal pathway, α-linolenic acid metabolism, alanine/aspartic acid and glutamate metabolism, arginine biosynthesis, and so on.

Peroxisomes are indispensable organelles in cells, which play a key role in redox signal transduction and lipid homeostasis. They contribute to many important metabolic processes, such as fatty acid oxidation, lipid biosynthesis, and free radical detoxification. Under the YM condition, 20 genes such as *PEX5, PEX3, PEX14, PEX16*, and *PXMP4* were upregulated and 3 genes, *AGXT, HMGL*, and *EPHX2*, were downregulated in the peroxisome metabolic pathway. A 3-hydroxy-3-methylglutaryl-CoA lyase (HMGL) catalyzes the formation of ketones from 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) in organisms. HMG-CoA is an important intermediate in the MVA pathway. The downregulation of HMGL makes more HMG-CoA flow to the MVA pathway, resulting in the synthesis of more carotenoids. The genes in the MVA pathway were overexpressed and located in the peroxisome of *Yarrowia lipolytica*. The engineering strain MP2 could produce MVA using fatty acids, and the yield was twice as high as that using the glucose in the cytoplasm (Jiang, 2019). It has not been reported that the increase of carotenoid synthesis is through the upregulation of the genes in peroxisome. Based on our transcriptome analysis, it was speculated that a strategy using peroxisome engineering might increase carotenoid synthesis in PR.

Studies have shown that fatty acid saturation affects the synthesis of astaxanthin (Liu et al., 2021), and fatty acid metabolism pathways are related to the synthesis of carotenoids. In our study, nine genes encoding long-chain acyl-CoA synthetase [EC:6.2.1.3], acyl-CoA oxidase [EC:1.3.3.6], acetyl-CoA acyltransferase [EC:2.3.1.16], and acyl-CoA dehydrogenase [EC:1.3.99] were upregulated in the fatty acid degradation pathway. Two genes encoding alcohol dehydrogenase 1 [EC:1.1.1.1] and aldehyde dehydrogenase (NAD+) [EC:1.2.1.3] were downregulated. The enhancement of the fatty acid degradation pathway should contribute to the synthesis of carotenoids.

In glycine, serine, and threonine metabolism, eight genes were upregulated including genes encoding 3-bisphosphoglycerate-independent phosphoglycerate mutase [EC:5.4.2.12], L-serine/L-threonine ammonia lyase [EC:4.3.1.17, 4.3.1.19], and primary-amine oxidase [EC:1.4.3.21]. *AGXT* for alanine-glyoxylate transaminase[EC:2.6.1.44]/serine-glyoxylate transaminase [EC:2.6.1.45]/serine-pyruvate

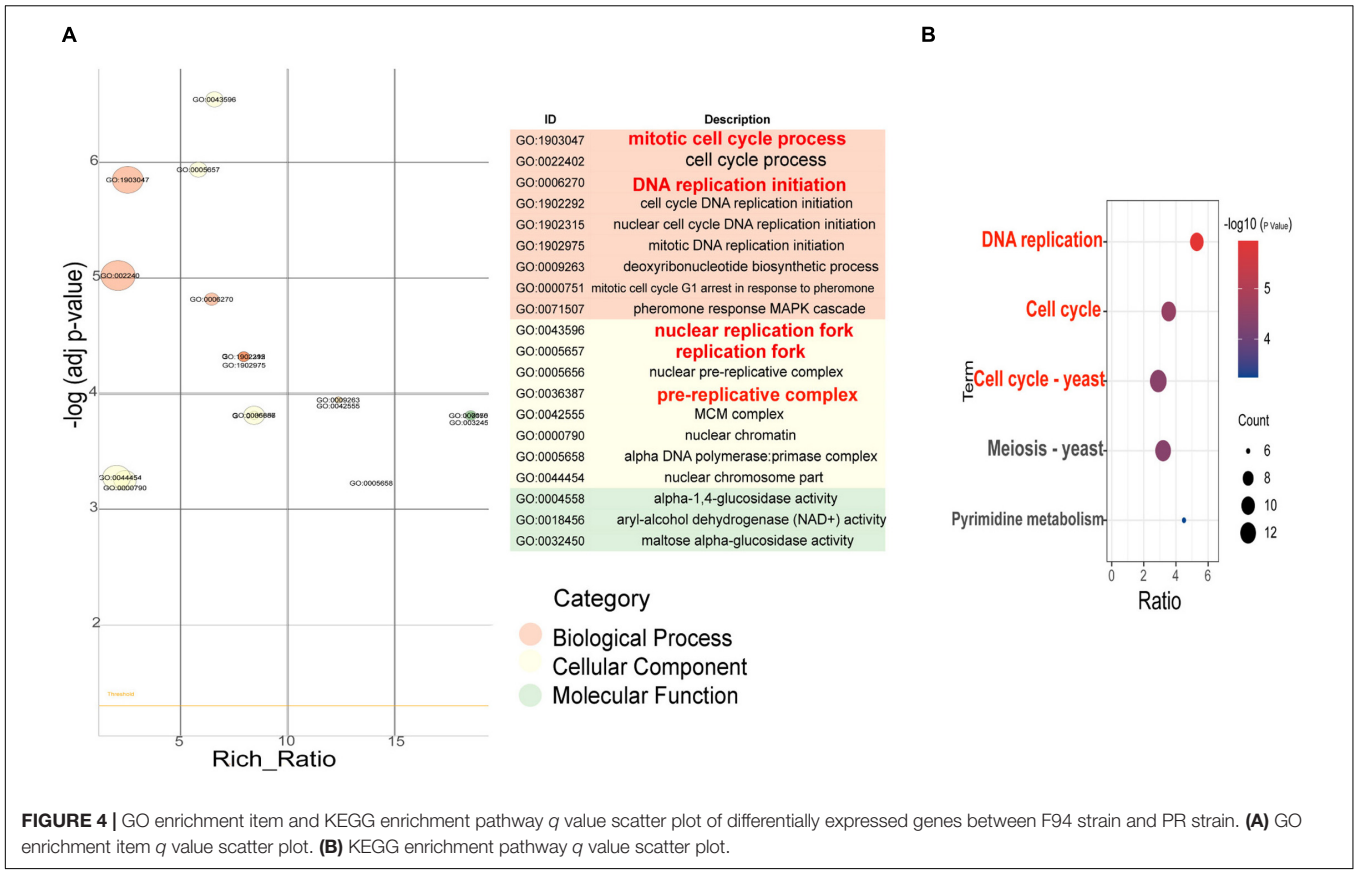



**FIGURE 4 |** GO enrichment item and KEGG enrichment pathway *q* value scatter plot of differentially expressed genes between F94 strain and PR strain. **(A)** GO enrichment item *q* value scatter plot. **(B)** KEGG enrichment pathway *q* value scatter plot.

transaminase [EC:2.6.1.51], gene for aminomethyltransferase [EC:2.1.2.10], and five other genes were downregulated. In alanine, aspartate, and glutamate metabolism, eight genes were downregulated such as the genes for argininosuccinate synthase [EC:6.3.4.5], alanine-glyoxylate transaminase/(R)-3-amino-2-methylpropionate-pyruvate transaminase [EC:2.6.1.44 2.6.1.40], argininosuccinate lyase [EC:4.3.2.1], and succinate-semialdehyde dehydrogenase/glutarate-semialdehyde dehydrogenase [EC:1.2.1.16 1.2.1.79 1.2.1.20].

Coincidentally, in the analysis of the transcription level of the mutant strain F94, the transcription level of argininosuccinate lyase was also downregulated. Amino acid metabolism is closely related to acetyl-CoA. Since acetyl-CoA is the precursor of the MVA pathway, we speculate that the carotenoid synthesis in PR may improve by changing the amino acid pathway, for example, glycine, serine, and threonine metabolism; alanine, aspartate, and glutamate metabolism; and arginine and proline metabolism, as mentioned above.

## The Effect of *GST1* Gene on the Growth Performance and Carotenoid Synthesis of *Phaffia rhodozyma*

Based on the intersection clustering of the four groups of differential genes in transcriptome analysis, 22 common differentially expressed genes were obtained from the Venn plot result (**Supplementary Figure 4**). Then, after combining the genomic SNP variation information (**Supplementary Table 7**), it was found that the gene *GST1* encoding glutathione S-transferase [EC: 2.5.1.18] had SNP mutations at four sites, and its expression level presented enhancement. In addition, the transcription level of *GST*1 in the YM medium is also upregulated, indicating that the difference between the mediums can be reversed to influence carotenoid synthesis. Comprehensively considering the probable effect of the above factors on the physiological and biochemical properties of PR, the *GST1* gene was chosen for further investigation.

In genome resequencing analysis, four mutations were found in *GST1* (A0317) in mutant strain F94, one of which was synonymous mutation and the other three sites were missense mutations, which were the 164 site of Pro mutating to Leu, 171 site of Thr to Lys, and 194 site of Ala to Val. In transcriptome analysis, the *GST1* level of strain F94 was upregulated by 2.42 times compared with that of the PR strain. KEGG metabolic pathway analysis showed that glutathione S-transferase was involved in the metabolism of xenobiotics by cytochrome P450. In eukaryotes, glutathione S-transferase is a superfamily, which is a multifunctional enzyme that participates in detoxification by catalyzing the binding of electrophilic compounds with glutathione (Jakoby, 1978). In terms of the taxonomic structure, the classification of glutathione S-transferase from bacteria, mammals, and higher plants is relatively clear. However, due to the diversity of fungal glutathione S-transferase, it is difficult to classify, and there is less research in fungus.

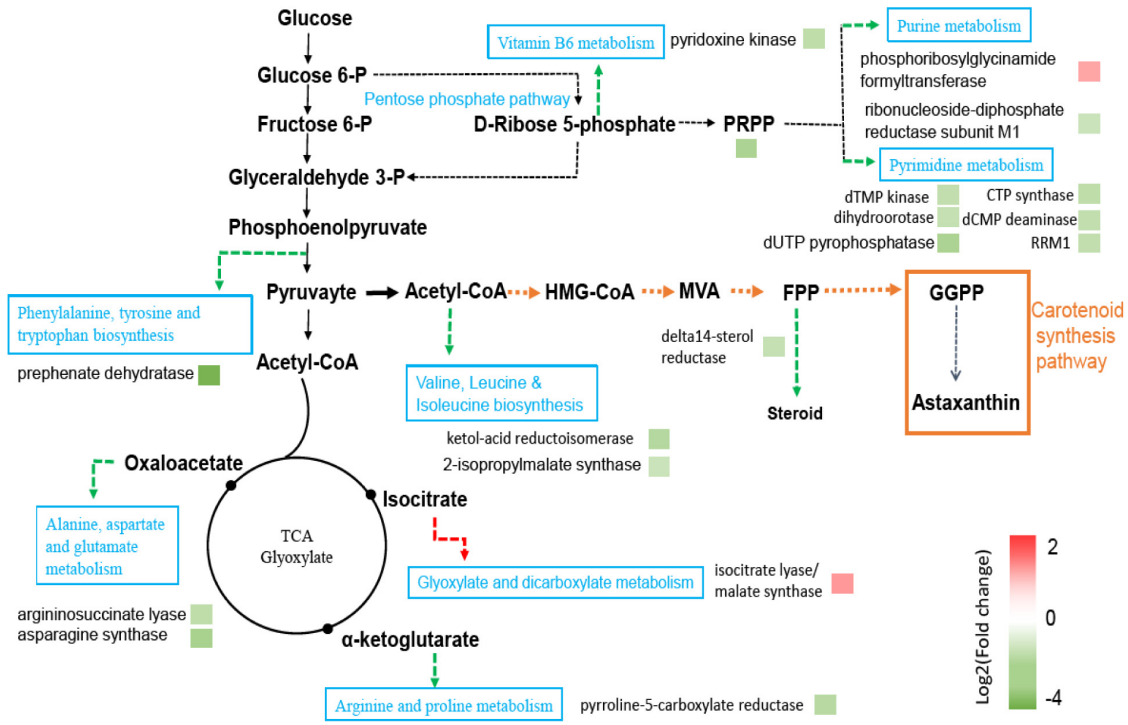



**FIGURE 5 |** Pathways of significantly differentially expressed genes between F94 strain and PR strain. Protein names were inferred from KEGG orthologs. Upregulation of genes encoding proteins were shown in red and downregulation of genes encoding proteins were in green.

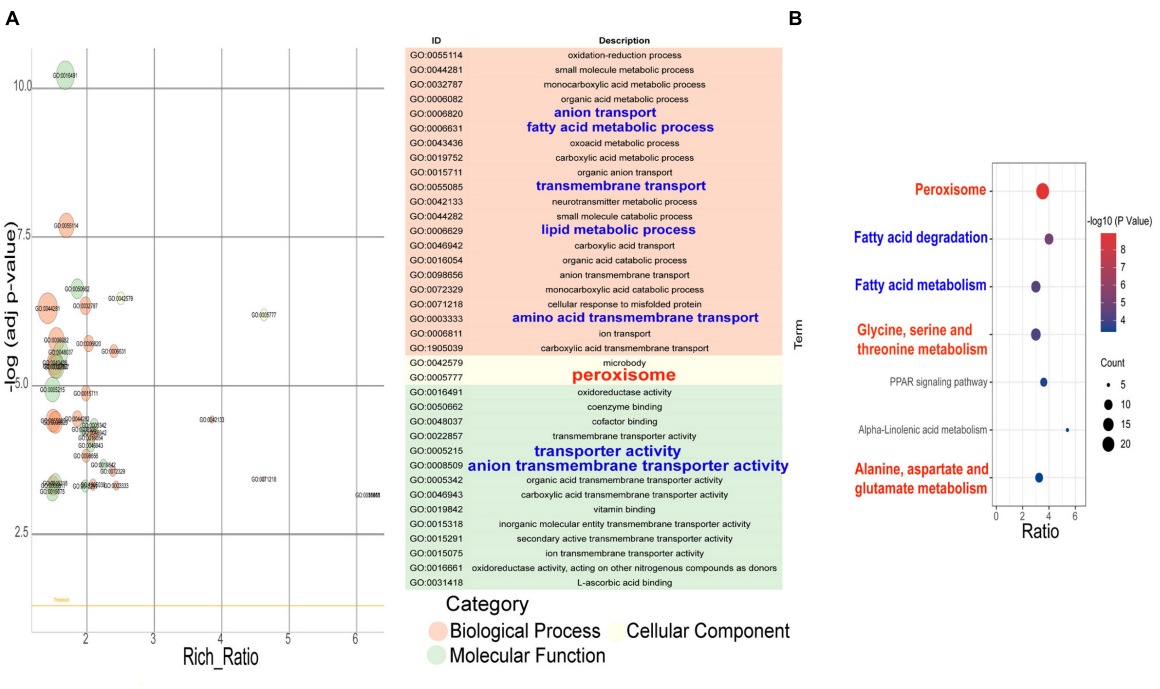



**FIGURE 6 |** GO enrichment item and KEGG enrichment pathway *q* value scatter plot of differentially expressed genes that the strains under the YM medium compares with the YPD medium. **(A)** GO enrichment item *q* value scatter plot. **(B)** KEGG enrichment pathway *q* value scatter plot.

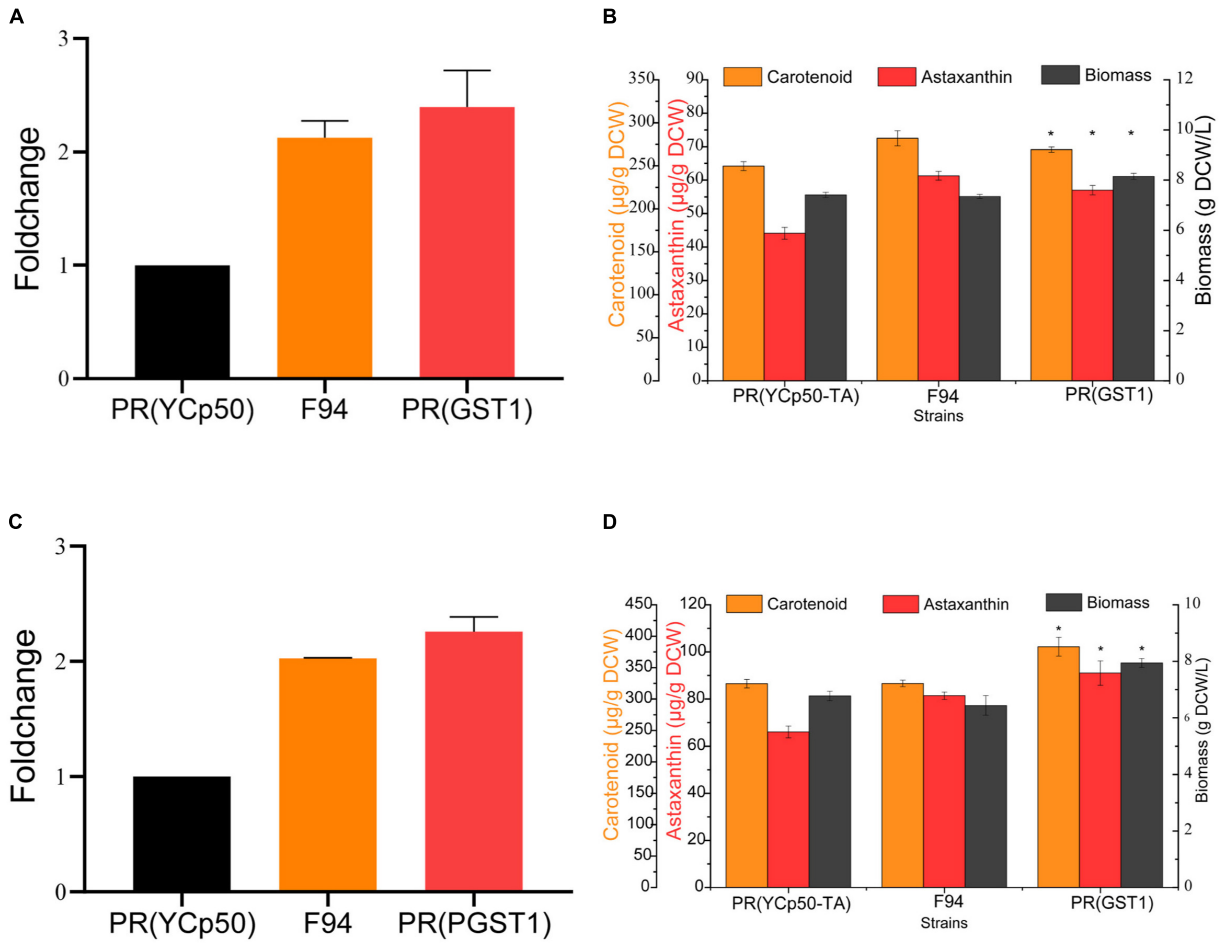



**FIGURE 7 |** Analysis of *GST1* transcription level and the growth and carotenoid synthesis ability of *GST1* overexpression strain. **(A)** Analysis of GST1 transcription level of PR (YCp-TA), F94, and PR (GST1) strains in YPD media. **(B)** Analysis of the growth and carotenoid synthesis ability of PR (YCp-TA), F94, and PR (GST1) strains in YPD media. **(C)** Analysis of *GST1* transcription level of PR (YCp-TA), F94, and PR (GST1) strains in YM media. **(D)** Analysis of the growth and carotenoid synthesis ability of PR (YCp-TA), F94, and PR (GST1) strains in YM media. Statistical analysis was performed using Student's *t*-test (*$p < 0.05$).

Gtt1p, Gtt2p, Gto1p, Gto2p, and Gto3p are currently found in the model fungus *Saccharomyces cerevisiae*, which belongs to the glutathione S-transferase family. The research on glutathione S-transferase in PR is almost blank (Choi et al., 1998; Ma et al., 2009) till now.

In order to investigate the effect of *GST1* on PR performance, using the complementary DNA (cDNA) of the PR strain as a template, the *GST1* gene with no intron was cloned, and its overexpression strain PR (GST1) was further constructed. The results of qRT-PCR showed that the level of *GST1* transcription in F94, PR (GST1) strains was twice as high as that in the control strain PR (YCp-TA) with an empty vector (**Figures 7A,C**), and there was no difference in the *GST1* transcription level between F94 and PR (GST1) strains by the *T*-test. The fermentation results of these three strains in YPD and YM media showed that when fermented in the YPD medium for 84 h, the biomass of PR (GST1) increased by 10%, the carotenoid content increased by 17%, and the astaxanthin content increased by 29% than that of PR (YCp-TA) (**Figure 7B**). When fermented in the YM medium for 84 h, the biomass of PR (GST1) increased by 17%, compared to the control strain PR (YCp-TA), the content of carotenoids increased by 18%, and the astaxanthin content increased by 40% than that of the control strain (**Figure 7D**), indicating that *GST1* overexpression not only promoted the growth of PR but also enhanced the synthesis of carotenoid and astaxanthin. The results also showed that the overexpression of *GST1* in the wild strain could promote the strain to synthesize pigment, reaching to the level as the mutant strain F94, which indicates that glutathione S-transferase plays an important role in the synthesis of astaxanthin by PR.

The glutathione S-transferase family exists in almost all plants. There are as many as 90 genes encoding glutathione S-transferase in plants. Most genes are differentially expressed under stress induction, and they play a role in the enzymatic removal of reactive oxygen. Under salt, low temperature, drought, heavy metal stress, and other conditions, glutathione S-transferase can remove active oxygen and protect the plant cell membrane structure and protein activity. The ThGSTZ1 gene of the bristly

willow was ever overexpressed in *Arabidopsis thaliana*; it was found that under the salt stress condition, the transgenic plants had lower water loss rate, increased biomass, and significantly lower hydrogen peroxide and superoxide anion levels in the body than that of the wild type (Yang et al., 2014). In our study, the overexpression of glutathione S-transferase in the PR strain led to the increased biomass, carotenoid, and astaxanthin content. The increased growth was probably owing to the fact that glutathione S-transferase could eliminate those harmful substances produced by the metabolism of the cells in the later fermentation period; hence, the cells were not destroyed by toxic substances. Since astaxanthin is also an antioxidant, the glutathione S-transferase overexpression strain might mainly rely on glutathione to remove active oxygen; as a result, the astaxanthin consumption in the cells were reduced and its content increased.

## CONCLUSION

In summary, a new and significant target *GST1* gene was found through the genome resequencing and transcriptome sequencing analyses of the wild strain PR and evolutionary strain F94 obtained by ARTP mutagenesis. We have verified that *GST1* plays an important role in the synthesis of astaxanthin and growth regulation of *P. rhodozyma*. Its influence mechanism needs to be further studied. Based on the analysis of differentially expressed genes derived from the differential pigment synthesis by the *P. rhodozyma* in different media, it was discovered that the metabolic pathways in peroxisomes and synthesis pathways of amino acids may be related to the synthesis of carotenoids in *P. rhodozyma*. The discovery of the above crucial target gene and pathways is helpful to understand the pigment synthesis in *P. rhodozyma*. It suggests that further evolution and mining the key information of the genome and transcriptome are of great significance for an in-depth elaboration of the pigment synthesis and molecular regulation mechanism, which can hopefully guide the efficient synthesis of carotenoids in *P. rhodozyma*.

## 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 in the article/ **Supplementary Material**.

## AUTHOR CONTRIBUTIONS

ZS: designing experiments and implementation, data curation, writing original draft, review, and editing. XXH: data curation. HZ and XG: ARTP mutants screening assistance. YC and XL: conception and methodology. ZW: conception, supervision, writing, review, and editing. XPH: conception, supervision, and funding acquisition. All authors agreed to be accountable for the content of the work.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.837894/full#supplementary-material

## REFERENCES


Adams, A., Gottschling, D. E., Kaiser, C. A., and Stearns, T. (1997). *Methods in Yeast Genetics: Acoldspring Harbor Laboratory Course Manual*. ColdSpring Harbor, NY: ColdSpring Harbor Laboratory Press.

Ambati, R. R., Phang, S. M., Ravi, S., and Aswathanarayana, R. G. (2014). Astaxanthin: sources, extraction, stability, biological activities and its commercial applications—a review. *Mar. Drugs.* 12, 128–152. doi: 10.3390/md12010128

Breitenbach, J., Visser, H., Verdoes, J. C., and Sandmann, G. (2011). Engineering of geranylgeranyl pyrophosphate synthase levels and physiological conditions for enhanced carotenoid and astaxanthin synthesis in *Xanthophyllomyces dendrorhous*. *Biotechnol. Lett.* 33, 755–761. doi: 10.1007/s10529-010-0495-2

Chen, Y., Laurent, D., Michel, S., Verena, S., and Jens, N. (2013). Establishing a platform cell factory through engineering of yeast acetyl-CoA metabolism. *Metab. Eng.* 15, 48–54. doi: 10.1016/j.ymben.2012.11.002

Choi, J. H., Lou, W., and Vancura, A. (1998). A novel membrane-bound Glutathione S-Transferase functions in the stationary phase of the yeast *Saccharomyces cerevisiae*. *J. Biol. Chem.* 273, 29915–29922. doi: 10.1074/jbc.273.45.29915

Gutiérrez, M. S., Campusano, S., González, A. M., Gómez, M., Barahona, S., et al. (2019). Sterol regulatory element-binding protein (Sre1) promotes the synthesis of carotenoids and sterols in *Xanthophyllomyces dendrorhous*. *Front. Microbiol.* 10:586. doi: 10.3389/fmicb.2019.00586

Higuera-Ciapara, I., Félix-Valenzuela, L., and Goycoolea, F. M. (2006). Astaxanthin: a review of its chemistry and applications. *Critic. Rev. Food Sci. Nutr.* 46, 185–196. doi: 10.1080/10408690590957188

Jakoby, W. B. (1978). The GlutathioneS-Transferases: a group of multifunctional detoxification proteins. *Adv. Enzymol. Relat. Areas Mol. Biol.* 46, 383–414. doi: 10.1002/9780470122914.ch6

Jiang, G. L., Zhou, L. Y., Wang, Y. T., and Zhu, M. J. (2017). Astaxanthin from jerusalem artichoke: production by fed-batch fermentation using *Phaffia rhodozyma* and application in cosmetics. *Process Biochem.* 59, 509–522. doi: 10.1016/j.procbio.2017.08.013

Jiang, X. (2019). *Enhancing a-Farnesene Synthesis Via Peroxisomal Localization of the Mevalonate Pathway in Yarrowia Lipolytica*. Shandong: ShanDong University. [Ph.D thesis].

Johnson, E. A. (2003). *Phafa rhodozyma*, colorful odyssey. *Int. Microbiol.* 6, 169–174. doi: 10.1007/s10123-003-0130-3

Johnson, E. A., and Lewis, M. J. (1979). Astaxanthin formation by the yeast *Phafa rhodozyma*. *Microbiology* 115, 173–183. doi: 10.1099/00221287-115-1-173

Liu, S., Yi, H., Zhan, H., Wang, L., Wang, J. H., Li, Y., et al. (2021). Gibberellic acid-induced fatty acid metabolism and ABC transporters promote astaxanthin production in *Phaffia rhodozyma*. *J. Appl. Microbiol.* 32:390-400. doi: 10.1111/jam.15187

Ma, X. X., Jiang, Y. L., He, Y. X., Bao, R., Chen, Y., and Zhou, C. Z. (2009). Structures of yeast glutathione-S-transferase Gtt2 reveal a new catalytic type of GST family. *EMBO Rep.* 10, 1320–1326. doi: 10.1038/embor.2009.216

Miao, L., Chi, S., Tang, Y., Su, Z., Yin, T., Guan, G., et al. (2015). Astaxanthin biosynthesis is enhanced by high carotenogenic gene expression and decrease of fatty acids and ergosterol in a *Phaffia rhodozyma* mutant strain[J]. *Fems Yeast Res.* 11, 192–201. doi: 10.1111/j.1567-1364.2010.00705.x

Russell, D., and Sambrook, J. (2001). *Molecular Cloning: a Laboratory Manual*. New York: Cold Spring Harbor Laboratory Press.

Sandmann, G. (2015). Carotenoids of biotechnological importance. *Adv. Biochem. Eng.Biotechnol.* 148, 449–467. doi: 10.1007/10_2014_277

Sandmann, G., Pollmann, H., Gassel, S., and Breitenbach, J. (2021). *Xanthophyllomyces dendrorhous*, a versatile platform for the production of carotenoids and other acetyl-CoA-derived compounds.Adv. Exp. Med. Biol. 1261, 137-151. doi: 10.1007/978-981-15-7360-6_11

Sieiro, C., Poza, M., Miguel, T., and Villa, T. G. (2003). Genetic basis of microbial caroteno-genesis. *Int. Microbiol.* 6, 11–16. doi: 10.1007/s10123-003-0097-0

Verdoes, J. C., Krubasik, K. P., Sandmann, G., and Ooyen, A. J. (1999). Isolation and functional characterisation of a novel type of carotenoid biosynthetic gene from *Xanthophyllomyces dendrorhous*. *Mol. Genet. Genomics.* 262, 453–461. doi: 10.1007/s004380051105

Wang, W. l, Liu, M. T., Samia, F., Xue, Y. C., Wu, M. Q., Huang, X. X., et al. (2021). Effects of dietary *Phaffia rhodozyma* astaxanthin on growth performance, carotenoid analysis, biochemical and immune-physiological parameters,intestinal microbiota, and disease resistance in *Penaeus monodon*. *Front. Microbiol.* 12:762689. doi: 10.3389/fmicb.2021.762689

Yang, G., Wang, Y., Xia, D., Gao, C., and Yang, C. (2014). Overexpression of a *GST* gene (ThGSTZ1) from *Tamarix hispida* improves drought and salinity tolerance by enhancing the ability to scavenge reactive oxygen species. *Plant Cell Tissue Organ Cult.* 117, 148–153. doi: 10.1007/s11240-014-0424-5

**Conflict of Interest:** HZ and XL were employed by Beijing DaBeiNong Science and Technology Group Co., Ltd. (DBN).

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.

*Copyright © 2022 Shi, He, Zhang, Guo, Cheng, Liu, Wang and He. 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.*

frontiers | Frontiers in Microbiology

ORIGINAL RESEARCH
published: 06 April 2022
doi: 10.3389/fmicb.2022.815366

# Proteomics Reveal the Effect of Exogenous Electrons on Electroactive *Escherichia coli*

*Jiao Feng, Jia Feng, Chunqiu Li, Sheng Xu*, Xin Wang and Kequan Chen*

*State Key Laboratory of Materials-Oriented Chemical Engineering, College of Biotechnology and Pharmaceutical Engineering, Nanjing Tech University, Nanjing, China*

Microbial cells utilizing electricity to produce high-value fuels and chemicals are the foundation of the biocathodic bioelectrochemical system. However, molecular mechanisms of electron transfer and utilization have not been elucidated. In this work, *Escherichia coli* engineered by introducing the Mtr pathway from *Shewanella oneidensis* exhibited stronger electrochemical activity than control and could utilize exogenous electrons to stimulate metabolite profiles and boost succinate production in the bioelectrochemical system. Proteomic analysis and real-time PCR were performed to investigate the effect of exogenous electrons on electroactive *E. coli.* Bioinformatics analysis suggested that the proteins of molecular function associated with oxidoreductase activity, 4 iron, 4 sulfur([4Fe-4S]) cluster binding, iron-sulfur cluster binding, and metal cluster binding were positively affected by exogenous electrons. Moreover, mapping to the Kyoto Encyclopedia of Genes and Genomes pathway database showed that the up-regulated proteins were mainly involved in metabolic pathways of tricarboxylic acid cycle, pyruvate metabolism, and nitrogen metabolism pathway, providing support for the metabolic balance of microbial cells shifting toward reduced end-products due to electron utilization. Using a biochemical method, the *ompF*-overexpressed strain was employed to investigate the function of the channel protein. These findings provided a theoretical basis for further improving electron transfer and utilization efficiency, and contributed to the potential applications of the bioelectrochemical system.

**Keywords: *Escherichia coli*, exogenous electrons, proteomics, differentially expressed proteins, bioelectrochemical system**

OPEN ACCESS

***Edited by:***
*Niranjan Koirala, University of Macau, China*

***Reviewed by:***
*Wolfgang Buckel, University of Marburg, Germany*
*Isabelle Meynial-Salles, Institut National des Sciences Appliquées de Toulouse (INSA), France*

****Correspondence:***
*Sheng Xu*
*henryxu@njtech.edu.cn*

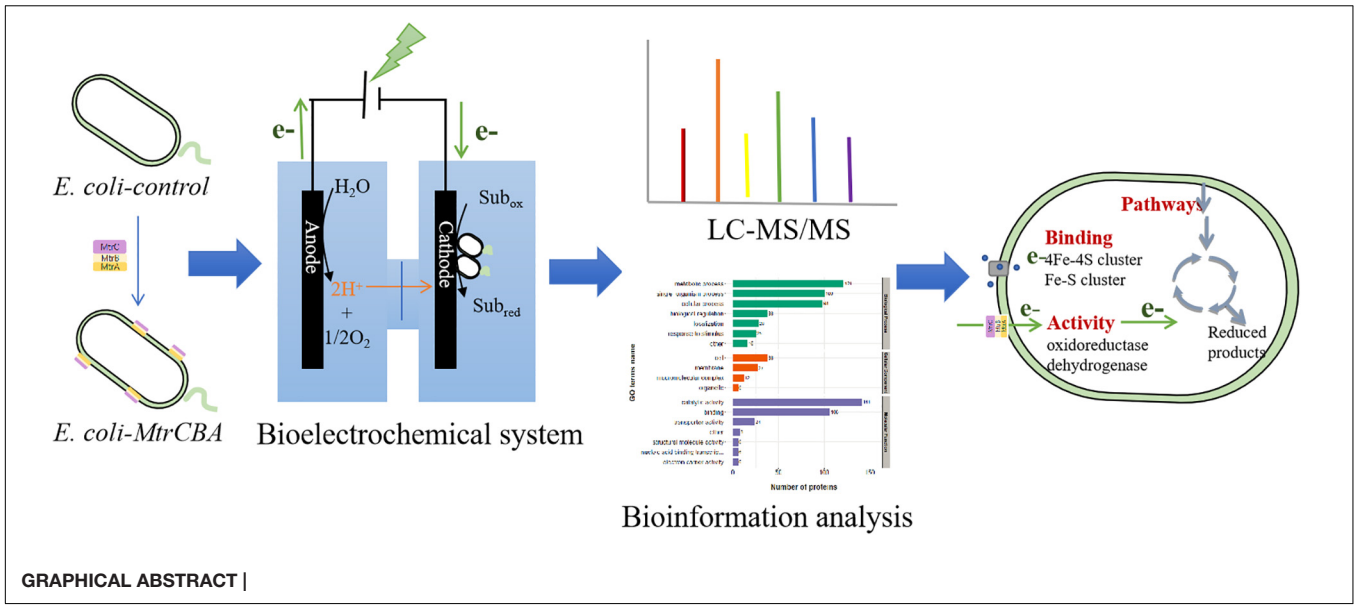
***Specialty section:***
*This article was submitted to Microbiotechnology, a section of the journal Frontiers in Microbiology*

***Received:*** *15 November 2021*
***Accepted:*** *17 February 2022*
***Published:*** *06 April 2022*

***Citation:***
*Feng J, Feng J, Li C, Xu S, Wang X and Chen K (2022) Proteomics Reveal the Effect of Exogenous Electrons on Electroactive Escherichia coli. Front. Microbiol. 13:815366. doi: 10.3389/fmicb.2022.815366*

## INTRODUCTION

In a bioelectrochemical system (BES), microorganisms exchange electrons directly or indirectly with solid electrodes via extracellular electron transfer to manufacture or utilize renewable resources (Kracke et al., 2015; Sun et al., 2020). Biocathodic BESs, e.g., those using microbial electrosynthesis (MES) or cathodic electro-fermentation (CEF), are considered to be a green, sustainable, and attractive novel technology. In such BESs, high-value biofuels and chemicals are converted by microbial catalysis from $CO_2$ or other substrates with electricity as the energy source, thereby achieving renewable energy storage in chemical bonds (Kracke et al., 2015; Patil et al., 2015;




**GRAPHICAL ABSTRACT |**

Jiang and Zeng, 2018; Kracke et al., 2018). One advantage of these BESs is that low-cost carbon sources or wastes can be utilized by microorganisms as substrates. Another main advantage is the ability of microorganisms to utilize clean electricity to boost biofuel and chemical generation.

Microorganisms are the catalytic center of biocathodic BESs, whose parameters, activities, and functions are related to chemical production, reactor performance, and efficiency (Kracke et al., 2015; Kokko et al., 2016; Kracke et al., 2018). Electrons are transferred from cathodes into microbial cells, and then participate as essential elements in intracellular redox reactions and microbial metabolism (Sun et al., 2020). This process is applied to break through intracellular redox or energy limitations and shift metabolite profiles toward increasing terminal products of interest (Kracke et al., 2015; Kracke et al., 2018; Wu et al., 2019). Therefore, the ideal biocathodic BES must require microbial cells to take up electrons and efficiently incorporate them into a target cellular metabolic pathway (Kracke et al., 2018). Increasing efforts have focused on creating or promoting electron transfer (Tefft and TerAvest, 2019; Wu et al., 2019; Zhao et al., 2020). However, the molecular mechanism of electron transfer and effects on microbial cells are not clear (Korth and Harnisch, 2019; Sun et al., 2020). Existing research data, especially of biocathodes, are insufficient to conduct a comprehensive analysis, hindering further engineering of microorganisms to use electrons for efficiently generating target products, thereby limiting practical applications of biocathodic BES.

Proteomics technology is the most common method for conducting large-scale analysis on proteomic profiles and screening different target proteins (Zhang et al., 2020). A complete map of expression states derived from quantitative protein information provides insights into biological processes and helps reveal biological mechanisms (Schmidt et al., 2016; Zhang et al., 2020). The proteomes of many typical model microbes have been supported by numerous resources and databases. The *Escherichia coli* proteome is by far the most well defined bacterial proteome and is thus extensively used for studying general features of the prokaryotic proteome (Soufi et al., 2015; Fortuin et al., 2020).

Exploiting *E. coli* to produce electrosynthetic chemicals by engineering the electron transfer pathway is an important research area in BESs, which could make up for some deficiencies of electrochemically active bacteria and broaden the application prospects of biocathodic BES (Kracke et al., 2015; Mayr et al., 2019; Wu et al., 2019). The Mtr pathway from *Shewanella oneidensis* MR-1 is one of the most extensively studied electron transfer pathways and has been successfully constructed in *E. coli* (Jensen et al., 2010; Ross et al., 2011; Wu et al., 2019). The major components of this pathway are MtrCAB complexes (encoded by the *mtrCAB* cluster) containing MtrC, MtrB, and MtrA proteins (Jensen et al., 2010). We previously introduced the Mtr pathway into *E. coli* cells by expressing *mtrCAB* from *S. oneidensis* and the engineered *E. coli* could use the Mtr pathway to transfer electrons from the cathode into *E. coli* cells (Feng et al., 2020a). However, the molecular mechanism of electron utilization on *E. coli* cells has not been elucidated.

Based on these previous findings, this work first compared the metabolite profiles and electrochemical activity of engineered *E. coli* and the control strain without the Mtr pathway in the BES driven by electricity. The engineered *E. coli* exhibited better electrochemical activity and could use exogenous electrons to stimulate metabolite profiles, increasing succinate production by 18%. Subsequently, real-time PCR (RT-PCR) and proteomics methods were used to comprehensively analyze the specific gene expression changes and the differentially expressed proteins (DEPs) of *E. coli*. Moreover, the channel protein OmpF was found to be up-regulated in the proteomic data of subcellular location,

**TABLE 1 |** Strains and plasmids.

| Strains/plasmids | Relevant characteristics | References/source |
|---|---|---|
| *E. coli* Trans1-T1 | Used for cloning | TransGen Biotech, Beijing, China |
| *E. coli* BA102 | Used as host strain | Feng et al., 2020a |
| *E. coli-control* | *E. coli* BA102 harboring plasmid pCWJ-*ccmA-H* | Feng et al., 2020a |
| *E. coli-MtrCBA* | *E. coli* BA102 harboring plasmids pBBR1MCS-5-*mtrCAB* and pCWJ-*ccmA-H* | Feng et al., 2020a |
| *E. coli-mtr-ompF* | *E. coli* BA102 harboring plasmids pBBR1MCS-5-*mtrCAB* and pCWJ-*ompF* | This work |
| *E. coli-mtr* | *E. coli* BA102 harboring plasmids pBBR1MCS-5-*mtrCAB* | This work |
| pBBR1MCS-5-*mtrCAB* | pBBR1MCS-5 carrying the *mtrCAB* gene of *S. oneidensis* MR-1 under constitutive *plac* control; $Gm^r$ | Feng et al., 2020a |
| pCWJ-*ccmA-H* | pCWJ carrying the *ccmABCDEFGH* gene of *E. coli* K12 under constitutive *ptrc* control; $Cm^r$ | Feng et al., 2020a |
| pCWJ-*ompF* | pCWJ carrying the *ompF* gene of *E. coli* BA102 under constitutive *ptrc* control; $Cm^r$ | This work |

and thus identified by further gene expression, showing the increased redox reactions between the electrode and the *E. coli* cells.

# MATERIALS AND METHODS

## Bacterial Strains and Culture Conditions

Bacterial strains and plasmids are listed in **Table 1**, and primers are in **Supplementary Material**. The recombinant strains used in this work were formed by transforming the corresponding recombinant plasmids (**Table 1**). The recombinant plasmid pCWJ-*ompF* was constructed by ligating *ompF* gene fragment into pCWJ between the *Eco*RI and *Sac*I sites. The *ompF* gene fragment was amplified by PCR using ompF-1 and ompF-2 as primers and the *E. coli* BA102 genome as the template.

A suspension of overnight-activated strains was inoculated into 100 ml of 2 × Yeast extract and Typeptone (2 × YT) medium in 500 ml flasks supplemented with the corresponding antibiotics and isopropyl-beta-D-thiogalactopyranoside (0.1 mM chloramphenicol for *E. coli-control*; 0.1 mM gentamicin for *E. coli-mtr*; and 0.1 mM chloramphenicol and 0.1 mM gentamicin for *E. coli-MtrCBA* and *E. coli-mtr-ompF*). The cultures were grown at 30°C with shaking at 200 rpm for 12–14 h. The suspended cells harvested by centrifugation were saved for further studies.

## Construction of Bioelectrochemical Reactors and Electrochemical Analysis

The harvested cells dispersed into the cathode chamber ($OD_{600}$ = 1) of the dual chamber "H" bioelectrochemical reactors with an internal volume of 100 mL and fine carbon felts as the cathode and anode electrodes. The anode electrolyte and cathode medium were described in our previous work (Feng et al., 2020a). The anodic electrolyte contained 25 mM NaCl, 15.7 mM $NaH_2PO_4{\cdot}2H_2O$, 7.0 mM $NaHPO_4{\cdot}12H_2O$, and 0.02% dithiothreitol (DTT) (filtered through a 0.22 nm syringe filter before use). The cathodic medium was composed of 47.8 mM $NaH_2PO_4{\cdot}2H_2O$, 88 mM $NaHPO_4{\cdot}12H_2O$, 130 mM $NaHCO_3$, and 5.6 g $L^{-1}$ yeast extract. Prior to operation, D-glucose was added to the cathodic medium. Cathodes were purged with $CO_2$, imposed at -0.8 V (vs. Ag/AgCl), and stirred using a magnetic stirrer at 200 rpm.

The cyclic voltammograms (CVs), electrochemical impedance spectroscopy (EIS), and the constant potential [−0.8 V (vs. Ag/AgCl)] were performed using an electrochemical instrument (PMC 1000/DC, AMETEK, Berwyn, PA, United States). The CV scan rate was 20 mV/s in the range from -800 to 600 mV (vs. Ag/AgCl). The EIS experiments (0.01–$10^5$ Hz) were carried out at 10 mV amplitude.

## Organic Acid Analysis

Organic acids were analyzed with high-performance liquid chromatography (HPLC) (Agilent 1290; Agilent Technologies, Santa Clara, CA, United States) via an HPX-87H column (300–7.8 mm, Bio-Rad, Hercules, CA, United States), equipped with an ultraviolet spectrophotometric detector (at 215 nm) and a refractive index detector. The mobile phase was 8 mM sulfuric acid and the flow rate was 0.5 mL/min.

## Real-Time PCR Analysis

RNA isolation was prepared using the RNAprep pure cell/bacteria kit (TIANGEN BIOTECH, Beijing, China). RNA concentration was determined using a NanoDrop 2000 spectrophotometer (Thermo Scientific, Waltham, MA, United States), and RNA purity was assessed based on the signals at 230, 260, and 280 nm. The integrity of extracted RNA was determined by agarose gel electrophoresis. RNA was transcribed into cDNA as an RT-PCR template using the PrimeScript RT Master Mix (Perfect Real Time) (Takara, Kusatsu, Shiga, Japan). RT-PCR was performed using SYBR Premix EX Taq (Takara, Kusatsu, Shiga, Japan) and a 7300 Plus Real-Time PCR System (Thermo Scientific, Waltham, MA, United States). *rpoD, ihfB,* and *recA* were used as housekeeping genes (Nitzschke and Bettenbrock, 2018), and gene expression was calculated by relative quantification applying the ΔΔCt method with three replicates. Reaction systems of RT-PCR analysis and gene-specific primers used for quantifying corresponding transcript levels are provided in **Supplementary Material**. The amplification conditions for RT-PCR analysis were: 95°C for 30 s, followed by 40 cycles each consisting of 95°C for 5 s, 60°C for 34 s, and 60°C for 1 min.

## Proteomic Analysis

### Protein Extraction, Trypsin Digestion, and Tandem Mass Tag Labeling

The bacterial sample was added to lysis buffer (8 M urea, 1% Protease Inhibitor Cocktail, 3 μM TSA, 50 mM NAM, and 2 mM EDTA) and sonicated on ice three times using a high intensity ultrasonic processor. After centrifugation (12,000 × g, 10 min, 4°C), the supernatant was collected into a new EP tube and the protein concentration was measured according to the manufacturer's instructions for the BCA kit. For digestion, the protein solution was reduced, alkylated, diluted, and digested by trypsin twice according to a standard protocol reported in previous studies (Li et al., 2018; Zhang et al., 2020). The peptide obtained by trypsin digestion was desalted using a Strata X C18 SPE column (Phenomenex, Torrance, CA, United States) and dried by vacuum. The peptide was reconstituted in 0.5 M triethylammonium bicarbonate (TEAB) and processed with a Tandem Mass Tag (TMT) kit. TMT kits enable multiplex relative quantitation by mass spectrometry (MS). Briefly, one unit of TMT reagent was thawed and reconstituted in acetonitrile. Then TMT label reagent was added to differentially label the sample with six TMT tags (Control group 1–3: 126, 127, and 128 label; MtrCBA group 1–3: 129, 130, and 131 label). After incubation for 2 h at room temperature, the peptide mixtures were pooled, desalted, and dried by vacuum centrifugation.

### High-Performance Liquid Chromatography Fractionation

The peptides were separated into fractions using high pH reverse-phase HPLC. Thermo Betasil C18 column (5 μm particles, 250 mm, and 10 mm i.d.) was used to perform the HPLC fractionation under a gradient of 8–32% acetonitrile (pH 9.0) within 60 min. Sixty fractions were collected and then combined to 10 fractions. After drying in a vacuum centrifuge, the fractions were dissolved in 0.1% formic acid for liquid chromatography tandem mass spectrometry (LC-MS/MS) analysis.

### Liquid Chromatography Tandem Mass Spectrometry Analysis

For the collected fractions, peptides were analyzed (three repeats per component) using an EASY-nLC 1000 UPLC system coupled online with Q Exactive™ Plus (Thermo, United States). A home-made reversed-phase analytical column (15 cm length, 75 μm i.d.) was used to separate the peptides. Mobile phases A (2% acetonitrile in 0.1% formic acid) and B (0.1% formic acid in 90% acetonitrile) were employed at a constant flow rate of 500 nL/min by the following method: linear gradient from 9% mobile phase B to 26% over 23 min; linear gradient to 38% mobile phase B in 9 min; 38–80% in 4 min and holding at 80% for 4 min. The mass spectrometry parameters were set as follows: the electrospray voltage was set as 2.0 kV; the m/z scan range was set as 350 to 1,800; and intact peptides were detected at 70,000 resolution in the Orbitrap. Peptides were then selected for MS/MS using NCE setting as 28 and the fragments were detected at 17,500 resolution in the Orbitrap. A data-dependent process was conducted that alternated between one MS scan followed by 20 MS/MS scans with 15.0 s dynamic exclusion. The fixed first mass was 100 m/z.

### Proteomics Data Analysis and Bioinformatics Methods

The resulting MS/MS data were searched against the *E. coli* K12 database (4,446 sequences) using Maxquant search engine (v.1.5.2.8). The parameters were as follows: cleavage enzyme was trypsin/P; missing cleavages was set to 2; first search peptide tolerance and main search peptide tolerance were 20 and 5 ppm, respectively; the fragment mass tolerance was 0.02 Da; fixed modification was carbamidomethyl, variable modifications were acetylation modification and oxidation on Met. The false discovery rate (FDR) was adjusted to below 1%.

Gene Ontology (GO) proteins were annotated using InterProScan and derived from the UniProt-GOA database,[1] which were classified according to three categories of GO annotation: biological process, cellular component, and molecular function. The protein pathway annotation was performed using Kyoto Encyclopedia of Genes and Genomes (KEGG) online tool KAAS (KEGG Automatic Annotation Server) based on the KEGG database. Subcellular localization predication software wolfpsort was used to predict subcellular localization. For functional enrichment, enrichment of GO analysis and pathway analysis were researched by employing a two-tailed Fisher's exact test to test the enrichment of the DEPs against all identified proteins, and the corrected $p$-value $< 0.05$ is considered significant. Based on DEP functional classification, further hierarchical clustering was conducted, then visualized by a heat map using the "heatmap.2" function from the "gplots" R-package.

# RESULTS

## Effect of the Exogenous Electrons on Metabolite Profiles

The biocathodic BES has been widely applied to convert electrons from electrodes into intracellular reducing power, promoting the production of target metabolites. In the succinate anaerobic fermentation pathway (see **Supplementary Material** for metabolic pathway), two molecules of NADH are produced from one molecule of glucose into two molecules of intermediate phosphoenolpyruvate (PEP), while four NADH molecules are needed to synthesize two molecules of succinate from two PEP molecules with $CO_2$ supplementation. Therefore, insufficient reducing power is the critical limiting factor in succinate yield (Zhu and Tang, 2017; Wu et al., 2019). In our work, succinate production in the BES was investigated and the succinate-producing strain *E. coli* BA102 was used as a host strain. This host strain was constructed to reduce the main byproduct by deleting *pflB* and *ldhA* genes. In addition, the pool of available PEP was increased by disabling the *ptsG* gene encoding an enzyme for the PEP-dependent phosphotransferase system to preserve enhanced succinate production (see **Supplementary Material** for metabolic pathway) (Cao et al., 2011; Singh et al., 2011; Feng et al., 2020b).

[1] http://www.ebi.ac.uk/GOA/

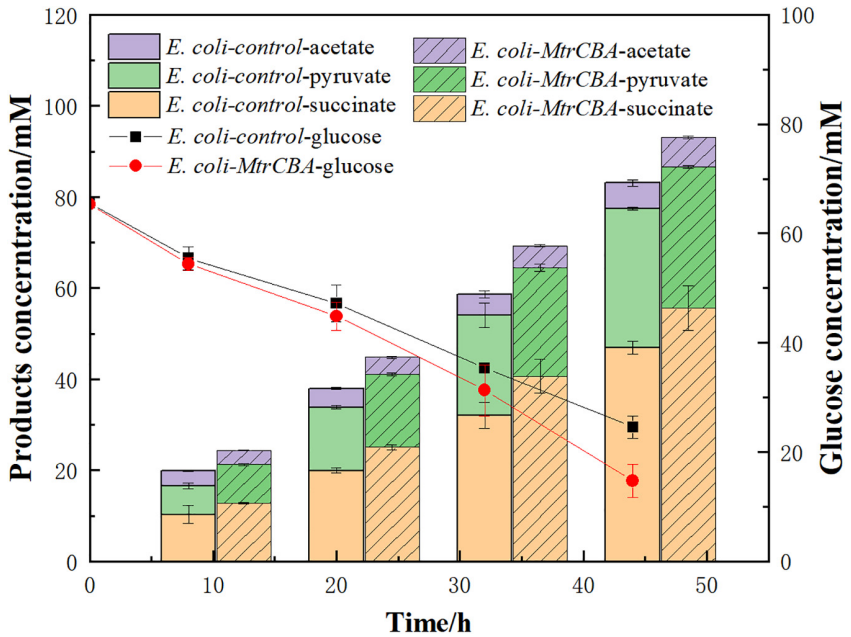



**FIGURE 1 |** Effects of the exogenous electrons on major metabolite profiles. Bioelectrochemical reactors inoculated with *E. coli-control* or *E. coli-MtrCBA* ($OD_{600}$ = 1) in cathode mediums were imposed at a constant potential [-0.8 V (vs. Ag/AgCl)].

*MtrCAB*-expressing *E. coli* strain *E. coli-Mtr* was constructed by introducing the Mtr pathway from *S. oneidensis* MR-1 and the engineered Mtr pathway allowed electron transfer from the cathode to *E. coli* cells, which realized the use of exogenous electrons (Feng et al., 2020a). To better understand the effect of exogenous electrons on metabolite profiles, *E. coli-MtrCBA* and *E. coli-control* were used in BES with glucose as the carbon source and $CO_2$ supplementation. As shown in **Figure 1**, *E. coli-MtrCBA* had faster glucose consumption and succinate production rates, and the succinate concentration of *E. coli-MtrCBA* was 18% higher than that of *E. coli-control* after 44 h. Pyruvate was accumulated as a result of gene knockouts. Production of the main by-products pyruvate and acetate showed no significant difference between *E. coli-MtrCBA* and *E. coli-control,* while higher levels of succinate/pyruvate and succinate/acetate were shown in *E. coli-MtrCBA* (increased by 17 and 2%, respectively). The results indicated that the intracellular metabolite profiles were stimulated and changed to facilitate the production of reduced end-products (succinate) in the BES driven by electricity (Wu et al., 2019).

## Electrochemical Activity of *Escherichia coli* Cells in the Bioelectrochemical System

To analyze the electrochemical performance of the BES, CV curves and nyquist plots of impedance were obtained from CV and EIS test to investigate the redox reaction characteristics between the electrode and the microbial cells. CVs of *E. coli-MtrCBA* and *E. coli-control* are presented in **Figure 2A**. Two oxidation peaks [ranging from +0.2 to +0.40 V (vs. Ag/AgCl) and −0.20 to 0.0 V (vs. Ag/AgCl)] and one reduction peak between −0.30 and −0.10 V (vs. Ag/AgCl) were observed, consistent with the broad potential of c-type cytochromes ranging from −0.45 to +0.20 V and the mediator molecules excreted by *E. coli* ranging from −0.45 to −0.3 V (Qiao et al., 2008; Ross et al., 2011; Kracke et al., 2015). This demonstrated that the redox peaks were caused by c-type cytochromes and endogenous redox species under the effect of exogenous electrons (Feng et al., 2020a). The higher peak area obtained from *E. coli-MtrCBA* compared with the control indicated that stronger redox reactions occurred in the BES inoculated with *E. coli-MtrCBA*. Moreover, the oxidation and reduction peaks of *E. coli-MtrCBA* shifted positively compared with the control group. Since proton transfer coupled with electron transfer is involved in the redox reaction, the CV result is affected by solution pH value. A previous report showed that the midpoint potential of the redox peak shifts positively with decreased pH (Qiao et al., 2008). More succinic acid produced by *E. coli-MtrCBA* led to a lower pH, resulting in a slightly positive shift in the redox peaks of *E. coli-MtrCBA*. In addition, the electron transfer efficiency at the cell-electrode interfaces was analyzed by EIS. EIS analysis reflected an improved electron transfer rate of *E. coli-MtrCBA* (**Figure 2B**) (Liu et al., 2012). These results confirmed that the engineered *E. coli* exhibited much better electrochemical activity and more effective biocatalytic behavior in BES compared to control.

## Real-Time PCR Analysis of Specific Gene Expression Changes

To analyze the effect of exogenous electrons transported into cells by intracellular metabolic pathways, the expression changes of

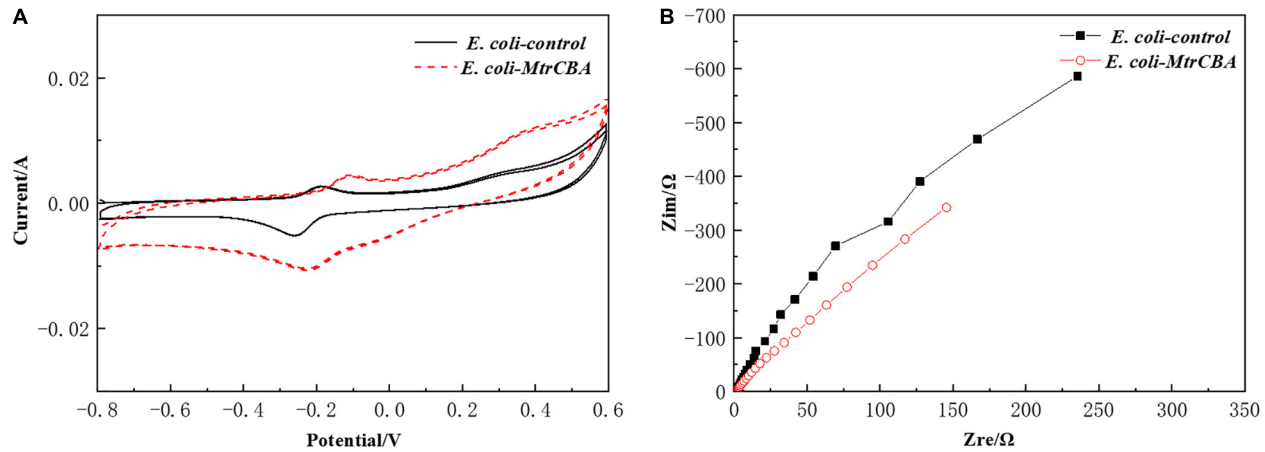



**FIGURE 2 |** Effects of exogenous electrons on electrochemical activity. Cyclic voltammogram (CV) analysis scanned at 10 mV/s in the range of -800–600 mV **(A)**. Electrochemical impedance spectroscopy scanned at 0.01−$10^5$ Hz with a 10 mV amplitude **(B)**.

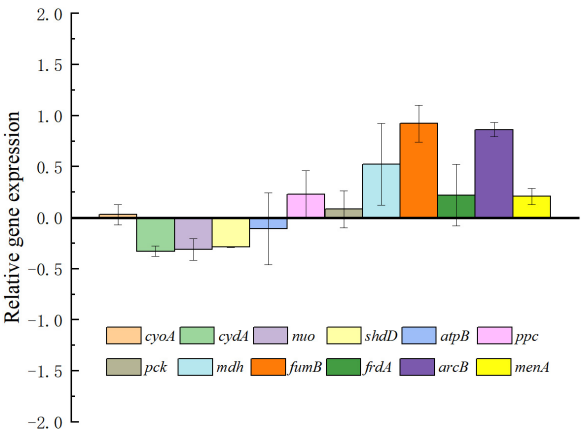



**FIGURE 3 |** Results of RT-PCR analysis. The transcription pattern was normalized to the reference genes for the expression of *E. coli-control*. Genes encoding cytochrome bo3 (*cyoA*), cytochrome bd (*cydA*), NADH: ubiquinone oxidoreductase (*nuo*) and succinate dehydrogenase (*sdhD*), ATP synthase (*atpB*), phosphoenolpyruvate carboxylase (*ppc*), phosphoenolpyruvate carboxykinase (*pck*), malate dehydrogenase (*mdh*), fumarate hydratase (*fumB*), fumarate reductase (*frdA*), sensor histidine kinase (*arcB*), and 1,4-dihydroxy-2-naphthoate octaprenyltransferase (*menA*) were analyzed. The unchanged relative gene expression level was calculated as 0. The Y-axis was formatted in a logarithmic scale with base 10.

important genes related to terminal oxidases, respiratory chain, and the tricarboxylic acid (TCA) cycle were investigated via RT-PCR. The relative expression levels for selected genes are shown in **Figure 3**. As expected, genes connected to the TCA cycle including *fum* and *mdh* were upregulated in *E. coli-MtrCBA*. This is consistent with more succinate being produced by *E. coli-MtrCBA* through accepting cathodic electrons, hinting at a more active succinate synthesis pathway. There was an obvious change in expression level of intracellular redox regulation system protein gene *arcB*. The Arc two-component system comprising ArcA and ArcB is the key element in the global transcriptional regulatory network affecting the expression of numerous genes associated with substrate metabolism and energy metabolism to adapt to various conditions (van Beilen and Hellingwerf, 2016; Teran-Melo et al., 2018). Van Beilen and Hellingwerf reported that reduced quinone species could activate ArcB as a redox-sensitive, membrane-embedded kinase (van Beilen and Hellingwerf, 2016). These results suggested that exogenous electrons might affect a wide range of intracellular gene expression and expressed proteins and then biological reactions and pathways.

## Comprehensive Proteomics Analysis of Electron Utilization in *Escherichia coli* Cells

The mechanism of electron transfer and utilization is complex and not fully understood. To reveal key intracellular proteins or pathways responding to exogenous electrons and further comprehensively analyze the effect of electron transfer and utilization, systematic proteomic analysis research was conducted.

### Tandem Mass Tag Analysis of Differentially Expressed Proteins

In total, 2,268 proteins were detected by TMT quantitative proteomics. Of these, 2,110 proteins were identified as quantifiable DEPs. Based on a change threshold of 1.2-fold, the results of protein expression levels are shown in the volcano plot in **Figure 4A**. Compared with the control, exogenous electrons from the cathode led to up-regulation of 152 proteins and down-regulation of 142 proteins in *E. coli-MtrCBA*.

### Gene Ontology Annotation and Enrichment Analysis

In order to acquire protein functional information, the DEPs were annotated and classified by GO annotation according to three categories (biological process, cellular component, and molecular function) and subcellular location (**Figures 4B,C**). Among them, the DEPs were respectively located in the

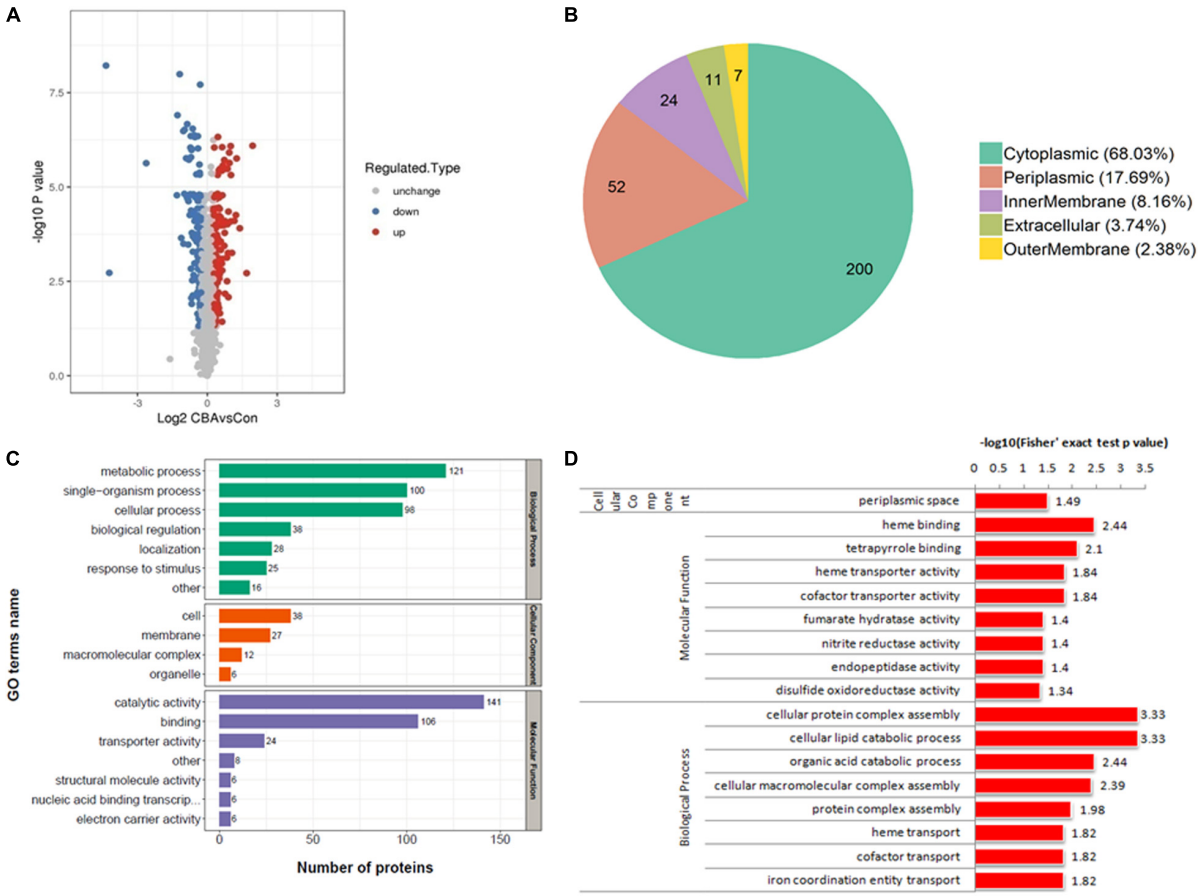



**FIGURE 4 |** Results of proteomic analysis. The volcano plot shows the upregulated (red) or downregulated (blue) proteins between *E. coli-MtrCBA/E. coli-control* **(A)**. Subcellular localization map of the differentially expressed proteins **(B)**. The differentially expressed proteins classified according to Gene Ontology (GO) annotation **(C)**. GO enrichment analysis of the differentially expressed proteins **(D)**. –Log 10 *P* value represents the negative logarithm of *p* value calculated by Student's *t*-test to base 10. Log 2 CBAvsCon represents the logarithm of the proteins' difference multiple to the base 2. –Log10 (Fisher's exact test *p* value) represents the negative logarithm of *p* value calculated by the Fisher's exact test to base 10.

cytoplasm (200 proteins, accounting for 68%), the periplasm (52 proteins, accounting for 18%), the innermembrane (24 proteins, accounting for 8%), the extracellular matrix (11 proteins, accounting for 4%), and the outermembrane (seven proteins, accounting for 2%).

Based on GO annotations and classification into three categories, GO enrichment analysis was performed using a two-tailed Fisher's exact test to identify significantly enriched functional groups ($p$-value $< 0.05$) of the DEPs in the case of exogenous electron transfer into *E. coli* cells through the Mtr pathway (**Figure 4D**). In terms of cellular components, the DEPs were highly significantly enriched in the perplasmic space. MtrCAB complexes that span the outer membrane allow electron transfer from the decaheme cytochrome MtrC located on the extracellular face of the outer membrane to the periplasmic decaheme cytochrome MtrA, so proteins in the periplasmic space are directly affected by electrons (Jensen et al., 2010). DEPs were highly significantly categorized into eight biological processes related to the metabolism of protein, lipid, and organic acid, including "cellular protein complex assembly," "cellular lipid catabolic process," "cellular macromolecular complex assembly," "protein complex assembly," and "organic acid catabolic process," and associated with GO terms about electrons, including "heme transport," "cofactor transport," and "iron coordination entity transport." Among molecular function-associated GO terms, the main enriched molecular functions of the DEPs were related to binding and transport, including "heme binding," "tetrapyrrole binding," "heme transporter activity," and "cofactor transporter activity." MtrC and MtrA, the important parts of MtrCAB complexes, are decaheme cytochrome c, and their maturations are required to load heme (Jensen et al., 2010). Cofactors, such as NADH, $MKH_2$, and FADH, are intracellular electron acceptors, which are used to store reducing power. In addition, "nitrite reductase activity" and "disulfide oxidoreductase activity," which functionally catalyze the reactions commonly involved in electrons, were significantly enriched.

## Enrichment-Based Clustering Analysis

On the basis of enrichment analysis, further cluster analysis was carried out to find correlations or new insights into protein functions and pathways. According to differential expression multiple, the DEPs were divided into four parts: Q1, Q2, Q3, and

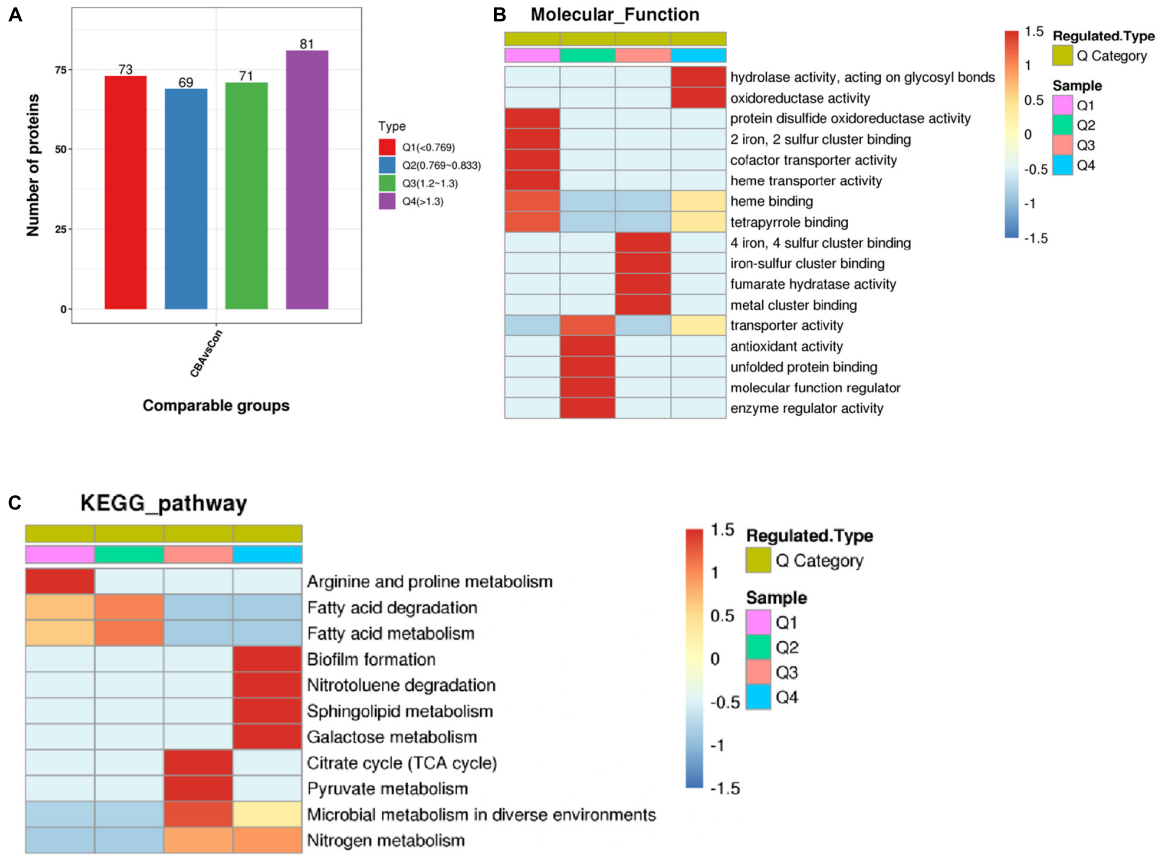



**FIGURE 5 |** Enrichment-based clustering analysis. The differentially expressed proteins were divided into four parts: Q1 (<0.769), Q2 (0.769–0.833), Q3 (1.2–1.3), and Q4 (>1.3) according to differential expression multiple **(A)**. Hierarchical clustering based on functional classification of differentially expressed proteins (Molecular Function) **(B)** and KEGG Pathway based on *Escherichia coli* K-12 MG1655 **(C)**. Color from blue to red represents the degree of enrichment from weak to strong, respectively.

Q4 (**Figure 5A**) and then clustered by hierarchical clustering. The results were visualized by heat maps.

A cluster analysis heat map based on molecular function enrichment is shown in **Figure 5B**. "Hydrolase activity" and "oxidoreductase activity" (Q4 > 1.3) were significantly enhanced by exogenous electrons transferred into *E. coli* cells through the Mtr pathway, functionally catalyzing the reactions commonly accompanied with electrons or protons. DEPs of molecular function associated with "4 iron, 4 sulfur([4Fe-4S]) cluster binding," "iron-sulfur (Fe-S) cluster binding," and "metal cluster binding" ($1.2 \leq Q3 \leq 1.3$) were up-regulated, while those associated with "2 iron, 2 sulfur ([2Fe-2S]) cluster binding" (Q4 < 0.769) were down-regulated (**Supplementary Material**). Fe-S clusters are versatile and essential biological cofactors biosynthesized by the Isc and Suf pathways in *E. coli* (Mettert and Kiley, 2015). Reducing power such as in the presence of the reducing agent DTT is absolutely required for mature [4Fe-4S] proteins, due to the requirement of electrons to generate [4Fe-4S] from [2Fe-2S] clusters (Chandramouli et al., 2007; Beilschmidt et al., 2017). Electrons were transferred from the cathode into *E. coli* cells through the Mtr pathway, resulting in a reductive intracellular environment and providing electrons for [4Fe-4S] generation. This further promoted [4Fe-4S] cluster binding and [4Fe-4S] protein maturation.

Fe-S cluster-containing proteins are partly reductases and dehydrogenases, in which Fe-S clusters are the relay station or conduit participating in electron transfer. Fe-S clusters also serve as prosthetic groups acting as binding sites or participating in enzyme catalysis. In addition, the up-regulation of "fumarate hydratase activity" suggested that it promoted formation of fumarate, the precursor of succinate. Fumarate hydratase is also a [4Fe-4S] cluster-containing protein, in which the [4Fe-4S] cluster is used for substrate binding and readily oxidized to an inactive [3Fe-4S] cluster (Flint et al., 1992). A reductive intracellular environment caused by exogenous electrons could promote binding and stable activity of [4Fe-4S] clusters. Unexpectedly, the DEPs were significantly enriched in terms of negative regulation for "heme transporter activity," "heme binding," and "cofactor transporter activity." Further analysis of the specific DEPs showed that CcmC and CcmF were the main DEPs in three terms related to the synthesis and maturation of the MtrCAB complexes. MtrCAB complexes spanning the outer membrane contain two cytochrome c MtrA and MtrC, whose maturation is required for cytochrome c maturation proteins

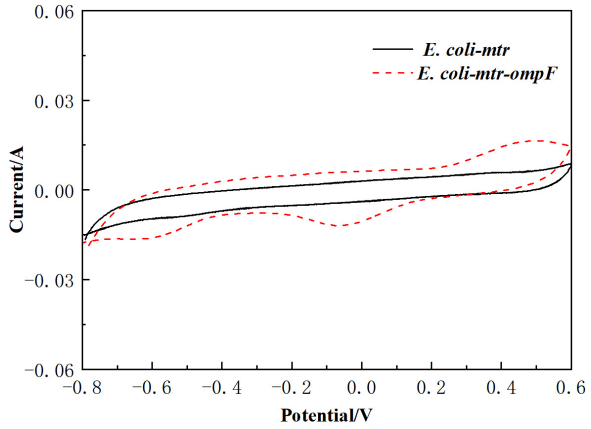



**FIGURE 6 |** CV analysis of *E. coli-mtr* and *E. coli-mtr-ompF* scanned from -800 to 600 mV.

CcmABCDEFGH (Jensen et al., 2010; Feng et al., 2020a). The polypeptides of MtrCAB and CcmABCDEFGH are secreted into the periplasm by the Sec secretion system, and then heme is translocated into the periplasm and loaded to form mature MtrC and MtrA by Ccm machinery. During synthesis and maturation of the MtrCAB complexes, there are multiple regulation points: excessive production or secretion of Mtr can inhibit and decrease Ccm, and inadequate heme production can lead to degradation of apocytochrome c or cytochrome c (Goldbeck et al., 2013). These might be the reasons why heme binding and transporter activity were down-regulated.

A cluster analysis heat map based on KEGG pathway enrichment is shown in **Figure 5C**. Among these metabolic pathways, DEPs are mainly connected with biofilm and cell membrane, such as the up-regulation of biofilm formation and sphingolipid metabolism ($Q > 1.3$), and down-regulation of fatty acid metabolism and fatty acid degradation ($0.769 \leq Q \leq 0.833$). These findings indicated that biofilm formation and cell membrane composition were significantly affected. The TCA cycle, pyruvate metabolism, and nitrogen metabolism pathway were promoted, which allowed the cells to produce more succinate (Feng et al., 2020a).

## The Channel Protein Promoted Redox Reactions Between the Electrode and *Escherichia coli* Cells

The results of electrochemical analysis confirmed that *E. coli-MtrCBA* exhibited much better bioelectrocatalytic activity and catalyzed stronger redox reactions compared with the control in the BES. The contributing factor was endogenous molecules under electrochemical stress (Feng et al., 2020a). Moreover, proteomic analysis indicated that the cell membranes were affected during the process of electron utilization in the BES. Combined with the data of subcellular location, seven of the significant DEPs were located in the outer membrane. Excluding uncharacterized proteins, the up-regulated outer membrane protein OmpF (*E.coli-MtrCBA/E.coli-control* Ratio = 1.653) was further identified. OmpF is a highly abundant primary porin in *E. coli*, which enables passive diffusion of hydrophilic molecules and influences outer membrane permeability (Duval et al., 2017).

To determine that the channel protein OmpF assisted with the redox reactions between the electrode and *E. coli* cells, *ompF* overexpressed strain *E. coli-mtr-ompF* and control *E. coli-mtr* were constructed and inserted into the BES to perform a CV test. As shown in **Figure 6**, significant redox peaks were observed from the BES inoculated with *E. coli-mtr-ompF*, indicating that OmpF might reduce transmission resistance of the cell membrane and improve outer membrane permeability to enhance the redox reactions. This was one of the reasons for the better bioelectrocatalytic activity exhibited by the engineered *E. coli-MtrCBA*.

# DISCUSSION

The process by which microorganisms utilize electrons from a cathode to produce high-value fuels and chemicals is the basis of the biocathodic BES. However, major knowledge gaps on molecular mechanisms of electron transfer and the effect on microbial cells have limited further engineering of microorganisms and practical applications of biocathodic BES. In order to utilize the electrons of cathodes, *E. coli* engineered with the Mtr pathway was constructed by expressing *mtrCAB* from *S. oneidensis* MR-1. The engineered *E. coli* exhibited much better electrochemical activity and more effective biocatalytic behavior than control in BES based on CV and EIS analyses. The Mtr pathway in *S. oneidensis* MR-1 is a typical bidirectional electron transfer pathway, in which electrons flow from cells into anodes or from cathodes into cells (Ross et al., 2011; Feng et al., 2020a). Moreover, the Mtr pathway has been successfully constructed in *E. coli* (Jensen et al., 2010; Goldbeck et al., 2013; Wu et al., 2019). The electrochemical performance of *E. coli-MtrCBA* indicated that our engineered *E. coli* cells could transfer electrons from the electrodes to cells to utilize exogenous electrons. Compared with *E. coli-control*, the electroactive *E. coli-MtrCBA* could boost succinate production by 18% and displayed higher levels of succinate/pyruvate and succinate/acetate (enhanced by 17 and 2% respectively) in BES driven by electricity. The results suggested that *E. coli-MtrCBA* accepted electrons from the cathode through the Mtr pathway and utilized the electrons to obtain excess reducing power, which led to an imbalance of the intracellular redox state and thus metabolite profiles were stimulated and changed to facilitate the production of reduced products (Wu et al., 2019).

Subsequently, systematic proteomic analysis research and RT-PCR were performed to reveal key intracellular proteins or pathways in response to exogenous electrons and comprehensively analyze the effect of electron utilization on *E. coli* cells. Compared with the control, 152 up-regulated and 142 down-regulated proteins were identified in *E. coli-MtrCBA*, respectively. Bioinformatics analysis showed that the proteins of molecular function associated with hydrolase activity, oxidoreductase activity, [4Fe-4S] cluster binding, Fe-S cluster binding, metal cluster binding, and fumarate hydratase activity were positively affected by exogenous electrons. In addition to the proteins that functionally catalyze reactions involving

electrons or protons, the Fe-S cluster was worth further study. Fe-S cluster-containing proteins are partly electron-accepting terminal oxidoreductases and dehydrogenases participating in multiple cellular processes, and Fe-S clusters are the relay station used to exchange electrons within the proteins (de Choudens and Barras, 2017; Feng et al., 2020a). Fe-S clusters can also function as prosthetic groups acting as the binding sites or participating in enzyme catalysis. Due to the requirement of electrons to generate [4Fe-4S] from [2Fe-2S] clusters, sufficient reducing power was absolutely required to generate mature [4Fe-4S] proteins and also stable activity of [4Fe-4S] clusters (Chandramouli et al., 2007; Beilschmidt et al., 2017).

Up-regulated proteins were mainly involved in the KEGG pathways of TCA, pyruvate metabolism, and nitrogen metabolism pathway, showing that the metabolic balance of microbial cells shifted toward reduced end-products due to electron utilization. Moreover, the DEPs enriched in sphingolipid metabolism and fatty acid metabolism indicated that cell membranes were affected during the process of electron utilization in BES. The *ompF* overexpressed strain was used to investigate the function of the channel protein OmpF located in the outer membrane. CV analysis showed that the up-regulated expression of *ompF* improved redox reactions between the electrode and the cells.

In conclusion, engineered *E. coli* transferred electrons from the cathode into the cells by the Mtr pathway and simultaneously up-regulated the channel protein to enhance redox reactions between the electrode and the cells to further improve electron transfer capacity. The electrons were transferred into *E. coli* cells, resulting in a reductive intracellular environment, which provided a basis to improve the maturation or stability of some proteins associated with "[4Fe-4S] cluster binding," "Fe-S cluster binding," and "metal cluster binding." Sufficient reducing power provided the cofactor and electron donor for the oxidoreductases and dehydrogenases. The DEPs induced by exogenous electrons were mapped to the KEGG pathway database, which showed that the electrons disturbed metabolism in multiple ways by biofilm formation, sphingolipid metabolism, fatty acid metabolism, TCA, pyruvate metabolism, and nitrogen metabolism pathway.

## CONCLUSION

In this work, *E. coli* engineered to contain the Mtr pathway exhibited better electrochemical activity and effective biocatalytic behavior. Based on electrochemical and metabolite profile analyses, the succinate yield increased by 18% showing the effect of exogenous electrons on *E. coli* cells. Subsequently, proteomics and RT-PCR were performed to investigate DEPs in *E. coli* in response to exogenous electrons. Based on bioinformatic analysis of the DEPs, combined with biochemical methods, potential proteins and metabolic pathways that might function in electron transfer and utilization of *E. coli* cells were identified. In future studies, more crucial gene targets for electron transfer and utilization need to be identified. Additionally, multi-level and large-scale connections should be explored. The detailed molecular mechanisms need much deeper research, as an in-depth understanding of molecular mechanisms is the basis for engineering microorganisms to improve electron utilization efficiency. This work revealed a global proteomic alteration of *E. coli* cells in response to exogenous electrons, which will aid in the understanding of molecular mechanisms on electron transfer and utilization, and provide a theoretical basis for improving electron transfer and utilization efficiency.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository and accession number can be found below: http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD030650.

## AUTHOR CONTRIBUTIONS

JiaoF: data curation, validation, investigation, writing–original draft, review, and editing, and funding acquisition. JiaF: data curation, formal analysis, and writing–review and editing. CL: data curation and writing–review and editing. SX: resources, project administration, funding acquisition, and writing–review and editing. XW and KC: resources and writing–review and editing. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was supported by the National Science Foundation of Young Scientists of China (Grant Nos. 22108123 and 21908099), the China Postdoctoral Science Foundation (2020M681569), and the Jiangsu Synergetic Innovation Center for Advanced Bio-Manufacture (XTD2208).

## ACKNOWLEDGMENTS

We thank the PTM Biolabs, Inc. for technical assistance.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.815366/full#supplementary-material

## REFERENCES

Beilschmidt, L. K., de Choudens, S. O., Fournier, M., Sanakis, I., Hograindleur, M. A., Clemancey, M., et al. (2017). ISCA1 is essential for mitochondrial Fe4S4 biogenesis *in vivo*. *Nat. Commun.* 8:15124. doi: 10.1038/Ncomms15124

Cao, Y. J., Cao, Y. G., and Lin, X. Z. (2011). Metabolically engineered *Escherichia coli* for biotechnological production of four-carbon 1,4-dicarboxylic

acids. *J. Ind. Microbiol. Biotechnol.* 38, 649–656. doi: 10.1007/s10295-010-0913-4

Chandramouli, K., Unciuleac, M. C., Naik, S., Dean, D. R., Huynh, B. H., and Johnson, M. K. (2007). Formation and properties of [4Fe-4S] clusters on the IscU scaffold protein. *Biochemistry* 46, 6804–6811. doi: 10.1021/bi6026659

de Choudens, S. O., and Barras, F. (2017). Genetic, biochemical, and biophysical methods for studying Fe-S proteins and their assembly. *Methods Enzymol.* 595, 1–32. doi: 10.1016/bs.mie.2017.07.015

Duval, V., Foster, K., Brewster, J., and Levy, S. B. (2017). A novel regulatory cascade involving BluR, YcgZ, and Lon controls the expression of *Escherichia coli* OmpF porin. *Front. Microbiol.* 8:1148. doi: 10.3389/Fmicb.2017.01148

Feng, J., Jiang, M. J., Li, K., Lu, Q. H., Xu, S., Wang, X., et al. (2020a). Direct electron uptake from a cathode using the inward Mtr pathway in *Escherichia coli*. *Bioelectrochemistry* 134:107498. doi: 10.1016/j.bioelechem.2020.107498

Feng, J., Lu, Q., Li, K., Xu, S., Wang, X., Chen, K., et al. (2020b). Construction of an electron transfer mediator pathway for bioelectrosynthesis by *Escherichia coli*. *Front. Bioeng. Biotechnol.* 8:590667. doi: 10.3389/fbioe.2020.590667

Flint, D. H., Emptage, M. H., and Guest, J. R. (1992). Fumarase a from *Escherichia coli*: purification and characterization as an iron-sulfur cluster containing enzyme. *Biochemistry* 31, 10331–10337. doi: 10.1021/bi00157a022

Fortuin, S., Nel, A. J. M., Blackburn, J. M., and Soares, N. C. (2020). Comparison between the proteome of *Escherichia coli* single colony and during liquid culture. *J. Proteomics* 228:103929. doi: 10.1016/J.Jprot.2020.103929

Goldbeck, C. P., Jensen, H. M., TerAvest, M. A., Beedle, N., Appling, Y., Hepler, M., et al. (2013). Tuning promoter strengths for improved synthesis and function of electron conduits in *Escherichia coli*. *ACS Synth. Biol.* 2, 150–159. doi: 10.1021/sb300119v

Jensen, H. M., Albers, A. E., Malley, K. R., Londer, Y. Y., Cohen, B. E., Helms, B. A., et al. (2010). Engineering of a synthetic electron conduit in living cells. *Proc. Natl. Acad. Sci. U.S.A.* 107, 19213–19218. doi: 10.1073/pnas.1009645107

Jiang, Y., and Zeng, R. J. (2018). Expanding the product spectrum of value added chemicals in microbial electrosynthesis through integrated process design-A review. *Bioresour. Technol.* 269, 503–512. doi: 10.1016/j.biortech.2018.08.101

Kokko, M. E., Makinen, A. E., and Puhakka, J. A. (2016). "Anaerobes in bioelectrochemical systems," in *Anaerobes in Biotechnology, Advances in Biochemical Engineering/Biotechnology*, eds R. Hatti-Kaul, G. Mamo, and B. Mattiasson (Berlin: Springer), 263–292. doi: 10.1007/10_2015_5001

Korth, B., and Harnisch, F. (2019). Modeling microbial electrosynthesis. *Bioelectrosynthesis* 167, 273–325. doi: 10.1007/10_2017_35

Kracke, F., Lai, B., Yu, S. Q., and Kromer, J. O. (2018). Balancing cellular redox metabolism in microbial electrosynthesis and electro fermentation - A chance for metabolic engineering. *Metab. Eng.* 45, 109–120. doi: 10.1016/j.ymben.2017.12.003

Kracke, F., Vassilev, I., and Kromer, J. O. (2015). Microbial electron transport and energy conservation - the foundation for optimizing bioelectrochemical systems. *Front. Microbiol.* 6:575. doi: 10.3389/fmicb.2015.00575

Li, F. D., Wang, Y. C., Li, Y., Yang, H. P., and Wang, H. M. (2018). Quantitative analysis of the global proteome in peripheral blood mononuclear cells from patients with new - onset psoriasis. *Proteomics* 18:e1800003. doi: 10.1002/Pmic.201800003

Liu, J., Yong, Y. C., Song, H., and Li, C. M. (2012). Activation enhancement of citric acid cycle to promote bioelectrocatalytic activity of arcA knockout *Escherichia coli* toward high-performance microbial fuel cell. *ACS Catal.* 2, 1749–1752. doi: 10.1021/cs3003808

Mayr, J. C., Grosch, J. H., Hartmann, L., Rosa, L. F. M., Spiess, A. C., and Harnisch, F. (2019). Resting *Escherichia coli* as chassis for microbial electrosynthesis: production of chiral alcohols. *ChemSusChem* 12, 1631–1634. doi: 10.1002/cssc.201900413

Mettert, E. L., and Kiley, P. J. (2015). How is Fe-S cluster formation regulated? *Annu. Rev. Microbiol.* 69, 505–526. doi: 10.1146/annurev-micro-091014-104457

Nitzschke, A., and Bettenbrock, K. (2018). All three quinone species play distinct roles in ensuring optimal growth under aerobic and fermentative conditions in *E. coli* K12. *PLoS One* 13:e0194699. doi: 10.1371/journal.pone.0194699

Patil, S. A., Gildemyn, S., Pant, D., Zengler, K., Logan, B. E., and Rabaey, K. (2015). A logical data representation framework for electricity-driven bioproduction processes. *Biotechnol. Adv.* 33, 736–744. doi: 10.1016/j.biotechadv.2015.03.002

Qiao, Y., Li, C. M., Bao, S. J., Lu, Z. S., and Hong, Y. H. (2008). Direct electrochemistry and electrocatalytic mechanism of evolved *Escherichia coli* cells in microbial fuel cells. *Chem. Commun.* 11, 1290–1292. doi: 10.1039/b719955d

Ross, D. E., Flynn, J. M., Baron, D. B., Gralnick, J. A., and Bond, D. R. (2011). Towards electrosynthesis in *Shewanella*: energetics of reversing the Mtr pathway for reductive metabolism. *PLoS One* 6:e16649. doi: 10.1371/journal.pone.0016649

Schmidt, A., Kochanowski, K., Vedelaar, S., Ahrne, E., Volkmer, B., Callipo, L., et al. (2016). The quantitative and condition-dependent *Escherichia coli* proteome. *Nat. Biotechnol.* 34, 104–110. doi: 10.1038/nbt.3418

Singh, A., Soh, K. C., Hatzimanikatis, V., and Gill, T. R. (2011). Manipulating redox and ATP balancing for improved production of succinate in *E. coli*. *Metab. Eng.* 13, 76–81. doi: 10.1016/j.ymben.2010.10.006

Soufi, B., Krug, K., Harst, A., and Macek, B. (2015). Characterization of the *E. coli* proteome and its modifications during growth and ethanol stress. *Front. Microbiol.* 6:103. doi: 10.3389/Fmicb.2015.00103

Sun, M., Zhai, L. F., Mu, Y., and Yu, H. Q. (2020). Bioelectrochemical element conversion reactions towards generation of energy and value-added chemicals. *Prog. Energy Combust. Sci.* 77:100814. doi: 10.1016/J.Pecs.2019.100814

Tefft, N. M., and TerAvest, M. A. (2019). Reversing an extracellular electron transfer pathway for electrode driven acetoin reduction. *ACS Synth. Biol.* 8, 1590–1600. doi: 10.1021/acssynbio.8b00498

Teran-Melo, J. L., Pena-Sandoval, G. R., Silva-Jimenez, H., Rodriguez, C., Alvarez, A. F., and Georgellis, D. (2018). Routes of phosphoryl group transfer during signal transmission and signal decay in the dimeric sensor histidine kinase ArcB. *J. Biol. Chem.* 293, 13214–13223. doi: 10.1074/jbc.RA118.003910

van Beilen, J. W. A., and Hellingwerf, K. J. (2016). All Three endogenous quinone species of *Escherichia coli* are involved in controlling the activity of the aerobic/anaerobic response regulator ArcA. *Front. Microbiol.* 7:1339. doi: 10.3389/Fmicb.2016.01339

Wu, Z. Q., Wang, J. S., Liu, J., Wang, Y., Bi, C., and Zhang, X. L. (2019). Engineering an electroactive *Escherichia coli* for the microbial electrosynthesis of succinate from glucose and CO2. *Microb. Cell Fact.* 18:15. doi: 10.1186/s12934-019-1067-3

Zhang, J., Li, Q. R., Liu, J. H., Lu, Y. H., Wang, Y., and Wang, Y. H. (2020). Astaxanthin overproduction and proteomic analysis of *Phaffia rhodozyma* under the oxidative stress induced by TiO2. *Bioresour. Technol.* 311: 123525. doi: 10.1016/j.biortech.2020.123525

Zhao, J., Li, F., Cao, Y., Zhang, X., Chen, T., Song, H., et al. (2020). Microbial extracellular electron transfer and strategies for engineering electroactive microorganisms. *Biotechnol. Adv.* 53:107682. doi: 10.1016/j.biotechadv.2020.107682

Zhu, L. W., and Tang, Y. J. (2017). Current advances of succinate biosynthesis in metabolically engineered *Escherichia coli*. *Biotechnol. Adv.* 35, 1040–1048. doi: 10.1016/j.biotechadv.2017.09.007

**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.

*Copyright © 2022 Feng, Feng, Li, Xu, Wang and Chen. 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.*

frontiers | Frontiers in Microbiology

ORIGINAL RESEARCH
published: 14 April 2022
doi: 10.3389/fmicb.2022.855736

# A Novel Efficient L-Lysine Exporter Identified by Functional Metagenomics

*Sailesh Malla*†, Eric van der Helm†, Behrooz Darbani†, Stefan Wieschalka, Jochen Förster†, Irina Borodina and Morten Otto Alexander Sommer**

*The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby, Denmark*

OPEN ACCESS

**Edited by:**
*Yu Wang,*
*Tianjin Institute of Industrial Biotechnology (CAS), China*

**Reviewed by:**
*Lothar Eggeling,*
*Helmholtz Association of German Research Centres (HZ), Germany*
*Hisashi Kawasaki,*
*The University of Tokyo, Japan*

***Correspondence:**
*Morten Otto Alexander Sommer*
*msom@bio.dtu.dk*
*Sailesh Malla*
*dksaim@chr-hansen.com*

**†Present addresses:**
*Sailesh Malla,*
*Strain Development Department, Discovery, R&D, Chr. Hansen (A/S), Hørshlom, Denmark*
*Eric van der Helm,*
*SNIPR BIOME & ODity.bio, Copenhagen, Denmark*
*Behrooz Darbani,*
*Department of Agroecology, Faculty of Technical Sciences, Aarhus University, Slagelse, Denmark*
*Jochen Förster,*
*Biosyntia ApS Fruebjergvej, Copenhagen, Denmark*

**Specialty section:**
*This article was submitted to Microbiotechnology, a section of the journal Frontiers in Microbiology*

**Received:** *15 January 2022*
**Accepted:** *23 March 2022*
**Published:** *14 April 2022*

**Citation:**
*Malla S, van der Helm E, Darbani B, Wieschalka S, Förster J, Borodina I and Sommer MOA (2022) A Novel Efficient L-Lysine Exporter Identified by Functional Metagenomics. Front. Microbiol. 13:855736. doi: 10.3389/fmicb.2022.855736*

Lack of active export system often limits the industrial bio-based production processes accumulating the intracellular product and hence complexing the purification steps. L-lysine, an essential amino acid, is produced biologically in quantities exceeding two million tons per year; yet, L-lysine production is challenged by efficient export system at high titers during fermentation. To address this issue, new exporter candidates for efficient efflux of L-lysine are needed. Using metagenomic functional selection, we identified 58 genes encoded on 28 unique metagenomic fragments from cow gut microbiome library that improved L-lysine tolerance. These genes include a novel L-lysine transporter, belonging to a previously uncharacterized EamA superfamily, which is further *in vivo* characterized as L-lysine exporter using *Xenopus oocyte* expression system as well as *Escherichia coli* host. This novel exporter improved L-lysine tolerance in *E. coli* by 40% and enhanced yield, titer, and the specific production of L-lysine in an industrial *Corynebacterium glutamicum* strain by 7.8%, 9.5%, and 12%, respectively. Our approach allows the sequence-independent discovery of novel exporters and can be deployed to increase titers and productivity of toxicity-limited bioprocesses.

**Keywords: amino acid, transporter, *E. coli*, *C. glutamicum*, *Xenopus* oocytes**

## INTRODUCTION

The global chemical industry is transitioning from reliance majorly on petrochemical processes to more sustainable bio-based production. This development holds promise to improve the sustainability of the chemical industry while also reducing the overall production costs of chemical products. In order to establish a cost-competitive bioprocess, titers of fermentations frequently exceed 100 $gl^{-1}$, which leads to a significant stress on the host organism. Indeed, a majority of industrial bioprocesses are limited in production due to several stresses, for example, osmotic stress, leading to product toxicity.

One of the most significant bio-based chemical products is L-lysine. The global bio-based L-lysine production is estimated to reach 3.0 million tons in 2022 corresponding to 5.6 billion USD of market value according to the current L-lysine market report (Elder, 2018). Industrial L-lysine bioprocesses entirely rely on *Corynebacterium glutamicum* and *Escherichia coli* production strains that achieve titers over 1.2 M (Becker and Wittmann, 2012; Lee and Kim, 2015).

It was observed that the L-lysine export rate is inhibited by 50% upon exceeding the extracellular concentration of 400 mM compared to that at 80 mM (Kelle et al., 1996), indicating the substantial inhibition of the L-lysine-specific export in industrial fermentation. Indeed, several studies have demonstrated the benefit of incorporating active efflux systems to address intracellular product accumulation in bio-based production (Malla et al., 2010; Hemberger et al., 2011; Borodina, 2019). Hence, L-lysine export system is an obvious target to maintain the producer organism at high lysine concentration as well as easing the downstream process. Despite its importance, there are only two identified lysine-specific exporters: (i) LysE as a member of the lysine efflux permease (LysE; **2.A.75**) family (Vrljic et al., 1999; Bellmann et al., 2001) and (ii) lysine outward permease (LysO or YbjE in *E. coli*; Vrljic et al., 1996; Pathania and Sardesai, 2015) achieved five-fold higher lysine export rate upon overexpression of LysE in *C. glutamicum* (Vrljic et al., 1996). Similarly, Yasueda and coworker have deployed this strategy for 10-fold improvement in L-lysine production by expressing a spontaneously mutated LysE from *C. glutamicum* in *Methylophilis methylotrophus* (Gunji and Yasueda, 2006). In addition to the rational engineering of the existing exporters, there is an urgent demand for new genetic building blocks to further improve L-lysine tolerance and production.

Functional metagenomic selection is an effective approach to discover novel genes and enzymes due to its ability to access the wide range of genetic elements present in a particular environmental niche (Sommer et al., 2009, 2010; Munck et al., 2015; Forsberg et al., 2016). Hence, we set out to use functional metagenomic selection to identify novel L-lysine transporter candidates from a cow fecal library, with the goal of improving industrial L-lysine production (**Figure 1A**). Using C13-labeled L-lysine and mRNA expression of the screened transporter in *Xenopus oocyte*, the transporter candidate was confirmed as L-lysine exporter. Expression of the metagenomic derived L-lysine transporter improved titer and productivity in both Gram-positive and Gram-negative production strains.

# MATERIALS AND METHODS

## Bacterial Strains, Growth Conditions, and Chemicals

All bacterial strains, vectors, and plasmids used in this study are listed in **Table 1**. All oligonucleotide primers (synthesized by Integrated DNA Technologies, Inc.) used are presented in **Table 2**. *E. coli* strains were routinely cultured at 37°C in Luria-Bertani (LB) broth or on agar supplemented with kanamycin (35 $\mu g\,mL^{-1}$) when necessary (hereafter referred as LB-km). For the L-lysine production in *E. coli* LB as well as M9 minimal media supplemented with 2 $gl^{-1}$ of yeast extract were used and the strains were cultivated in 96-deep well plate in 600 µl of media. *Corynebacterium glutamicum* strains were cultured in modified CGXII medium (Keilhauer et al., 1993) supplemented with 5% sucrose, 1% BHI, 0.5 mM IPTG, and kanamycin (when necessary), at 30°C and 250 RPM. Growth media were supplemented with 1.5% agar for plate assays.

DNA manipulations were carried out following standard protocols (Sambrook and Russell, 2001). All chemicals were purchased from Sigma-Aldrich (St. Louis, MO, United States). Restriction enzymes and T4 DNA ligase were purchased from Fermentas (Denmark) and New England Biolabs (Hertfordshire, United Kingdom). DNA sequencing was performed using an automated DNA sequence analyzer.

## Metagenomic Library Construction and Functional Screening for L-Lysine

A metagenomic expression library of total DNA extracted from a cow fecal sample was constructed as described previously (Sommer et al., 2009). Briefly, library construction involves (i) the isolation of total DNA from 5 g of fecal material using the PowerMax Soil DNA Isolation Kit (MO BIO Laboratories Inc.), (ii) fragmentation of extracted DNA into pieces of an average size of 2 kb by sonication using a Covaris E210 (Massachusetts, United States) followed by end-repair using the End-It end-repair kit (Epicentre), and (iii) blunt-end ligation into the pZE21 MCS1 expression vector with constitutive promoter pLteo-1(Lutz and Bujard, 1997) at the *HincII* site using the Fast Link ligation kit (Epicentre), and (iv) transformation of the ligated sample into electrocompetent *E. coli* top10 cells by the standard method (One Shot TOP10 Electrocomp Cells, Invitrogen).

After electroporation, cells were recovered in 1 ml of SOC medium for 1 h at 37°C, and the library was titered by plating out 1, 0.1, and 0.01 µl of recovered cells onto LB-km plates. The insert size distribution was estimated by gel electrophoresis of the PCR products obtained by amplifying inserts using primers annealing to the vector backbone flanking the *Hinc*II site. The average insert size for the library was 1.7 kb. The total size of the metagenomics library was determined by multiplying the average insert size by the number of colony-forming units (CFU) per ml ($5 \times 10^8$ bp). The remainder of the recovered cells were inoculated into 10 ml of LB-km liquid media and grown overnight; the library was frozen down in 15% glycerol and stored at −80°C.

The *E. coli* C4860 strain was used for the functional screening of L-lysine exporters. First, a plasmid prep of the metagenomics library was carried out from the *E. coli* top10 cells harboring the library. Then, 400 ng of metagenomics plasmid DNA was transformed into electrocompetent *E. coli* C4860 cells, and the library titer was determined as described above. On the basis of the determined library sizes and the titer of the library, $10^6$ cells (i.e., 100 µl of the library cells) were plated on LB-km agar supplemented with L-lysine at the selective concentration (14 $gl^{-1}$). Plates were incubated at 37°C, and the growth of colonies (likely lysine-tolerant clones) was assayed after 48–65 h of incubation.

The metagenomic inserts present in L-lysine-tolerant clones were Sanger sequenced using the pZE21_F and pZE21_R primer pair, which annealed to the vector backbone. The resulting raw sequencing chromatogram files were analyzed and functionally annotated using the deFUME web server (van der Helm et al., 2015): http://www.cbs.dtu.dk//services/deFUME/.

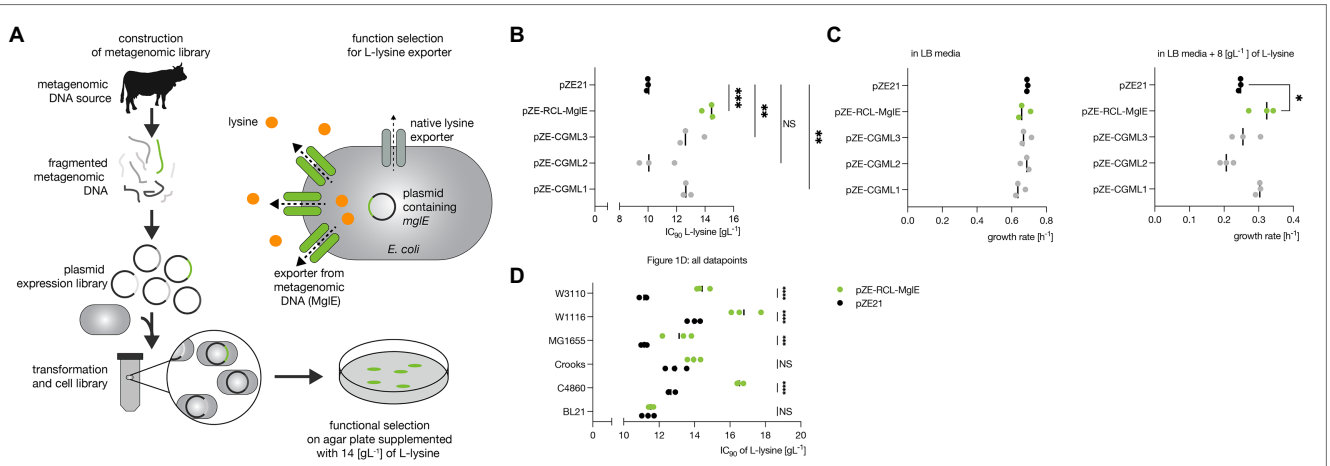



**FIGURE 1 |** Functional metagenomic screening and Lysine tolerance of the screened metagenomics clones. **(A)** Functional metagenomic screening for lysine transporters. Cow fecal metagenomic DNA library construction and its functional selection for L-lysine exporters. **(B)** L-lysine IC90 values for the selected *E. coli* C4860 resistant clones harboring putative transporters. A one-way ANOVA with Dunnett's multiple testing correction of the L-Lysine IC90 was applied to compare the metagenomic inserts vs. the empty vector (pZE21). **(C)** The growth rates of *E. coli* C4860 harboring pZE-RCL-MglE and pZE21 (control) in LB-Km and LB-Km supplemented with 8 $gl^{-1}$ of L-lysine. A one-way ANOVA with Dunnett's correction for multiple testing was used to compare the growth rate of the metagenomic gene inserts vs. the empty vector (pZE21). A significantly different of growth rate ($p = 0.0375$) was found between MglE and the empty vector control pZE21 at 8 $gl^{-1}$ of L-lysine. **(D)** Improved L-lysine tolerance (IC90 values) by expression of metagenomics insert carrying MglE transporter in industrially relevant *E. coli* strains. A two way ANOVA with Bonferroni correction for multiple testing of the IC90 values was applied per background strain comparing the empty vector (pZE21) against MglE metagenomic insert (pZE-RCL-MglE). In C4860, MG1655, W1116 and W3110 there was a significant difference between the empty vector and MglE ($p = <0.0001$). For BL21 and Crooks, no significant difference was found ($p > 0.05$).

## Minimum Inhibitory Concentration and IC90 Determination

For minimum inhibitory concentration (MIC) determination, the *E. coli* strains were cultured in LB liquid media at 37°C overnight, and then approximately $1 \times 10^4$ cells were inoculated from the overnight cultures into LB liquid media and grown at 30°C and 300 RPM in 96-well microtiter plates containing 150 µl of medium per well. MICs were determined using a logarithmic L-lysine (or chemical) concentration gradient with two-fold serial dilutions. Endpoint absorbance measurements ($A_{600nm}$) were taken with a plate reader (Synergy H1, BioTek) after 24 or 48 h of incubation and were background-subtracted. Growth inhibition of the *E. coli* strains was plotted against L-lysine (chemical) concentration with a polynomial interpolation between neighboring data points using R software.[1] The percentage of inhibition was calculated using the formula: $1 - [A_{600nm\ lysine\ (or\ chemical)}/A_{600nm\ control}]$. The inhibitory concentration was defined as the lowest concentration of the chemical that inhibited 90% of the growth of the strain tested (IC90).

## Growth Experiments

Single colonies of the tolerant clone(s) harboring transporter homologues were grown overnight at 37°C with shaking in liquid LB-km medium. The OD values at 600 nm [$(OD)_{600nm}$] were determined by 10-fold dilution. Then, the cultures were diluted to adjust the $(OD)_{600nm}$ to 0.1, and 5 µl of each culture was inoculated in 150 µl fresh media containing L-lysine (0 and 8 $gl^{-1}$) and kanamycin in a 96-well microtiter plate. The plate was incubated at 37°C for 24 h in an automated spectrophotometer (ELx808, BioTek) that recorded the $(OD)_{630nm}$ at an interval of 60 min. The data were subsequently retrieved and analyzed to determine growth rates. The growth rate data are the average of triplicate experiments, with error bars representing the standard error of the mean (SEM).

## PCR and *Escherichia coli* Transformation

PCR was performed in a total volume of 50 µl under the following DNA amplification conditions: 95°C for 5 min, followed by 30 cycles of 95°C for 30 s, 50°C–65°C for 30 s, and 72°C for 1 min, and finally 72°C for 5 min. Electrocompetent or chemically competent *E. coli* cells were transformed with the ligation mixture using a standard protocol and plated onto LB-km agar plates for selection.

## *In silico* Analysis

The amino acid sequence of MglE (**M**etagenomics **g**ene for **l**ysine **E**xport) was analyzed using Interpro,[2] its 2D membrane topology was predicted using Phobius (Käll et al., 2007) and visualized using Protter (Omasits et al., 2014), and the phylogenetic tree was plotted using iTOL. The UniProtKB database (UniProt Consortium, 2014) was accessed on June 1 2016 to query the closest homologs of MglE. Pfam version 29.0 (December 2015; Finn et al., 2014) was used to extract the LysE family (Pfam id: PF01810) members. The Maximum Likelihood Phylogenetic tree was constructed using CLC Main Workbench 7 with the neighbor-joining construction method.

[1]http://www.r-project.org

[2]http://www.ebi.ac.uk/interpro/

**TABLE 1** | Bacterial strains and plasmids used in this study.

| Strains/Plasmids | Description | Source/reference |
|---|---|---|
| **Strains** | | |
| *Escherichia coli* | | |
| DH5α | General cloning host | Invitrogen |
| Top10 | Highly efficient host for cloning and plasmid propagation | Invitrogen |
| BL21(DE3) | *ompT hsdT hsdS* ($r_B^-$ $m_B^-$) *gal* (DE3) | Novagen |
| C4860 | Host for phage PhiX174, sup-, lambda-, F-. (now known as DSM 13127) | DSMZ |
| MG1655 | $F^-$ $I^-$ ilvG$^-$ rfb-50 rph-1 | Blattner et al., 1997 |
| W3110 | $F^-$ lambda$^-$ IN(rrnD-rrnE)1 rph-1 | Hayashi et al., 2006 |
| W1116 | strain W and ATCC 9637 | Archer et al., 2011 |
| Crooks | *E. coli* ATCC 8739 | DSMZ |
| BL21/pZE21 | *E. coli* BL21 carrying pZE21 | This study |
| C4860/pZE21 | *E. coli* C4860 carrying pZE21 | This study |
| MG1655/pZE21 | *E. coli* MG1655 carrying pZE21 | This study |
| W3110/pZE21 | *E. coli* W3110 carrying pZE21 | This study |
| W1116/pZE21 | *E. coli* W1116 carrying pZE21 | This study |
| Crooks/pZE21 | *E. coli* Crooks carrying pZE21 | This study |
| BL21(DE3)/pZE-RCL-MglE | *E. coli* BL21 carrying pZE-RCL-MglE | This study |
| C4860/pZE-RCL-MglE | *E. coli* C4860 carrying pZE-RCL-MglE | This study |
| MG1655/pZE-RCL-MglE | *E. coli* MG1655 carrying pZE-RCL-MglE | This study |
| W3110/pZE-RCL-MglE | *E. coli* W3110 carrying pZE-RCL-MglE | This study |
| W1116/pZE-RCL-MglE | *E. coli* W1116 carrying pZE-RCL-MglE | This study |
| Crooks/pZE-RCL-MglE | *E. coli* Crooks carrying pZE-RCL-MglE | This study |
| C4860/pZE-MglE | *E. coli* C4860 carrying pZE-MglE | This study |
| C4860/pZE0-P-MglE-T | *E. coli* C4860 carrying pZE0-P-MglE-T | This study |
| C4860/pZE-YbjE | *E. coli* C4860 carrying pZE-YbjE | This study |
| C4860/pZE-Gene1 | *E. coli* C4860 carrying pZE-Gene1 | This study |
| C4860/pZE-Gene2 | *E. coli* C4860 carrying pZE-Gene2 | This study |
| C4860/pZE-Gene3 | *E. coli* C4860 carrying pZE-Gene3 | This study |
| C4860/pZE-Gene4 | *E. coli* C4860 carrying pZE-Gene4 | This study |
| W3110/pSIJ8 | *E. coli* W3110 carrying pSIJ8 | This study |
| W3110:: ΔldcC/pSIJ8 | *ldcC* deletion mutant of *E.coli* W3110 carrying pSIJ8 | This study |
| W3110::ΔldcC.ΔcadA/pSIJ8 | *ldcC* and *cadA* deletion mutant of *E.coli* W3110 (DMLC) carrying pSIJ8 | This study |
| DMLC | *ldcC* and *cadA* deletion mutant of *E.coli* W3110 | This study |
| DMLC/pZE21 | *E.coli* DMLC carrying pZE21 | This study |
| DMLC/pZE-RCL-MglE | *E.coli* DMLC carrying pZE-RCL-MglE | This study |
| Keio strain | | |
| b0186 | *ldcC* replaced by kanamycin cassette *E.coli* K-12 BW25113 | Baba et al., 2006 |
| b4131 | *cadA* replaced by kanamycin cassette in *E.coli* K-12 BW25113 | Baba et al., 2006 |
| *Corynebacterium glutamicum* | | |
| VL5 | Industrial L-lysine production strain of *C. glutamicum* | Vitalys |
| VL5/pEKEx2 | *C. glutamicum* VL5 carrying pEKEx2 | This study |
| VL5/pEK-RCL-MglE | *C. glutamicum* VL5 carrying pEK-RCL-MglE | This study |
| **Plasmids and vectors** | | |
| pZE21 | *E.coli* expression vector with pLtet promoter, pBR322 origin of replication. and kanamycin cassette | Lutz and Bujard, 1997 |
| pSIJ8 | pKD46 based vector with rhaRS-prha-FLP and ampicillin cassette | Jensen et al., 2015 |
| pUSER016 | pNB1u based vector with T7 promoter for *in vitro* transcription | Darbani et al., 2016 |
| pEKEx2 | IPTG inducible *C. glutamicum* expression vector with kanamycin cassette from pUC4 K; Ptrc, lacI, pUC18 mcs, and pBL1 origin of replications. | Eikmanns et al., 1991 |
| pZE-RCL-MglE | pZE21 carrying L-lysine resistant metagenomic insert with MglE operon from cow fecal sample | This study |
| pZE-MglE | pZE21 carrying *mglE* gene | This study |
| pZE0-P-MglE-T | Promoter less pZE21 carrying native promoter and *mglE* gene and native terminator region | This study |
| pUSER016-pT7-MglE | Codon-optimized MglE gene cloned into pUSER016 for functional studies in *Xenopus* oocytes | This study |
| pEK-RCL-MglE | pEKEx2 carrying L-lysine resistant metagenomic insert with MglE operon from cow fecal sample | This study |
| pZE-YbjE | pZE21 carrying *ybjE* gene from *E.coli* BL21(DE3) | This study |
| pZE-Gene1 | pZE21 carrying gene from *Bacteroides coprophilus* CAG:333 having 82% identity with *mglE* | This study |
| pZE-Gene2 | pZE21 carrying codon-optimized gene from *Bacteroides barnesiae* having 78% identity with *mglE* | This study |
| pZE-Gene3 | pZE21 carrying gene from *Bacteroides* sp. CAG:770 having 54% identity with *mglE* | This study |
| pZE-Gene4 | pZE21 carrying codon-optimized gene from *Alistipes finegoldii* CAG:68 having 43% identity with *mglE* | This study |

*The gene sequences of mglE, ybjE, gene1, gene2, gene3 and gene4, are provided in the* **Supplementary Information**.

**TABLE 2** | Oligonucleotides used in this study.

| Primers | Oligonucleotide sequences (5′-3′) | Restriction site |
|---|---|---|
| pZE21_F | ATCAGTGATAGAGATACTGAGCAC | |
| pZE21_R | TTTCGTTTTATTTGATGCCTCTAG | |
| MglE_F | AAA*GGTACC*ATGAGAAATTTAAGTAAAAAAGC | *KpnI* |
| MglE_R | TTA*AAGCTT*CTATCTCTTTGTTACTGAAATCAT | *HindIII* |
| RCL-MglE-F | CGG*GGTACC*AAGAGGAGAAAGGTACCGGGC | *KpnI* |
| RCL-MglE-R | C*GAGCTC*CCGGGCTGCAGGAATTCG | *SacI* |
| YbjE_F | AAA *GGTACC* ATGTTTTCTGGGCTGTTAATCA | *KpnI* |
| YbjE_R | TAT *GGATCC* TTACGCAGAGAAAAAGGCGAT | *BamHI* |
| linear_cassette F | TGCAAGGCGATTAAGTTGGGTAACGC | |
| linear_cassette R | CCATGATTACGCCAAGCTATTTAGGTGACAC | |
| **For user cloning** | | |
| pZE21_User_F | ATAAGACGG**U** ATCGATAAGCTTGATATCGAATTCC | |
| pZE21_User_R | AGGGTACCC**U** CGAGGTGAAGACGAAAGGGCCTCG | |
| P-MglE-T_User_F | AGGGTACCC**U** GAAGTCAAGCATCTCAAAAAACTAC | |
| P-MglE-T_ User_R | ACCGTCTTA**U** TTACGATTTTACTACGGAGTATTAA | |
| pZE21_User_F4 | ATAAGCT**U** GATATCGAATTCCTGCAGCCC | |
| pZE21_User_R4 | AATGAAT**U** CGGTCAGTGCGTCCTGCTGAT | |
| **For gene deletion and confirmation** | | |
| IdcC_F1 | TCAGCGCCTGATGAGCTACG | |
| IdcC_F2 | AGTTCTGAAAAAGGGTCACTTC | |
| IdcC_R | TCGCAATATGGTGAACCTGTT | |
| CadA_F1 | TGAAGTACTCCCAGATTTGGATC | |
| CadA_F2 | CGGCTGTGAGGGTGTTTTCA | |
| CadA_R | TTAATTAAAAGTATTTTCCGAGGCTCC | |

*Restriction sites are indicated by underlined and italics.*

## Transport Assays in *Xenopus* Oocytes

The *Xenopus laevis* oocytes were obtained from EcoCyte Bioscience (Germany). Oocytes were kept in Kulori buffer (pH 7.4) and at 18°C. A linear cassette (including T7 promoter, *mglE* as the gene of interest, and 3'UTR) was amplified from the plasmid pUSER016-T7-MglE using Phusion Hot Start polymerase (Thermo Fisher Scientific) and the primers linear_ cassette F and linear_cassette R. The linear cassette was used as template for *in vitro* transcription. Capped cRNAs was synthesized using the mMESSAGE mMACHINE T7 Transcription Kit (AM1344; Thermo Fisher). The quality and quantity of the RNA were determined by Agilent 2,100 Bioanalyzer. For microinjection of cRNA and 13C-labeled L-lysine, we used the RoboInject (Multi Channel Systems, Reutlingen, Germany) automatic injection system (Darbani et al., 2019). Injection needles with opening of 25 µm were used (Multi Channel Systems, MCS GmbH). For expression in oocytes, 25 ng of the *in vitro* transcribed cRNA for the MglE transporter was microinjected into the oocytes 3 days prior to transport assays. To perform the export assay, 50 nl of the 13C-labeled L-lysine stock solution was injected into the oocytes to obtain estimated internal concentrations of 6 mM, assuming an after-injection dilution factor of 20 (Darbani et al., 2016). Following four washing steps, each batch of 20 oocytes was incubated for 180 min in 90 µl Kulori buffer at pH 5 or pH 7.4. After incubation, 70 µl of the medium was collected from each batch with intact oocytes and added onto 70 µl of 60% MeOH before LC–MS analysis. The oocytes were washed four times with Kulori buffer pH 7.4 and intracellular metabolites were extracted in 30% MeOH to analyzed on LC–MS.

## *Escherichia coli* Double Deletion Mutant Construction

The chromosomal lysine decarboxylases *ldcC* and *cadA* in *E. coli* W3110 were successively knocked out by PCR targeting to create *E.coli* DMLC strain. The gene disruption process was carried out as described previously using an ampicillin-resistant pSIJ8 helper plasmid containing both the λ Red and FLP systems (Jensen et al., 2015). The primer pair IdcC_F2/IdcC_R (situated 64 bp away from the start/stop codon *ldcC* gene) was used to amplify the kanamycin cassette from the genomic DNA of the *ldcC* inframe knocked out Keio strain b0186 whereas the primer pair CadA_F2/CadA_R (situated 56 bp away from the start/stop codon *cadA* gene) was used to amplify the kanamycin cassette from the genomic DNA of the *cadA* inframe knock-out Keio strain b4131. These PCR products were used to delete *ldcC* and *cadA* in *E. coli* W3110 strain. The detailed process of construction of *E. coli* DMLC is given in **Supplementary Material**.

## Plasmid Construction

### Subcloning the Metagenomic Insert Implicated in Tolerance Phenotypes

The recombinant plasmids pZE-MglE and pZE0-P-MglE-T were constructed as described below. The construction of recombinant plasmids was verified by both restriction mapping and direct nucleotide sequencing of the respective genes in the recombinant plasmids.

Using the oligonucleotide pair MglE_F/MglE_R and the pZE-RCL-MglE plasmid as template DNA, the exact *orf* of the putative carboxylate/amino acid transporter homologue (referred to as MglE) was amplified. The amplicon was purified with a Qiagen PCR purification kit and digested with the restriction enzymes *Kpn*I and *Hind*III (NEB, United Kingdom), followed by ligation using T4 DNA ligase (Fermentas, Denmark) into the corresponding restriction sites of the multiple cloning site (MCS) of the pZE21 vector to construct the recombinant plasmids pZE-MglE.

To construct the recombinant plasmid pZE0-P-MglE-T, USER cloning was applied. The pZE21-vector backbone without its promoter (pZE0) was amplified using the oligonucleotide pair pZE21_User_F/pZE21_User_R. The putative carboxylate/amino acid transporter homologue and the native promoter and terminator sequences (P-MglE-T fragment) were amplified from the pZE-RCL-MglE plasmid using the primer pair P-MglE-T_User_F/P-MglE-T_ User_F. The amplified PCR products were ligated into the recombinant plasmid pZE0-P-MglE-T using the recommended standard USER cloning protocol (NEB, United Kingdom).

### Codon-Optimized mglE Plasmids Construction for Oocytes Expression Study

For functional analysis of MglE in *Xenopus* oocytes, we constructed recombinant plasmid pUSER016-pT7-MglE by USER cloning. The pUSER016 vector backbone was digested with *Pac*I for 18 h and the ends were further nicked with Nt.BbvCI

for 2 h. The codon-optimized *mglE* (sequence provided in **Supplementary Data**) for *Xenopus laevis* was synthetized along with N-terminal linker sequence—GGCTTAA**U** and C-terminal linker sequences—**A**TTAAACC (in the complementary sequence, T is replaced by U for pairing with **A**) including uracil with compatible USER sites easing direct cloning into linearized and nicked pUSER016 vector by USER cloning (USER, NEB), resulting in plasmid pUSER016-pT7-MglE.

### Recombinant Plasmids With *Escherichia coli* Lysine Exporter and Homologous Genes of mglE

The lysine exporter, *ybjE* (900 bp, GenBank accession no. CAQ31402), from *E. coli* BL21 (DE3) was amplified using oligonucleotides YbjE-F and YbjE-R. The PCR product of *ybjE* was cloned into pZE21 vector excised with *Kpn*I and *BamH*I restriction enzymes to construct pZE-YbjE expression plasmid.

Similarly, four *mglE* homologous genes: (i) Gene1 (891 bp, 82% identity, GenBank accession no. CDC57518) from *Bacteroides coprophilus* CAG:333, (ii) Gene2 (891 bp, 78% identity, GenBank accession no. WP_018711839) from *Bacteroides barnesiae*, (iii) Gene3 (870 bp, 54% identity, GenBank accession no. CDC66277) from *Bacteroides* sp. CAG:770, and (iv) Gene4 (912 bp, 43% identity, GenBank accession no. CCZ76555) from *Alistipes finegoldii* CAG, were synthetized along with N-terminal linker sequence—AATTCAT**U** AAAGAGGAGAAAGGTACC and C-terminal linker sequence—GTCGACGGTATCG **A**TAAGCTT (in the complementary sequence, T is replaced by U for pairing with **A**) including uracil easing direct user cloning. These four synthetized gene fragments were cloned into the linear PCR fragment of pZE21 vector amplified by using oligonucleotides pZE21_User-F4 and pZE21_User-R4 (**Table 2**) to construct pZE-Gene1, pZE-Gene2, pZE-Gene3 and pZE-Gene4 recombinant plasmids, respectively.

## Construction of Recombinant Industrial *Corynebacterium glutamicum* Strains

The full metagenomics insert derived from pZE-RCL-MglE plasmid was amplified using the primer pair RCL-MglE-F/RCL-MglE-R. The resulting PCR product was cloned into pEKEx2 vector, using the *Kpn*I and *Sac*I restriction sites, yielding expression plasmid pEK-RCL-MglE. The newly constructed expression plasmid was introduced into the industrial L-lysine producing strain *C. glutamicum* VL5 by electroporation and concomitant selection on kanamycin containing 2xTY-Agar. Successful construction of the new strain, named *C. glutamicum* VL5/pEK-RCL-MglE, was verified by re-isolation of the plasmid, followed by control digest and PCR on the length of the insert. The same procedure was performed to introduce the empty pEKEx2 expression vector into *C. glutamicum* VL5, yielding control strain *C. glutamicum* VL5/pEKEx2.

## Analysis of L-Lysine in Supernatant of Bacterial Culture

A calibration curve of an authentic standard was generated using concentrations ranging from 0.4 to 77 mg $L^{-1}$. The accurate mass of L-lysine from 20-fold diluted samples (supernatants containing L-lysine for quantification) was analyzed using an LC–MS Fusion (Thermo Fisher Scientific, United States) with positive electrospray ionization (ESI+). The final concentration was adjusted for the dilution factor. Bracketing calibration was used for the quantification of the external concentration. For the quantification of L-lysine, an LC–MS/MS, EVOQ (Bruker, Fremont United States) was used with multiple reaction monitoring (MRM) transition in positive ionization mode (ESI+), with the quantified transition 147 → 84 (CE 10) and qualifier transition 147 → 130 (CE 7). The significance of the specific lysine production was calculated using a Student *t*-test.

# RESULTS AND DISCUSSION

## Identification of Metagenomic L-Lysine Tolerance Genes

The gut microbiota from industrial livestock animals are conceivably enriched with the microbes capable to survive at higher L-lysine concentration which is supplied as food additives. A metagenomic library derived from cow feces was constructed and used to find candidate genes that could lead to a higher L-lysine tolerance in *E. coli* C4860 strain. The resulting library was screened for L-lysine tolerance by plating on LB agar plates supplemented with selective concentrations of L-lysine (14 $gl^{-1}$; **Figure 1A**; section "Materials and Methods"). Colonies appeared only on the plates with the metagenomic library but not on the control plates with *E. coli* C4860 harboring empty vector. Eighty L-lysine-tolerant clones were selected for further analysis.

Metagenomic inserts present in those L-lysine-tolerant clones were PCR amplified, sequenced, and annotated using deFUME (van der Helm et al., 2015). In total, 28 unique clones containing 58 complete or partial open reading frames were identified in this analysis. The sequence analysis revealed that resistant clones contained *orf*s homologous to genes encoding L-lysine modification/degradation enzymes, membrane proteins, signaling proteins, and hypothetical proteins (**Supplementary Table S1**).

## Selection of Transporter Candidates

For the purpose of improving industrial production strains, novel efflux systems constitute more relevant genetic building blocks than those involved in degradation and modification, which would counteract the objective of L-lysine production. Accordingly, we focused our subsequent analysis on *orfs* encoding potential efflux systems. We identified six unique metagenomic inserts encoding putative membrane proteins, some of which are annotated in GenBank as hypothetical proteins (**Supplementary Table S1**). Four inserts were selected for further testing based on the absence of putative degradation enzymes flanking the putative transporters on the metagenomic insert.

We determined the L-lysine IC90 values for *E. coli* transformants where each of the four selected metagenomic inserts have been introduced. The L-lysine IC90 values for these transformants ranged from 10.44 ± 1.277 to 14.25 ± 0.415

$gl^{-1}$, corresponding to a more than 40% (value of $p = 0.0001$) increase in the L-lysine IC90 for the clone harboring pZE-RCL-MglE compared to the empty vector control strain (**Figure 1B**). All of these strains have similar growth profiles in LB medium, whereas the growth rates varied in the liquid LB supplemented with $8\,gl^{-1}$ of L-lysine. At $8\,gl^{-1}$ of L-lysine supplementation, the growth rate of the highest tolerant metagenomic clone harboring pZE-RCL-MglE was 30% higher (value of $p = 0.0375$) than that of the empty vector control (**Figure 1C**).

## Lysine Tolerance by MglE in *Escherichia coli* Strains

To test whether the pZE-RCL-MglE plasmid carrying the *mglE* gene could confer L-lysine tolerance to industrially relevant *E. coli* strains, we introduced the plasmid into various *E. coli* strains; BL21 (DE3), MG1655, Crooks, W1116, and W3110 and IC90 values of L-lysine were determined for all of the constructed *E. coli* strains. Interestingly, the IC90 values of L-lysine for most of these *E. coli* strains (except Crooks and BL21) were increased ($p = <0.0001$) upon introduction of the pZE-RCL-MglE plasmid (**Figure 1D**). Of particular interest, the L-lysine IC90 was increased by 30% ($p < 0.0001$) in W3110, which is a widely used background strain for L-lysine production (Imaizumi et al., 2005). These data demonstrate that the mechanism of L-lysine tolerance mediated by pZE-RCL-MglE is effective across the industrially relevant *E. coli* strains and accordingly should be generally applicable to industrial *E. coli*-based L-lysine fermentations.

## Protein Sequence Analysis of the Metagenomic Insert Carrying MglE

Sequence analysis of the 1.6 kb PCR amplicon from pZE-RCL-MglE metagenomics insert contained an open reading frame *mglE* flanked by native promoter and terminator sequences (**Figure 2A**). The MglE protein consists of 297 amino acids, and its closest homolog is a hypothetical protein from *Bacteroides coprophilus* (GenBank accession no. WP_008140691) with 82% identity at the amino acid level. The MglE protein is a member of RhaT/EamA-like transporter family of the drug/metabolite transporter (DMT; **2.A.7**) superfamily. The MglE contains two copies of the EamA domain, which is found in transporters belonging to the EamA family, at 9–143 aa (pfam (Finn et al., 2014) e-value: $3e^{-11}$) and 152–292 aa (pfam e-value $9.6e^{-12}$). The members of the EamA family are diverse, and most of their functions are unknown (Franke et al., 2003). Nevertheless, a few proteins belonging to EamA family are well characterized as exporters, such as PecM exports a pigmented compound indigoidine in *Erwinia chrysanthemi* (Rouanet and Nasser, 2001) and YdeD export metabolites of the cysteine pathway in *E. coli* (Daßler et al., 2000). The predicted two-dimensional topology of the MglE transporter possesses six cytoplasmic regions, five periplasmic regions, and 10 transmembrane domains, with both the N- and C-terminals in the cytoplasmic region (**Figure 2B**).

Phylogenetic analysis shows that the closest MglE homologs (>50% sequence identity) are present in the phyla bacteroidetes and are annotated as uncharacterized proteins (**Figure 2C**). Members of the LysE family (pfam id: PF01810) are mainly (>5 species per phyla) found in the phyla: Proteobacteria, Bacteroidetes, Firmicutes, Actinobacteria. The full list of phyla that contain species with a LysE domain are Proteobacteria: 102, Bacteroidetes: 73, Firmicutes: 45, Actinobacteria: 28, Chloroflexi: 1, Cyanobacteria: 3, Spirochaetes: 1, Verrucomicrobia: 1, and Gemmatimonadetes: 1. Interestingly, nine genomes containing the MglE homolog also contain additional LysE family (pfam id: PF01810, TC code: **2.A.75**) genes (**Figure 2C** orange dot).

## Effect of Regions in the Metagenomics Insert on L-Lysine Tolerance

To investigate the effects of promoter region of the metagenomic clone on lysine tolerance, the recombinant plasmids pZE-MglE (*mglE* cloned into pZE21), and pZE0-P-MglE-T (MglE along with its native promoter and terminator cloned into pZE21 without the pLteo-1 vector promoter) were subcloned from pZE-RCL-MglE plasmid. The constructed plasmids were transformed into *E. coli* C4860, and IC90 values of L-lysine were determined (**Figure 2D**). Although these recombinant C4860 strains grew in the same extent in LB media, they showed different L-lysine tolerance. We observed that the IC90 for the strain harboring the pZE0-P-MglE-T plasmid was higher than that of the strain harboring the pZE-MglE plasmid, indicating that the native promoter resulted in higher lysine tolerance than the vector promoter. However, the synergistic effect of these two promoters was better still (i.e., the IC90 values of the pZE-RCL-MglE plasmid were higher than those of the strains harboring pZE-MglE or the pZE0-P-MglE-T plasmids). Hence, the lysine tolerance of the *E. coli* C4860 strain was the highest upon harboring the full metagenomic insert (i.e., pZE-RCL-MglE).

## MglE Exports L-Lysine When Expressed in *Xenopus* Oocytes

For functional analysis, MglE was expressed in *Xenopus* oocytes and the export of C13-labeled L-lysine was measured. The water-injected oocytes were used as negative control. L-lysine export assay was performed by injecting C13-labeled L-lysine to obtain a final cytosolic concentration of 6 mM in the oocytes. The oocytes were incubated in Kulori buffers with two different pH (pH 5 and 7.4) for 3 h, and L-lysine concentrations were measured in the buffer and within the oocytes. MglE-expressing oocytes resulted in 6.6-fold and 8.5-fold higher extracellular concentrations of C13-labeled L-lysine and natural C12 L-lysine, respectively, than control oocytes, when buffer with pH 5 was used (**Figure 3Ai**). The intracellular concentrations of C13-labeled and natural C12 L-lysine were correspondingly 2.1-fold and 1.4-fold lower than in control oocytes (**Figure 3Ai**). On the other hand, there was no significant change in L-lysine quantities when oocytes were incubated at pH 7.4 (**Figure 3Aii**). This indicates a proton-gradient dependent export mechanism for MglE transporter.

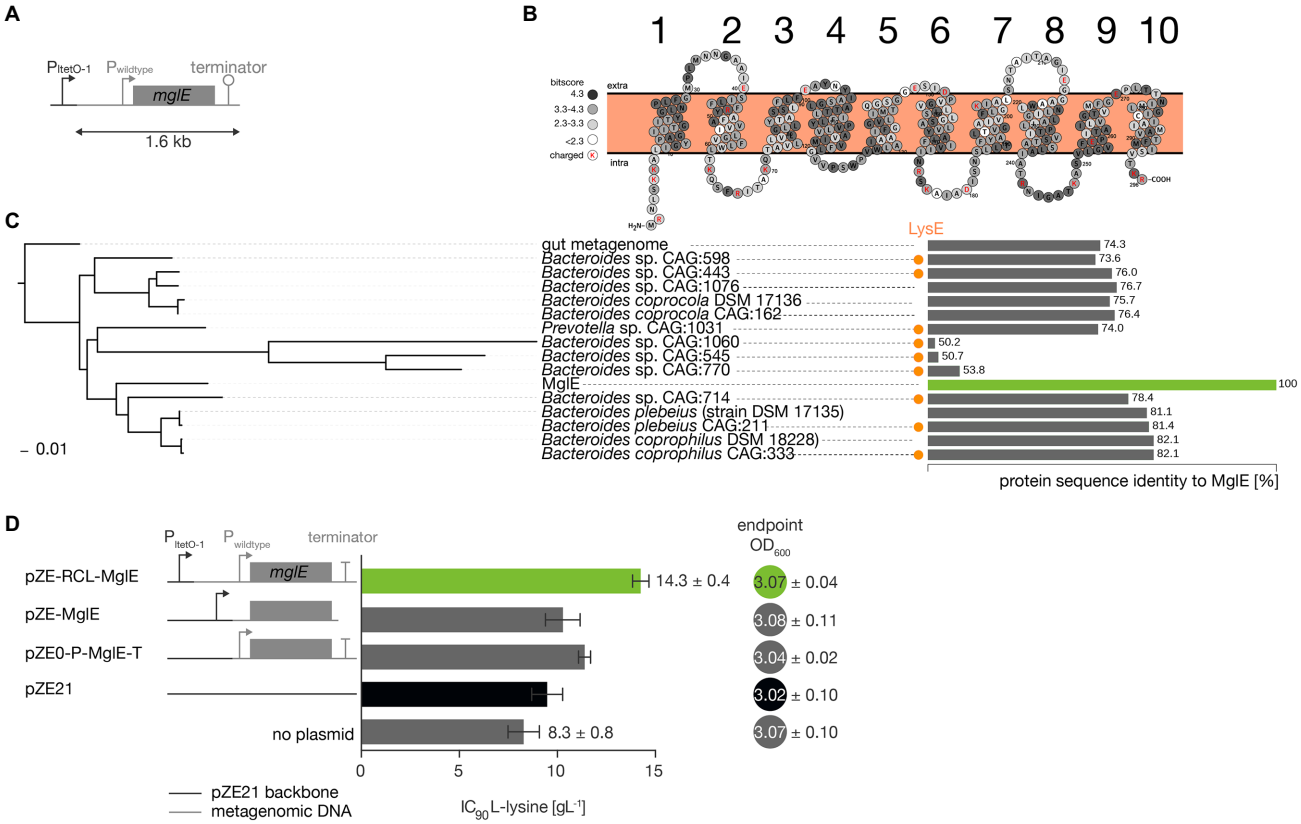



**FIGURE 2 |** *In silico* analysis of the metagenomic insert contained in pZE-RCL-MglE plasmid and L-lysine tolerance of its subclones. **(A)** The metagenomic insert present in the pZE-RCL-MglE plasmid is shown. P: promoter region, MglE: membrane bound transporter protein, and T: terminator sequence. **(B)** The membrane topology of the MglE transporter protein showing 6 cytoplasmic regions, 5 periplasmic regions, and 10 transmembrane α-helices. The 15 BLASTp hits with a > 50% sequence identity were aligned to MglE and the conservation (bitscore) for each residues was calculated using the Biopython package (Cock et al., 2009). **(C)** A Maximum Likelihood phylogenetic protein tree containing all the significant (sequence identity >50%) BLASTp hits in the UniProtKB database against MglE. The percentage identity of the 15 hits is shown as horizontal bars. The presence of a LysE family member in the genomes is shown with an orange dot. The highest homology with MglE is found in *Bacteroides coprophilus* CAG:333, with 82% sequence identity. **(D)** The strategy for subcloning the MglE transporter with or without its native promoter and the resultant L-lysine tolerance phenotypes for the *E. coli* C4860 strain are shown. The biomass of the strains (at 24 h) cultured in LB media was also shown in terms of OD in closed circles.

## MglE Assists L-Lysine Export in *Escherichia coli*

To analyze the function of MglE for L-lysine efflux, *E. coli* DMLC strain was constructed by knocking out two lysine decarboxylase genes, constitutive (*ldcC*) and acid-inducible (*cadA*). These deletions should prevent L-lysine degradation, thereby leading to increased L-lysine production (Kikuchi et al., 1997). Then, the pZE-RCL-MglE plasmid and pZE21 empty vector were transformed into *E. coli* DMLC, yielding *E. coli* DMLC/pZE-RCL-MglE and *E. coli* DMLC/pZE21, respectively. Subsequently, these strains along with *E. coli* W3110/pZE21, were cultured in LB media as well as M9 minimal media supplemented with yeast extract in 96 deep well plates with culture volume of 600 µl.

The bacterial growth and extracellular L-lysine concentrations were measured after 24 h.

In both media, upon knocking out *ldcC* and *cadA*, there were higher L-lysine concentration in supernatant as compared to the wild-type strain and further the extracellular L-lysine concentration was increased upon expression of MglE operon in the deletion mutant. However, the difference of extracellular L-lysine concentration between *E. coli* DMLC and *E. coli* DMLC/pZE-RCL-MglE is not significant in LB media as compared to M9 minimal media because of the presence of higher amount of L-lysine in LB media.

We found that in LB media the absolute extracellular L-lysine titer was increased from 744 to 806 $mgl^{-1}$ in the presence of the metagenomic insert consisting MglE operon that is, 8.3% higher L-lysine production in the *E. coli* DMLC/pZE-RCL-MglE strain as compared to *E. coli* DMLC/pZE21 in LB. However, the biomass of the strain expressing MglE was decreased by about 50% as compared to the control strain (**Figure 3Bi**). Furthermore, the L-lysine production was analyzed in minimal media supplementing yeast extract. In minimal media, the extracellular L-lysine accumulation was increased from 6 to 30 $mgl^{-1}$ ($p < 0.0001$, t-test) whereas there was 12% reduced in biomass upon expressing the MglE metagenomic fragment (**Figure 3Bii**). The difference in extracellular L-lysine concentration among the strain expressing MglE and the control

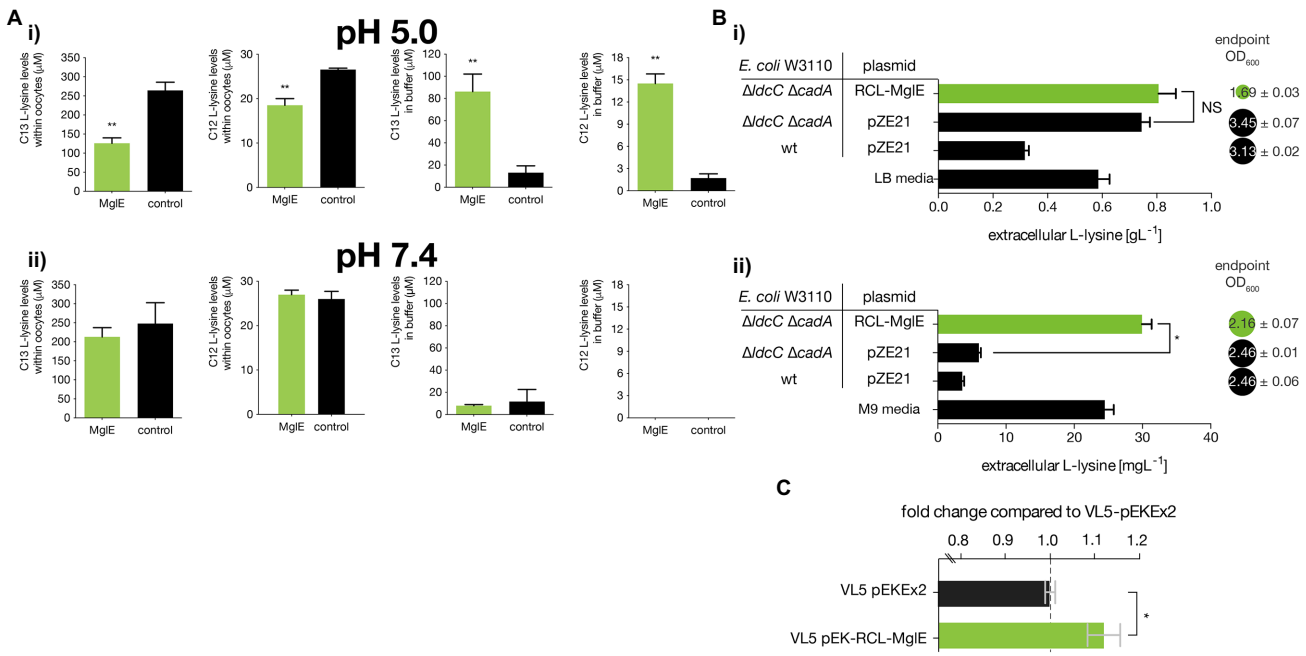



**FIGURE 3 |** Functional characterization of MglE transporter activity in *Xenopus* oocytes, *E.coli* and *Corynebacterium glutamicum*. **(A)** L-lysine export assay of MglE transporter in *Xenopus* oocytes in Kulori buffer (i) at pH 5 and (ii) at pH 7.4. The bars represent lysine contents in the oocytes or the buffer (means ± SD, 3–4 biological replicates each involving 20 oocytes). Significant changes are in comparison with the controls marked by asterisks (** $p < 0.01$; Fishers one-way ANOVA). **(B)** The extracellular L-lysine concentration of reference *E. coli* W3110 and its isogenic mutant strain DMLC along with the cell $OD_{600nm}$ values confirming the exporter activity of the MglE protein in (i) LB media and (ii) in M9 minimal media supplemented with yeast extract (* $p < 0.0001$, *t*-test). **(C)** Fold enhancement of specific production of L-lysine by MglE protein in the industrial *C. glutamicum* VL5 strain compared to the empty vector control. Error bars are s.e.m., * denotes $p = 0.023$, n = 4.

strain is prominent in minimal media (*$p < 0.0001$, *t*-test) than in LB media ($p = 0.1997$, *t*-test). Hence, the above results demonstrate that MglE assists exporting L-lysine from *E. coli*.

The reduced biomass might be due to the combined effect of expression of membrane protein in high copy number and active export of L-lysine by MglE which could have resulted in less available energy for the cell to build biomass. It is worth noting that a big proportion of total cellular energy-demand is for the transport machinery which has accordingly been under evolutionary selection toward a higher energetic efficiency (Darbani et al., 2018).

## Improvement of L-Lysine Productivity in Industrial *Corynebacterium glutamicum* Strain

Expression of a functionally active gene(s) from one strain to another often requires special optimizations. For the practical application in industrial L-lysine bioprocesses, retaining the activity of the discovered exporter into the production host, *C. glutamicum*, is very essential. To test whether the MglE can boost the industrial L-lysine bioprocesses, we cloned and expressed it in *C. glutamicum* production strain.

The full metagenomic insert carrying the MglE along with its native promoter and terminator was amplified from the pZE-RCL-MglE plasmid and cloned into pEKEx2 expression vector to construct pEK-RCL-MglE expression plasmid. The constructed plasmid was introduced into *C. glutamicum* VL5 strain (Materials and methods), a producer strain currently used for industrial L-lysine production. Using sucrose as a carbon source, L-lysine production from the constructed *C. glutamicum* recombinant strain was analyzed at 30 h. The *C. glutamicum* VL5 parent strain and VL5 transformed with empty plasmid (i.e., VL5/pEKEx2) produced 26.59 and 27.13 $gl^{-1}$ of L-lysine respectively, whereas 29.72 $gl^{-1}$ of L-lysine titer was produced by the strain expressing MglE operon (i.e., VL5/pEX-RCL-MglE). Furthermore, L-lysine yield with respect to amount of sucrose consumed were calculated for these three strains which resulted 0.76, 0.77, and 0.83 $gg^{-1}$ for the strains VL5, VL5/pEKEx2, and VL5/pEK-RCL-MglE, respectively. Hence, expression of MglE operon into *C. glutamicum* VL5 improved the L-lysine yield and titer by 7.8% and 9.5%, respectively in comparison to the empty vector control. Unlike in *E. coli* strains, expression of the full metagenomics insert harboring MglE did not have any growth inhibition effect as the vector control and the recombinant strains both have nearly equal biomass. Notably, upon incorporation of the MglE operon into *C. glutamicum*, the specific L-lysine production was improved by 12% ± 0.07% relative to the empty vector control (**Figure 3C**). By expressing the operon in this highly efficient producer strain, we showed not only the functionality of the MglE in a Gram-positive species but also the benefit of the efflux system on an already highly optimized producer strain.

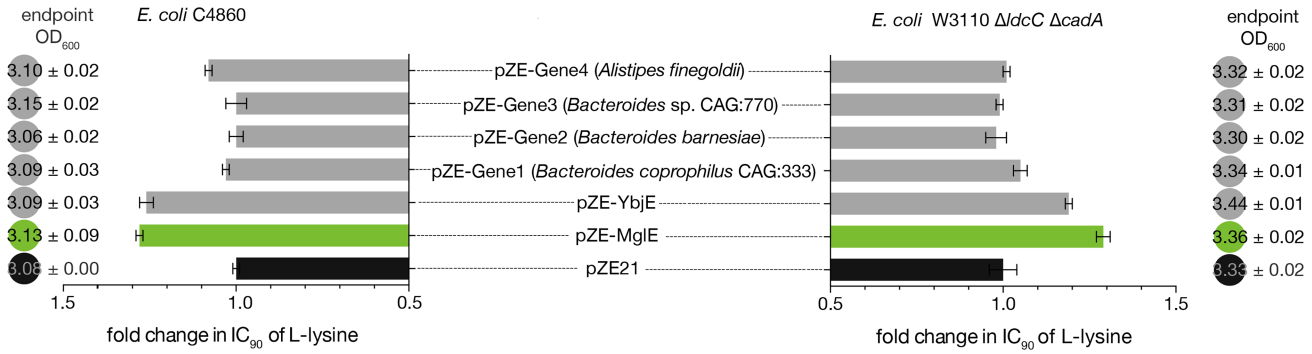



**FIGURE 4 |** Comparison of lysine tolerance (presented in relative IC90 values) displayed by MglE, YbjE (Lysine-specific exporter from *E. coli*), and four synthetic MglE homologs. The biomass of the strains (at 24 h) cultured in LB media was also shown in closed circles.

## Comparison of L-Lysine Tolerance Conferred by MglE and Its Homologous Genes

The YbjE transporter from *E. coli* is one of the functionally characterized lysine-specific exporters. Based upon the homology search, four genes in the NCBI database having different homologies (ranging from 43% to 82% in amino acid level) with MglE were also selected and gene synthetized. To find out the lysine tolerance of MglE as compared to that of YbjE and the four selected MglE homologous proteins (with 82%, 78%, 54%, and 43% amino acid sequence identity), recombinant expression plasmids pZE-YbjE, pZE-Gene1, pZE-Gene2, pZE-Gene3, and pZE-Gene4 were constructed. The nucleotide sequences of these genes are provided in the **Supplementary Information**. The constructed plasmids, pZE-MglE and pZE21 vector control were introduced into *E. coli* C4860 and *E. coli* DMLC strains and determined the IC90 of L-lysine for the constructed recombinant strains as described in early experiments.

Interestingly, we found that the expression of MglE provided similar L-lysine tolerance in *E. coli* C4860 strain as that of YbjE whereas in *E. coli* DMLC strain, MglE displayed even higher tolerance than YbjE. On the other hand, all of the four MglE homologs showed lower lysine tolerance than MglE in both strains (**Figure 4**).

## CONCLUSION

We deployed functional metagenomic selections to identify novel genetic building blocks encoding product efflux systems. Through applying this approach to the problem of L-lysine toxicity at high concentration, we discovered MglE, which is confirmed as a novel type of L-lysine exporter that belongs to the EamA superfamily and this gene could not have been identified by only sequence analysis or homology to known L-lysine transporters. We demonstrated the benefits of MglE in both Gram-negative (*E. coli*) and Gram-positive (*C. glutamicum*) strains and showed that MglE expression improves the productivity of an industrial L-lysine production strain. If incorporated into industrial-scale bioprocesses, this discovery has the potential to enhance L-lysine production by about 10%–12%, representing an increased profit on the order of 200–500 million USD per year. Our approach of using functional metagenomics to identify novel transporters is independent of prior knowledge of transporter gene sequences and can be generally applied to most of the toxic compounds/chemicals production by fermentation.

## PROTEIN ACCESSION NUMBERS

The sequence information of full metagenomics insert harboring *mglE* is deposited in the NCBI GenBank database with the accession number KU708839.

## 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 in the article/**Supplementary Material**.

## AUTHOR CONTRIBUTIONS

MS and SM conceived the study and drafted the manuscript. SM designed and performed the *E. coli* experiments. BD and IB conceived the experiments on *Xenopus* oocytes. BD performed the experiments. SW performed *C. glutamicum* experiments. EH performed *in silico* analysis. MS, IB, and JF led the research teams. EH, BD, SW, JF, and IB contributed to the writing. All authors contributed to the article and approved the submitted version.

## FUNDING

The research leading to these results was funded by the Novo Nordisk Foundation (grant agreement no. NNF10CC1016517) and the European Union Seventh Framework Programme (FP7-KBBE-2013-7-single-stage) under grant agreement no. 613745, Promys. This work was further supported by the European Union Seventh Framework Programme-ITN (FP7/2012/317058 to EH). IB acknowledges the financial support

of the European Research Council under the European Union's Horizon 2020 research and innovation programme (YEAST-TRANS project, grant agreement no. 757384).

## ACKNOWLEDGMENTS

SW acknowledges funding from Højteknologifonden. The authors would like to thank Vitalys I/S for providing the VL5 strain for lab-scale testing and Christian Munk for helping with R-script to calculate IC90 values.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.855736/full#supplementary-material

## REFERENCES

Archer, C. T., Kim, J. F., Jeong, H., Park, J., Vickers, C. E., Lee, S., et al. (2011). The genome sequence of *E. coli* W (ATCC 9637): comparative genome analysis and an improved genome-scale reconstruction of *E. coli*. *BMC Genomics* 12:9. doi: 10.1186/1471-2164-12-9

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., et al. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* 2:2006.0008. doi: 10.1038/msb4100050

Becker, J., and Wittmann, C. (2012). Systems and synthetic metabolic engineering for amino acid production—the heartbeat of industrial strain development. *Curr. Opin. Biotechnol.* 23, 718–726. doi: 10.1016/j.copbio.2011.12.025

Bellmann, A., Vrljić, M., Patek, M., Sahm, H., Kramer, R., and Eggeling, L. (2001). Expression control and specificity of the basic amino acid exporter LysE of *Corynebacterium glutamicum*. *Microbiology* 147, 1765–1774. doi: 10.1099/00221287-147-7-1765

Blattner, F. R., Plunkett, G. I., Bloch, A. C., Perna, T. N., Burland, V., Riley, M., et al. (1997). The complete genome sequence of *Escherichia coli* K-12. *Science* 277, 1453–1462. doi: 10.1126/science.277.5331.1453

Borodina, I. (2019). Understanding metabolite transport gives an upper hand in strain development. *Microb. Biotechnol.* 12, 69–70. doi: 10.1111/1751-7915.13347

Cock, P. J. A., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., et al. (2009). Biopython: freely available python tools for computational molecular biology and bioinformatics. *Bioinformatics* 25, 1422–1423. doi: 10.1093/bioinformatics/btp163

Darbani, B., Kell, D. B., and Borodina, I. (2018). Energetic evolution of cellular Transportomes. *BMC Genomics* 19:418. doi: 10.1186/s12864-018-4816-5

Darbani, B., Motawia, M. S., Olsen, C. E., Nour-Eldin, H. H., Møller, B. L., and Rook, F. (2016). The biosynthetic gene cluster for the cyanogenic glucoside dhurrin in *Sorghum bicolor* contains its co-expressed vacuolar MATE transporter. *Sci. Rep.* 6, 1–8. doi: 10.1038/srep37079

Darbani, B., Stovicek, V., van der Hoek, S. A., and Borodina, I. (2019). Engineering energetically efficient transport of dicarboxylic acids in yeast *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.* 116, 19415–19420. doi: 10.1073/pnas.1900287116

Daßler, T., Maier, T., Winterhalter, C., and Böck, A. (2000). Identification of a major facilitator protein from *Escherichia coli* involved in efflux of metabolites of the cysteine pathway. *Mol. Microbiol.* 36, 1101–1112. doi: 10.1046/j.1365-2958.2000.01924.x

Eikmanns, B. J., Kleinertz, E., Liebl, W., and Sahm, H. (1991). A family of *Corynebacterium glutamicum/Escherichia coli* shuttle vectors for cloning, controlled gene expression, and promoter probing. *Gene* 102, 93–98. doi: 10.1016/0378-1119(91)90545-M

Elder, M. (2018). World markets for fermentation ingredients, BCC Research: Market Research Reports. FOD020E. Available at: https://www.bccresearch.com/market-research/food-and-beverage/world-markets-for-fermentation-ingredients-fod020e.html (Accessed January 08, 2022).

Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., et al. (2014). Pfam: the protein families database. *Nucleic Acids Res.* 42, D222–D230. doi: 10.1093/nar/gkt1223

Forsberg, K. J., Patel, S., Witt, E., Wang, B., Ellison, T. D., and Dantas, G. (2016). Identification of genes conferring tolerance to lignocellulose-derived inhibitors by functional selections in soil metagenomes. *Appl. Environ. Microbiol.* 82, 528–537. doi: 10.1128/AEM.02838-15

Franke, I., Resch, A., Daßler, T., Maier, T., and Bock, A. (2003). YfiK from *Escherichia coli* promotes export of *O*-acetylserine and cysteine. *J. Bacteriol.* 185, 1161–1166. doi: 10.1128/JB.185.4.1161

Gunji, Y., and Yasueda, H. (2006). Enhancement of l-lysine production in methylotroph *Methylophilus methylotrophus* by introducing a mutant LysE exporter. *J. Biotechnol.* 127, 1–13. doi: 10.1016/j.jbiotec.2006.06.003

Hayashi, K., Morooka, N., Yamamoto, Y., Fujita, K., Isono, K., Choi, S., et al. (2006). Highly accurate genome sequences of Escherichia coli K-12 strains MG1655 and W3110. *Mol. Syst. Biol.* 2, 1–5. doi: 10.1038/msb4100049

Hemberger, S., Pedrolli, D. B., Stolz, J., Vogl, C., Lehmann, M., and Mack, M. (2011). RibM from *Streptomyces davawensis* is a riboflavin/roseoflavin transporter and may be useful for the optimization of riboflavin production strains. *BMC Biotechnol.* 11:119. doi: 10.1186/1472-6750-11-119

Imaizumi, A., Takikawa, R., Koseki, C., Usuda, Y., Yasueda, H., Kojima, H., et al. (2005). Improved production of L-lysine by disruption of stationary phase-specific *rmf* gene in *Escherichia coli*. *J. Biotechnol.* 117, 111–118. doi: 10.1016/j.jbiotec.2004.12.014

Jensen, S. I., Lennen, R. M., Herrgård, M. J., and Nielsen, A. T. (2015). Seven gene deletions in seven days: fast generation of *Escherichia coli* strains tolerant to acetate and osmotic stress. *Sci. Rep.* 5:17874. doi: 10.1038/srep17874

Käll, L., Krogh, A., and Sonnhammer, E. L. L. (2007). Advantages of combined transmembrane topology and signal peptide prediction-the Phobius web server. *Nucleic Acids Res.* 35, W429–W432. doi: 10.1093/nar/gkm256

Keilhauer, C., Eggeling, L., and Sahm, H. (1993). Isoleucine synthesis in *Corynebacterium glutamicum*: molecular analysis of the ilvB-ilvN-ilvC operon. *J. Bacteriol.* 175, 5595–5603. doi: 10.1128/jb.175.17.5595-5603.1993

Kelle, R., Laufer, B., Brunzema, C., Weuster-Botz, D., Krämer, R., and Wandrey, C. (1996). Reaction engineering analysis of L-lysine transport by *Corynebacterium glutamicum*. *Biotechnol. Bioeng.* 51, 40–50. doi: 10.1002/(SICI)1097-0290(19960705)51:1<40::AID-BIT5>3.0.CO;2-0

Kikuchi, Y., Kojima, H., Tanaka, T., Takatsuka, Y., and Kamio, Y. (1997). Characterization of a second lysine decarboxylase isolated from *Escherichia coli*. *J. Bacteriol.* 179, 4486–4492. doi: 10.1128/jb.179.14.4486-4492.1997

Lee, S. Y., and Kim, H. U. (2015). Systems strategies for developing industrial microbial strains. *Nat. Biotechnol.* 33, 1061–1072. doi: 10.1038/nbt.3365

Lutz, R., and Bujard, H. (1997). Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res.* 25, 1203–1210. doi: 10.1093/nar/25.6.1203

Malla, S., Niraula, N. P., Liou, K., and Sohng, J. K. (2010). Self-resistance mechanism in *Streptomyces peucetius*: overexpression of drrA, drrB and drrC for doxorubicin enhancement. *Microbiol. Res.* 165, 259–267. doi: 10.1016/j.micres.2009.04.002

Munck, C., Albertsen, M., Telke, A., Ellabaan, M., Nielsen, P. H., and Sommer, M. O. A. (2015). Limited dissemination of the wastewater treatment plant core resistome. *Nat. Commun.* 6:8452. doi: 10.1038/ncomms9452

Omasits, U., Ahrens, C. H., Müller, S., and Wollscheid, B. (2014). Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics* 30, 884–886. doi: 10.1093/bioinformatics/btt607

Pathania, A., and Sardesai, A. A. (2015). Distinct paths for basic amino acid export in *Escherichia coli*: YbjE (LysO) mediates export of L-lysine. *J. Bacteriol.* 197, 2036–2047. doi: 10.1128/JB.02505-14

Rouanet, C., and Nasser, W. (2001). The PecM protein of the phytopathogenic bacterium *Erwinia chrysanthemi*, membrane topology and possible involvement in the efflux of the blue pigment. *J. Mol. Microbiol. Biotechnol.* 3, 309–318.

Sambrook, J., and Russell, D. W. (2001). *Molecular Cloning: A Laboratory Manual. 3rd Edn.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Sommer, M. O. A., Church, G. M., and Dantas, G. (2010). A functional metagenomic approach for expanding the synthetic biology toolbox for biomass conversion. *Mol. Syst. Biol.* 6:360. doi: 10.1038/msb.2010.16

Sommer, M. O. A., Dantas, G., and Church, G. M. (2009). Functional characterization of the antibiotic resistance reservoir in the human microflora. *Science* 325, 1128–1131. doi: 10.1126/science.1176950

UniProt Consortium (2014). UniProt: a hub for protein information. *Nucleic Acids Res.* 43, D204–D212. doi: 10.1093/nar/gku989

van der Helm, E., Geertz-Hansen, H. M., Genee, H. J., Malla, S., and Sommer, M. O. A. (2015). deFUME: dynamic exploration of functional metagenomic sequencing data. *BMC. Res. Notes* 8, 328–330. doi: 10.1186/s13104-015-1281-y

Vrljic, M., Garg, J., Bellmann, A., Wachi, S., Freudi, R., Malecki, M., et al. (1999). The LysE superfamily: topology of the lysine exporter LysE of *Corynebacterium glutamicum*, a paradyme for a novel superfamily of transmembrane solute translocators. *J. Mol. Microbiol. Biotechnol.* 1, 327–336.

Vrljic, M., Sahm, H., and Eggeling, L. (1996). A new type of transporter with a new type of cellular function: L-lysine export from *Corynebacterium glutamicum*. *Mol. Microbiol.* 22, 815–826. doi: 10.1046/j.1365-2958.1996.01527.x

**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.

*Copyright © 2022 Malla, van der Helm, Darbani, Wieschalka, Förster, Borodina and Sommer. 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.*

frontiers | Frontiers in Microbiology

ORIGINAL RESEARCH
published: 18 April 2022
doi: 10.3389/fmicb.2022.872397

# Improving the Yield and Quality of Daptomycin in *Streptomyces roseosporus* by Multilevel Metabolic Engineering

*Zhong-Yuan Lyu[1,2], Qing-Ting Bu[1,2], Jiao-Le Fang[1,2], Chen-Yang Zhu[1,2], Wei-Feng Xu[1,2], Lie Ma[1,2], Wen-Li Gao[1,2], Xin-Ai Chen[1,2] and Yong-Quan Li[1,2]**

[1] First Affiliated Hospital and Institute of Pharmaceutical Biotechnology, Zhejiang University School of Medicine, Hangzhou, China, [2] Zhejiang Provincial Key Laboratory for Microbial Biochemistry and Metabolic Engineering, Hangzhou, China

Daptomycin is a cyclic lipopeptide antibiotic with a significant antibacterial action against antibiotic-resistant Gram-positive bacteria. Despite numerous attempts to enhance daptomycin yield throughout the years, the production remains unsatisfactory. This study reports the application of multilevel metabolic engineering strategies in *Streptomyces roseosporus* to reconstruct high-quality daptomycin overproducing strain L2797-VHb, including precursor engineering (i.e., refactoring kynurenine pathway), regulatory pathway reconstruction (i.e., knocking out negative regulatory genes *arpA* and *phaR*), byproduct engineering (i.e., removing pigment), multicopy biosynthetic gene cluster (BGC), and fermentation process engineering (i.e., enhancing $O_2$ supply). The daptomycin titer of L2797-VHb arrived at 113 mg/l with 565% higher comparing the starting strain L2790 (17 mg/l) in shake flasks and was further increased to 786 mg/l in 15 L fermenter. This multilevel metabolic engineering method not only effectively increases daptomycin production, but can also be applied to enhance antibiotic production in other industrial strains.

**Keywords:** ***Streptomyces roseosporus*, daptomycin, titer, quality, multi-level metabolic engineering**

OPEN ACCESS

***Edited by:***
*Niranjan Koirala, Gandaki Province Academy of Science and Technology, Nepal*

***Reviewed by:***
*Sanjay Kumar Singh Patel, Konkuk University, South Korea*
*Gang Liu, Institute of Microbiology (CAS), China*

***Correspondence:**
*Yong-Quan Li*
*lyq@zju.edu.cn*

***Specialty section:***
*This article was submitted to Microbiotechnology, a section of the journal Frontiers in Microbiology*

***Received:*** *09 February 2022*
***Accepted:*** *11 March 2022*
***Published:*** *18 April 2022*

***Citation:***
*Lyu Z-Y, Bu Q-T, Fang J-L, Zhu C-Y, Xu W-F, Ma L, Gao W-L, Chen X-A and Li Y-Q (2022) Improving the Yield and Quality of Daptomycin in Streptomyces roseosporus by Multilevel Metabolic Engineering. Front. Microbiol. 13:872397. doi: 10.3389/fmicb.2022.872397*

## INTRODUCTION

Daptomycin is a cyclic lipopeptide antibiotic produced by *Streptomyces roseosporus* (*S. roseosporus*) *via* non-ribosomal peptide synthetases (NRPSs; Debono et al., 1988). It is the best substituent to vancomycin, due to its high antibacterial activity to Gram-positive pathogens, including vancomycin-resistant *Staphylococcus aureus* (*S. aureus*), methicillin-resistant *S. aureus*, and vancomycin-resistant *Enterococci* (Akins and Rybak, 2001). Therefore, how to increase daptomycin production has attracted more attention of researchers.

Daptomycin biosynthetic gene cluster (BGC) belongs to the classical NRPS pathway. Daptomycin synthesis begins with the activation of decanoic acid by DptE, which then transfers the acid onto DptF. The condensation reaction between acid and tryptophan is catalyzed by DptA. Next, under the action of DptA, DptBC, and DptD, the remaining 12 amino acids are condensed and connected to the being synthesized peptide chain in turn. Finally, under the catalysis of the thioesterase domain contained in DptD, through the condensation of Thr4 and Kyn13, the peptide chain is cyclized to produce daptomycin (**Supplementary Figure 1**;

Miao et al., 2005; Wittmann et al., 2008). Kynurenine (Kyn) is one of the main non-proteinogenic amino acid precursors in daptomycin biosynthesis, which is an intermediate produced in the tryptophan degradation pathway. In *S. roseosporus*, tryptophan degrades to generate Kyn, catalyzed by tryptophan-2,3-dioxygenase (DptJ and TDO) and N-formyl kynurenine formamidase, successively. Then, Kyn is converted to anthranilate by kynureninase encoded by the *kyn* gene (Liao et al., 2013) or *orf3244* (predicted by us in this study). According to our prediction in this study, kynurenine is also a substrate for the catalysis of the protein encoded by gene *orf3242*. Overexpression of *dptJ* and disruption of *kyn* made the production of daptomycin an increase of 110 and 30%, respectively (Liao et al., 2013). Decanoic acid (DA) is another important precursor in the production of daptomycin (Wittmann et al., 2008). Daptomycin yield was increased by 40% compared with the wild-type through developing a DA-resistant (DAR) *S. roseosporus via* a sequential adaptation method (Lee et al., 2016). Oxygen is also a pivotal factor in the synthesis of secondary metabolites by *Streptomyces*. Sufficient oxygen helps in better growth and differentiation of mycelium, thereby facilitating product accumulation. Increasing stirring speed or airflow was usually used to improve the dissolved oxygen in the culture broth (Wentzel et al., 2012). However, because of filamentous nature of *Streptomyces*, higher agitation speed may damage mycelium itself and then affect the production of secondary metabolites (Ferraiuolo et al., 2021). Vitreous hemoglobin (VHb) is an oxygen-binding protein that helps microorganisms to grow under low dissolved oxygen conditions. Its heterologous expression successfully increased metabolite productivity by enhancing oxygen utilization efficiency (Horng et al., 2010; Zhu et al., 2011, Li et al., 2016; Mironczuk et al., 2019). This is a great alternative to boost the oxygen supply to prevent hyphae damage.

Furthermore, daptomycin biosynthesis is under the control of several regulators. Three cluster-situated regulators, namely, DptR1, DptR2, and DptR3, are all required for daptomycin production (Wang et al., 2014; Zhang et al., 2015, Yu et al., 2020). Additionally, other factors far away from the daptomycin BGC also play a regulatory role. arpA, a homolog of the A-factor receptor, negatively controls morphological development and daptomycin production in *S. roseosporus*. Removal of *arpA* can increase daptomycin production (Mao et al., 2015). In 2018, a new transcriptional regulator phaR was identified in *S. roseosporus*, which negatively regulates the expression of daptomycin BGC. Deletion of *phaR* led to an approximately 43% increase of daptomycin production in fed-batch fermentation (Luo et al., 2018b). Moreover, the daptomycin BGC has been captured and used to improve the production of daptomycin in its native or heterologous hosts. For instance, daptomycin BGC was cloned from *S. roseosporus* NRRL 15998 and introduced to the strain itself, leading to an increase of daptomycin yield (Du et al., 2015). In addition, a 65-kb daptomycin BGC was expressed heterologously in *Streptomyces coelicolor* (*S. coelicolor*) M511 and the production of daptomycin and its derivatives reached 28.9 mg/l (Choi et al., 2019). Besides, it will also be accompanied by the output of byproducts in the secondary metabolic process, which will affect the separation and purification process, resulting in an increase in industrial costs (Wang et al., 2019). Notably, pigments are sometimes found as byproducts of the antibiotic biosynthesis process. Reseachers improved the fidaxomicin production by removing the orange pigment byproduct in an industrial strain *Actinoplanes deccanensis* YP-1 (Li et al., 2021c).

Although many strategies have been applied to enhance the production of daptomycin, most of them are only confined to the manipulation of single factor, such as regulation or precursor supply. Actually, the biosynthesis of daptomycin is influenced by multiple (external and internal) factors. Therefore, to achieve a higher yield of daptomycin, various favorable factors need to be considered and integrated. Here, a combinatorial metabolic engineering measure was implemented to construct daptomycin-overproducing strains (**Figure 1**). To this end, the supply of precursor Kyn was enhanced by deletion of *orf3242* and *orf3244* (level 1). Next, pleiotropic regulator genes, including *arpA* and *phaR*, were knocked out to release the negative regulation (level 2). Then, partial knockout of genes participating in pigment biosynthesis blocked the pigment production (level 3) and this is the first time to remove red pigment for daptomycin production. Subsequently, extra copy of daptomycin BGC was integrated into L2796 (level 4). Finally, heterologous VHb was expressed in L2797 to alleviate the low dissolved oxygen limitation during large-scale fermentation with daptomycin production up to 786 mg/l (level 5). Such a multilevel design and integration may be offering a method to improve the quality and yield of products in many industrial bacteria.

# MATERIALS AND METHODS

## Strains and Media

*Streptomyces roseosporus* L2790 was a daptomycin-producing strain, which was stored in our laboratory. *Escherichia coli* (*E. coli*) TG1 (Novagen) was a general cloning host. *E. coli* ET12567/pUZ8002 and ET12567/pUB307 were used to introduce plasmid into *Streptomyces*. *E. coli* GB05RedTrfA (containing pSC101-BAD-ETgA-tet) was used to clone the daptomycin BGC (Wang et al., 2016).

Solid and shake-flask fermentation medium was described previously (Kieser et al., 2000; Wang et al., 2014). In fed-batch fermentation, the primary seed medium contained 3% Tryptic Soy Broth (TSB) and 3% maltodextrin (MD), abbreviated as TSB-MD. The secondary seed medium contained 6% maltodextrin, 2.5% soybean powder, 1.5% glucose, 0.08% $(NH_4)_2Fe(SO_4)_2{\cdot}6H_2O$, 0.1% Yeast extract, 0.5% calcium carbonate, and 0.2% molasses. The fermentation medium contained 7.2% maltodextrin, 1.2% yeast powder, 1.0% glucose, 0.08% $(NH_4)_2Fe(SO_4)_2{\cdot}6H_2O$, 0.72% molasses, and 0.1% defoamer GPE.

## Plasmid Construction

The plasmids and primers used in this article are shown in **Supplementary Tables 1, 2**, respectively. Primer pairs 1/2 and 3/4 were used to amplify the left and right homologous arms,

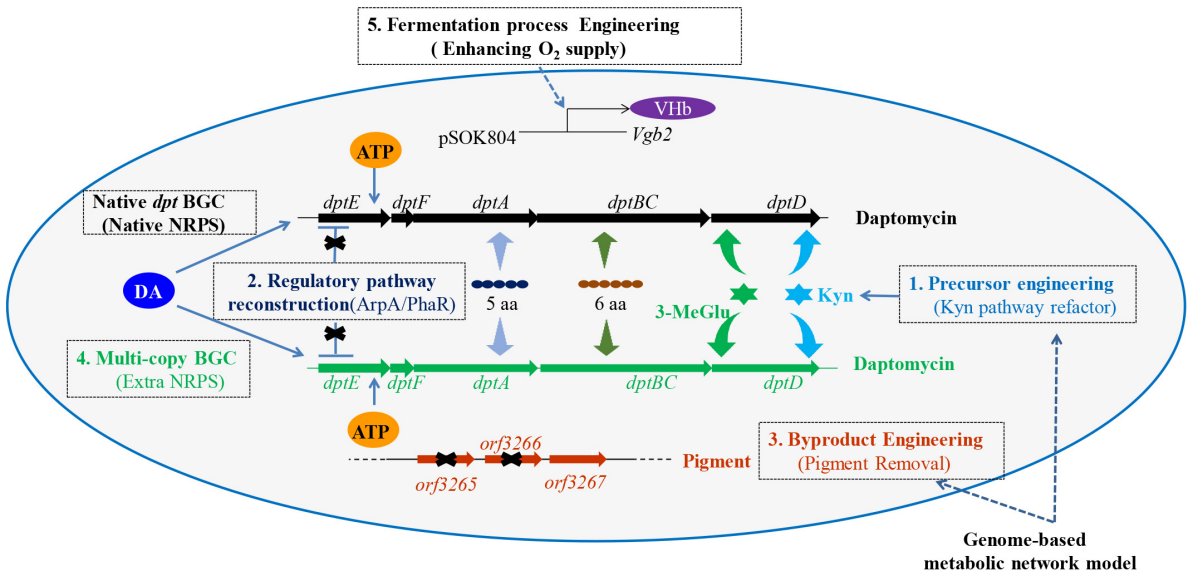



**FIGURE 1 |** The overview map of metabolic engineering strategies for improving daptomycin production and quality in *Streptomyces roseosporus* (*S. roseosporus*) L2790. In this study, the multilevel engineering strategies included enhancement of Kyn precursor supply (level 1), reconstruction of the regulatory pathway (level 2), byproduct engineering (level 3), multicopy of BGC (level 4), and fermentation process engineering (level 5). Kyn, kynurenine; 3-MeGlu, 3′-methylglutamic acid; 6 aa, six amino acids (Orn 6, Asp 7, D-Ala 8, Asp 9, Gly 10, D-Ser 11); 5 aa, five amino acids (Trp 1, D-Asn 2, Asp 3, Thr 4, Gly 5); DA, decanoic acid; NRPS, non-ribosomal peptide synthetase; BGC, biosynthetic gene cluster; VHb, *Vitreoscilla* hemoglobin; *Vgb2*, a gene encoding protein VHb; pSOK804, an overexpression vector; *dptE*, *dptF*, *dptA*, *dptBC*, and *dptD* are the genes locating in daptomycin BGC and being involved in the biosynthesis of daptomycin; for level 2, the "X" indicates a block of negative regulation; for level 3, the "X" indicates deletion of the relevant genes.

which were cloned into pKC1139 to get the plasmid pKC1139-Δ*orf3244*. Similar to the above construction process, we got pKC1139-Δ*arpA* and pKC1139-Δ*orf3265*-Δ*orf3266*. Deletion of *orf3242* was implemented by using the CRISPR/Cpf1-mediated gene-editing system. The software sgRNA Scorer version 2.0 was used to design the spacer sequences, which target the deletion sequences with Protospacer Adjacent Motif (PAM) TTV [the PAM of Cpf1 where T is thymine and V is nucleotide A(adenine), C(cytidine), or G(guanine)] at its 5' end (Chari et al., 2017). The primer pair 5/6 was used to amplify the CRISPR RNA (CrRNA) cassettes. Primer pairs 7/8 and 9/10 were used to amplify two homologous arms. Then, assembled homologous arms were obtained by overlapped PCR. Finally, the ClonExpress II One Step Cloning Kit (Vazyme Biotech Corporation, China) was used to connect *Nde*I/*Spe*I-digested pKCCpf1 and the overlapping PCR products to get plasmid pKCCpf1-Δ*orf3242*. Similar to the construction of plasmid pKCCpf1-Δ*orf3242*, we got pKCCpf1-Δ*phaR*. The primer pair 11/12 was used to amplify *dptJ*, which was cloned into pIJ8661, an integrative plasmid containing the strong promoter *ermEp**, to get pIJ8661-*dptJ*. Primers 13 and 14 were used to amplify *orf3245* with its native ribosome binding site (RBS), which was cloned into pIJ8661-*dptJ* to get pIJ8661-*dptJ-orf3245*. Primers 15 and 16 were used to amplify *orf3243* with its native RBS, which was cloned into pIJ8661-*dptJ-orf3245* to get pIJ8661-*dptJ-orf3245-orf3343*. The plasmid pIJ8661-*dptJ-orf3245-orf3343* was digested by *Nde*I/*Kpn*I to get *Nde*I-*dptJ-orf3245-orf3343-Kpn*I, which was ligated to *Nde*I/*Kpn*I-digested pSOK804 by the T4 DNA ligase (Vazyme Biotech Corporation Ltd., China) to get plasmid pSOK804-*ermEp**-*dptJ-orf3245-orf3343* at last. The primer pair 30/31 was used to amplify ampicillin (AMP) resistance gene fragment by PCR with the plasmid pUT18 as a template. The primer pair 32/33 was used to amplify the attP-integrase fragment with plasmid pSET152-*vgb2* as a template. The attP-φ31 integrase-AMP cassette was obtained by fusion PCR and then it was electroporated into arabinose-induced *E. coli* GB2005RedTrfA and inserted into plasmid pBeloBac*:*dpt* (**Supplementary Table 1**) by *in vivo* linear-circular homologous recombination (LCHR), generating the plasmid 701DIAA. The primer pair 34/35 was used to amplify homologous recombination fragment (-SmR-) containing spectinomycin resistance (*SmR*) gene and then it was inserted into 701DIAA by *in vivo* LCHR, generating the final plasmid 702DIAAS. Recombinant plasmids 701DIAA and 702DIAAS were identified by PCR (**Supplementary Figure 2**). The primer pair 36/37 was used to amplify *Vitreoscilla* hemoglobin (*vgb2*) gene with pSET152-*vgb2* as template and then the *vgb2* fragment was ligated to *Nde*I/*Kpn*I-digested pSOK804 by T4 DNA ligase. All the fragments were amplified with KOD (the trade name of DNA polymerase from the hyperthermophilic Archaeon *Thermococcus kodakaraensis* KOD1) plus-neo DNA polymerase (Toyobo), a high-fidelity DNA polymerase.

## Building of Metabolic Network Models Based on Genome

Genomic analyses of *S. roseosporus* L2790 were performed by the Rapid Annotations using Subsystems Technology (RAST)

Server[1] (Aziz et al., 2008). According to analysis results, the putative Kyn metabolic network model was built. In addition, we searched for genes perhaps participating in pigment synthesis in the Spreadsheet downloaded from the RAST Server.

## Construction and Fermentation of *Streptomyces roseosporus* Strains

The plasmid pKC1139-Δ*orf3244* was introduced into *S. roseosporus* L2790 by conjugation to get L2791. The deletion plasmid pKCCPf1-Δ*orf3242* was introduced into L2791 by conjugation to get L2792. Similarly, plasmid pKCCPf1-*phaR* and pKC1139-*arpA* were used to knock out gene *phaR* and *arpA* in sequence in L2792 to get L2795. The plasmid pKC1139-Δ*orf3265*-Δ*orf3266* was introduced into L2795 by conjugation to get L2796. The strains L2791, L2792, L2795, and L2796 were all obtained by using the in-frame deletion strategy. Furthermore, the plasmid pSOK804-*dptJ-sro3245-orf3343* was introduced into L2792 to get strain L2792a. Plasmid 702DIAAS was introduced into L2796 to get L2797. Plasmid pSOK804-*vgb2* was introduced into L2797 to get L2797-VHb, which can heterologously express VHb protein in L2797. Strains L2792a, L2797, and L2797-VHb were all obtained by insertion of target genes into the genome *via* site-specific integrase, respectively. All the recombinant strains were identified by PCR (**Supplementary Figure 3**). Shake-flask and scale-up fermentation were performed according to **Supplementary Fermentation Process Document**.

## Analysis of Daptomycin by High-Performance Liquid Chromatography

The C18 reverse-phase column (Zorbax 300SB-C18, 5 μm, 4.6 mm × 250 mm) from Agilent (Wang et al., 2014) was used to analyze daptomycin by high-performance liquid chromatography (HPLC) at a rate of 1 ml/min with mobile phases, including solution A ($H_2O$ containing 0.1% Formic acid) and solution B (acetonitrile). The HPLC procedure and samples preparation were described previously (Luo et al., 2018a). UV detection wavelength was 215 nm.

## Creation of Heatmap

Programs, such as Local Basic Local Alignment Search Tool (BLAST) and TBtools, were used to build a heatmap. In brief, EC 2.6.1.7 may be a kynurenine aminotransferase or kynurenine–oxoglutarate transaminase according to sequence data for EC numbers in the Kyoto Encyclopedia for Genes and Genomes (KEGG). In the National Center for Biotechnology Information (NCBI) database, we chose a kynurenine aminotransferase AM113472 from *Micromonospora* spp. *ML1* and a kynurenine–oxoglutarate transaminase NP_012475 from *Saccharomyces cerevisiae* S288C for local BLAST. In addition, an aspartate aminotransferase AAF10201 from *Deinococcus radiodurans* R1, which had the closest match with AM113472 (Lombo et al., 2006), was also chosen for local BLAST. A heatmap was built by using TBtools according to the local BLAST results.

[1]https://rast.nmpdr.org/

## Analysis of Vitreous Hemoglobin by Carbon Monoxide-Difference Spectrum

Carbon monoxide (CO)-difference spectrum was used to detect the activity of VHb (Zhu et al., 2011; Li et al., 2016). The mycelia of L2797 and L2797-VHb were broken with an ultrasonic homogenizer (JY92-IIDN, Ningbo Scientz Biotechnology Corporation Ltd.) on ice for 15 min with a cycle of a 3-s work and a 5-s pause. The working power was 35%. The CO-difference spectrum was obtained by scanning in the 400–550 nm range using a UV-visible spectrophotometer (Evolution 220, Thermo Fisher Scientific Corporation Ltd.).

## Concentration Measurement of Tryptophan and Kynurenine by High-Performance Ligquid Chromatography

In shake-flask fermentation progress, 1 ml of culture was harvested each 24 h. The mycelia were washed and then disrupted by sonication with 500 μl phosphate buffer (50 mM, pH = 7.0). The content of total protein was determined by the Bradford method. Tryptophan and Kyn in the lysate were analyzed by HPLC (1260 Infinity, Agilent Technologies) with a reverse-phase column (Zorbax 300SB-C18, 5 μm, 4.6 mm × 250 mm; Agilent Technologies) with solution A (15 mM sodium acetate, pH = 4.0) and solution B (100% acetonitrile) at the ratio of A:B = 92:8. The detected wavelength was 360 nm for Kyn and 280 nm for tryptophan. The flow rate was 1.0 ml/min. 100 μl of cleared lysate was injected for HPLC analysis.

# RESULTS AND DISCUSSION

## Improve Daptomycin Titer by Enhancing Kynurenine Supply

Kynurenine is a non-proteinogenic amino acid, which is an important precursor for the synthesis of daptomycin. Daptomycin production was improved by partially modifying the Kyn pathway (Liao et al., 2013). We measured the concentration of Kyn and tryptophan in the cytoplasm of strain L2790 (**Supplementary Figure 4**). The results showed that the concentration of tryptophan varied from 0.08 to 0.15 nmol/μg (amino acid/total protein), while the concentration of Kyn was 0.003 nmol/μg in 24 h and was too low to be detected in other time points. It suggested that the supply of Kyn was seriously inadequate or the kynurenine was drained out rapidly as soon as it was generated. Therefore, we attempted to systematically enhance the metabolic flux to supply more Kyn.

The L-Kyn metabolic network based on the KEGG metabolic analysis in the strain L2790 was established using the RAST Server (**Figure 2A**; Aziz et al., 2008). At least five proteins were found to be involved in its metabolic pathway (**Supplementary Table 2**). Among them, tryptophan-2,3-dioxygenase, including DptJ and Orf3245, converts tryptophan to N′-formyl-L-kynurenine (NFK). NFK was converted to Kyn by an NFK formamidase (Orf3243) or formylanthranilate by L-Kynureninase (Orf3244). L-Kyn can be turned to

anthranilate catalyzed by Orf3244, 3-hydroxy-L-kynurenine by an oxidoreductase (Orf3242), or 4-(2-aminophenyl)-2,4-dioxobutanoate by a kynurenine aminotransferase (EC2.6.1.7), which may be one of three putative proteins, namely, Orf3256, Orf2688, or Orf2371, according to a heatmap (**Supplementary Figure 5**).

To improve Kyn supply, we firstly knocked out *orf3244* and *orf3242* in turn. Disruption of *orf3244* in L2790 got strain L2791, in which daptomycin concentration reached 30.0 mg/l with an increase of 74% compared with L2790 (**Figure 2B**). Disruption of *orf3242* in L2791, generating strain L2792, further improved daptomycin production to 36.3 mg/l, 1.21-fold of strain L2791 (**Figure 2B**). Besides, gene *dptJ*, *orf3245*, and *orf3243* were involved in kynurenine biosynthesis according to L-Kyn metabolic network (**Figure 2A**). Therefore, we co-overexpressed *dptJ*, *orf3245*, and *orf3243* under the control of a strong promoter *ermEp** in L2792, generating strain L2792a. However, daptomycin production of L2792a did not increase (**Figure 2B**).

## Engineering of Regulatory Pathway in Daptomycin Biosynthesis

ArpA and PhaR are two known pleiotropic regulators in *S. roseosporus*. ArpA indirectly and negatively regulates the transcription of daptomycin BGC through the A factor signaling pathway, while PhaR direct negatively regulates the transcription level of the gene cluster by binding to the transcription initiation region of the daptomycin BGC (Mao et al., 2015; Luo et al., 2018b). Thus, these two regulatory genes were knocked out to remove their negative effect on daptomycin biosynthesis. The resultant strain L2795 produced 68 mg/l daptomycin, an increase of 90%, compared with L2792 (**Figure 3A**).

## Blocking the Red Pigment Synthesis to Eliminate the Byproduct in *Streptomyces roseosporus*

Red pigment is observed as the main byproduct accompanied by the daptomycin production *in S. roseosporus* L2795 (**Figure 3B**), which seriously affected the following separation and purification process, thereby reducing daptomycin quality and increasing the production cost. The RAST Server was used to find the pathway of pigment biosynthesis and five genes could be involved in the synthesis of the pigment (**Supplementary Table 2**). Among them, *orf5781*, a putative indigoidine synthase A-like protein-encoding gene, might not be related to red pigment synthesis and its loci were also far away from the other four genes, namely, *orf3259*, *orf3265*, *orf3266*, and *orf3267* (**Supplementary Table 2**). According to the gene function annotated by the RAST Server, these four genes are involved in red pigment biosynthesis. Orf3259 is polyketide cyclase WhiE VII, Orf3265 is polyketide chain length factor WhiE-CLF, Orf3266 is polyketide beta-ketoacyl synthase WhiE-KS, and Orf3267 is polyketide cyclase WhiE II. They are all the typical type II PKS and adjacent to each other. It is reported that many pigment-type natural products are synthesized by type II PKS (Staunton and Weissman, 2001). To prove this idea, *orf3265* and *orf3266* were knocked out to block the synthesis of red pigment, generating strain L2796. Compared with the parent strain L2795, the aerial mycelium of L2796 turned white (**Figure 3B**) and, meanwhile, the fermentation titer of daptomycin in shake flasks was increased by 10% (**Figure 3A**), indicating that deletion of byproduct pathway may reduce the metabolic burden and facilitate the production of target daptomycin.

## Improving Daptomycin Production by Duplication of *dpt* Biosynthetic Gene Cluster

Increasing the copy number of BGC is an effective way to increase the yield of natural products (Li et al., 2021c). A previous study showed that daptomycin BGC with only 65 kb length was enough to be heterologously expressed in daptomycin non-producing strain *S. coelicolor* M511 and successfully led to the production of 28.9 mg/l daptomycin (Choi et al., 2019). pBeloBac*:*dpt* is a plasmid containing a large size of daptomycin BGC of 128 kb. We first reduced its size by removing redundant genes (spanning about 60 kb) to decrease the metabolic burden and increase the efficiency of conjugation into hosts. We tried to simplify the plasmid by LCHR. The *dptR1* gene locates downstream in daptomycin BGC. A previous study showed that daptomycin production decreased no matter what *dptR1* was deleted or overexpressed (Yu et al., 2020). Therefore, the simplification of the downstream gene sequence started from *dptR1* and, meanwhile, integrase φ31 and AMP-resistant gene were introduced by cassette attP-φ31 integrase-AMP. The upstream *dptP* gene of the *dpt* BGC is benefit to daptomycin resistance (Zhang et al., 2020). Therefore, we select a suitable position (near gene *orf6577*) upstream of the *dptP* gene as the starting position to simplify the upstream gene sequence. Linear recombination fragment—SmR—was used to simplify the upstream region of the daptomycin BGC. Finally, the resulting plasmid 702DIAAS was introduced into L2796 to get L2797. The production of daptomycin in L2797 reached 105 mg/l, an increase of 40%, compared with the parental strain L2796 (**Figure 3A**).

## Fermentation Process Engineering to Enhance Daptomycin Biosynthesis

To test the ability of L2797 to produce daptomycin in the fermenter, we conducted fed-batch fermentation in a 15-L fermenter with 9 L working volume. The production of daptomycin reached 635 mg/l. However, throughout the fermentation process, although the stirring speed has been gradually increased to a maximum speed of 450 rpm, the dissolved oxygen value still often touched around 0. Enough oxygen supply is critical for energy metabolism, primary metabolite, and secondary metabolite production. Even though higher stirring speed may ensure sufficient oxygen, the filamentous differentiation of some *Streptomyces*, such as *S. roseosporus*, will be damaged under high stirring shear force, which is pivotal for the secondary metabolite production with high titer (Ferraiuolo et al., 2021). Hence, we try to express heterologous *Vitreoscilla hemoglobin* (VHb) protein in *S. roseosporus* to enhance the utilization efficiency of oxygen under low dissolved oxygen conditions to improve daptomycin

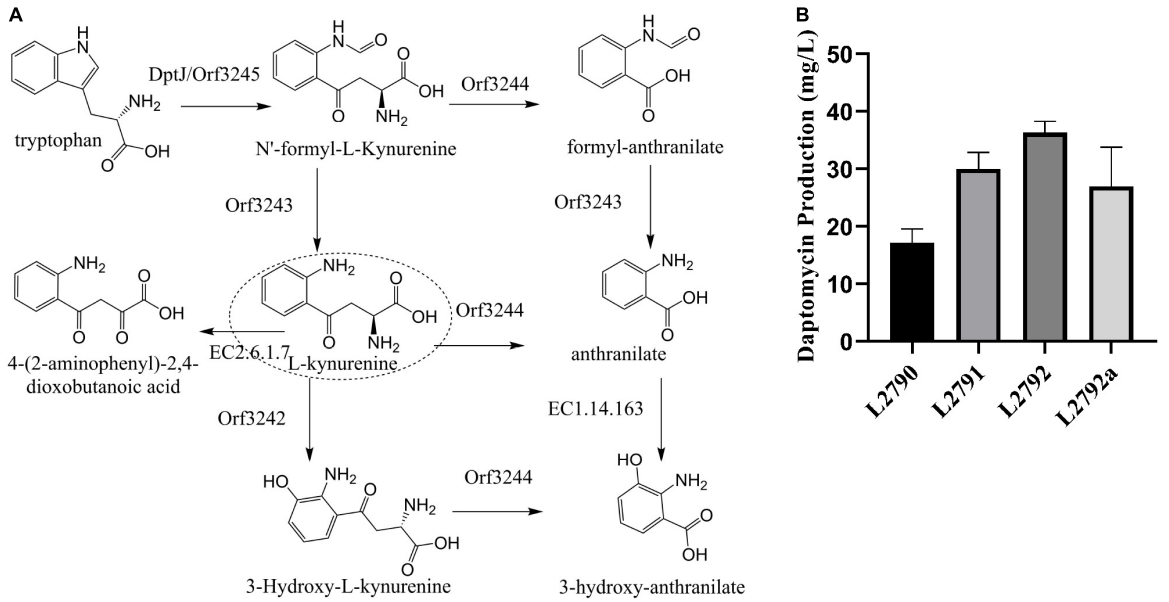



**FIGURE 2 |** Improving daptomycin production by enhancing Kyn precursor supply. **(A)** The putative L-Kyn metabolic network in strain L2790 according to the RAST Server. DptJ/Orf3245, tryptophan-2,3-dioxygenase; Orf3244, L-kynureninase; Orf3243, N′-formyl-L-kynurenine formidase; Orf3242, oxidoreductase; EC 2.6.1.7, a putative kynureninase aminotransferase. **(B)** The daptomycin production of different strains obtained in a stepwise manner. The culture samples were collected at 144 h and daptomycin production was analyzed by high-performance liquid chromatography (HPLC). The fermentations were performed in triplicate. L2790, initial strain; L2791, the resultant strain by deleting *orf3244* in L2790; L2792, the resultant strain by deleting *orf3242* in L2791; L2792a, the resultant strain by co-overexpressing *dptJ*, *orf3245*, *orf3243* driven by *ermEp** in L2792.

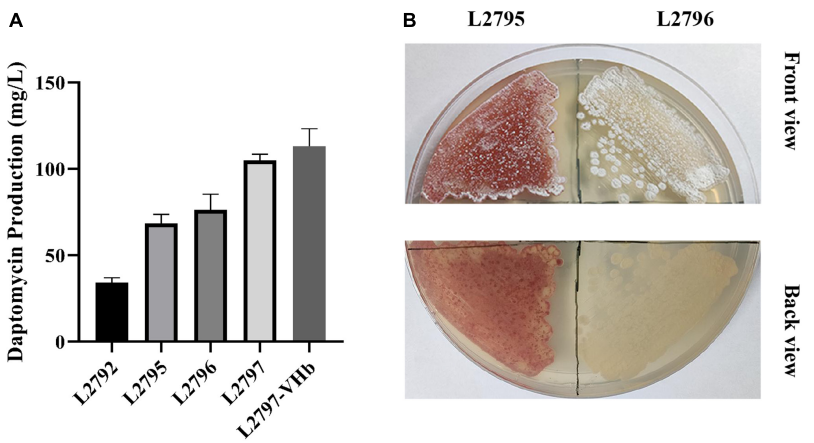



**FIGURE 3 |** Improving daptomycin production and quality by regulatory pathway reconstruction, pigment removal, multicopy BGC, and heterologous expression of VHb in a stepwise manner. **(A)** The daptomycin production of different strains obtained in a stepwise manner. The culture samples were collected at 144 h and daptomycin production was analyzed by HPLC. The cultures were performed in triplicate. L2792, the resultant strain by deleting *orf3244* and *orf3242* in the initial strain L2790; L2795, the resultant strain by deleting gene *phaR and arpA* in L2792; L2796, the resultant strain by deleting gene *orf3265 and orf3266* in L2795; L2797, the resultant strain by introducing the extra copy of daptomycin BGC into L2796; L2797-VHb, the resultant strain by introducing a *vgb2* gene into L2797. **(B)** Phenotype of L2795 and L2796 cultured on R5 medium.

production. According to the results of the CO-difference spectrum absorbance assay, an obvious absorption peak at 420 nm was observed in strain L2797-VHb compared with L2797 (**Figure 4A**). These results showed that the recombinant strain L2797-VHb expressed functional VHb successfully. The growth curve indicated that L2797-VHb exhibited better oxygen utilization efficiency than L2797 under simulated low dissolved oxygen state (**Supplementary Figure 6**). As we expected, the daptomycin titer of L2797-VHb was further increased to 113 mg/l under normal conditions described previously in shake flasks (**Figure 3A**). In the fed-batch fermentation of 15 L fermenter, the ability of the resultant strain L2797-VHb to produce daptomycin reached as high as 786 mg/l, an increase of 24%, compared with the parent strain L2797 (**Figure 4B**).

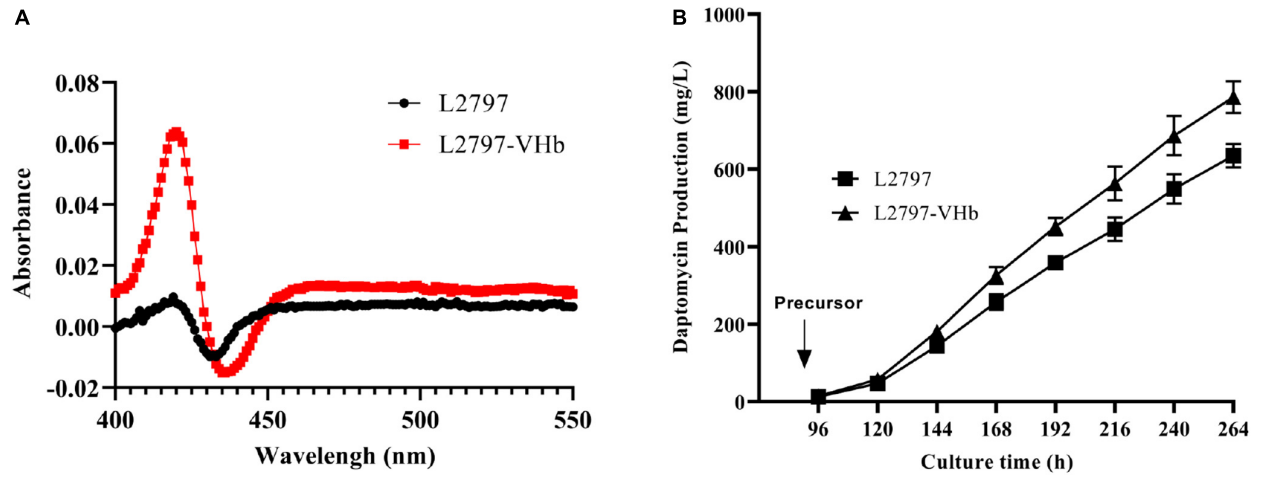



**FIGURE 4 |** Improving daptomycin production by fermentation process engineering. **(A)** Carbon monoxide (CO)-difference spectra analysis of L2797-VHb and L2797. **(B)** The daptomycin production curves of L2797-VHb and L2797 during the fermentation process in the 15-L fermenter. The fermentation samples were collected every 24 h from 96 to 264 h and the daptomycin production was analyzed by HPLC. The fermentation for each strain was performed in triplicate. Addition of precursor started from 96 h.

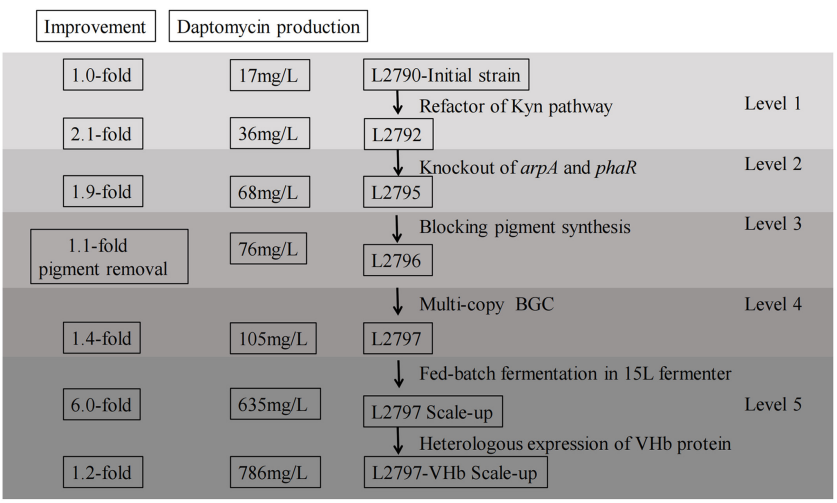



**FIGURE 5 |** Diagram summarizing multilevel metabolic engineering strategies implemented to daptomycin production and quality. Counting from the right, the first column, the level number for each metabolic engineering approach; the second column, the specific implementation content for each step of combinatorial metabolic engineering strategies from the beginning to the end; the third column, daptomycin production caused by each metabolic engineering approach; the last column, the corresponding fold changes caused by each metabolic engineering approach.

## DISCUSSION

Various metabolic engineering strategies have been developed to improve the production of natural products (Bu et al., 2021). As one of the natural products, commercially available daptomycin (e.g., Cubicin) is widely used in anti-inflammation, skin infection, and bacteremia, etc., bringing huge market demands for this antibiotic. The global market size of daptomycin is expected to be 5.1 billion USD and is expected to register a growth of 4.4% from 2020 to 2027 (Market Research Future, 2021). Although many efforts have been made to enhance the yield of daptomycin for cost reduction, such as refactoring regulatory pathway, enhancing precursor supply, or traditional mutagenesis, most of these strategies are clearly targeted at a single factor, which cannot improve the production of daptomycin at global or systematic level (**Supplementary Table 4**). The daptomycin fermentation titer remains deficient or the red pigment is not removed. Combined with our previous studies on daptomycin biosynthesis (Mao et al., 2015; Luo et al., 2018b), we designed a combinatorial engineering strategy to further advance the biosynthetic efficiency of daptomycin.

With the help of genomic analysis, we first built a Kyn metabolic network model. According to the model, we improved daptomycin production to 36.3 mg/l with an increase of 111% by knocking out kynurenine-related genes, such as *orf3344* and *orf3242* (**Figure 2B**). However, co-overexpression of other related

genes, such as *dptJ*, *orf3245*, and *orf3243,* made daptomycin production decrease by 26% (**Figure 2B**) and this may be due to the excessive accumulation of some intermediates, which affect the further synthesis of L-Kyn (Ji et al., 2021) or are toxic to the cell itself, thereby affecting the accumulation of daptomycin. The underlying mechanisms may be complicated. Like Kyn, decanoic acid is also an important precursor substance in the biosynthesis of daptomycin. Excess decanoic acid is toxic to cells (Lee et al., 2016). Therefore, during the fed-batch fermentation process in a 15 L fermenter, we used a constant rate to add decanoic acid to keep its concentration at a low level in the culture. In addition, during the shake-flask fermentation process, we found that the growth of the bacteria was stagnant and it was difficult to enter the secondary metabolism if we added decanoic acid too early, such as cultivating at 24 h (data not shown). Therefore, we speculate that enzymes involved in the metabolism of decanoic acid may only gradually begin to be expressed during the transition from primary metabolism to secondary metabolism. If we compare and analyze the transcriptome before and after the addition of capric acid, we may be able to find which enzymes are mainly involved in its metabolism and transport and improve the tolerance and utilization efficiency of decanoic acid by modifying these enzymes.

On the basis of refactoring the Kyn pathway, we deleted two negative regulators, namely, ArpA and PhaR, at the same time, which made daptomycin titer increase by 88% and reach 68 mg/l (**Figure 3A**). Although knocking out *arpA* or *phaR* can increase the production of daptomycin, they played different roles in the growth and differentiation of hyphae (Mao et al., 2015; Luo et al., 2018b). Whether there is a cross-talking effect mechanism from both the regulatory genes on mycelial growth and daptomycin accumulation, it could be further investigated through transcriptome analysis. There may exist a sophisticated and cascaded regulatory network, which can facilitate us to fine-tune it for further improvement of daptomycin production.

Red pigment is a visible impurity in the production of daptomycin. Usually, many pigment biosynthetic pathways belong to type II PKSs in *Streptomyces*. Short-chain acyl-CoAs, such as malonyl-CoA, methylmalonyl-CoA, and ethylmalonyl-CoA, are important precursors of many type II PKS pathways (Lin et al., 2020; Bednarz et al., 2021). Acetyl-CoAs are the main elements for the formation of short-chain acyl-CoA precursor pool. As we all know, acetyl-CoAs also participate in primary metabolisms, especially the tricarboxylic acid (TCA) cycle, for the supply of energy responsible for cell growth and primary/secondary metabolisms, such as gene expression, amino acid synthesis, and natural product biosynthesis. Here, amino acid synthesis is pivotal for daptomycin biosynthesis because many amino acids are direct precursors of daptomycin. Therefore, disruption of pigment biosynthesis can block the consumption of short-chain acyl-CoA precursors and enrich more acetyl-CoAs toward primary metabolism, thereby supplying more energy or precursors (especially amino acids) for daptomycin production. Besides, pigment production will negatively affect the purification of target metabolites. For example, in the production of amphotericin B, a ceramic membrane is usually used to remove pigments (Kai et al., 2020). During the extraction of daptomycin, we also observed that a large number of red pigments are accompanied. Furthermore, the biological activity of red pigments also remains unknown and it is required to remove them for the purity of commercial daptomycin. We next removed it in strain L2795 obtained by enhancing the Kyn pathway and deleting two negative regulators. The daptomycin titer of the resulting strain L2796 reached 76 mg/l with an increase of 10% and its red pigment was removed thoroughly (**Figures 3A,B**). In addition to the red pigment, the production process of daptomycin is also accompanied by other byproducts, such as homologs of daptomycin. Although deleting branched-chain α-keto acid dehydrogenase (BKD) enzyme complex could remove homologs of daptomycin containing branched-chain fatty acids (BCFAs), it introduced new homologs of daptomycin with straight-chain fatty acids (Ji et al., 2021). Therefore, in future studies, other methods should be considered to completely solve the problem of daptomycin homologs, such as the directed evolution of DptE, an acyl AMP ligase, on improving its substrate selectivity. Besides, according to the anti-SMASH annotation of genome, there are other 27 non-target BGCs, except daptomycin BGC, which may compete precursor and/or energy flux with daptomycin pathway and result in metabolic burden. Transcriptome can be conducted to determine the active BGCs and then we can delete them to enrich metabolic flux toward the daptomycin pathway and, meanwhile, simplify metabolic background for the improvement of daptomycin production in yield and quality.

Finally, we further increased daptomycin to 105 mg/l with an increase of 40% by introducing an extra copy of daptomycin BGC, which implied that multicopy of BGC can efficiently enhance the biosynthesis of daptomycin (**Figure 3A**). According to previous studies, the overall transcriptional level of dpt BGC is too low to meet the efficient production of daptomycin in *S. roseosporus* (Ji et al., 2021). Recently, Lu's team developed an integrase-mediated multicopy system used for the integration of 3–5 copies BGCs into *Streptomyces* genomes. For example, the start and engineered 5-oxomilbemycin-producing strains, such as KF200, KF201, KF202, and KF203, harboring 1–4 copies of 5-oxomilbemycin BGCs by advanced Multiplex Site-specific Genome Engineering (aMSGE) system, respectively, can make a stepwise increase in 5-oxomilbemycin production (from 2,228, 4,415, 5,592 to 6,368 mg/l; Li et al., 2019). Previous studies also suggested that tandem amplification of BGCs could increase the copy number (4–12 copies), thereby enhancing the production of natural products, such as actinorhodin, bleomycin, and validamycin A (Murakami et al., 2011; Zhou et al., 2014, Li et al., 2021a,b). These studies strongly imply that increasing the cluster copies can further enhance the production of target metabolites. From the perspective of gene expression, multicopy is actually equivalent to overexpressing every gene in the BGC and the former is a more efficient strategy than the latter because upregulating all the genes needs to design different promoters to balance their expressions for avoiding toxic intermediates or incomplete conversion. Of course, the yields of metabolites may depend more on the tolerance of engineered strains to specialized metabolites. In the future, we will first investigate the resistance level of *S. roseosporus* against daptomycin and then determine the possible copy number of *dpt* BGC or upregulate the expression of transporters to increase

the tolerance of engineered strains for the maximized production of daptomycin. We can also effectively combine the above engineering strategies to obtain a high-performance cell factory with higher daptomycin production in future studies.

Here, we make a comparison among different engineering strategies for the improvement of daptomycin production (**Supplementary Table 4**). We can see that our strategy not only remove red pigment but also efficiently enhance the production of daptomycin, which reached as high as 786 mg/l in the 15 L fermenter. Although the addition of sodium decanoate can substantially increase the yield of daptomycin sharply, it is a surfactant. In our shake-flask experiments, adding sodium decanoate to the fermentation broth resulted in a lot of foam, which may cause the fermentation broth to escape and become contaminated during large-scale fermentation. Therefore, we chose decanoic acid as the feed precursor instead of sodium decanoate. However, in the future, we can try to explore a new process that can greatly increase the production of daptomycin through sodium caprate and eliminate the generation of foam at the same time.

In conclusion, this study revealed an integrated metabolic engineering strategy for daptomycin production in *S. roseosporus* for the first time. Step by step, the combinatorial approach was employed to create daptomycin overproducing strains (**Figure 5**). Daptomycin production was increased by roughly 5-fold using rational design and refactoring at the molecular level and an additional about 7-fold using fermentation process engineering. This study not only boosts daptomycin productivity but also enhances its quality with low byproducts. Importantly, the engineering strategies used here are expected to significantly reduce expenditure for daptomycin industrial manufacturing.

## DATA AVAILABILITY STATEMENT

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

## AUTHOR CONTRIBUTIONS

Z-YL contributed to conceptualization, methodology, validation, investigation, data curation, formal analysis, visualization, writing—original draft, and writing, reviewing and editing the manuscript. Q-TB and J-LF contributed to methodology, data curation, formal analysis, software, 15 L fermentation, and writing, reviewing and editing the manuscript. C-YZ and W-FX contributed to methodology, data curation, and 15 L fermentation. LM, W-LG, and X-AC contributed to methodology, data curation, software, and writing, reviewing and editing the manuscript. Y-QL contributed to conceptualization, project administration, funding acquisition, supervision, and writing, reviewing, and editing the manuscript. All authors read and approved final version of the manuscript.

## FUNDING

This study was supported by the National Key R&D Program of China (grant number 2019YFA09005400) and the National Natural Science Foundation of China (grant numbers 31730002 and 2170057).

## ACKNOWLEDGMENTS

We thank Hua-Rong Tan, Wen Liu, and Cun-Jiang Song for providing us with relevant plasmids.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.872397/full#supplementary-material

## REFERENCES

Akins, R. L., and Rybak, M. J. (2001). Bactericidal activities of two daptomycin regimens against clinical strains of glycopeptide intermediate-resistant *Staphylococcus aureus*, vancomycin-resistant *Enterococcus faecium*, and methicillin-resistant *Staphylococcus aureus* isolates in an in vitro pharmacodynamic model with simulated endocardial vegetations. *Antimicrob. Agents Chemother.* 45, 454–459. doi: 10.1128/AAC.45.2.454-459.2001

Aziz, R. K., Bartels, D., Best, A. A., Dejongh, M., Disz, T., Edwards, R. A., et al. (2008). The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9:75–75. doi: 10.1186/1471-2164-9-75

Bednarz, B., Millan-Oropeza, A., Kotowska, M., Wiat, M., and Pawlik, K. (2021). Coelimycin synthesis activatory proteins are key regulators of specialized metabolism and precursor flux in *Streptomyces coelicolor* A3(2). *Front. Microbiol.* 12:616050. doi: 10.3389/fmicb.2021.616050

Bu, Q. T., Li, Y. P., Xie, H., Li, J. F., and Li, Y. Q. (2021). Rational engineering strategies for achieving high-yield, high-quality and high-stability of natural product production in actinomycetes. *Metab. Eng.* 67, 198–215. doi: 10.1016/j.ymben.2021.06.003

Chari, R., Yeo, N. C., Chavez, A., and Church, G. M. (2017). sgRNA scorer 2.0: a species-independent model to predict CRISPR/Cas9 activity. *ACS Synth. Biol.* 902:6b00343. doi: 10.1021/acssynbio.6b00343

Choi, S., Nah, H. J., Choi, S., and Kim, E. S. (2019). Heterologous expression of daptomycin biosynthetic gene cluster via *Streptomyces* artificial chromosome vector system. *J. Microbiol. Biotechnol.* 29, 1931–1937. doi: 10.4014/jmb.1909.09022

Debono, M., Abbott, B. J., Molloy, R. M., Fukuda, D. S., Hunt, A. H., Daupert, V. M., et al. (1988). Enzymatic and chemical modifications of lipopeptide antibiotic A21978C-the synthesis and evaluation of daptomycin (Ly146032). *J. Antibiot.* 41, 1093–1105. doi: 10.7164/antibiotics.41.1093

Du, D., Wang, L., Tian, Y., Liu, H., Tan, H., and Niu, G. (2015). Genome engineering and direct cloning of antibiotic gene clusters via phage varphiBT1 integrase-mediated site-specific recombination in *Streptomyces*. *Sci. Rep.* 5:8740. doi: 10.1038/srep08740

Ferraiuolo, S. B., Cammarota, M., Schiraldi, C., and Restaino, O. F. (2021). *Streptomycetes* as platform for biotechnological production processes of drugs. *Appl. Microbiol. Biotechnol.* 105:2. doi: 10.1007/s00253-020-11064-2

Horng, Y. T., Chang, K. C., Chien, C. C., Wei, Y. H., Sun, Y. M., and Soo, P. C. (2010). Enhanced polyhydroxybutyrate (PHB) production via the coexpressed *phaCAB* and *vgb* genes controlled by arabinose P promoter in *Escherichia coli*. *Lett. Appl. Microbiol.* 50, 158–167. doi: 10.1111/j.1472-765X.2009.02772.x

Ji, C. H., Kim, H., Je, H. W., Kwon, H., Lee, D., and Kang, H. S. (2021). Top-down synthetic biology approach for titer improvement of clinically important

antibiotic daptomycin in *Streptomyces roseosporus*. *Metab. Eng.* 69, 40–49. doi: 10.1016/j.ymben.2021.10.013

Kai, H., Bo, Z., Zysa, B., Xue, C., Zqla, B., and Ygza, B. (2020). Enhanced amphotericin B production by genetically engineered *Streptomyces nodosus*. *Microbiol. Res.* 242:126623. doi: 10.1016/j.micres.2020.126623

Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F., Hopwood, D. A., Charter, K., et al. (2000). *Practical Streptomyces Genetics*. Norwich: John Innes Foundation.

Lee, S. K., Kim, H. R., Jin, Y. Y., Yang, S. H., and Suh, J. W. (2016). Improvement of daptomycin production via increased resistance to decanoic acid in *Streptomyces roseosporus*. *J. Biosci. Bioeng.* 122, 427–433. doi: 10.1016/j.jbiosc.2016.03.026

Li, H. J., Zhang, D. H., Yue, T. H., Jiang, L. X., Yu, X., Zhao, P., et al. (2016). Improved polysaccharide production in a submerged culture of *Ganoderma lucidum* by the heterologous expression of *Vitreoscilla* hemoglobin gene. *J. Biotechnol.* 217, 132–137. doi: 10.1016/j.jbiotec.2015.11.011

Li, Y. P., Bu, Q. T., Li, J. F., Xie, H., Su, Y. T., Du, Y. L., et al. (2021c). Genome-based rational engineering of *Actinoplanes deccanensis* for improving fidaxomicin production and genetic stability. *Bioresour. Technol.* 330:124982. doi: 10.1016/j.biortech.2021.124982

Li, H., Gao, W., Cui, Y., Pan, Y., and Liu, G. (2021a). Remarkable enhancement of bleomycin production through precise amplification of its biosynthetic gene cluster in *Streptomyces verticillus*. *Sci. China Life Sci.* 6, 1–9. doi: 10.1007/s11427-021-1998-8

Li, H., Pan, Y., and Liu, G. (2021b). Multiplying the heterologous production of spinosad through tandem amplification of its biosynthetic gene cluster in *Streptomyces coelicolor*. *Microb. Biotechnol.* 2021:13965. doi: 10.1111/1751-7915.13965

Li, L., Wei, K., Liu, X., Wu, Y., Zheng, G., Chen, S., et al. (2019). aMSGE: advanced multiplex site-specific genome engineering with orthogonal modular recombinases in actinomycetes. *Metab. Eng.* 2019:1. doi: 10.1016/j.ymben.2018.12.001

Liao, G., Wang, L., Liu, Q., Guan, F., Huang, Y., and Hu, C. (2013). Manipulation of kynurenine pathway for enhanced daptomycin production in *Streptomyces roseosporus*. *Biotechnol. Prog.* 29, 847–852. doi: 10.1002/btpr.1740

Lin, C. Y., Zhang, Y., Wu, J. H., Xie, R. H., and Zhao, G. R. (2020). Regulatory patterns of Crp on monensin biosynthesis in *Streptomyces cinnamonensis*. *Microorganisms* 8:271. doi: 10.3390/microorganisms8020271

Lombo, F., Velasco, A., Castro, A., De La Calle, F., Brana, A. F., Sanchez-Puelles, J. M., et al. (2006). Deciphering the biosynthesis pathway of the antitumor thiocoraline from a marine actinomycete and its expression in two streptomyces species. *Chembiochem* 7, 366–376. doi: 10.1002/cbic.200500325

Luo, S., Chen, X. A., Mao, X. M., and Li, Y. Q. (2018b). Transposon-based identification of a negative regulator for the antibiotic hyper-production in *Streptomyces*. *Appl. Microbiol. Biotechnol.* 102, 6581–6592. doi: 10.1007/s00253-018-9103-5

Luo, S., Chen, X. A., Mao, X. M., and Li, Y. Q. (2018a). Regulatory and biosynthetic effects of the bkd gene clusters on the production of daptomycin and its analogs A21978C1-3. *J. Ind. Microbiol. Biotechnol.* 45, 271–279. doi: 10.1007/s10295-018-2011-y

Mao, X. M., Luo, S., Zhou, R. C., Wang, F., Yu, P., Sun, N., et al. (2015). Transcriptional regulation of the daptomycin gene cluster in *Streptomyces roseosporus* by an autoregulator, AtrA. *J. Biol. Chem.* 290, 7992–8001. doi: 10.1074/jbc.M114.608273

Market Research Future (2021). *Daptomycin Market: By Indication (Complicated Skin Structure Infections, Bacteremia, others), Age Group (Pediatric, Adult), Strength (350mg, 500mg) and Region (Americas, Europe, Asia-Pacific, Middle East & Africa) - Forecast to 2027*. Pune: Market Research Future.

Miao, V., Coeffet-Legal, M. F., Brian, P., Brost, R., Penn, J., Whiting, A., et al. (2005). Daptomycin biosynthesis in *Streptomyces roseosporus*: cloning and analysis of the gene cluster and revision of peptide stereochemistry. *Microbiology* 151, 1507–1523. doi: 10.1099/mic.0.27757-0

Mironczuk, A. M., Kosiorowska, K. E., Biegalska, A., Rakicka-Pustulka, M., Szczepanczyk, M., and Dobrowolski, A. (2019). Heterologous overexpression of bacterial hemoglobin VHb improves erythritol biosynthesis by yeast *Yarrowia lipolytica*. *Microb. Cell Fact.* 18:176. doi: 10.1186/s12934-019-1231-9

Murakami, T., Burian, J., Yanai, K., and Thompson, C. J. (2011). A system for the targeted amplification of bacterial gene clusters multiplies antibiotic yield in *Streptomyces coelicolor*. *Proc. Natl. Acad. Sci.* 108, 16020–16025. doi: 10.1073/pnas.1108124108

Staunton, J., and Weissman, K. J. (2001). Polyketide biosynthesis: a millennium review. *Nat. Prod. Rep.* 18, 380–416. doi: 10.1039/a909079g

Wang, F., Ren, N. N., Luo, S., Chen, X. X., Mao, X. M., and Li, Y. Q. (2014). DptR2, a DeoR-type auto-regulator, is required for daptomycin production in *Streptomyces roseosporus*. *Gene* 544, 208–215. doi: 10.1016/j.gene.2014.04.044

Wang, H. L., Li, Z., Jia, R. N., Hou, Y., Yin, J., Bian, X. Y., et al. (2016). RecET direct cloning and Red alpha beta recombineering of biosynthetic gene clusters, large operons or single genes for heterologous expression. *Nat. Protoc.* 11, 1175–1190. doi: 10.1038/nprot.2016.054

Wang, Y., Ling, C., Chen, Y., Jiang, X., and Chen, G. Q. (2019). Microbial engineering for easy downstream processing. *Biotechnol. Adv.* 37, 107365. doi: 10.1016/j.biotechadv.2019.03.004

Wentzel, A., Bruheim, P., Overby, A., Jakobsen, O. M., Sletta, H., Omara, W. A., et al. (2012). Optimized submerged batch fermentation strategy for systems scale studies of metabolic switching in *Streptomyces coelicolor* A3(2). *BMC Syst. Biol.* 6:59. doi: 10.1186/1752-0509-6-59

Wittmann, M., Linne, U., Pohlmann, V., and Marahiel, M. A. (2008). Role of DptE and DptF in the lipidation reaction of daptomycin. *FEBS J.* 275, 5343–5354. doi: 10.1111/j.1742-4658.2008.06664.x

Yu, G., Hui, M., Li, R., and Zhang, S. (2020). Pleiotropic regulation of daptomycin synthesis by DptR1, a LuxR family transcriptional regulator. *World J. Microbiol. Biotechnol.* 36:173. doi: 10.1007/s11274-019-2771-1

Zhang, D., Wang, X. X., Ye, Y., He, Y., He, F. Q., Tian, Y. Q., et al. (2020). Label-free proteomic dissection on *dptP*-deletion mutant uncovers dptP involvement in strain growth and daptomycin tolerance of *Streptomyces roseosporus*. *Microb. Biotechnol.* 2020:13736. doi: 10.1111/1751-7915.13736

Zhang, Q., Chen, Q., Zhuang, S., Chen, Z., Wen, Y., and Li, J. (2015). A MarR family transcriptional regulator, DptR3, activates daptomycin biosynthesis and morphological differentiation in *Streptomyces roseosporus*. *Appl. Environ. Microbiol.* 81, 3753–3765. doi: 10.1128/AEM.00057-15

Zhou, T. C., Kim, B. G., and Zhong, J. J. (2014). Enhanced production of validamycin A in *Streptomyces hygroscopicus* 5008 by engineering validamycin biosynthetic gene cluster. *Appl. Microbiol. Biotechnol.* 98, 7911–7922. doi: 10.1007/s00253-014-5943-9

Zhu, H., Sun, S., and Zhang, S. (2011). Enhanced production of total flavones and exopolysaccharides via *Vitreoscilla* hemoglobin biosynthesis in *Phellinus igniarius*. *Bioresour. Technol.* 102, 1747–1751. doi: 10.1016/j.biortech.2010.08.085

**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.

*Copyright © 2022 Lyu, Bu, Fang, Zhu, Xu, Ma, Gao, Chen and Li. 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.*

frontiers | Frontiers in Microbiology

ORIGINAL RESEARCH
published: 10 May 2022
doi: 10.3389/fmicb.2022.902990

# Activation and Characterization of Lanthomicins A–C by Promoter Engineering in *Streptomyces chattanoogensis* L10

*Xiao-Fang Liu*[1,2], *Jun-Xiao Wang*[1], *Xin-Ai Chen*[1,2], *Yu Liu*[3] *and Yong-Quan Li*[1,2]*

[1] First Affiliated Hospital and Institute of Pharmaceutical Biotechnology, Zhejiang University School of Medicine, Hangzhou, China, [2] Zhejiang Provincial Key Laboratory for Microbiol Biochemistry and Metabolic Engineering, Hangzhou, China, [3] College of Life Science, Zhejiang University, Hangzhou, China

The emergence of drug resistance highlights the importance of new drug discovery. Microbial secondary metabolites encoded in biosynthetic gene clusters (BGCs) are a prolific source of drugs, whereas most of these BGCs are cryptic. Thus, taking strategies to activate these cryptic BGCs is of great importance for potential drug discovery. In this work, three novel pentangular polyphenols lanthomicin A–C were identified by activating a cryptic aromatic polyketide BGC through promoter engineering combined with optimization of fermentation conditions. We further confirmed the involvement of lanthomicin (*ltm*) BGC in biosynthesis by CRISPR-Cpf1-assisted gene editing. Based on functional analysis of homologous genes, a putative biosynthetic pathway was proposed for the three lanthomicins. Particularly, lanthomicin A showed antiproliferative activity with $IC_{50}$ 0.17 μM for lung cancer cell line A-549. The discovery of lanthomicins brings new members to the pentangular polyphenol subclade of aromatic polyketide and demonstrates the potential of *Streptomyces* as a source for drug discovery.

**Keywords: cryptic gene cluster, lanthomicin, antiproliferative activity, CRISPR-Cpf1, *Streptomyces chattanoogensis* L10**

OPEN ACCESS

***Edited by:***
*Sailesh Malla,*
*Chr. Hansen, Denmark*

***Reviewed by:***
*Guoqing Niu,*
*Southwest University, China*
*Gang Liu,*
*Institute of Microbiology (CAS), China*

***Correspondence:**
*Yong-Quan Li*
*lyq@zju.edu.cn*

***Specialty section:***
*This article was submitted to Microbiotechnology, a section of the journal Frontiers in Microbiology*

***Received:*** *23 March 2022*
***Accepted:*** *11 April 2022*
***Published:*** *10 May 2022*

***Citation:***
*Liu X-F, Wang J-X, Chen X-A, Liu Y and Li Y-Q (2022) Activation and Characterization of Lanthomicins A–C by Promoter Engineering in Streptomyces chattanoogensis L10. Front. Microbiol. 13:902990. doi: 10.3389/fmicb.2022.902990*

## INTRODUCTION

*Streptomyces* genus has been one of the richest sources of bioactive natural products since the early 1950s (Baral et al., 2018). Indeed, a vast majority of compounds originating from this genus are clinically used in medicine, such as daptomycin (Vilhena and Bettencourt, 2012), doxorubicin (Arcamone et al., 1969), and FK506 (Rath, 2013). However, the drug discovery reached a bottleneck based on the classic bioactivity-guided approaches, while the drug resistance became a major threat to modern healthcare (Strachan and Davies, 2017). One promising approach to address this question comes with the development of genome sequencing technology. Genome mining of *Streptomyces* has revealed that they have far greater potential to produce specialized metabolites than have ever been estimated, based on the plethora of BGCs identified (Kalkreuter et al., 2020). Since it is by now well-established that most natural product BGCs are silent or poorly expressed under standard laboratory conditions that the cognate products do not accumulate to detectable levels (Nett et al., 2009). Therefore, the key question remaining to answer is how to access these hidden potentials.

In the light of huge sequence data in the genomic era, several main strategies have been developed in an attempt to harness this tremendous chemical diversity. These include promoter engineering (Li et al., 2017), diversification of culture conditions (Zhang et al., 2014), ribosome engineering (Wang et al., 2008; Thong et al., 2016), manipulation of regulators (Chen et al., 2017), and others. Among them, promoter engineering has been proved as a powerful and efficient method to activate or improve the biosynthesis of interest products (Li et al., 2015, 2017). *Streptomyces chattanoogensis* L10 is an industrial producing strain of natamycin, which is synthesized by type I polyketide synthases and widely used as an antifungal agent in both drug therapy and the food industry (Aparicio et al., 2016; Meena et al., 2021). Researchers have also activated the cryptic BGCs of chattamycins (Zhou et al., 2015) and anthrachamycin (Li et al., 2019) in *S. chattanoogensis* L10 by overexpressing regulatory gene and ribosome engineering, respectively. Both of them fell into the angucycline clade of aromatic polyketides. It seems that *S. chattanoogensis* L10 is inclined to produce polyketides, suggesting the potential to be developed as chassis for expressing exogenous BGCs (Bu et al., 2019).

In this study, we identified and activated a cryptic BGC of lanthomicin (*ltm*) by promoter engineering combined with optimization of fermentation conditions. Three newly synthesized lanthomicins were purified and their chemical structures were elucidated. We found that the lanthomicins represent new members of the pentangular polyphenols family. Cytotoxicity assay revealed that lanthomicin A possesses potential antiproliferative activity.

# MATERIALS AND METHODS

## Strains and Media

The strain *S. chattanoogensis* L10 was used as a starting strain for genome mining. *Escherichia coli* DH5α was used as a general plasmid-cloning host. *E. coli* ET12567/pUZ8002 (Kieser et al., 2000) was used to introduce plasmids into *Streptomyces* by interspecies conjugation. *S. chattanoogensis* L10 and recombinant strains were grown at 30°C on YMG solid medium (yeast extract 0.4%, malt extract 1%, glucose 0.4%, agar 2%, and $CaCO_3$ 0.2%) for genetic screen and spore preparation (**Supplementary Table 1**). A total of 3% tryptic soy broth was used for vegetative growth of *Streptomyces* for genomic DNA extraction. MS solid medium (mannitol 2%, soybean flour 2%, and agar 2%) was used for exoconjugants growth. YEME liquid medium (yeast extract 0.3%, malt extract 0.3%, tryptone 0.5%, and glucose 4%) was used to produce lanthomicins. *E. coli* strains were cultured in a Luria-Bertani medium. When required, media were supplemented with antibiotics at the following final concentrations: chloramphenicol (25 μg/ml), nalidixic acid (25 μg/ml), and kanamycin (50 μg/ml). Spectinomycin was used in a final concentration of 50 μg/ml for *E. coli*, while 200 μg/ml for *Streptomyces*.

## General DNA Manipulation

Polymerase chain reaction (PCR) was performed with KOD FX DNA polymerase (TOYOBO, Japan) for high-fidelity cloning and with rTaq DNA polymerase (Takara) for colony screening according to the instructions of the manufacturers. Gene fragments were purified by a gel extraction kit (Magen Biotech Co.) and plasmids were recovered from a plasmid extraction kit (Shanghai Generay Biotech Co.). DNA concentration was measured by a Nanodrop Lite spectrophotometer (Thermo Scientific Co.). Genomic DNA isolation of *S. chattanoogensis* L10 and its derivatives was performed as described (Kieser et al., 2000). *Streptomyces* colony PCR was also performed for rapid genotype validation: mycelium from single clones was scraped with a sterilized toothpick and ultrasonication in FTPS (10 mM Tris-HCl; 0.5 mM EDTA; and 0.1 mM KCl) solution, 1 μl of the solution was then used as a template for reactions.

## Plasmid Construction

Primers used in this study were listed in **Supplementary Table 3**. Genomic DNA isolated from the L10 strain was used as a template for PCR amplification. The fidelity of amplified fragments was confirmed by DNA sequencing. pSET152 and pKC1139 were digested with *Sac*I. The spectinomycin resistance gene was cloned by PCR with a primer pair spec-F/R using plasmid pIJ778 (Gust et al., 2003) as a template and then inserted into the above-digested vectors by seamless cloning, generating pSET152-spec and pKC1139-spec. For the purpose of constructing a universal vector to overexpress genes, pSET152-spec was digested with *Xba*I and *Eco*RV. A 97 bp PCR fragment of the *kasO** promoter amplified with primers *kasO**p-F/R was inserted into pSET152-spec, affording plasmid pSET152-spec-*kasO**. Detailed procedures for plasmid construction were described in **Supplementary Methods**. All the plasmids used in this work were listed in **Supplementary Table 2**.

## High-Performance Liquid Chromatography (HPLC) and Liquid Chromatography-Mass Spectrometry (LC-MS) Analysis

Fermentation broth of *S. chattanoogensis* L10 and recombinant strains from YEME were mixed with two volumes of methanol. After centrifugation for 5 min at 14,000 rpm (Eppendorf Centrifuge 5424R), the supernatant was filtered for high-performance liquid chromatography (HPLC) analysis, which was performed on an Agilent Liquid Chromatograph 1,260 (Agilent, Inc) equipped with a 4.6 × 150 mm Extend-C18 column, with a linear gradient of 5–100% MeCN-$H_2O$ in 30 min followed by 100% MeCN for 5 min and 5% MeCN for another 5 min in a flow rate of 1 ml/min. LC-MS analysis was performed in an Agilent 1200 HPLC system (Agilent, Santa Clara, CA, USA) and a Termo Finnigan LCQDeca XP Max LC/MS system (Termo Finnigan, Waltham, MA, USA). Agilent Extend-C18 was used as the column, and $H_2O$ (0.1% formic acid) and acetonitrile were used as the mobile phases A and B with a linear gradient from 5% to 100% (v/v) B over 30 min.

## Cultivation, Extraction, and Isolation of Compounds

To accumulate enough compounds, L10 strains were cultivated in a 15 L fermenter at 30°C in YEME media for 7 days.

When the fermentation was stopped, the fermentation broth was first adjusted pH to be lower than 4.0 with 2M hydrochloric acid solution and extracted three times using an equal volume of ethyl acetate. After removing the solvent in vacuo, the extract was fractioned by silica gel chromatography. Then, the fractions containing target compounds were subjected to Sephadex LH-20 column chromatography eluted with pure methanol. Purified extract was further refined by semipreparative liquid chromatography on an Agilent Eclipse XDB-C18 column (5 μm, 9.4 × 250 mm) using acetonitrile and 0.1% HCOOH (v/v) as eluent, yielding 5.4 mg of **1**, 7 mg of **3**, and 11.4 mg of **4**.

## Antiproliferative Assay

The cytotoxicity of lanthomicins A, B, and C was evaluated using the lung cancer cell line A-549, the breast carcinoma cell line MCF-7, the liver cancer cell line HepG2, and the colon carcinoma cell line HCT-116. HCT-116 cells were grown in McCoy's 5A Modified Medium (Gibco) supplemented with 10% fetal bovine serum (FBS). A-549, MCF-7, and HepG2 cells were grown in a DMEM medium (Gibco) supplemented with 10% FBS. Cells in log phase growth were harvested by trypsinization. Trypsinized cells were seeded onto 96-well plates (20,000 cells/ml) and incubated overnight at 37°C in the presence of 5% $CO_2$. Cells were transferred to a fresh medium supplemented with lanthomicins A–C (in dimethyl sulfoxide (DMSO)) in different final concentrations and incubated for 72 h to detect the cell viability using CCK8-based colorimetric assays. Doxorubicin was used as a positive control and DMSO was used as a negative control. Cell viability was recorded based on the percent stain present in each well relative to no drug DMSO control wells.

# RESULTS

## Characterization of the Lanthomicin Gene Cluster From *S. chattanoogensis* L10

To get insight into the BGCs of *S. chattanoogensis* L10, we reanalyzed the completely sequenced genome (CGMCC 2644) using antiSMASH. It was found to possess 37 BGCs, namely, 8 polyketide synthases (PKSs), 2 nonribosomal peptide synthases (NRPSs), and one PKS-NRPS hybrid. A few terpene and lassopeptide BGCs were identified by the presence of corresponding synthase genes (**Supplementary Table 4**). Among the predicted BGCs, a type II PKS gene cluster (*ltm*, cluster 1) (**Figure 1C**) was shown to have about 50% homologies to known gene clusters of *xan* and *arx*, which are responsible for the biosynthesis of xantholipin and arixanthomycin, respectively (Zhang et al., 2012; Kang and Brady, 2014). Hence, we supposed the gene cluster *ltm* may be involved in the biosynthesis of pentangular polyphenols which belong to one subclade of the aromatic PKS family. Phylogenetic analysis also found that ketosynthase beta subunit ($KS_\beta$) LtmB from *ltm* gene cluster fell into the same subclade with a 26-carbon skeleton as that of *xan* and *arx* gene clusters (**Figure 1B**). The core skeleton of this subfamily was synthesized by minimal type II PKS (min-PKS), cyclases, and ketoreductases (KRs) in similar steps. Then the core structure was modified by various tailoring enzymes, which contributed to structural diversity. Examples of compounds from this subclade were listed in **Figure 1A** and distinctive structures were shown in color.

According to bioinformatic prediction, the *ltm* gene cluster consists of 34 genes (**Table 1**), including three FAD-dependent monooxygenases, among which LtmG3 showed considerable homology to XanO4 (70% identity) and Arx30 (64% identity). Phylogenetic analysis implied LtmG3 could be classified into an atypical Baeyer–Villiger monooxygenase branch (**Supplementary Figure 1**). LtmO, which strongly resembled the asparagine synthase homolog XanA (56% identity) and the amidotransferase homolog Arx5 (44% identity), also possessed five conserved amino acid sites for ligand binding and catalyzing (**Supplementary Figure 2**). Three methyltransferases were there, among them LtmMT2 showed the highest (68%) identity to XanM3 which directed the remethylation of C17 hydroxyl after a cryptic demethylation step catalyzed by XanO4 (Kong et al., 2016). Two cytochrome P450 oxidases LtmP1 and LtmP2 were found in the *ltm* cluster. Three enzymes, LtmD4, LtmD5, and LtmJ, whose exact functions were uncharacterized all distributed among these clusters.

## Activation of *ltm* by Promoter Knock-In and Variating Fermentation Environment

As **Figure 1C** shows, two pathway-specific regulatory genes are in the *ltm* cluster, a SARP family regulator-encoding gene *ltmR1* and a TetR family regulator-encoding gene *ltmR2*. Usually, the TetR family regulators play a negative control in antibiotic biosynthesis, while SARP family regulators showed a positive role (Krause et al., 2020). Therefore, we attempted to overexpress the activator gene *ltmR1* under the control of the strong promoter *kasO*∗p and delete the repressor gene *ltmR2* based on homologous recombination. The resulting recombinant strains L10-OE-R1 and L10-ΔR2 were cultivated for analysis of SMs as compared with starter strain L10. However, the metabolite profiles did not show any differential peaks (**Figure 2C** II and III), indicating that overexpression and/or deletion of the single regulatory gene was not sufficient to activate *ltm* cluster.

To induce the expression of the *ltm* gene cluster, we further try to overexpress the core genes involved in polycyclic skeleton biosynthesis that would bypass the real regulatory pathway. The structural genes (*ltmA-C*), polyketide cyclase genes (*ltmD1-D3*), and a 3-oxoacyl-acyl carrier protein (ACP) reductase gene (*ltmF1*) seemed to be located in a single multicistron (**Figure 1C**), we cloned this multicistronic cassette (*ltmF1-D3*) into the integrative vector pSET152-spec-*kasO**, in which the strong promoter *kasO**p initiated their transcription. The resulting plasmid pXF03 was introduced into L10 to yield L10-OE-F1D3. As expected, overexpression of core gene cassette led to the production of brown pigments both on solid YMG plate and in liquid YEME media, suggesting the activation of gene cluster *ltm* successfully (**Figure 2A**; **Supplementary Figure 3**). Metabolites analysis of fermentation broth by HPLC showed that there are two new compounds produced in the overexpressed strain L10-OE-F1D3, lanthomicin A (**1**) and compound **2** (**Figure 2C** IV).

Inspired by the above results, we found that just downstream of the seven-gene multicistron, three monooxygenase genes

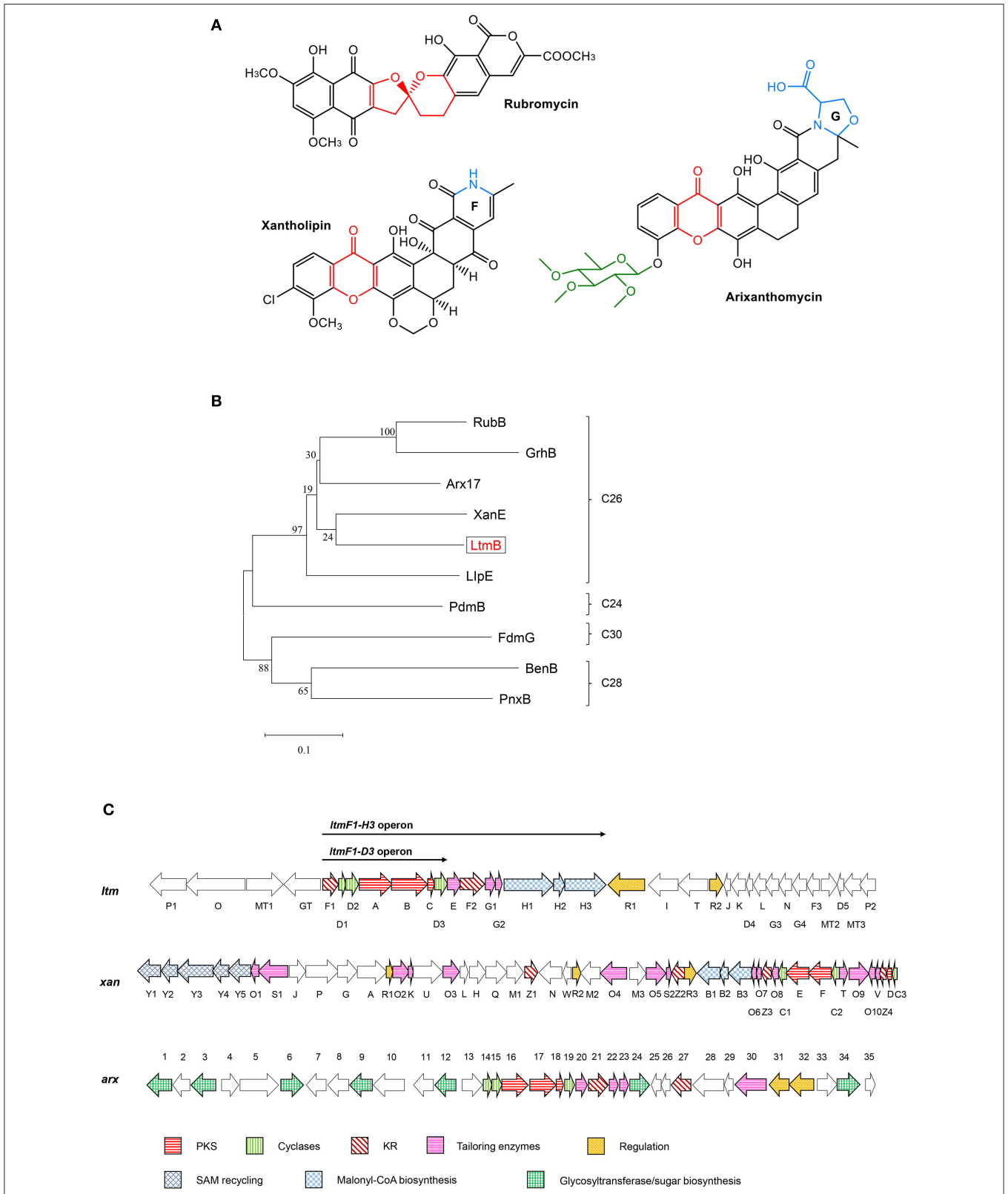



**FIGURE 1 |** Comparative analysis of the lanthomicin biosynthetic gene cluster. **(A)** Examples of compounds belong to the pentangular polyphenol subfamily. **(B)** The maximum likelihood tree of LtmB with homologous $KS_{\beta}$ proteins from the well-studied pentangular polyphenol clusters. These $KS_{\beta}$ units could be divided into four groups according to the carbon chain length of their products. **(C)** *ltm* cluster was indicated by comparison with *xan* and *arx* clusters. The operons for promoter engineering were also indicated in the figure.

**TABLE 1 |** Deduced functions of genes in the lanthomicin biosynthesis gene cluster and protein homologs annotation.

| Gene product | aa | Proposed function | Most similar protein (acc.number) | Identical aa (%) | Homologous gene in *xan*, *arx* | Identity/similarity (%) |
|---|---|---|---|---|---|---|
| ORF11 | 62 | Transposase | WP_161968789.1 | 69 | | |
| LtmP1 | 403 | Cytochrome P450 | WP_137814202.1 | 83 | XanO2 | 35/52 |
| LtmO | 620 | Asparagine synthase | WP_137814203.1 | 85 | XanA, Arx5 | 56/67, 44/58 |
| LtmMT1 | 328 | O-methyltransferase | WP_107082228.1 | 76 | Arx6 | 39/58 |
| LtmGT | 354 | Glycosyltransferase | WP_137814140.1 | 76 | XanG, Arx3 | 55/66, 42/58 |
| LtmF1 | 208 | 3-oxoacyl-ACP reductase | WP_137814141.1 | 86 | XanS2, Arx4 | 53/64, 55/66 |
| LtmD1 | 111 | Cyclase | WP_137814142.1 | 87 | XanC3, Arx14 | 67/82, 80/92 |
| LtmD2 | 145 | Cyclase | WP_137814143.1 | 82 | XanC2, Arx15 | 68/79, 70/76 |
| LtmA | 422 | Beta-ketoacyl synthase alpha | WP_137814144.1 | 90 | XanF, Arx16 | 80/88, 78/87 |
| LtmB | 397 | Ketosynthase chain-length factor | WP_137814129.1 | 87 | XanE, Arx17 | 73/80, 76/83 |
| LtmC | 84 | Acyl carrier protein | WP_137814128.1 | 72 | XanD, Arx18 | 40/60, 51/73 |
| LtmD3 | 152 | Cyclase | WP_137814127.1 | 87 | XanC1, Arx19 | 65/76, 70/82 |
| LtmE | 153 | Monooxygenase | WP_049718470.1 | 79 | XanO8, Arx20 | 66/77, 64/73 |
| LtmF2 | 250 | 3-oxoacyl-ACP reductase | WP_137814126.1 | 85 | XanZ3, Arx21 | 65/82, 64/78 |
| LtmG1 | 108 | Monooxygenase | WP_116198857.1 | 86 | XanO7, Arx22 | 58/74, 67/80 |
| LtmG2 | 92 | Monooxygenase | WP_137814125.1 | 84 | XanO6, Arx23 | 62/72, 66/76 |
| LtmH1 | 571 | carboxyl transferase alpha | WP_125308160.1 | 77 | XanB3 | 67/76 |
| LtmH2 | 174 | biotin carboxyl carrier protein | WP_137814199.1 | 76 | XanB2, Arx35 | 53/60, 49/62 |
| LtmH3 | 469 | biotin carboxylase | WP_137814198.1 | 88 | XanB1 | 74/83 |
| LtmR1 | 613 | SARP family transcriptional regulator | WP_198535620.1 | 73 | | |
| LtmI | 540 | FAD-dependent monooxygenase | WP_200827518.1 | 74 | XanO5, Arx10 | 45/58, 41/57 |
| LtmT | 477 | MFS transporter | WP_137814146.1 | 75 | | |
| LtmR2 | 228 | TetR family transcriptional regulator | WP_158879437.1 | 86 | | |
| LtmJ | 149 | Monooxygenase | WP_043782530.1 | 78 | XanO10, Arx25 | 72/81, 63/78 |
| LtmK | 292 | Dehydrogenase | WP_137814130.1 | 79 | XanS1, Arx13 | 53/67, 50/63 |
| LtmD4 | 128 | CurD-like protein | WP_049718476.1 | 86 | XanV, Arx26 | 76/87, 80/89 |
| LtmL | 493 | Peptide permease | KNB50326.1 | 73 | XanQ, Arx28 | 52/68, 49/64 |
| LtmG3 | 541 | FAD-dependent monooxygenase | WP_137814132.1 | 89 | XanO4, Arx30 | 70/82, 64/77 |
| LtmN | 286 | Reductase | WP_137814133.1 | 88 | XanZ1, Arx33 | 57/70, 45/64 |
| LtmG4 | 421 | FAD-dependent monooxygenase | WP_192909516.1 | 83 | XanO4 | 26/39 |
| LtmF3 | 245 | 3-oxoacyl-ACP reductase | WP_137814135.1 | 88 | XanZ4, Arx27 | 65/77, 67/79 |
| LtmMT2 | 336 | Methyltransferase | WP_137814136.1 | 86 | XanM3 | 68/80 |
| LtmD5 | 122 | CurD-like protein | WP_137814137.1 | 83 | XanT, Arx29 | 63/79, 61/75 |
| LtmMT3 | 351 | Methyltransferase | WP_162834049.1 | 80 | XanM3 | 45/62 |
| LtmP2 | 467 | Cytochrome P450 | WP_043782514.1 | 82 | XanO2 | 39/55 |
| ORF47 | 191 | Transcriptional regulator NovG | GCB88085.1 | 70 | | |

(*ltmE, ltmG1-G2*), another one 3-oxoacyl-ACP reductase gene (*ltmF2*), and three genes (*ltmH1-H3*) serving the synthesis of malonyl-CoA extender unit, are arranged in the same direction as the *ltmF1-D3* gene cassette (**Figure 1C**). We hypothesized that all these genes might be transcribed together to form a bigger transcript. So we replaced the native promoter with the strong promoter *kasO**p, which was inserted upstream of the *ltmF1* gene by homologous recombination (**Figure 2B**). Finally, the generated strain *S. chattanoogensis* XF1 was confirmed by PCR (**Supplementary Figure 4**). The corresponding fermentation result revealed a new compound lanthomicin B (**3**) was synthesized (**Figure 2C** V). By changing the fermentation environment, we found that fermenting in a 15 L fermentor could produce several new homologs of **3** compared with a flask, and two of them displayed a distinct UV absorption spectrum (**Figure 2C** VI; **Supplementary Figure 5**). The difference in absorption spectra might imply structural change and we would like to identify the chemical alteration of lanthomicin C (**4**) with a higher yield. We also conducted the quantitative real-time PCR (qRT-PCR) experiment, the transcriptional levels of genes involved in the activated multicistron were significantly upregulated compared to the control strain L10, while genes distributed at two sides were almost the same transcriptional levels as the control strain (**Figure 3**).

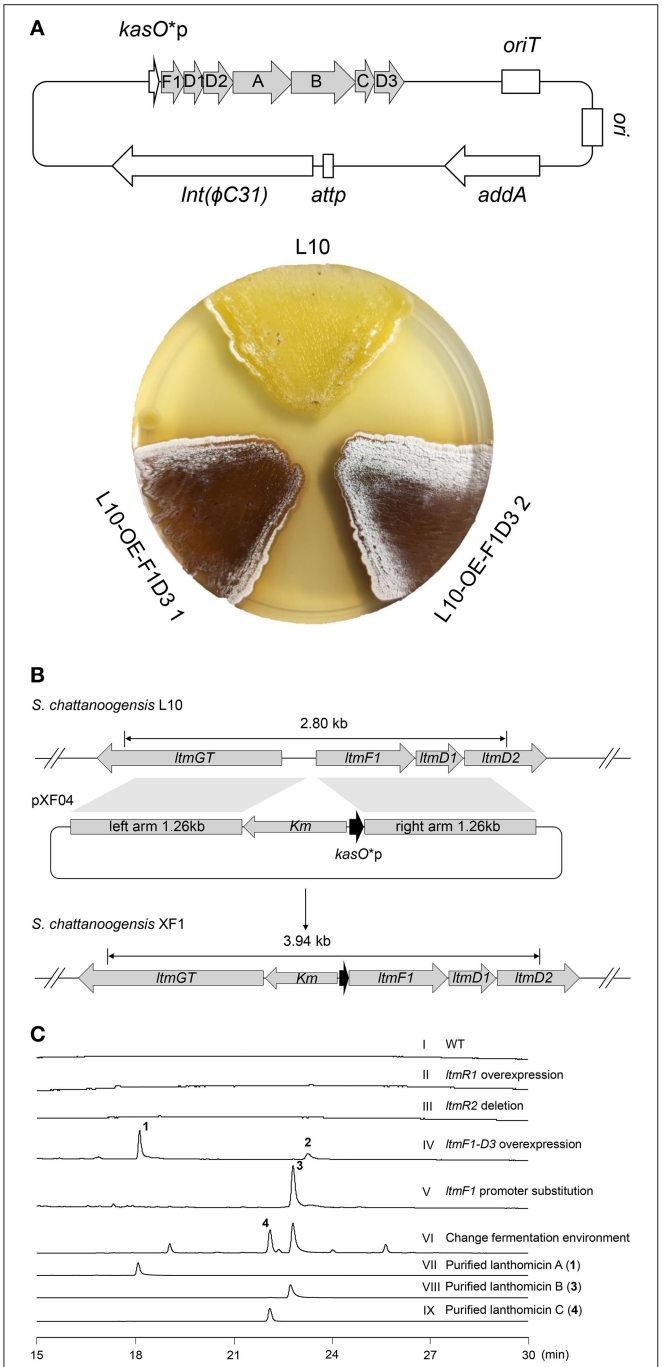



**FIGURE 2 |** Activation of cryptic lanthomicin gene cluster by promoter knock-in and changing cultivating environment. **(A)** Lanthomicins were produced on a solid YMG plate by overexpression of *ltmF1-D3* cassette. **(B)** Schematic of a promoter replacement strategy that a *Km-kasO**p cassette was inserted upstream of the *ltmF1* gene and replaced the native promoter by homologous recombination. **(C)** HPLC analysis of the metabolites from different strains (UV at 480 nm): (I) *S. chattanoogensis* L10; (II) *S. chattanoogensis* L10-OE-R1; (III) *S. chattanoogensis* L10-ΔR2; (IV) *S. chattanoogensis* L10-OE-F1D3; (V) *S. chattanoogensis* XF1; (VI) *S. chattanoogensis* XF1 (changing fermentation environment); (VII) isolated lanthomicin A; (VIII) isolated lanthomicin B; and (IX) isolated lanthomicin C.

## Confirmation of Lanthomicins Biosynthesis Gene Cluster *ltm*

Traditional genetic manipulation for *S. chattanoogensis* L10 was proved to be intricate and time-consuming (Liu et al., 2015). The CRISPR-Cpf1 system has been widely used in many bacterial species (Meliawati et al., 2021). In the genus *Streptomyces*, this system has been applied successfully to conduct precise genome editing by homology-directed repair (Li et al., 2018). To confirm that the *ltm* gene cluster is responsible for the biosynthesis of the three polyphenol lanthomicins, the functional domain of *ltmA* was used as the target for cleavage. First, the CRISPR-Cpf1 editing plasmid pKCCpf1 was introduced into *S. chattanoogensis* XF1. As **Figure 4A** shows, the transformation efficiency is very low and we suppose the expression level of Cpf1 is lethal to the host cell. So the inducible promoter *tipAp* was adopted to direct Cpf1 expression. When a gene-specific spacer was introduced into pKCCpf1(*tipAp*) to obtain pKCCpf1(*tipAp*)-ltmAspacer, the target position of *ltmA* in all exconjugants was destroyed, resulting in the abolishment of colored metabolites as compared with the strain XF1 containing pKCCpf1(*tipAp*) (**Supplementary Figure 6**). The addition of homologous arms (HAs) improved conjugation efficiency significantly and the successfully deleted colonies were found as expected (**Figure 4A**; **Supplementary Figure 7**). The metabolic profiles were performed to prove that lanthomicin B was not detected in the resulting strain *S. chattanoogensis* XF1-ΔA by HPLC analysis (**Figure 4B** II). For the complementation experiment, the *ltmA* gene was randomly inserted into the genome of mutant strain XF1-ΔA affording strain XF1-ΔA-OE-A, which restored the ability to produce lanthomicin B (**Figure 4B** III). These results suggested that the *ltm* gene cluster was involved in the biosynthesis of lanthomicins.

## Isolation and Characterization of Lanthomicins

The structure of lanthomicins was elucidated based on NMR data (**Supplementary Table 5**; **Supplementary Figures 13–33**). Lanthomicin A (**1**) was obtained as murrey amorphous powder. The molecular formula was determined as $C_{25}H_{18}O_9$ on the basis of the anion peak at *m/z* 461.0879 [M-H]- by high-resolution mass spectrometry (HRMS). Careful comparison of the NMR data of **1** and metabolite 4, which has been activated and isolated from an engineered *Streptomyces viridochromogenes* strain in 2017 (Zhang et al., 2017), revealed that they share a similar dihydrobenzo[α] naphthacenequinone core. The only difference between them was that the methyl group at C-22 in metabolite 4 was replaced by the acetonyl group in **1**. The acetonyl fragment was confirmed by the HMBC correlations from H3-25 (δ 2.17) to C-24 (δ 207.79) and C-23 (δ 53.32). The HMBC correlations from H2-23 (δ 2.68) to C-1 (δ 70.89), C-22 (δ 42.96), and C-2 (δ 52.17) revealed that the acetonyl moiety was attached to C-1. Hence, the structure of **1** was established (**Figure 5**). HRMS of compound **2** predicted a molecular formula $C_{26}H_{16}O_{10}$ with exact mass 487.0667 [M-H]-.

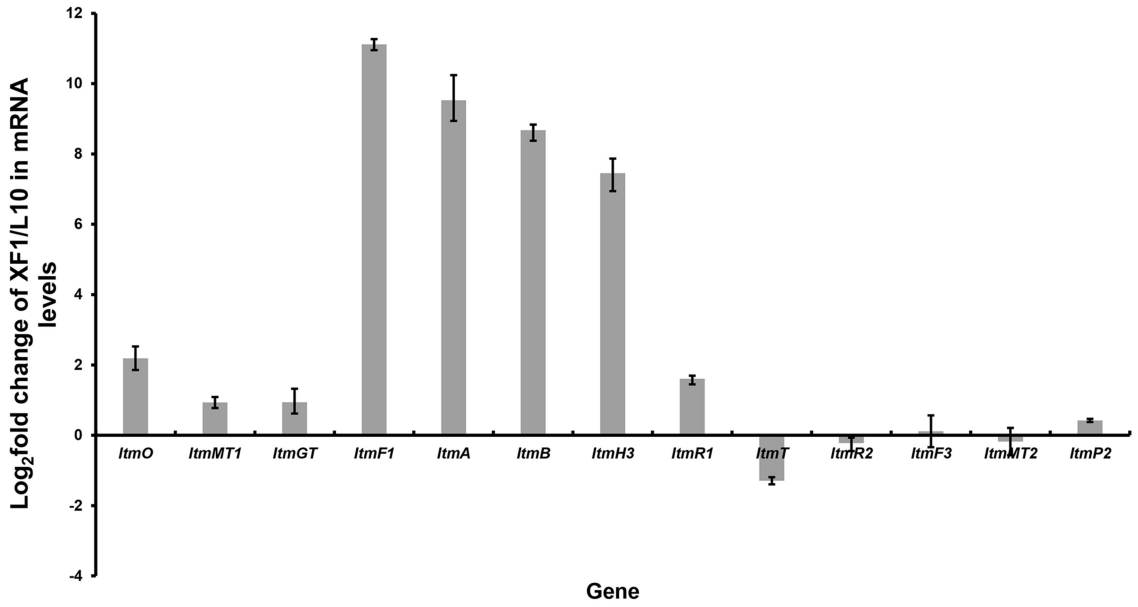



**FIGURE 3 |** Quantitative real-time PCR experiment. By promoter replacement, gene transcription levels from *ltmF1* to *ltmH3* in *S. chattanoogensis* XF1 were significantly upregulated, while genes distributed at two sides were almost the same transcriptional levels as the control strain L10. Fold change was indicated as $\log_2$ (XF1/L10).

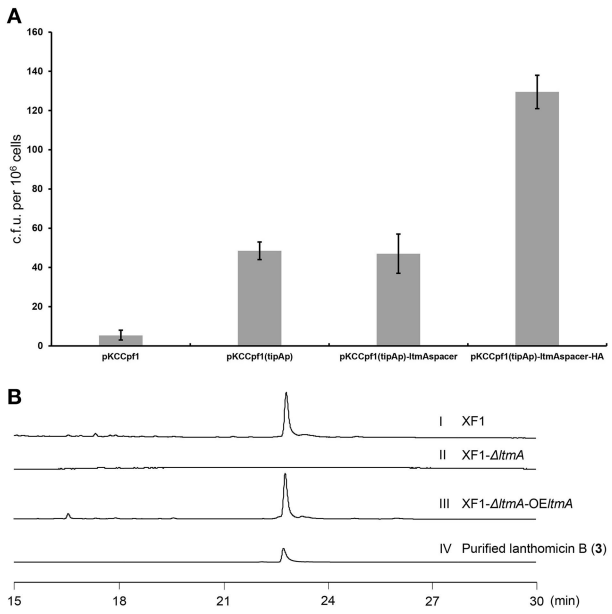



**FIGURE 4 |** Confirmation of lanthomicins biosynthesis gene cluster *ltm* by the CRISPR-Cpf1-based gene-editing system. **(A)** Transformation efficiency was evaluated by counting the transformant number per $10^6$ viable spores. **(B)** HPLC analysis of lanthomicin B production from *ltmA* deletion and complement strains (UV at 480 nm): (I) *S. chattanoogensis* XF1; (II) *S. chattanoogensis* XF1-ΔA; (III) *S. chattanoogensis* XF1-ΔA-OE-A; and (IV) isolated lanthomicin B.

Lanthomicin B (**3**) was obtained as a deep orange powder and its molecular formula was $C_{26}H_{18}O_9$ with an anion peak at 473.0876 [M-H]- based on the HRMS data. By literature searching, three main differences existed between **3** and hexaricin F (Gao et al., 2018), the substituents of C-6, C-11, and C-17 were -H, -OH, and -OH in **3** and -OH, -H, and -$OCH_3$ in hexaricin F, respectively. The structure was further confirmed by the new peak H-6 (δ 9.02, s) with correlations to C-5 (δ 141.18) and C-7 (δ 130.20), the disappearance of H-11, and the HMBC correlations from OH-17 (δ 12.51) to C-16 (δ 112.88) and C-18 (δ 131.89). Thus, the structure of **3** was established.

Lanthomicin C (4) was obtained as scarlet amorphous powder with molecular formula $C_{28}H_{22}O_{11}$ according to the anion peak at *m/z* 533.1087 [M-H]-. The COSY correlations of H-11 (7.30, d)/H-12 (7.34, d) indicated the difference in ring A between **3** and **4**, and this was further confirmed by the HMBC correlations from HO-10 (δ 12.14) to C-9 (δ 113.42) and C-11 (δ 129.29). Another main COSY difference occurred in H-19 (δ 4.68)/H-20 (δ 3.51, 2.53), the lack of one hydrogen confirmed the oxhydryl introduced at C-19. And for **4**, two extra molecules of methyl were linked by HMBC correlations from δ 3.41 to C-1 (δ 170.18) and C-25 (δ 106.72), from δ 3.89 to C-17 (δ 152.57). Henceforth, the structure of **4** was finished.

## Antitumor Bioactivity of Lanthomicins

It has been reported that compounds synthesized by type II PKS often showed better inhibitory activity against tumor cell lines (Gan et al., 2015; Woo et al., 2016; Jiang et al., 2018). Therefore, we performed *in vitro* cytotoxicity assays for lanthomicins toward cancer cell lines using doxorubicin as a positive control. The results indicated that lanthomicin A showed antiproliferative activity toward lung cancer cell line A-549

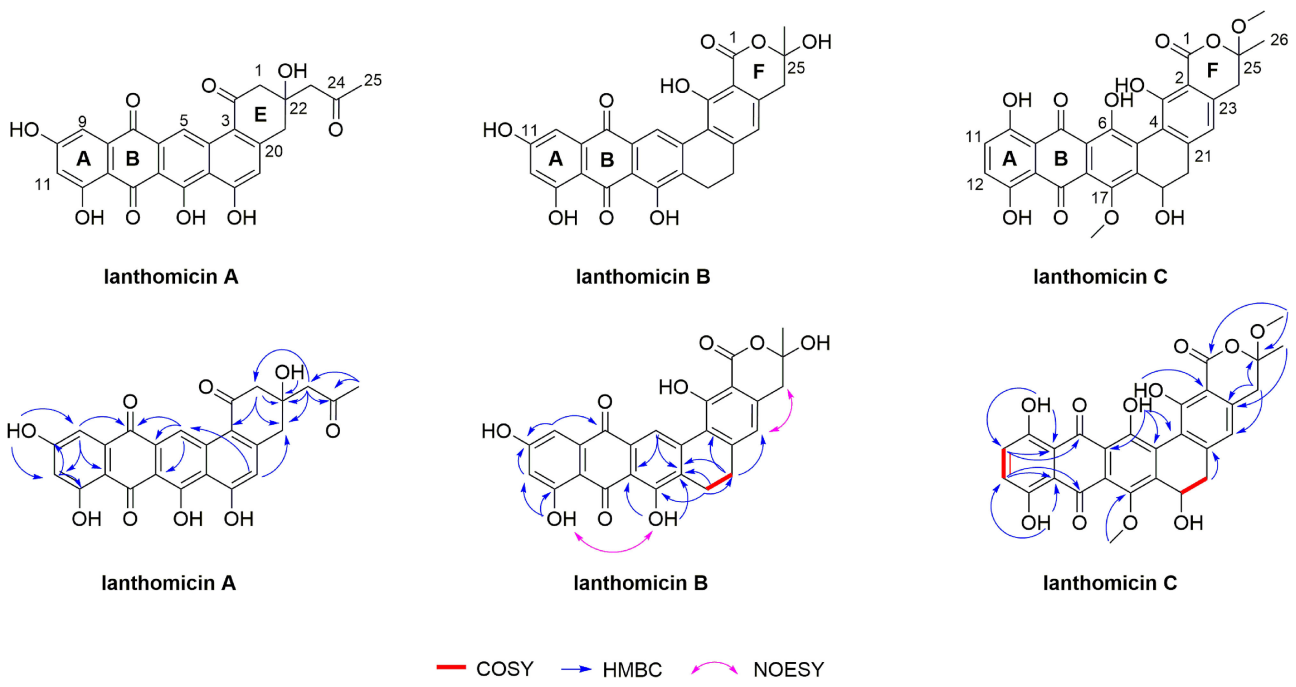



**FIGURE 5 |** Planar structures of lanthomicin A-C. Key COSY, HMBC, and NOESY correlations of lanthomicin A–C are indicated.

with $IC_{50}$ 0.17 μM, while lanthomicin B and C demonstrated nonsignificant antiproliferative activity against all the cell lines (**Table 2**; **Supplementary Figure 8**).

## Proposal of a Biosynthetic Pathway for Lanthomicins

The biosynthesis of the lanthomicins can be rationalized on the basis of successive activation of silent genes and functional analysis by homologous comparison. The key enzymes predicted to be involved in this process are shown in **Figure 1C**. The biosynthesis probably begins with condensing malonate units by the minimal type II PKSs LtmABC (Ketosynthase, $KS_\alpha$; chain length factor or $KS_\beta$; ACP). Three enzymes LtmH1-H3 seem to be responsible for the biosynthesis of malonyl-CoA from the carboxylation of acetyl-CoA. Three KRs were found in *ltm* cluster and a phylogenetic analysis implied that LtmF1 might belong to C17 KRs, while LtmF2 was thus the likely C19 KR (**Supplementary Figure 9**). Thus LtmF1 was probably involved in the reduction of the keto group at C17 of the nascent polyketide chain to a hydroxyl group. The cyclases LtmD1-D3 were supposed to further modify the reduced skeleton directing the cyclization through intramolecular aldol condensations to form the angular 5-ring aromatic structure. Activation of the multicistronic cassette *ltmF1D1D2ABCD3* could produce lanthomicin A and compound **2** (**Figure 6** route 1). Three monooxygenases LtmE, LtmG1-G2, and a predicted C19 KR LtmF2 were also activated by promoter substitution and the fermentation profile changed leading to the discovery of lanthomicin B. LtmF2 was thus likely to catalyze the reduction of the 19-ketone to a hydroxyl group. We proposed that three monooxygenases might take responsibility for the closure of ring F in lanthomicin B. Lanthomicin C showed more methylation and hydroxylation than lanthomicin B. LtmF3, which was predicted as a C11 KR, thus might catalyze reduction in C11 carbonyl. LtmJ showed 55% identity to FdmM, which hydroxylated the C6 site during the biosynthesis of fredericamycin A (Chen et al., 2009), might also play a similar role during the biosynthesis (**Figure 6** route 2).

**TABLE 2 |** Antitumor activity of lanthomicins.

| Tumor cell lines | $IC_{50}$ (μM/L) | | | |
|---|---|---|---|---|
| | Doxorubicin | Lanthomicin A | Lanthomicin B | Lanthomicin C |
| A-549 (lung) | 0.02 | 0.17 | >100 | >100 |
| MCF-7 (breast) | 0.62 | 5.98 | >100 | >100 |
| HepG2 (liver) | 4.38 | 34.34 | >100 | >100 |
| HCT-116 (colon) | 10.58 | 82.27 | >100 | >100 |

## DISCUSSION

Secondary metabolites embedded in cryptic biosynthetic gene clusters (BGCs) with potential clinical value are well belongings of filamentous microorganisms, especially the genus *Streptomyces*. To activate cryptic clusters in *S. chattanoogensis* L10, we consider that natamycin is the main product of this

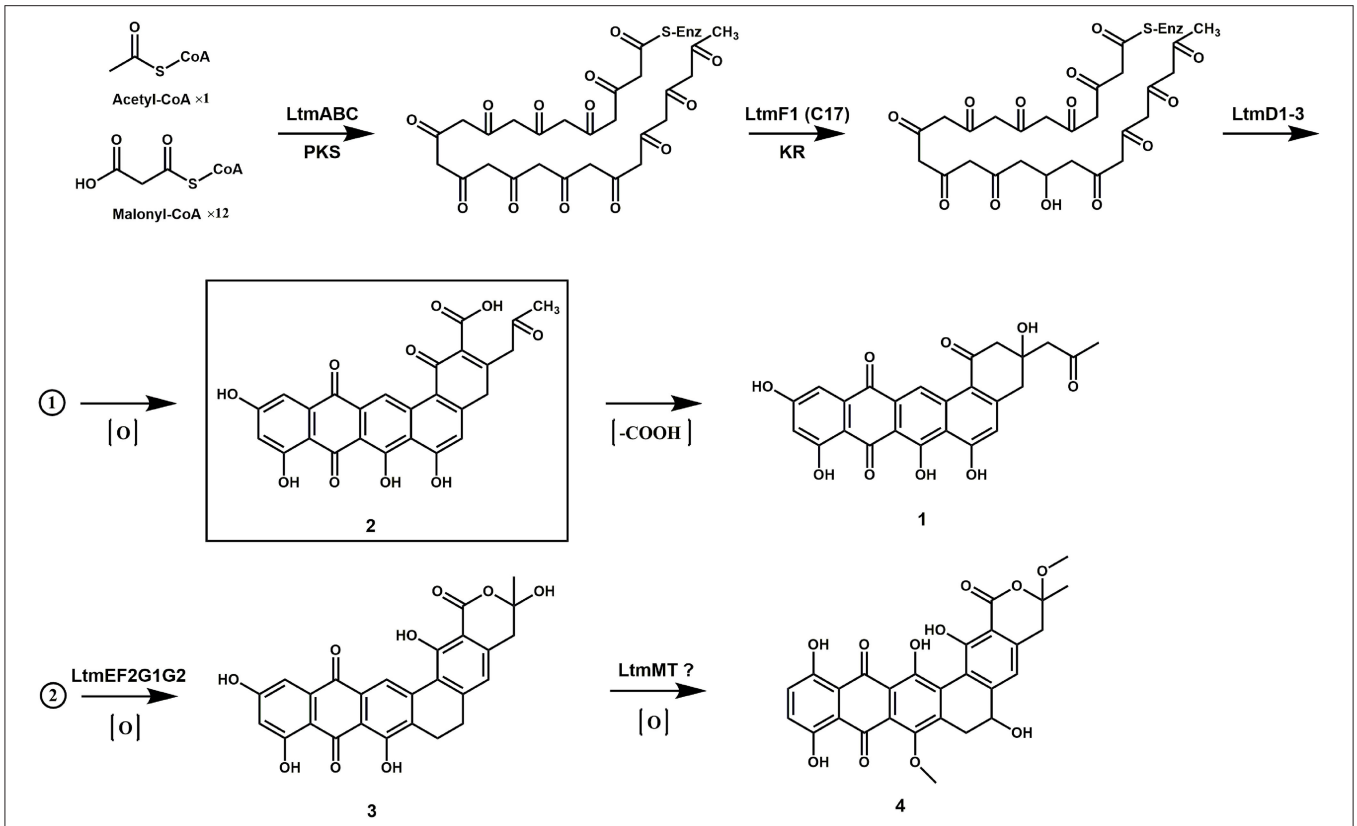



**FIGURE 6 |** The proposed biosynthetic pathway of lanthomicins.

strain and that blocking the main products in *Streptomyces* accelerates new compounds discovery (Culp et al., 2019). So we knocked out the first PKS gene in natamycin BGC to block its production, and we found that no new secondary metabolites were produced. We then used the natamycin-inactivated strain for activation. As mentioned above, 11 putative BGCs containing PKSs or NRPSs were found in the genome of *S. chattanoogensis* L10. Except for natamycin, chattamycin, and lanthomicin gene clusters, we also activated cluster 20 by a regulatory means. Recently, azoxymycins were found to be synthesized by linking two precursors with an azoxy bond (Guo et al., 2015) and how this unusual chemical connection could be formed was deeply interpreted (Guo et al., 2019). Several terpene and polypeptide BGCs were also found and bioactive polypeptide antibiotics could be identified by uncovering these cryptic pathways (Daniel-Ivad et al., 2017).

In this work, we found the *ltm* gene cluster with homology to *xan* and *arx* gene clusters, the corresponding natural products belong to the pentangular polyphenols subclade of aromatic PKS. The chemical diversity of this subclade is dependent on the postsynthetic modifications by some enzymes like methyltransferase and glycosyltransferase, especially some unknown oxygenases. It is worth mentioning that the pentangular polyphenols often display an inhibitory effect on tumor cells (Winter et al., 2013). To activate the cryptic *ltm* gene cluster, we adopt a combined strategy. We first overexpressed the putative positive regulator, disrupted the repressor gene, and combined the both. However, all these attempts failed to activate the gene cluster. Then we constitutively expressed the core gene cassette, resulting in the production of lanthomicin A and compound **2**. The further knock-in of strong promoter *kasO**p in the upstream of *ltmF1* along with changing fermentation environment ultimately led to the discovery of lanthomicin B and lanthomicin C. QRT-PCR showed that the transcriptional levels from *ltmF1* to *ltmH3* were strikingly improved and these genes seemed to be cotranscribed. Then, we successfully applied the CRISPR-Cpf1-based gene-editing tool to knock out the core skeleton gene *ltmA* in *ltm* gene cluster. The *ltmA* gene knockout and complement experiments further confirmed that the *ltm* gene cluster is responsible for the biosynthesis of the three polyphenol lanthomicins.

In this study, three novel pentangular polyphenols lanthomicin A–C were characterized by structure elucidation and antiproliferative activity evaluation. Lanthomicin A showed a characteristic in the E ring that the carboxyl group seemed to be lost and this situation was also found in the biosynthesis of pradimicin A (Zhan et al., 2008). Lanthomicin A also showed antiproliferative activity toward all of the tested cancer cell lines, with $IC_{50}$ values of 0.17 and 5.98 μM in A-549 and MCF-7 cells, respectively. Lanthomicin C showed more methylation

and hydroxylation than others and reduction in C11 carbonyl was assumed to be catalyzed by the predicted C11 KR LtmF3 under a changed fermentation environment. There are no clues about which enzymes may be responsible for the formation of bonds at C5-C18 and C2-C23 (C1-C2 for lanthomicin A). Zhan and coworkers showed a monooxygenase PdmH worked collaboratively with two cyclases to form the 5-ring structure (Zhan et al., 2008). Hence, further studies were needed to uncover the cyclization in lanthomicin biosynthesis.

## 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 in the article/**Supplementary Material**.

## AUTHOR CONTRIBUTIONS

X-FL: conceptualization, methodology, validation, investigation, data curation, formal analysis, visualization, writing—original draft, and writing—review and editing. J-XW: methodology, validation, and formal analysis. X-AC: formal analysis and writing—review and editing. YL: data curation and formal analysis. Y-QL: conceptualization, project administration, funding acquisition, supervision, and writing—review and editing. All authors read and approved the final manuscript.

## FUNDING

This work was supported by the National Key Research and Development Program (Nos. 2018YFA0903200 and 2019YFA09005400) and the National Natural Science Foundation of China (Nos. 31730002 and 31520103901).

## ACKNOWLEDGMENTS

We thank Prof. Yin-Hua Lu, School of Life and Environmental Sciences, Shanghai Normal University, for providing the plasmid CRISPR-Cpf1.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.902990/full#supplementary-material

## REFERENCES

Aparicio, J. F., Barreales, E. G., Payero, T. D., Vicente, C. M., de Pedro, A., and Santos-Aberturas, J. (2016). Biotechnological production and application of the antibiotic pimaricin: biosynthesis and its regulation. *Appl. Microbiol. Biotechnol.* 100, 61–78. doi: 10.1007/s00253-015-7077-0

Arcamone, F., Cassinelli, G., Fantini, G., Grein, A., Orezzi, P., Pol, C., et al. (1969). Adriamycin, 14-hydroxydaunomycin, a new antitumor antibiotic from *S.* peucetius var caesius. Biotechnol. *Bioeng.* 11, 1101–1110. doi: 10.1002/bit.260110607

Baral, B., Akhgari, A., and Metsä-Ketel,ä, M. (2018). Activation of microbial secondary metabolic pathways: avenues and challenges. *Synth. Syst. Biotechnol.* 3, 163–178. doi: 10.1016/j.synbio.2018.09.001

Bu, Q. T., Yu, P., Wang, J., Li, Z. Y., Chen, X. A., Mao, X. M., et al. (2019). Rational construction of genome-reduced and high-efficient industrial *Streptomyces* chassis based on multiple comparative genomic approaches. *Microb. Cell. Fact.* 18, 16. doi: 10.1186/s12934-019-1055-7

Chen, R. D., Zhang, Q. B., Tan, B., Zheng, L. J., Li, H. X., Zhu, Y. G., et al. (2017). Genome mining and activation of a silent PKS/NRPS gene cluster direct the production of totopotensamides. *Org. Lett.* 19, 5697–5700. doi: 10.1021/acs.orglett.7b02878

Chen, Y. H., Wendt-Pienkoski, E., Rajski, S. R., and Shen, B. (2009). In vivo investigation of the roles of FdmM and FdmM1 in fredericamycin biosynthesis unveiling a new family of oxygenases. *J. Biol. Chem.* 284, 24735–24743. doi: 10.1074/jbc.M109.014191

Culp, E. J., Yim, G., Waglechner, N., Wang, W., Pawlowski, A. C., and Wright, G. D. (2019). Hidden antibiotics in actinomycetes can be identified by inactivation of gene clusters for common antibiotics. *Nat. Biotechnol.* 37, 1149–1154. doi: 10.1038/s41587-019-0241-9

Daniel-Ivad, M., Hameed, N., Tan, S., Dhanjal, R., Socko, D., Pak, P., et al. (2017). An engineered allele of afsQ1 facilitates the discovery and investigation of cryptic natural products. *ACS. Chem. Biol.* 12, 628–634. doi: 10.1021/acschembio.6b01002

Gan, M. L., Liu, B., Tan, Y., Wang, Q., Zhou, H. X., He, H. W., et al. (2015). Saccharothrixones A-D, tetracenomycin-type polyketides from the marine-derived actinomycete *Saccharothrix sp*. 10-10. *J. Nat. Prod.* 78, 2260–2265. doi: 10.1021/acs.jnatprod.5b00577

Gao, C. Z., Guo, Z. Y., Lu, X. Z., Chen, H. Y., Liu, L. W., Yu, Z. G., et al. (2018). Hexaricins, pradimicin-like polyketides from a marine sediment-derived *Streptosporangium sp*. and their antioxidant effects. *J. Nat. Prod.* 81, 2069–2074. doi: 10.1021/acs.jnatprod.8b00397

Guo, Y. Y., Li, H., Zhou, Z. X., Mao, X. M., Tang, Y., Chen, X., et al. (2015). Identification and biosynthetic characterization of natural aromatic azoxy products from *Streptomyces chattanoogensis* L10. *Org. Lett.* 17, 6114–6117. doi: 10.1021/acs.orglett.5b03137

Guo, Y. Y., Li, Z. H., Xia, T. Y., Du, Y. L., Mao, X. M., and Li, Y. Q. (2019). Molecular mechanism of azoxy bond formation for azoxymycins biosynthesis. *Nat. Commun.* 10, 4420. doi: 10.1038/s41467-019-12250-1

Gust, B., Challis, G. L., Fowler, K., Kieser, T., and Chater, K. F. (2003). PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proc. Natl. Acad. Sci.* 100, 1541–1546. doi: 10.1073/pnas.0337542100

Jiang, L., Pu, H., Xiang, J. X., Su, M., Yan, X. H., Yang, D., et al. (2018). Huanglongmycin A-C, cytotoxic polyketides biosynthesized by a putative type II polyketide synthase from *Streptomyces sp*. CB09001. *Front. Chem.* 6, 254. doi: 10.3389/fchem.2018.00254

Kalkreuter, E., Pan, G., Cepeda, A. J., and Shen, B. (2020). Targeting bacterial genomes for natural product discovery. *Trends. Pharmacol. Sci.* 41, 13–26. doi: 10.1016/j.tips.2019.11.002

Kang, H. S., and Brady, S. F. (2014). Arixanthomycins A-C: Phylogeny-guided discovery of biologically active eDNA-derived pentangular polyphenols. *ACS Chem. Biol.* 9, 1267–1272. doi: 10.1021/cb500141b

Kieser, T., Bibb, M. J., Butter, M. J., Chater, K. F., and Hopwood, D. A. (2000). *Practical Streptomyces Genetics*. Norwich: The John Innes Foundation.

Kong, L. X., Zhang, W. K., Chooi, Y. H., Wang, L., Cao, B., Deng, Z. X., et al. (2016). A multifunctional monooxygenase XanO4 catalyzes xanthone formation in xantholipin biosynthesis via a cryptic demethoxylation. *Cell. Chem. Biol.* 23, 508–516. doi: 10.1016/j.chembiol.2016.03.013

Krause, J., Handayani, I., Blin, K., Kulik, A., and Mast, Y. (2020). Disclosing the potential of the SARP-type regulator PapR2 for the activation of antibiotic gene clusters in *Streptomycetes*. *Front. Microbiol.* 11, 225. doi: 10.3389/fmicb.2020.00225

Li, L., Wei, K. K., Zheng, G. S., Liu, X. C., Chen, S. X., Jiang, W. H., et al. (2018). CRISPR-Cpf1-assisted multiplex genome editing and transcriptional repression in *Streptomyces*. *Appl. Environ. Microbiol.* 84, e00827–e00818. doi: 10.1128/AEM.00827-18

Li, S., Li, Y., Lu, C., Zhang, J., Zhu, J., Wang, H., et al. (2015). Activating a cryptic ansamycin biosynthetic gene cluster to produce three new naphthalenic octaketide ansamycins with n-pentyl and n-butyl side chains. *Org. Lett.* 17, 3706–3709. doi: 10.1021/acs.orglett.5b01686

Li, S., Wu, X., Zhang, L., Shen, Y., and Du, L. (2017). Activation of a cryptic gene cluster in *Lysobacter enzymogenes* reveals a module/domain portable mechanism of nonribosomal peptide synthetases in the biosynthesis of pyrrolopyrazines. *Org. Lett.* 19, 5010–5013. doi: 10.1021/acs.orglett.7b01611

Li, Z. Y., Bu, Q. T., Wang, J., Liu, Y., Chen, X. A., Mao, X. M., et al. (2019). Activation of anthrachamycin biosynthesis in *Streptomyces chattanoogensis* L10 by site-directed mutagenesis of *rpoB*. *J. Zhejiang. Univ-Sci. B.* 20, 983–994. doi: 10.1631/jzus.B1900344

Liu, S. P., Yuan, P. H., Wang, Y. Y., Liu, X. F., Zhou, Z. X., Bu, Q. T., et al. (2015). Generation of the natamycin analogs by gene engineering of natamycin biosynthetic genes in *Streptomyces chattanoogensis* L10. *Microbiol. Res.* 173, 25–33. doi: 10.1016/j.micres.2015.01.013

Meena, M., Prajapati, P., Ravichandran, C., and Sehrawat, R. (2021). Natamycin: a natural preservative for food applications-a review. *Food. Sci. Biotechnol.* 30, 1481–1496. doi: 10.1007/s10068-021-00981-1

Meliawati, M., Schilling, C., and Schmid, J. (2021). Recent advances of Cas12a applications in bacteria. *Appl. Microbiol. Biotechnol.* 105, 2981–2990. doi: 10.1007/s00253-021-11243-9

Nett, M., Ikeda, H., and Moore, B. S. (2009). Genomic basis for natural product biosynthetic diversity in the actinomycetes. *Nat. Prod. Rep.* 26, 1362–1384. doi: 10.1039/b817069j

Rath, T. (2013). Tacrolimus in transplant rejection. *Expert. Opin. Pharmacother.* 14, 115–122. doi: 10.1517/14656566.2013.751374

Strachan, C. R., and Davies, J. (2017). The whys and wherefores of antibiotic resistance. *Cold. Spring. Harb. Perspect. Med.* 7, a025171. doi: 10.1101/cshperspect.a025171

Thong, W. L., Shin-ya, K., Nishiyama, M., and Kuzuyama,. T. (2016). Methylbenzene-containing polyketides from a *Streptomyces* that spontaneously acquired rifampicin resistance: structural elucidation and biosynthesis. *J. Nat. Prod.* 79, 857–864. doi: 10.1021/acs.jnatprod.5b00922

Vilhena, C., and Bettencourt, A. (2012). Daptomycin: a review of properties, clinical use, drug delivery and resistance. *Mini. Rev. Med. Chem.* 12, 202–209. doi: 10.2174/1389557511209030202

Wang, G. J., Hosaka, T., and Ochi, K. (2008). Dramatic activation of antibiotic production in *Streptomyces coelicolor* by cumulative drug resistance mutations. *Appl. Environ. Microbiol.* 74, 2834–2840. doi: 10.1128/AEM.02800-07

Winter, D. K., Sloman, D. L., and Porco, J. A. Jr. (2013). Polycyclic xanthone natural products: structure, biological activity and chemical synthesis. *Nat. Prod. Rep.* 30, 382–391. doi: 10.1039/c3np20122h

Woo, C. M., Li, Z. W., Paulson, E. K., and Herzon, S. B. (2016). Structural basis for DNA cleavage by the potent antiproliferative agent (-)-lomaiviticin A. *Proc. Natl. Acad. Sci.* 113, 2851–2856. doi: 10.1073/pnas.1519846113

Zhan, J. X., Watanabe, K. J., and Tang, Y. (2008). Synergistic actions of a monooxygenase and cyclases in aromatic polyketide biosynthesis. *Chembiochem* 9, 1710–1715. doi: 10.1002/cbic.200800178

Zhang, M. M., Wong, F. T., Wang, Y. J., Luo, S. W., Lim, Y. H., Heng, E., et al. (2017). CRISPR-Cas9 strategy for activation of silent *Streptomyces* biosynthetic gene clusters. *Nat. Chem. Biol.* doi: 10.1038/nchembio.2341

Zhang, W. J., Li, S. M., Zhu, Y. G., Chen, Y. C., Chen, Y. L., Zhang, H. B., et al. (2014). Heronamides D-F, polyketide macrolactams from the deep-sea-derived *Streptomyces sp*. SCSIO 03032. *J. Nat. Prod.* 77, 388-91. doi: 10.1021/np400665a

Zhang, W. K., Wang, L., Kong, L. X., Wang, T., Chu, Y. W., Deng, Z. X., et al. (2012). Unveiling the post-PKS redox tailoring steps in biosynthesis of the type II polyketide antitumor antibiotic xantholipin. *Cell. Chem. Biol.* 19, 422–432. doi: 10.1016/j.chembiol.2012.01.016

Zhou, Z. X., Xu, Q. Q., Bu, Q. T., Guo, Y. Y., Liu, S. P., Liu, Y., et al. (2015). Genome mining-directed activation of a silent angucycline biosynthetic gene cluster in *Streptomyces* chattanoogensis. *ChemBioChem.* 16, 496–502. doi: 10.1002/cbic.201402577

**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.

*Copyright © 2022 Liu, Wang, Chen, Liu and Li. 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.*

frontiers | Frontiers in Microbiology

ORIGINAL RESEARCH
published: 17 June 2022
doi: 10.3389/fmicb.2022.911791

# Effects of Inoculation With *Acinetobacter* on Fermentation of Cigar Tobacco Leaves

*Tianfei Zheng[1,2,3], Qianying Zhang[4], Qiaoyin Wu[1,2,3], Dongliang Li[4]*, Xinying Wu[1,2,3], Pinhe Li[4], Quanwei Zhou[4], Wen Cai[4], Juan Zhang[1,2,3]* and Guocheng Du[1,2,3]*

[1]School of Biotechnology, Jiangnan University, Wuxi, China, [2]Key Laboratory of Industrial Biotechnology, Ministry of Education, School of Biotechnology, Jiangnan University, Wuxi, China, [3]Science Center for Future Foods, Jiangnan University, Wuxi, China, [4]Cigar Fermentation Technology Key Laboratory of China Tobacco, China Tobacco Sichuan Industrial Co., Ltd., Chengdu, China

OPEN ACCESS

**Edited by:**
Niranjan Koirala,
Gandaki Province Academy of Science and Technology, Nepal

**Reviewed by:**
Nikolaos Remmas,
Democritus University of Thrace, Greece
Anupama Shrestha,
Madan Bhandari University of Science and Technology, Nepal
Jelena Popović-Djordjević,
University of Belgrade, Serbia

***Correspondence:**
Dongliang Li
360188228@qq.com
Juan Zhang
zhangj@jiangnan.edu.cn

**Specialty section:**
This article was submitted to Microbiotechnology, a section of the journal Frontiers in Microbiology

**Received:** 03 April 2022
**Accepted:** 23 May 2022
**Published:** 17 June 2022

**Citation:**
Zheng T, Zhang Q, Wu Q, Li D, Wu X, Li P, Zhou Q, Cai W, Zhang J and Du G (2022) Effects of Inoculation With Acinetobacter on Fermentation of Cigar Tobacco Leaves. Front. Microbiol. 13:911791. doi: 10.3389/fmicb.2022.911791

Metabolic activity of the microbial community greatly affects the quality of cigar tobacco leaves (CTLs). To improve the quality of CTLs, two extrinsic microbes (*Acinetobacter* sp. 1H8 and *Acinetobacter indicus* 3B2) were inoculated into CTLs. The quality of CTLs were significantly improved after fermentation. The content of solanone, 6-methyl-5-hepten-2-one, benzeneacetic acid, ethyl ester, cyclohexanone, octanal, acetophenone, and 3,5,5-trimethyl-2-cyclohexen-1-one were significantly increased after inoculated *Acinetobacter* sp. 1H8. The inoculation of *Acinetobacter* sp. 1H8 enhanced the normal evolutionary trend of bacterial community. The content of trimethyl-pyrazine, 2,6-dimethyl-pyrazine, and megastigmatrienone were significantly increased after inoculated *Acinetobacter indicus* 3B2. The inoculation of *Acinetobacter indicus* 3B2 completely changed the original bacterial community. Network analysis revealed that *Acinetobacter* was negatively correlated with *Aquabacterium*, positively correlated with *Bacillus*, and had significant correlations with many volatile flavor compounds. This work may be helpful for improving fermentation product quality by regulating microbial community, and gain insight into the microbial ecosystem.

**Keywords: cigar tobacco leaves, inoculation, microbial community, flavor, *Acinetobacter***

## INTRODUCTION

Traditional fermentation processes, such as Chinese liquor, cigar, and fermented vegetable, are mainly driven by complex microbial communities (Song et al., 2017; Liu et al., 2019; Smyth et al., 2019). However, these processes are long and uncontrolled (Jin et al., 2017; Wu et al., 2021). Modern artificial fermentation has been trying to control these processes by adjusting the fermentation conditions or adding specially formulated starters to the natural fermentation system (Shilei et al., 2019). These methods were widely used in the traditional multi-species fermentation industry because of their various advantages, including shortening fermentation times, improving product qualities, and improving safety. For example, the pH condition considerably affected hydrogen fermentation, hydrogen gas was efficiently produced with unconditioned anaerobic sludge when the pH was adjusted to 6.0 throughout the culture period (Kawagoshi et al., 2005). When *Debaryomyces hansenii* and *Yarrowia lipolytica* were incorporated

into the cheese as part of the starter, the cheese could significantly shorten the ripening time while maintaining a good strong flavor (Ferreira and Viljoen, 2003). The inoculation of *Bacillus licheniformis* could affect the microbial community structure and enzyme activity of Daqu, and increased the content of pyrazines and aromatic compounds (Wang et al., 2017). Additionally, the inoculation of *Saccharomyces uvarum* and *Saccharomyces servazzii* significantly improved the quality of Chinese liquor (Wu et al., 2016). However, adjusting fermentation conditions is sometimes not effective, because some traditional fermentation processes have been adjusted thousands of times in their long history. Exogenous addition of microorganisms does not always produce positive results, because inoculated microorganisms cannot promote the evolution of the community towards favorable product fermentation (Wu et al., 2016).

The main reason for the failure of externally added microbes is that microbial interactions between the inoculated microbes (extrinsic) and the native microbes (intrinsic) lead to incomplete structure and function of microbial communities. Microbial interactions, such as symbiotic, synergistic, predation, parasitic, and competitive, can greatly affect the metabolic activity of the microbial communities (Ghosh et al., 2016; Nawaz et al., 2022; Pierce and Dutton, 2022). Rational addition of exogenous microorganisms to activate interactions between microorganisms could strengthen the metabolic capacity of the microbial community. However, it is difficult to determine whether the addition of microbes can promote the fermentation of product unless the microbial interaction is deeply studied. Therefore, it is very important to select suitable microorganisms and study the effects of externally added microbes on the structure and function of microbial communities.

Cigars are one of the oldest traditional tobacco fermented products (Reid et al., 1937). Flavor-producing microbes play an important role in tobacco fermentation (Liu et al., 2021). In our previous work, we found that *Acinetobacter* were the important producers of the aldehydes and ketones. And we isolated two flavor producers, *Acinetobacter* sp. 1H8 and *Acinetobacter indicus* 3B2 from high-quality cigar tobacco leaves (CTLs). We speculated that their addition could improve the quality of some ordinary CTLs. In this study, we investigated whether these two extrinsic stains could improve the quality and flavor of the CTLs and their effects on the structure and function of bacterial communities.

# MATERIALS AND METHODS

## Strains

The strains used in this study were isolated from the CTLs and deposited in the China General Microbiological Culture Collection Center (CGMCC): *Acinetobacter* sp. 1H8 CGMCC NO.23678 and *Acinetobacter indicus* 3B2 CGMCC NO.23679.

## Inoculation Experiment

The fermentation medium for culturing strains was prepared with sucrose 20 g/l, peptone 20 g/l, yeast powder 15 g/l, $K_2HPO_4$ 1.5 g/l $NaH_2PO_4$ 3 g/l, and autoclaved at 121°C for 15 min. Strains were inoculated by sterile loops in 250-ml flasks with 50 ml fermentation medium, and cultured at 220 rpm 30°C for 36 h. Then the fermentation broth was inoculated into CTLs with 30% inoculation amount (v/g) and fermented in a biochemical incubator. All fermentations were conducted without agitation at 30°C for 15 days. The unfermented CTLs and uninoculated CTLs were prepared as Control 1 and Control 2, respectively. All experiments were performed in triplicate. The quality of CTLs was blindly assessed by three professional tasters. With 10–20 years of testing experience, these tobacco tasters have conducted sensory evaluation on more than 2,000 cigar samples, and can accurately, consistently, and repeatedly evaluate cigars.

## Volatile Flavor Compound Analysis

Volatile flavor compounds (VFCs) in CTLs were analyzed by headspace solid phase microextraction-gas chromatography–mass spectrometry (HS-SPME-GC–MS). CTLs were dried at 40°C and pulverized by a grinder. 1.5-g powder was placed in a 10 ml glass vial and extracted by headspace solid-phase microextraction (50/30 μm DVB/CAR/PDMS fibre, Supelco, Bellefonte, PA, United States) at 60°C for 30 min. After extraction, volatile flavor compounds (VCFs) were identified using a Pegasus BT GC-TOFMS (LECO Co., St. Joseph, MI, United States), with a DB-5MS column (60 m × 0.25 mm id × 0.25 μm film thickness). Helium C-60 was used as a carrier gas with a flow rate of 1 ml/min, and the injector port was heated to 250°C. The oven temperature was fixed at 40°C for 2 min, increased to 250°C at a rate 10°C/min, and then held for 5 min. Meanwhile, the transfer line and ion source temperatures were maintained at 280°C and 210°C, respectively. Electron impact (EI) was used as the ionization mode, with an EI voltage of 70 eV, and a mass scan range of 33–400 m/z was used for full-scan mode with an acquisition rate of 10 scans/s.

The acquired GC–MS raw data from the QC sample was analyzed with the Automatic Mass Spectral Deconvolution and Identification System (AMDIS) to verify individual analytes' presence and deconvolute the co-eluting peaks as previously described. The parameters for the deconvolution were set as follows: component width = 20, adjacent peaks subtraction = 2, resolution = high, sensitivity = low, and shape requirements = medium. The parameters for peak detection were set at the default values. The detected compounds with an abundance of <1,000 and a signal-to-noise value of <50 were removed to discard identification errors or duplication. Compounds were identified by comparing the mass spectra and retention index with those in the NIST08 Mass Spectral Database, the Agilent Fiehn Metabolomics Retention Time Locked (RTL) Library, and an in-house library focused on tobacco metabolites. Standards for constructing the in-house library were prepared as described above. A SCAN + SIM method and an in-house automatic integration method were then established in the Agilent MSD Chem Station and applied to quantify the selective ion traces as previously described. Manual corrections were performed to guarantee the accuracy of the integration. A three-dimensional matrix was generated, including the sample

information, peak retention time, and relative peak intensities. Internal standards and any known artificial peaks, such as peaks caused by noise, column bleed, and derivatization procedures, were removed from the matrix.

## Bacterial Community Analysis

To collect microbes from the CTLs, ~5-g CTLs were added to a 250-ml flask containing 100-ml filtered normal saline (0.9% NaCl, pH 7) and oscillated at 4°C, 220 rpm for 4 h. CTLs were removed by gauze, and microbes were collected by centrifuging at 7,000 × *g* for 15 min. Total microbial genomic DNA was extracted using the DNeasy PowerSoil Kit (QIAGEN, Inc., Netherlands), following the manufacturer's instructions, and stored at −20°C until further analysis. The quality of extracted DNA was measured by agarose gel electrophoresis. PCR amplification of the bacterial and archaeal 16S rRNA genes V4–V5 region was performed using the forward primer 515F (Parada et al., 2016; 5′-GTGCCAGCMGCCGCGGTAA-3′) and the reverse primer 907R (Morales and Holben, 2009; 5′-CCGTCAATTCMTTTRAGTTT-3′). Sample-specific 7-bp barcodes were incorporated into the primers for multiplex sequencing. The final PCR volume was 25 μl and consisted of 5 μl of Q5 reaction buffer (5×), 5 μl of Q5 High-Fidelity GC buffer (5×), 0.25 μl of Q5 High-Fidelity DNA Polymerase (5 U/μl), 2 μl (2.5 mM) of dNTPs, 1 μl 515F primer (10 μM; final: 0.4 μM), 1 μl 806R primer (10 μM, final: 0.4 μM), 2 μl of DNA Template, and 8.75 μl nuclease-free water. The thermocycler conditions were: initial denaturation at 98°C for 2 min followed by 28 cycles consisting of denaturation at 98°C for 15 s, annealing at 55°C for 30 s, and extension at 72°C for 30 s, with a final extension of 5 min at 72°C. PCR products were purified with Agencourt AMPure Beads (Beckman Coulter, IN) and quantified using the PicoGreen dsDNA Assay Kit (Invitrogen, Carlsbad, CA, United States). After the individual quantification step, amplicons were pooled in equal amounts, and paired-end 2 × 300 bp sequencing was performed using the Illumina MiSeq platform with MiSeq Reagent Kit v3.

The 16S rRNA gene sequences were processed using QIIME 2 (Bolyen et al., 2019). Briefly, raw sequencing reads with exact matches to the barcodes were assigned to respective samples and identified as valid sequences. The low-quality sequences were filtered through the following criteria (Chen and Jiang, 2014): sequences that had a length of <150 bp, sequences that had average Phred scores of <20, sequences that contained ambiguous bases, and sequences that contained mononucleotide repeats of >8 bp. Paired-end reads were assembled using FLASH (Magoc and Salzberg, 2011). After chimera detection, the remaining high-quality sequences were clustered into amplicon sequence variants (ASVs). Taxonomic classification was performed using the plugin q2-feature-classifier using the classify-sklearn method (Pedregosa et al., 2011) and the pre-trained SILVA version 132 database (Quast et al., 2013) with 99% similarity. Alpha diversity indexes were calculated by using the command "qiime diversity alpha-rarefaction." PICRUSt (Phylogenetic Investigation of Communities by Reconstruction of Unobserved States) function prediction was performed using bacterial 16S rRNA sequencing data to determine the abundance of microbial functional genes in the KEGG metabolic pathway (Langille et al., 2013).

## Statistical Analysis

Heat maps and cluster analyses were performed in the statistical environment R v. 4.0.0. The Galaxy[1] was used for linear discriminant analysis effect size (LEfSe) analysis to assess significant differences of CTLs with different treatments. Additionally, the correlation between the representative bacteria and core VFCs based on Spearman's correlation coefficients ($p < 0.05$, $|r| > 0.3$), network analysis was performed by using Gephi software.

# RESULTS

## Bacterial Composition Varied After Inoculation and Fermentation

To explore the influence of the inoculation of *Acinetobacter* on original bacterial community, bacteria in CTLs with different treatments were sequenced. The high-throughput sequencing generated 642,151 sequence reads from sequence reads from 12 CTLs. After the quality control processing, including filtered, denoised, merged, and nonchimeric, there were 483,897 high-quality sequences, with an average of 40,324 sequences, and obtained 1,914 ASVs. The abundance of bacterial taxa was shown in **Figure 1**, *Proteobacteria*, *Firmicutes*, and *Bacteroidetes* were predominant phyla, *Aquabacterium*, *Bacillus*, *Acinetobacter*, *Muribaculaceae*, and *Pseudomonas* were predominant genera. There was no doubt that fermentation changed the composition of bacterial communities. At phyla level, the abundance of *Proteobacteria* and *Firmicutes* increased from 59.9%5 and 9.98% to 77.59% and 15.58% after fermentation, respectively, while the abundance of *Bacteroidetes* decreased from 23.10% to 2.16%. At genera level, the abundance of *Aquabacterium* increased from 14.26% to 41.08%, while the abundance of *Muribaculaceae* and *Pseudomonas* decreased from 17% and 14.04% to <0.01%. Similarly, external microorganisms also changed the composition of bacterial communities. At phyla level, the abundance of *Proteobacteria* increased from 59.95% to 95.61% after inoculation with *Acinetobacter* sp. 1H8 and decreased to 37.63% after inoculation with *Acinetobacter indicus* 3B2, while the abundance of *Bacteroidetes* decreased from 15.58% to 3.1% after inoculation with *Acinetobacter* sp. 1H8 and increased to 37.63% after inoculation with *Acinetobacter indicus* 3B2. At genera level, the abundance of *Aquabacterium* increased from 41.08% to 59.01% after inoculation with *Acinetobacter* sp. 1H8 and decreased to 27.61% after inoculation with *Acinetobacter indicus* 3B2. The abundance of *Bacillus* and *Acinetobacter* increased from 1.83% and 0.62% to 33.68% and 22.43% after inoculation with *Acinetobacter indicus* 3B2, respectively.

The alpha diversity of the bacterial community was evaluated by indices including Chao1 index, Shannon index, and Simpson index, in which the first and the latter two represent richness

[1]https://huttenhower.sph.harvard.edu/galaxy/

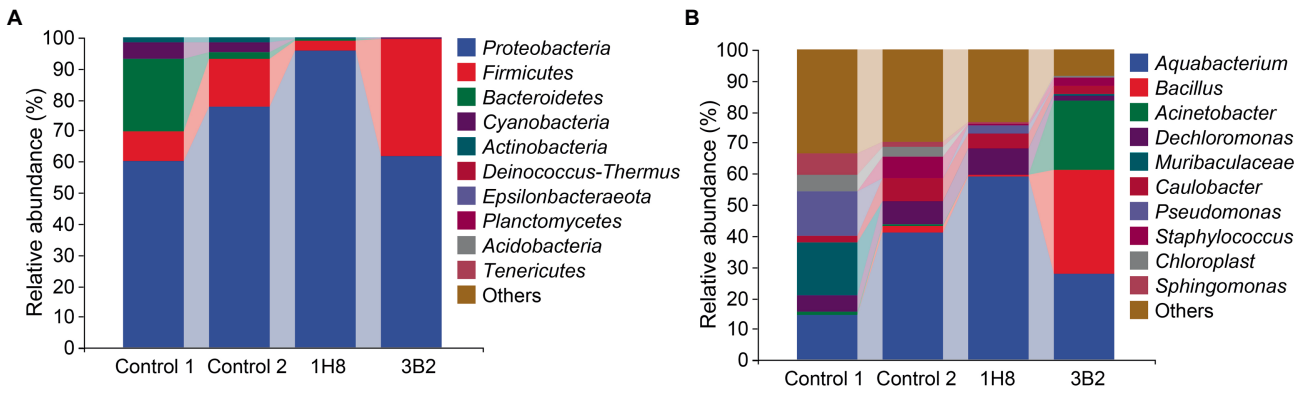



**FIGURE 1 |** Bacterial communities in cigar tobacco leaves. The top 10 predominant bacterial phyla **(A)** and genera **(B)**.

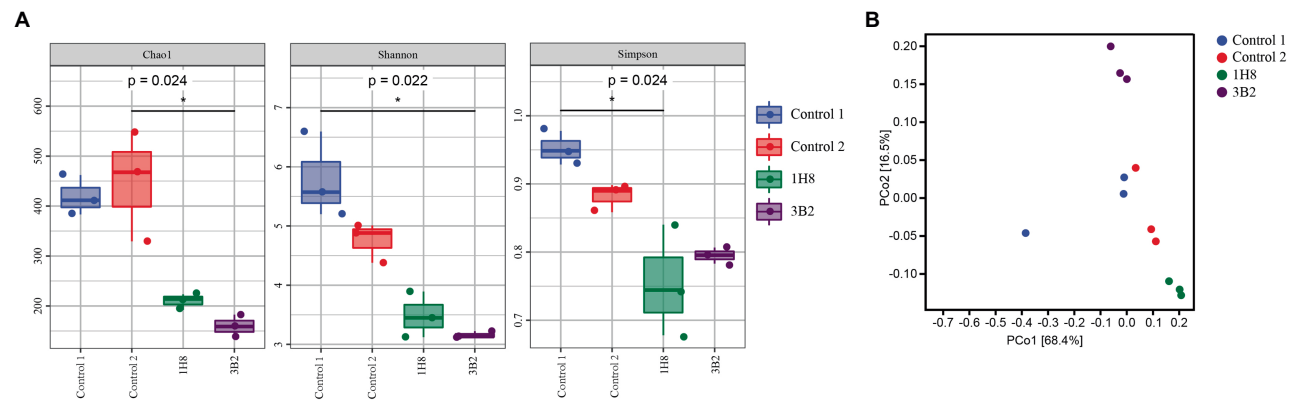



**FIGURE 2 |** Diversity of bacterial communities in cigar tobacco leaves. Alpha diversity was determined based on the Chao1 index, the Shannon index, and the Simpson index **(A)**, Beta diversity was measured by weighted UniFrac distance **(B)**.

and diversity, respectively (**Figure 2A**). They all showed external microorganisms significantly reduced the richness and evenness of original bacterial community. Unconstrained principal coordinate analysis (PCoA) of Weighted-unifrac distance revealed that the microbiota of CTLs with different treatment formed three distinct clusters (**Figure 2B**), which separated along the second coordinate axis. As expected, the inoculation of *Acinetobacter* indisputably altered the structure of bacterial community.

## Significant Different Microbes After Inoculation and Fermentation

To explore the different bacteria among CTLs with different treatments, LEfSe analysis was conducted to reveal the significant differences below the level of phylum (**Figure 3**). The circles from inner to outer represents bacteria classification from phylum to genus levels, and corresponding colors in every group denotes bacteria taxa with a significant difference. Notably, 81 different bacteria appeared in the LDA threshold of 2 judging by statistically significant differences ($p < 0.05$), which consist of 5 phyla, 4 classes, 18 orders, 25 families, and 29 genera. In detail, 17 genera were significantly enriched in uninoculated CTLs, such as *Rhodococcus*, *Blastococcus*, *Promicromonospora*, and *Sanguibacter*. Eight genera were significantly enriched in unfermented CTLs, such as *Lactobacillus*, *Sphingobium*, and *Rhizopus*. Two genera were significantly enriched in CTLs inoculated *Acinetobacter indicus* 3B2, such as *Bacillus* and *Acinetobacter*. Only *Aquabacterium* was significantly enriched in CTLs inoculated *Acinetobacter* sp. 1H8.

## Changes in Bacterial Metabolic Pathways After Inoculation and Fermentation

PICRUSt analysis results showed that there were more functional genes annotated to biosynthesis, degradation/utilization/assimilation, generation of precursor metabolite and energy, glycan pathways, macromolecule modification, and metabolic clusters (**Figure 4**). The cluster analysis identified the CTLs inoculated *Acinetobacter* sp. IH8 and uninoculated CTLs were classified into one cluster with the similar abundance of bacterial metabolic pathway, this may be because they have similar bacterial community. The addition of external microorganisms not only changed the structure of bacterial community, but also changed the function of bacterial community. The abundance of a large number of bacterial metabolic pathways decreased significantly after

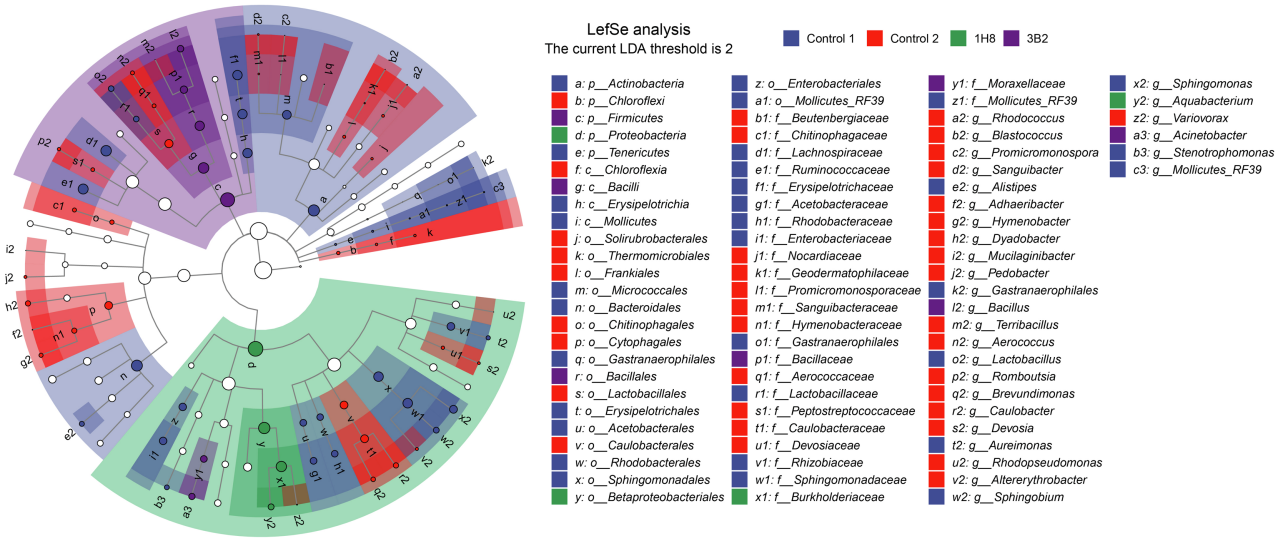



**FIGURE 3 |** Evolutionary branch map of the bacteria with significantly different in cigar tobacco leaves with different treatments.

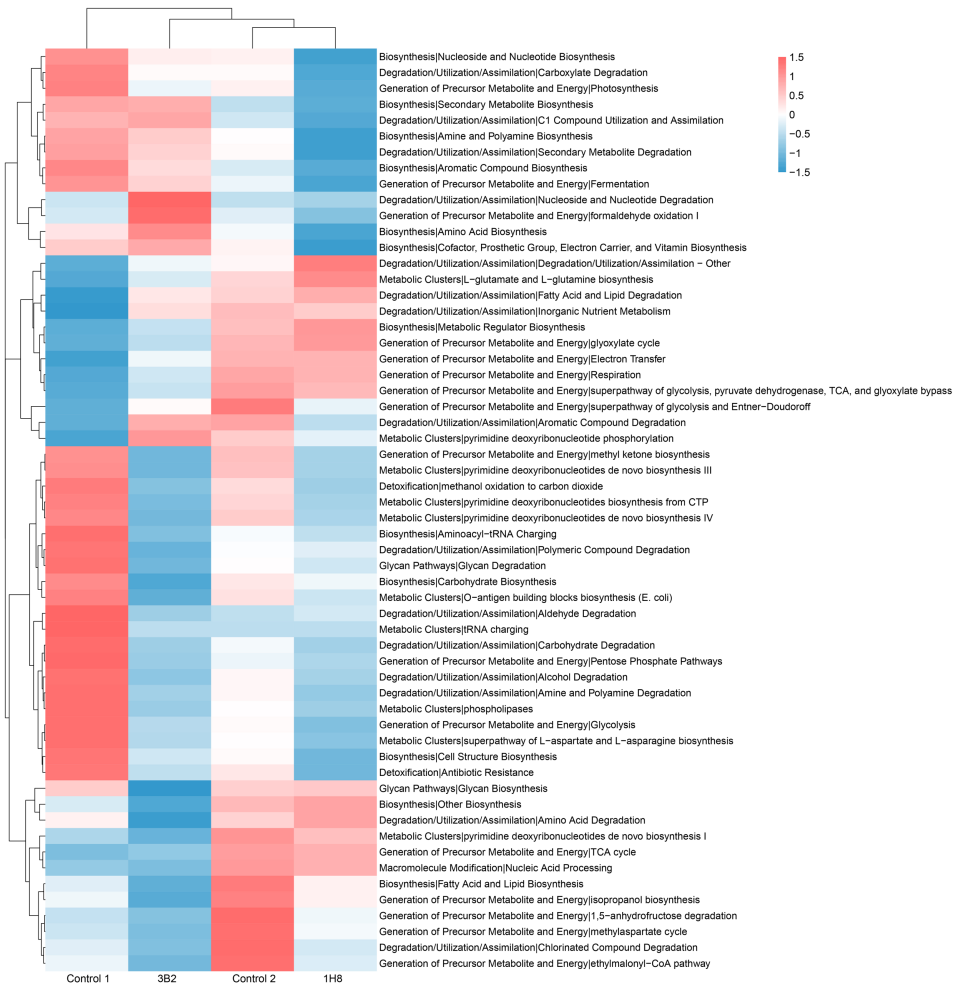



**FIGURE 4 |** The relative abundance of bacterial metabolic pathway in cigar tobacco leaves.

inoculation. LEfSe analysis was conducted to identify the significant different bacterial metabolic pathways in CTLs with different treatments (**Figure 5**). Metabolic pathways involved in biosynthesis, including fatty acid and lipid biosynthesis, amino acid biosynthesis, nucleoside and nucleotide biosynthesis, aromatic compound biosynthesis, and carbohydrate biosynthesis, were significantly enriched in unfermented CTLs. Metabolic pathways involved in degradation, including aromatic compound degradation, fatty acid and lipid degradation, nucleoside and nucleotide degradation, were significantly enriched in CTLs inoculated *Acinetobacter* sp. IH8. Metabolic pathways involved in cell growth, including cell structure biosynthesis, inorganic nutrient metabolism, amino acid biosynthesis, and TCA cycle, were significantly enriched in CTLs inoculated *Acinetobacter indicus* 3B2.

## Interactions Between Extrinsic and the Intrinsic Microbes

Microbial interactions are the main factors shaping community structure (Mould and Hogan, 2021). To elucidate interactions between the inoculated strains and the native bacteria, an association network was established based on bacterial abundance (Kumar et al., 2019). As showed in **Figure 6**, *Acinetobacter* was negative correlative with *Aquabacterium*, and was positive correlative with *Bacillus*. There must be a non-cooperative relationship between *Acinetobacter* and *Aquabacterium*, such as competition, parasitism, predation, and antagonism. There must also be a cooperative relationship of mutualism and symbiosis between *Acinetobacter* and *Bacillus*. The inoculated *Acinetobacter* 1H8 was inhibited by *Aquabacterium*, while inoculated *Acinetobacter indicus* 3B2 inhibited *Aquabacterium* and promoted *Bacillus*. They in turn affected other microbiotas, such as *Staphylococcus*, *Aerococcus*, *Lutispora*, and *Zoogloea*. In addition to the negative correlation between *Acinetobacter* and *Aquabacterium*, other interactions were positive.

## Changes in the Profiles of Volatile Flavor Compounds After Inoculation and Fermentation

*Acinetobacter* were found to be the main producer of produce aldehydes and ketones in CTLs in our previous studies. They could produce flavor-related aldehydes and ketones in a simple synthetic medium, such as benzaldehyde, phenylacetaldehyde, 4-hydroxy-3-ethoxy-benzaldehyde, and 3,5,5-trimethyl-2-cyclohexene-1-one. There is no doubt that they are able to produce more products in CTLs because they have more nutrients available. To investigate the effect of extrinsic strains on the metabolic activity of the original microbial community, the VFCs in CTLs were determined. In total, 37 VFCs were selected among hundreds of compounds based on content and co-detection for further analysis, which included 4 alcohols, 9 aldehydes, 13 ketones, 3 esters, 3 pyrazines, 3 alkene,1 furan, and 1 acid. As shown in **Figure 7A**, the cluster analysis divided unfermented and fermented CTLs into two clusters, the vast majority of VFCs have changed after fermentation. For example, benzeneacetaldehyde, benzaldehyde, and 3,7,11,15-tetramethyl-2-hexadecen-1-ol were obviously increased in uninoculated CTLs after fermentation. Megastigmatrienone, 2-ethyl-1-hexanol, tetramethyl-pyrazine, trimethyl-pyrazine,

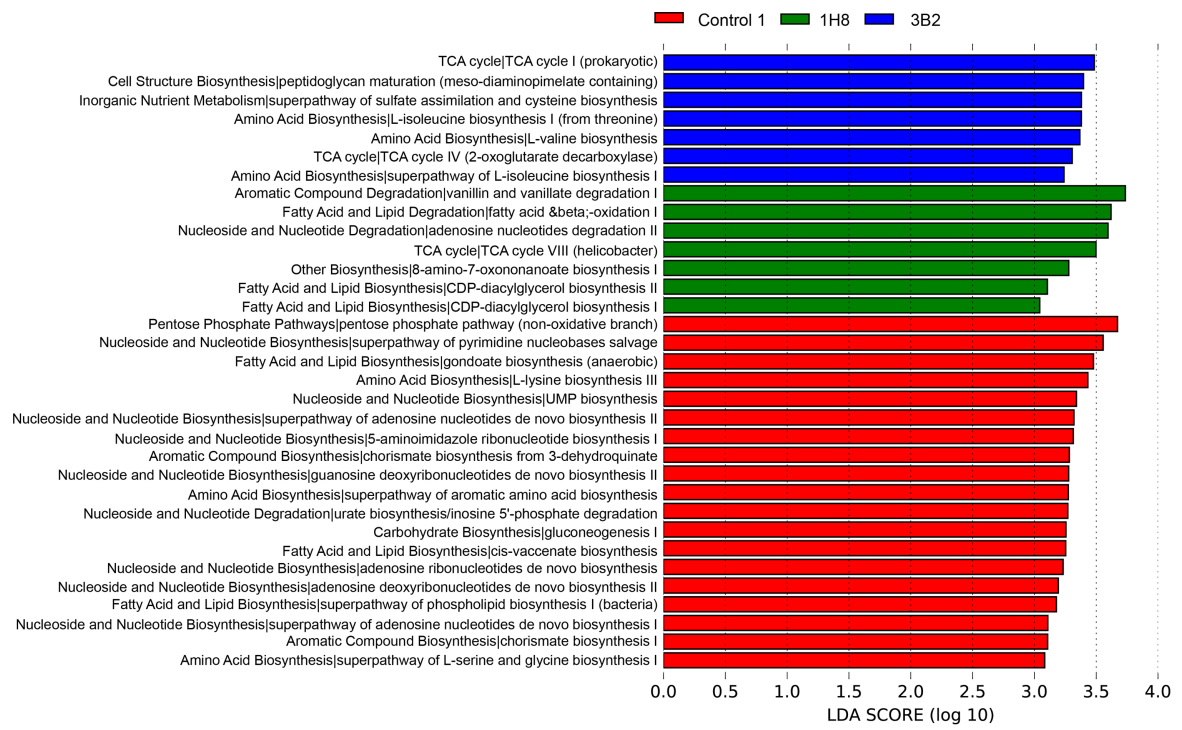



**FIGURE 5 |** Significantly different bacterial metabolic pathway in cigar tobacco leaves with different treatments.

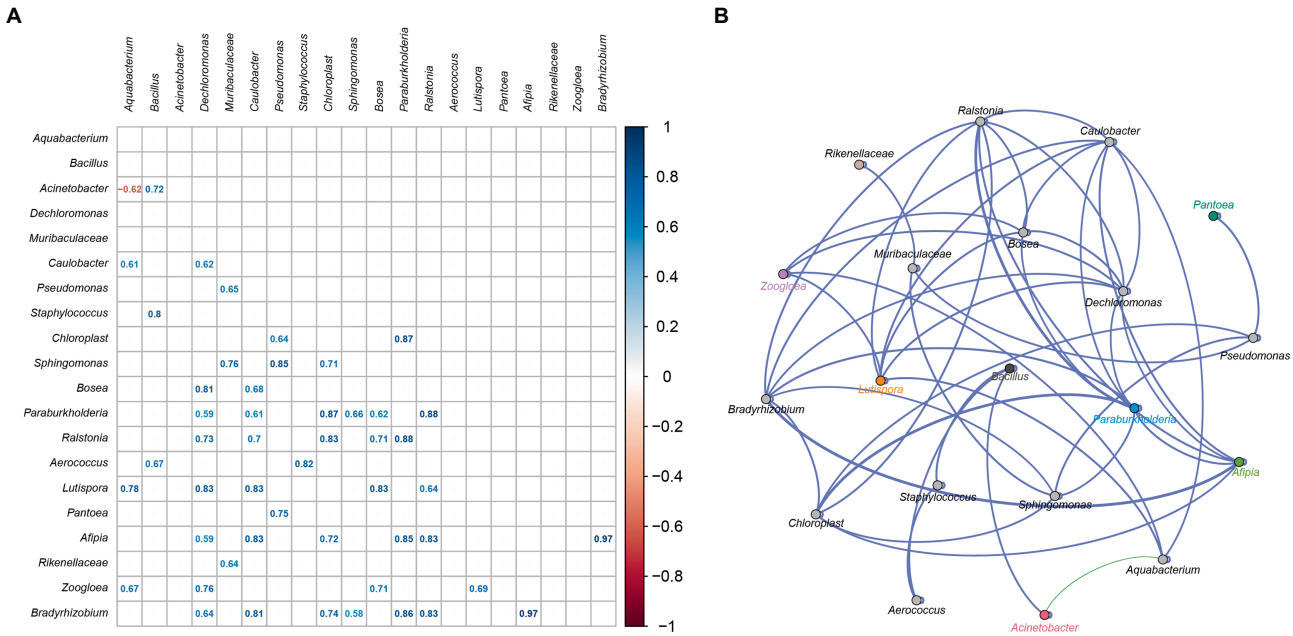



**FIGURE 6 |** The correlation between bacterial community. Heatmap analysis based on Spearman's correlation coefficients between representative bacterial taxa **(A)**. Co-occurrence networks of representative bacterial taxa. Green line means positive correlation; red line means negative correlation **(B)**.

neophytadiene, pyrazine, and 2,6-dimethyl were obviously increased in CTLs inoculated *Acinetobacter indicus* 3B2. 3,5,5-Trimethyl-2-cyclohexen-1-one, heptanal, solanone, megastigmatrienone 4, octanoic acid, ethyl ester, 3,5-octadien-2-one, and 6-methyl-5-hepten-2-one were obviously increased in CTLs inoculated *Acinetobacter* sp. 1H8. While ketoisophorone, acetic acid, 2-methyl-furan, 6-methyl-3,5-heptadiene-2-one, farnesyl acetone, limonene, (E)-6,10-dimethyl-5,9-undecadien-2-one, and (E)-3,7-dimethyl-2,6-octadienal were decreased after fermentation. VFCs of inoculated CTLs varied significantly compared to uninoculated CTLs.

LEfSe analysis was used to identify the significant different VFCs in CTLs with different treatments (**Figure 7B**). Among the 37 VFCs, 26 VFCs appeared in the LDA threshold of 2 judging by statistically significant differences ($p < 0.05$). In detail, six VFCs were significantly enriched in uninoculated CTLs, such as acetic acid, ethyl ester, benzeneacetaldehyde, and benzeneethanol. Eight VFCs were significantly enriched in unfermented CTLs, such as (E)-6,10-dimethyl-5,9-undecadien-2-one, acetic acid, and 6-methyl-3,5-heptadiene-2-one. Four VFCs were significantly enriched in CTLs inoculated *Acinetobacter indicus* 3B2, including neophytadiene, tetramethyl-pyrazine, 2,6-dimethyl-pyrazine, and megastigmatrienone. Seven VFCs were significantly enriched in CTLs inoculated *Acinetobacter* sp. 1H8, such as solanone, 6-methyl-5-hepten-2-one, and benzeneacetic acid, ethyl ester.

## Correlation Analysis of the Predominant Bacteria and VFCs

The changes of volatile flavor compounds were mainly caused by microbial metabolism, so it was very important to determine the relationship between bacteria and VFCs. The relationships between VFCs (*n*:37) and representative bacteria (*n*: 20) were analyzed by Spearman's correlation coefficients and visualized by Gephi. As shown in **Figure 8**, *Acinetobacter* were positively related to acetic acid, ethyl ester, 2-ethyl-1-hexanol, and megastigmatrienone, and negatively related to 6-methyl-5-hepten-2-one, hexanal, 3-methyl−butanal, acetophenone, benzeneacetic acid, ethyl ester, octanal, 3,5-octadien-2-one, and 2,6-dimethyl-2,6-octadiene. *Aquabacterium* and *bacillus*, which were significantly associated with *Acinetobacter*, were also related to many volatile flavor compounds. *Aquabacterium* was positively related to most volatile flavor compounds, while *Bacillus* was negatively related to most volatile flavor compounds.

## Changes in Flavor of Cigar Tobacco Leaves After Inoculation and Fermentation

Not surprising, the quality of CTLs was significantly changed after inoculation and fermentation. Surprisingly, however, the tasters gave the same evaluation to CTLs inoculated different strains. They felt the bean fragrance, mellow, smooth, aftertaste, sweetness, cleanliness, and aftertaste were significantly improved, and the impurities and irritation were reduced. By analyzing the function of VFCs, it was found that solanone may be the main force to improve the quality of CTLs, because it could reduce impurities and irritation, and increase the mellow, fluency, lingering, sweetness, cleanliness, and aftertaste of CTLs. Bean flavor might be composed of different combinations and proportions of VFCs.

# DISCUSSION

This work revealed that the quality of CTLs were improved by the addition of extrinsic microbes. Changes in the quality

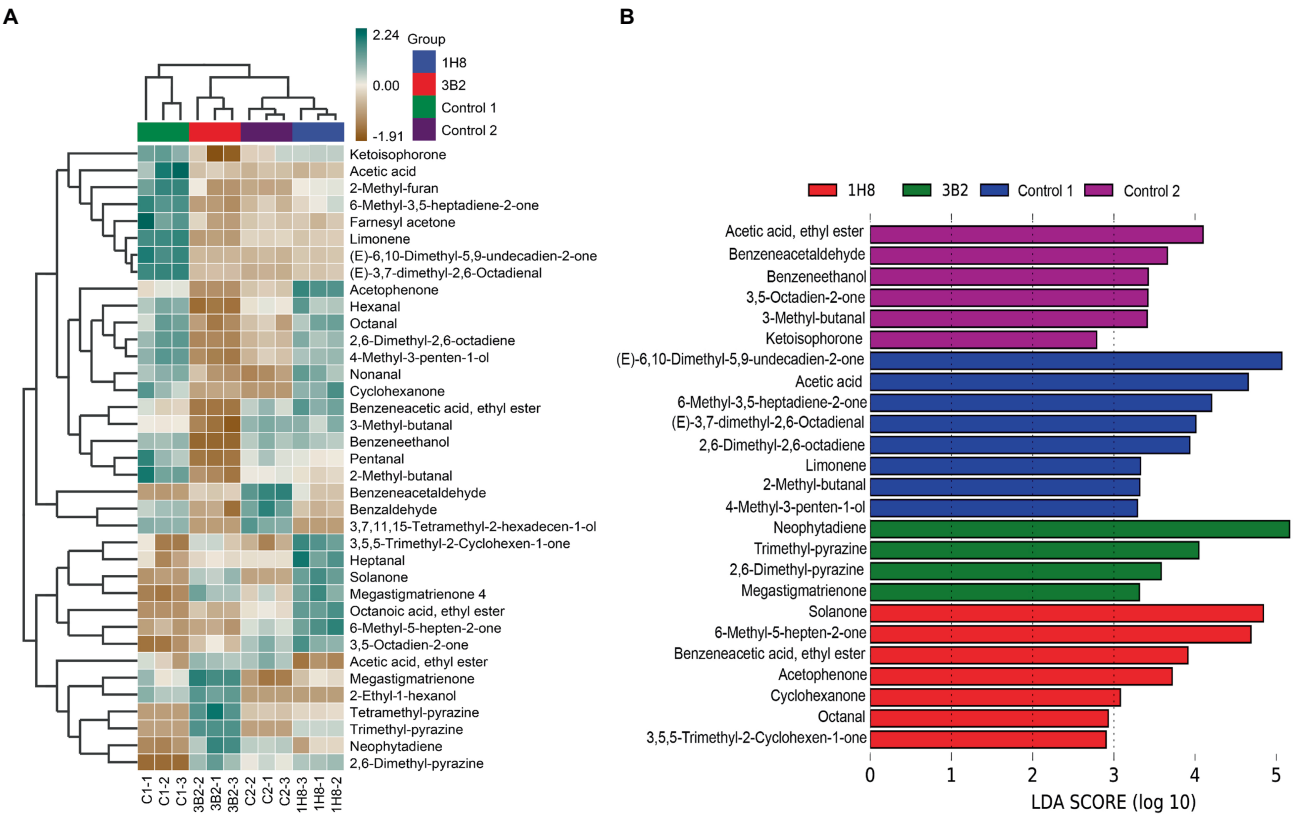



**FIGURE 7 |** Volatile flavor compounds in the cigar tobacco leaves. Hierarchical clustering of volatile flavor compounds in the cigar tobacco leaves **(A)**, Different volatile flavor compounds in different cigar tobacco leaves **(B)**.

of CTLs were not only related to the flavor production ability of the extrinsic microbes, but also to the interaction between the external and internal microbes. Microbial interactions were important forces in reconstructing the microbial community (Romdhane et al., 2022). The interactions between the external and internal microbes changed the original microbial community. Meanwhile, changes in microbial community led to variations in VFCs and quality of CTLs. Therefore, it is important to gain deep insight into these changes and the underlying reasons for these changes.

In this work, two extrinsic microbes (*Acinetobacter* sp. 1H8 and *Acinetobacter* indicus 3B2) were inoculated into CTLs, they exerted different effects on the original microbial community. The abundance of *Acinetobacter* remained stable in uninoculated CTLs after fermentation. When *Acinetobacter* sp. 1H8 was inoculated in CTLs, this strain was completely inhibited by the native microbiotas, the abundance of *Acinetobacter* was significantly reduced after fermentation. The inoculation of *Acinetobacter* sp. 1H8 caused the endogenous microbes to turn from competition to cooperation to compete together against the extrinsic microbes to suppress the extrinsic perturbation. Normally, the microbial community remains stable due to the complete suppression of extrinsic microbes by endogenous microbes (Wu et al., 2016). However, although *Acinetobacter* sp. 1H8 grew poorly, it still influenced the overall metabolic activity of the intrinsic microbes. *Aquabacterium* proliferated wildly upon stimulation by exogenous microorganisms, their abundance was significantly increased. Its influence has had a positive effect. From the perspective of the overall change of microbial community, the inoculation of *Acinetobacter* sp. 1H8 enhanced the normal evolutionary trend of bacterial community. Similar result was also found in in the cocultures of *Metschnikowia pulcherrima* and *S. cerevisiae*, the positive effect leads to increases in the types and amounts of metabolites, such as fatty acids, ethyl esters and acetates, and terpinol (Sadoudi et al., 2012). It may be helpful to promote the succession of microbial community, accelerated the fermentation process, and shortened the fermentation time (Zhang et al., 2021). When *Acinetobacter indicus 3B2* was inoculated in CTLs, they were successfully colonized CTLs, the abundance of *Acinetobacter* was significantly increased after fermentation, meanwhile, they also caused an increase in the abundance of *Bacillus* and a decrease in the abundance of *aquabacterium*. The inoculation of *Acinetobacter indicus* 3B2 greatly changed structure of the microbial community in CTLs. *Acinetobacter indicus 3B2* might be considered keystone specie in CTLs, which has an extremely high impact on a particular ecosystem (Rottjers and Faust, 2019). It was also critical for the overall structure and function of an ecosystem (Banerjee et al., 2018). The inoculation of *Acinetobacter* also significantly changed the metabolic functions in bacterial

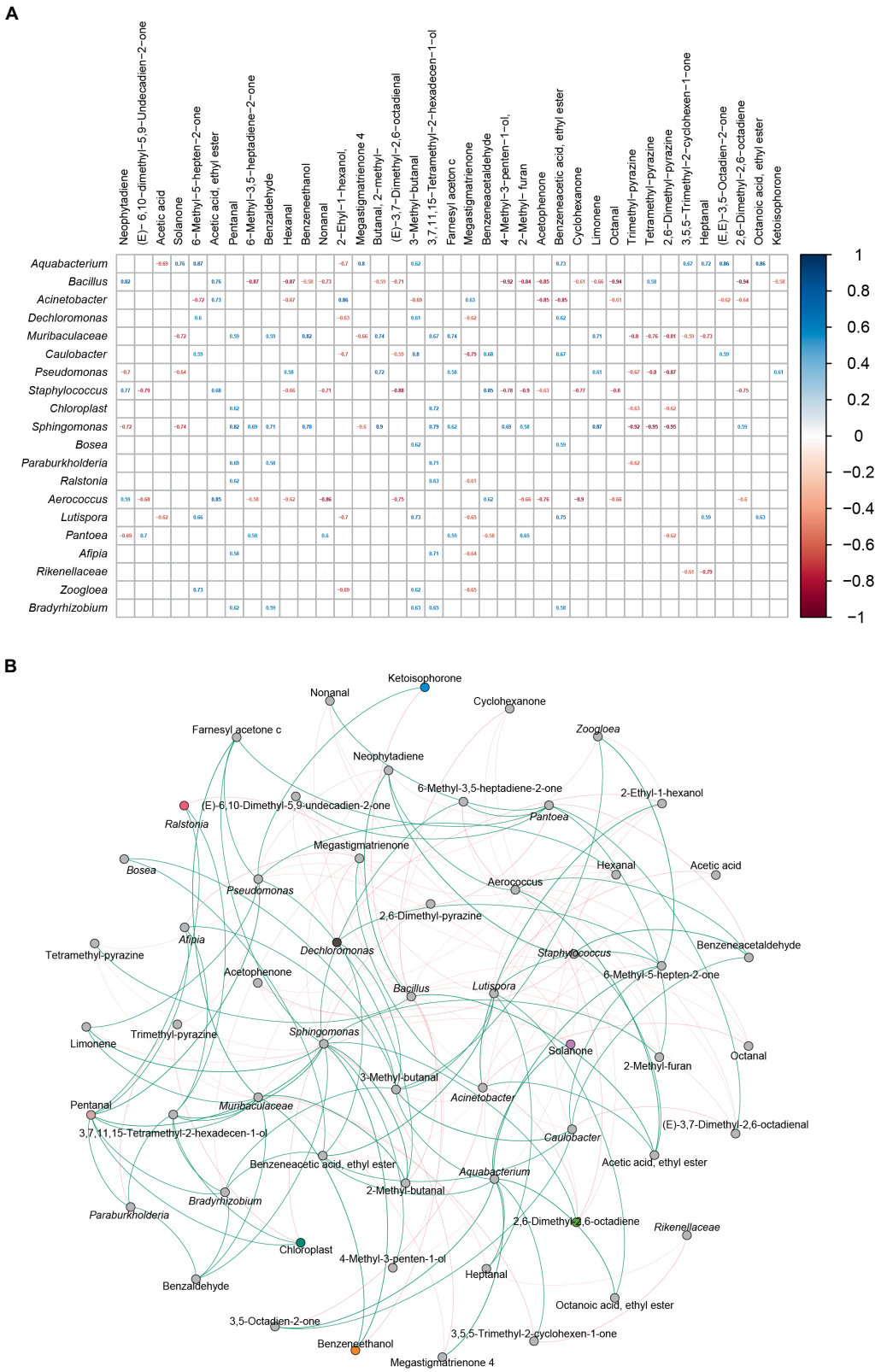



**FIGURE 8 |** The correlation between bacterial community and volatile flavor compounds. Heatmap analysis based on Spearman's correlation coefficients between representative bacteria and core volatile flavor compounds **(A)**. Co-occurrence networks of representative bacterial taxa and core volatile flavor compounds. Green line means positive correlation; red line means negative correlation **(B)**.

communities. Metabolic pathways involved in degradation were significantly enriched in CTLs inoculated *Acinetobacter* sp. IH8. The inoculation of *Acinetobacter* sp. IH8 significantly increased the degradation ability of macromolecular substances in microbial community, it would increase precursors of VFCs. The inoculation of *Acinetobacter indicus* 3B2 significantly promoted the growth of some functional microbiotas, such as *bacillus* and *Acinetobacter*, which were found to be the main producers of aldehyde and ketones in CTLs in our previous study. Microbial interaction analysis found *Acinetobacter* was negative correlative with *Aquabacterium*, and was positive correlative with *Bacillus*.

*Acinetobacter* were found to be the main producer of produce aldehydes and ketones in CTLs in our previous studies. They could produce flavor-related aldehydes and ketones in a simple synthetic medium, such as benzaldehyde, phenylacetaldehyde, 4-hydroxy-3-ethoxy-benzaldehyde, and 3,5,5-trimethyl-2-cyclohexene-1-one. There is no doubt that they are able to produce more products in CTLs because they have more nutrients available. Their inoculation also greatly changed volatile flavor compound profile of CTLs. The inoculation of *Acinetobacter* sp. 1H8 significantly increased the content of solanone, 6-methyl-5-hepten-2-one, benzeneacetic acid, ethyl ester, cyclohexanone, octanal, acetophenone, and 3,5,5-trimethyl-2-cyclohexen-1-one. The inoculation of *Acinetobacter indicus* 3B2 significantly increased the content of trimethyl-pyrazine, 2,6-dimethyl-pyrazine, and megastigmatrienone. These VFCs have an important contribution to the flavor of CTLs. For example, solanone may be the main force to reduce impurities and irritation, and increase the mellow, fluency, lingering, sweetness, cleanliness, and aftertaste of CTLs. Its content was greatly increased in CTLs inoculated both microorganisms. The increase of pyrazine would enhance baked, roasted, rosy, and honey-like aroma. However, due to the increase and decrease of a variety of VFCs, the mixed flavor compounds produce a new flavor. Bean flavor might be composed of a variety of VFCs, and there were differences in the composition of the two CTLs. This may be the main reason why CTLs inoculated different microorganisms showed a flavor characteristic.

## CONCLUSION

When some traditional fermented products that are produced by spontaneous fermentation unable to meet the demands of consumers, inoculating extrinsic microbes may improve these products. In this work, we demonstrate that the inoculation of two extrinsic microbes (*Acinetobacter* sp. 1H8 and *Acinetobacter indicus* 3B2) improved the quality and flavor of CTLs. Inoculated microbes can not only exert their own metabolic ability, but also affect the structure and function of native microbial community. We revealed that the interaction between exogenous microorganisms and native microbes, the different effects of exogenous microorganisms on original microbial community, the association between microbes and VFCs, and the formation mechanism of tobacco flavor. Collectively, our present work has proved the effect of inoculated microorganisms in traditional fermentation industry and elucidated changes in the overall structure and function of microbial communities after inoculation. These results suggest that controlling the microbial community could significantly improve the quality and safety of fermentation products, and this work may provide a good way to gain insight into the microbial ecosystem of traditional fermentation.

## 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: NCBI SRA BioProject, accession no: PRJNA813020.

## AUTHOR CONTRIBUTIONS

TZ: conceptualization, data curation, formal analysis, methodology, software, and writing—original draft. QiZ, QW, and PL: investigation, methodology, and resources. XW, QuZ, and WC: methodology, resources, and project administration. JZ, GD, and DL: funding acquisition, supervision, and writing—review and editing. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was supported by the National Key Research and Development Program of China (2019YFC1605800), the China National Tobacco Corporation 2020 Major Science and Technology Project 110202001040(XJ-02), and the Major projects on constructing the "mellow and sweet fragrance styles" of Chinese-Stylistic Tobacco (ctx201905).

## REFERENCES

Banerjee, S., Schlaeppi, K., and van der Heijden, M. G. A. (2018). Keystone taxa as drivers of microbiome structure and functioning. *Nat. Rev. Microbiol.* 16, 567–576. doi: 10.1038/s41579-018-0024-1

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N., 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

Chen, H., and Jiang, W. (2014). Application of high-throughput sequencing in understanding human oral microbiome related with health and disease. *Front. Microbiol.* 5:508. doi: 10.3389/fmicb.2014.00508

Ferreira, A. D., and Viljoen, B. C. (2003). Yeasts as adjunct starters in matured Cheddar cheese. *Int. J. Food Microbiol.* 86, 131–140. doi: 10.1016/S0168-1605(03)00252-6

Ghosh, S., Chowdhury, R., and Bhattacharya, P. (2016). Mixed consortia in bioprocesses: role of microbial interactions. *Appl. Microbiol. Biotechnol.* 100, 4283–4295. doi: 10.1007/s00253-016-7448-1

Jin, G. Y., Zhu, Y., and Xu, Y. (2017). Mystery behind Chinese liquor fermentation. *Trends Food Sci. Technol.* 63, 18–28. doi: 10.1016/j.tifs.2017.02.016

Kawagoshi, Y., Hino, N., Fujimoto, A., Nakao, M., Fujita, Y., Sugimura, S., et al. (2005). Effect of inoculum conditioning on hydrogen fermentation and pH effect on bacterial community relevant to hydrogen production. *J. Biosci. Bioeng.* 100, 524–530. doi: 10.1263/jbb.100.524

Kumar, M., Ji, B., Zengler, K., and Nielsen, J. (2019). Modelling approaches for studying the microbiome. *Nat. Microbiol.* 4, 1253–1267. doi: 10.1038/s41564-019-0491-9

Langille, M. G. I., Zaneveld, J., Caporaso, J. G., McDonald, D., Knights, D., Reyes, J. A., et al. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat. Biotechnol.* 31, 814–821. doi: 10.1038/nbt.2676

Liu, Z., Li, J., Wei, B., Huang, T., Xiao, Y., Peng, Z., et al. (2019). Bacterial community and composition in Jiang-shui and Suan-cai revealed by high-throughput sequencing of 16S rRNA. *Int. J. Food Microbiol.* 306:108271. doi: 10.1016/j.ijfoodmicro.2019.108271

Liu, F., Wu, Z. Y., Zhang, X. P., Xi, G. L., Zhao, Z., Lai, M., et al. (2021). Microbial community and metabolic function analysis of cigar tobacco leaves during fermentation. *Microbiology* 10:e1171. doi: 10.1002/mbo3.1171

Magoc, 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

Morales, S. E., and Holben, W. E. (2009). Empirical testing of 16S rRNA gene PCR primer pairs reveals variance in target specificity and efficacy not suggested by in silico analysis. *Appl. Environ. Microbiol.* 75, 2677–2683. doi: 10.1128/AEM.02166-08

Mould, D. L., and Hogan, D. A. (2021). Intraspecies heterogeneity in microbial interactions. *Curr. Opin. Microbiol.* 62, 14–20. doi: 10.1016/j.mib.2021.04.003

Nawaz, M. Z., Subin Sasidharan, R., Alghamdi, H. A., and Dang, H. (2022). Understanding interaction patterns within deep-sea mcrobial communities and their potential applications. *Mar. Drugs* 20:108. doi: 10.3390/md20020108

Parada, A. E., Needham, D. M., and Fuhrman, J. A. (2016). Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ. Microbiol.* 18, 1403–1414. doi: 10.1111/1462-2920.13023

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: machine learning in python. *J. Mach. Learn. Res.* 12, 2825–2830.

Pierce, E. C., and Dutton, R. J. (2022). Putting microbial interactions back into community contexts. *Curr. Opin. Microbiol.* 65, 56–63. doi: 10.1016/j.mib.2021.10.008

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, 590–596. doi: 10.1093/nar/gks1219

Reid, J. J., McKinstry, D. W., and Haley, D. E. (1937). The fermentation of cigar-leaf tobacco. *Science* 86:404. doi: 10.1126/science.86.2235.404.a

Romdhane, S., Spor, A., Aubert, J., Bru, D., Breuil, M. C., Hallin, S., et al. (2022). Unraveling negative biotic interactions determining soil microbial community assembly and functioning. *ISME J.* 16, 296–306. doi: 10.1038/s41396-021-01076-9

Rottjers, L., and Faust, K. (2019). Can we predict keystones? *Nat. Rev. Microbiol.* 17:193. doi: 10.1038/s41579-018-0132-y

Sadoudi, M., Tourdot-Marechal, R., Rousseaux, S., Steyer, D., Gallardo-Chacon, J. J., Ballester, J., et al. (2012). Yeast-yeast interactions revealed by aromatic profile analysis of sauvignon Blanc wine fermented by single or co-culture of non-Saccharomyces and Saccharomyces yeasts. *Food Microbiol.* 32, 243–253. doi: 10.1016/j.fm.2012.06.006

Shilei, W., Qun, W., Yao, N., Jianfeng, W., and Yan, X. (2019). Construction of synthetic microbiota for reproducible flavor compound metabolism in Chinese light-aroma-type liquor produced by solid-state fermentation. *Appl. Environ. Microbiol.* 85, e03090–e03018. doi: 10.1128/AEM.03090-18

Smyth, E. M., Chattopadhyay, S., Babik, K., Reid, M., Chopyk, J., Malayil, L., et al. (2019). The bacterial communities of little cigars and cigarillos are dynamic over time and varying storage conditions. *Front. Microbiol.* 10:2371. doi: 10.3389/fmicb.2019.02371

Song, Z., Du, H., Zhang, Y., and Xu, Y. (2017). Unraveling core functional microbiota in traditional solid-state fermentation by high-throughput amplicons and metatranscriptomics sequencing. *Front. Microbiol.* 8:1294. doi: 10.3389/fmicb.2017.01294

Wang, P., Wu, Q., Jiang, X., Wang, Z., Tang, J., and Xu, Y. (2017). *Bacillus licheniformis* affects the microbial community and metabolic profile in the spontaneous fermentation of Daqu starter for Chinese liquor making. *Int. J. Food Microbiol.* 250, 59–67. doi: 10.1016/j.ijfoodmicro.2017.03.010

Wu, Q., Kong, Y., and Xu, Y. (2016). Flavor profile of Chinese liquor is altered by interactions of intrinsic and extrinsic microbes. *Appl. Environ. Microbiol.* 82, 422–430. doi: 10.1128/AEM.02518-15

Wu, Q., Zhu, Y., Fang, C., Wijffels, R. H., and Xu, Y. (2021). Can we control microbiota in spontaneous food fermentation?-Chinese liquor as a case example. *Trends Food Sci. Technol.* 110, 321–331. doi: 10.1016/j.tifs.2021.02.011

Zhang, Z. J., Li, X., Hu, X., Zhang, S., Li, A. N., Deng, Y. C., et al. (2021). Downward aeration promotes static composting by affecting mineralization and humification. *Bioresour. Technol.* 338:125592. doi: 10.1016/j.biortech.2021.125592

**Conflict of Interest:** QiZ, DL, PL, QuZ, and WC were employed by China Tobacco Sichuan Industrial Co., Ltd.

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.

*Copyright © 2022 Zheng, Zhang, Wu, Li, Wu, Li, Zhou, Cai, Zhang and Du. 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.*

frontiers | Frontiers in Microbiology

ORIGINAL RESEARCH
published: 07 July 2022
doi: 10.3389/fmicb.2022.906345

# Discovery, Yield Improvement, and Application in Marine Coatings of Potent Antifouling Compounds Albofungins Targeting Multiple Fouling Organisms

*Weiyi She[1,2,3], Wei Ye[1,2], Aifang Cheng[1,2], Wenkang Ye[1,2,3], Chunfeng Ma[4], Ruojun Wang[1,2], Jinping Cheng[1,2], Xuan Liu[1,2], Yujing Yuan[1,2], Sin Yu Chik[1,2], Jessie James Limlingan Malit[1,2], Yanhong Lu[2,3], Feng Chen[5]* and Pei-Yuan Qian[1,2]**

[1] Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Guangzhou, China, [2] Department of Ocean Science and Hong Kong Brach of Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Hong Kong University of Science and Technology, Hong Kong, China, [3] SZU-HKUST Joint PhD Program in Marine Environmental Science, Shenzhen University, Shenzhen, China, [4] Faculty of Materials Science and Engineering, South China University of Technology, Guangzhou, China, [5] Institute for Advanced Study, Shenzhen University, Shenzhen, China

OPEN ACCESS

**Edited by:**
Niranjan Koirala,
Gandaki Province Academy of Science and Technology, Nepal

**Reviewed by:**
Monica Butnariu,
Banat University of Agricultural Sciences and Veterinary Medicine, Romania
Kshitiz Raj Shrestha,
St. Xavier's College, Maitighar, Nepal

***Correspondence:**
Feng Chen
sfchen@szu.edu.cn
Pei-Yuan Qian
boqianpy@ust.hk

**Specialty section:**
This article was submitted to Microbiotechnology, a section of the journal Frontiers in Microbiology

**Received:** 28 March 2022
**Accepted:** 06 June 2022
**Published:** 07 July 2022

**Citation:**
She W, Ye W, Cheng A, Ye W, Ma C, Wang R, Cheng J, Liu X, Yuan Y, Chik SY, Limlingan Malit JJ, Lu Y, Chen F and Qian P-Y (2022) Discovery, Yield Improvement, and Application in Marine Coatings of Potent Antifouling Compounds Albofungins Targeting Multiple Fouling Organisms. Front. Microbiol. 13:906345. doi: 10.3389/fmicb.2022.906345

Marine biofouling caused huge economic losses of maritime industries. We aim to develop high-efficient, less-toxic, and cost-effective antifoulants to solve the problems of biofouling. In this study, we described the antifouling compounds albofungin and its derivatives (albofungin A, chrestoxanthone A, and chloroalbofungin) isolated from the metabolites of bacterium *Streptomyces chrestomyceticus* BCC 24770, the construction of high-yield strains for albofungin production, and application of albofungin-based antifouling coatings. Results showed that these albofungins have potent antibiofilm activities against Gram-positive and Gram-negative bacteria and anti-macrofouling activities against larval settlement of major fouling organisms with low cytotoxicity. With the best antifouling activity and highest yield in bacterial culture, albofungin was subsequently incorporated with hydrolyzable and degradable copolymer to form antifouling coatings, which altered biofilm structures and prevented the settlement of macrofouling organisms in marine environments. Our results suggested that albofungins were promising antifouling compounds with potential application in marine environments.

**Keywords: antifouling activity, marine coatings, albofungin yield improvement, *Streptomyces chrestomyceticus*, cluster-situated regulator**

## INTRODUCTION

Marine biofouling is the natural colonization of undesirable micro and macro-organisms on submerged artificial surfaces. These organisms are roughly bifurcated into microfoulers such as bacteria and diatoms and macrofoulers such as barnacles, mussels, tubeworms, and bryozoans (Callow and Callow, 2002). The bacterial biofilm is composed of diverse bacterial colonies of

cells that form a community. It is a multidimensional (3-D) aggregation of bacteria attached to a surface and enclosed in an extracellular polymeric matrix, which is composed of polysaccharides, proteins, environmental DNA, and phospholipids (Srinivasan et al., 2021). The colonization through biofilm formation by microfoulers attracts macrofoulers to attach, such as the tubeworm *Hydroides elegans*, which prefer to settle on biofilm surfaces (Ralston and Swain, 2009). Compared to planktonic bacteria, biofilm bacteria are 10–1,000 times more resistant to antibiotics (Stewart and Costerton, 2001). Nearly all marine structures in seawater are colonized with biofilm-forming bacteria. Biofoulers colonize ship hulls, underwater pipelines, and marinas, causing a substantial economic loss in marine operations and environmental problems such as the introduction of invasive species (Hellio and Yebra, 2009). For example, ships need additional power due to the extra burden from biofouler attachment, leading to high fuel consumption and heavy engine stress (Bixler and Bhushan, 2012).

Various antifouling technologies have been developed to prevent biofouling, including the addition of antifouling agents such as copper pyrithione, chlorothalonil, zinc pyrithione, and SeaNine 211 to marine paints (Qian et al., 2013). Besides these, antifouling paint containing biocide tributyltin (TBT) is highly efficient in preventing the settlement and growth of biofoulers but it is toxic toward non-target organisms and persistent in marine environments; hence, its application on ships has been banned by the International Maritime Organization (IMO) (Champ, 2000). Thus, there is an urgent need to develop non-toxic, cost-effective, and environmentally friendly antifoulants.

In recent years, antifouling compounds derived from various natural sources have been discovered and considered to be degradable and low toxic to the marine environment (Qi and Ma, 2017). Especially, bacterial sources of bioactive compounds are preferable as they can be reproduced and scaled up for ensuring product supply for commercialization (Heidarian et al., 2019). It is undeniable that natural products isolated from bacterial fermentation are vast resources for the exploitation of antifouling compounds. For example, butenolide isolated from the marine bacterium *Streptomyces albidoflavus* prevents the settlement of dominant fouling organisms (barnacles, tubeworms, and bryozoans) with a low toxic effect (Xu et al., 2010); 3,3-Diindolylmethane isolated from *Pseudovibrio denitrificans* exhibits antifouling activities against barnacles and bryozoans with equivalent field performance to that of the commercial antifouling agent SeaNine 211 (Wang et al., 2015). Previously reported natural products with antifouling activity also include fatty acids, lactones, terpenes, steroids, benzenoids, phenyl ethers, polyketides, alkaloids, nucleosides, and peptides (Wang et al., 2017). However, it is difficult to develop marine natural product-based antifouling coatings because of the quick release of antifouling compounds, the complicated procedure for chemical synthesis, and the low yields of the antifoulants (Sisson et al., 2013). To solve problems related to compound release control, the self-polishing copolymer, which uses silyl acrylate as a carrier of antifoulants, generates a self-renewing surface through hydrolysis and degradation and thus, controls antifouling compound release (Bressy et al., 2010; Xie et al., 2019). To solve the compound supply issue, chemical synthesis or biosynthesis of target antifouling compounds are common practices. Yet, there have been no studies that address these challenges in a holistic (systematic) way.

Albofungin and its derivatives were identified as secondary metabolites from various *Streptomyces* species, and they belong to polycyclic isoquinolone–xanthone family. Albofungin and chloroalbofungin have been crystallized and their structures were determined by single-crystal X-ray diffraction (Ye et al., 2020). Our previous study has reported remarkable antibacterial activities against "ESKAPE pathogens" and antitumor activities of albofungin and its derivatives isolated from the bacterium *S. chrestomyceticus* BCC 24770 (She et al., 2021). In the present study, the antifouling activities of four albofungins were firstly evaluated. Since the albofungin biosynthesis pathways and their regulatory genes in *S. chrestomyceticus* have been already known (She et al., 2021), it is a natural extension of our current work to develop high-yield engineered strains to improve albofungin production through manipulation of its biosynthesis gene cluster in *S. chrestomyceticus*. Based on their structure-activity relationship and production analyses, the albofungin-based coatings were further prepared and evaluated in marine environments.

# MATERIALS AND METHODS

## General Experimental Procedures

The strains and plasmids are listed in **Table 1**. *Staphylococcus* sp. Z01 and *Micrococcus* sp. Z02 were isolated from marine biofilm grown on Petri dishes (Corning Inc., New York, United States) in a subtidal zone as described in Wang et al. (2020). The 16S rRNA genes amplicon was performed using 8F/1492R primers, followed by Sanger sequencing in BGI (Beijing, China). BLAST searches on the NCBI 16S ribosomal RNA sequences database and EzBioCloud database was performed on the obtained sequences to identify the taxonomy of the isolates. The 16S rRNA gene sequence of *Micrococcus* sp. has 94.7% of similarity *to Micrococcus yunnanensis*, and that of *Staphylococcus* sp. has 96.2% of similarity *to Staphylococcus warneri*.

Albofungins were isolated and purified as previously described (She et al., 2021). The purity of tested compounds was confirmed by high-performance liquid chromatography (95% purity, HPLC, Waters 2695, Milford, MA, United States), and their structures were determined by the Bruker NMR spectrometers (Bruker, Billerica, MA, United States) and X-ray crystallography as previously described (Ye et al., 2020; She et al., 2021).

## Assessment of Biofilm Formation by MTT Assay

The ability to form static biofilms of the 6 marine bacteria was tested. Marine bacteria were cultured overnight in marine broth at 22°C (*Pseudomonas pachastrellae* MCCC 1A01390, *Sulfitobacter pontiacus* MCCC 1A04899, and *Psychrobacter nivimaris* MCCC 1A11723) or 30°C (*Staphylococcus aureus* B04, *Staphylococcus* sp. Z01, and *Micrococcus* sp. Z02) with 220 rpm

**TABLE 1 |** Bacteria strains and plasmids.

| Strain or plasmid | Characteristics | References |
|---|---|---|
| *Escherichia coli* TOP10 | Cloning host | O'Sullivan et al., 2001 |
| *E. coli* ET12567/pUZ8002 | *Streptomyces* conjugation | Guan and Pettis, 2009 |
| MCCC 1A01390 *Pseudomonas pachastrellae* | Biofilm formation strain | Marine Culture Collection of China (MCCC) |
| MCCC 1A04899 *Sulfitobacter pontiacus* | Biofilm formation strain | Marine Culture Collection of China (MCCC) |
| MCCC 1A11723 *Psychrobacter nivimaris* | Biofilm formation strain | Purchased from the Marine Culture Collection of China (MCCC) |
| *Micrococcus* sp. Z02 | Biofilm formation strain | This study |
| *Staphylococcus aureus* B04 | Biofilm formation strain | Culture collection of our laboratory |
| *Staphylococcus* sp. Z01 | Biofilm formation strain | This study |
| *Streptomyces* strains | | |
| *Streptomyces chrestomyceticus* BCC 24770 | Parental strain to produce albofungin | |
| 24770/pPWW-*alb45* | Overexpression strain | This study |
| 24770/pPWW-*alb22* | Overexpression strain | This study |
| **Plasmids** | | |
| pPWW50a | Cloning and expression vector | Malit et al., 2021 |
| pPWW-*alb45* | pPWW50a with *alb45* expression under the control of the strong constitutive promoter *ermE**p | This study |
| pPWW-*alb22* | pPWW50a with *alb22* expression under the control of the strong constitutive promoter *ermE**p | This study |

agitation and then diluted into approximately $10^7$ CFU $mL^{-1}$ in marine broth supplemented with 1% of glucose. Afterward, 200 μL of the diluted solution were added to each well of a 96-well plate (Corning Inc., New York, United States) and was then incubated at 22 or 30°C for 24 h. The culture medium, planktonic cells and loosely adhered bacteria were removed by dual washing with a phosphate-buffered saline (PBS) buffer, and the firmly attached bacteria were incubated with 20 μL of MTT (5 mg $mL^{-1}$) at 37°C for 3 h. The supernatant was discarded, and formazan was dissolved in 150 μL of 100% DMSO (Sigma-Aldrich, St. Louis, United States). Absorbance was measured using the Multiskan$^{TM}$ FC microplate photometer (Thermo Fisher Scientific, Waltham, United States) under 570 nm, and *Staphylococcus aureus* ATCC 43300 was used as a positive control. All the experiments were performed in triplicate.

## Antibacterial Assay and Antibiofilm Assay

The antibacterial activities of albofungins against the marine bacteria were tested following specific protocols (She et al., 2020). Marine broth was used as the test medium for the marine bacteria. The plate was kept at 22 or 30°C overnight. The minimum inhibitory concentration (MIC) which means the lowest concentration that the drug prevents the visible growth of bacteria was used to evaluate antibacterial assay.

The marine bacteria which have successfully formed biofilms were further used for antibiofilm assays. Different concentrations of albofungin compounds were added to each well. Biofilm formation was assessed through the MTT assay as previously described. The minimum biofilm inhibitory concentration ($MBIC_{90}$) which refers to the lowest concentration of a drug to effectively inhibit 90% of the biofilm formation was calculated. Data were analyzed using one-way ANOVA to detect significant differences and standard deviation (SD) was calculated using GraphPad Prism 9. All the assays were performed in triplicate.

## Collection, Culturing, and Anti-larval Settlement Bioassay of Barnacle *Amphibalanus amphitrite* Larvae and Bryozoan *Bugula neritina* Larvae for Anti-larval Settlement Bioassay

*A. amphitrite* adults were collected from Pak Sha Wan Pier, Hong Kong (22°38′N, 114°28′E) and after keeping in dark for 24 h, a light source (LED lamp, 1,500 lumens) was used to stimulate the larval release. Within 1 h, the larvae were collected and cultured in 0.22 μm filtered seawater with a daily diet of *Chaetoceros gracilis* Schutt at $1 \times 10^6$ cells $mL^{-1}$ until their growth into cyprids, which were used for the anti-larval settlement bioassay. The bryozoan *B. neritina* adults were collected from a fish farm in Pak Shek Kok, Hong Kong (22°43′N, 114°20′E) and kept in flow-through seawater for no more than 7 days before use. The larvae were released within 30 min before the bioassay as described (Xu et al., 2010).

The bioassay was conducted in a 24-well polystyrene tissue culture plate with 15–20 larvae in each well. 0.1% of DMSO in filtered seawater (FSW), and 0.625, 1.25, 2.5, 5, 10, 20, and 40 μg $mL^{-1}$ albofungin concentrations were tested in triplicate. In each well, 1 mL of FSW containing 15–20 larvae and 1 mL of albofungin solution of different concentrations. The plate was kept for 48 h at 25°C (bryozoan *B. neritina* larvae were kept for 3 h). The wells with 0.1% of DMSO in FSW served as a negative control, and butenolide was used as a positive control. The number of attached, swimming, and dead larvae were counted under an Olympus optical microscope (Olympus Corporation, Tokyo, Japan). The settlement rate was calculated as the ratio of settled larvae to the total number of larvae in each well, and the death rate was calculated as the ratio of dead or missing larvae to the total number of larvae in each well. Half maximal effective concentration ($EC_{50}$) and half lethal concentration ($LC_{50}$) were determined for each compound, and the ratio of $LC_{50}/EC_{50}$ was used to evaluate the toxicity of the antifoulant. Experiments were performed in three independent batches of larvae. Data were analyzed by one-way ANOVA to detect

significant differences in the larval settlement, and SD was calculated by GraphPad Prism 9.

## Construction of Activator Overexpressed *Streptomyces chrestomyceticus* and Analysis of Albofungin Production

The overexpression plasmids were constructed as follows. Gene sequences were obtained by PCR using the primers listed in **Supplementary Table 1** with the genomic DNA of *S. chrestomyceticus* as the template. Each PCR amplicon was ligated into the linear vector pPWW50a digested with *Nde*I and *Spe*I under the control of strong constitutive promoter *ermE**p. All the constructed plasmids were confirmed by DNA sequencing, introduced into *E. coli* ET12567/pUZ8002, and conjugated to *S. chrestomyceticus* BCC 24770 following a previous protocol (Gust et al., 2003). The conjugants were collected and grown on a selective plate containing apramycin (25 μg $mL^{-1}$) and nalidixic acid (25 μg $mL^{-1}$). After 3 days, total DNA was extracted using Chelex 100 resin (Bio-Rad, Hercules, United States) and PCR amplification for the positive conjugants 24770/pPWW-*alb22* and 24770/pPWW-*alb45* (**Supplementary Figure 15**, using primers pPWW50a-check-F and pPWW50a-check-R). BCC24770 with the empty plasmid was used as a negative control (24770/pPWW). The positive conjugants of *S. chrestomyceticus* 24770/pPWW-*alb22* and *S. chrestomyceticus* 24770/pPWW-*alb45* and parental strain *S. chrestomyceticus* 24770/pPWW were inoculated into seed medium (4 g $L^{-1}$ glucose, 4 g $L^{-1}$ yeast extract, 10 g $L^{-1}$ malt extract, and pH being adjusted to 7.0–7.4) with apramycin (50 μg $mL^{-1}$) and grown for 2 days. Afterward, 1% of the preculture was added into the fermentation medium (4 g $L^{-1}$ glucose, 4 g $L^{-1}$ yeast extract, 10 g $L^{-1}$ malt extract, and pH being adjusted to 7.0–7.4) with or without apramycin (50 μg $mL^{-1}$) and grown for 9 days in the 250 mL shaking flasks (Corning Inc., New York, United States) at 220 rpm. The fermentation products were further analyzed by HPLC. The production of albofungin was calculated according to the standard curve (see **Supplementary Figure 14**).

*S. chrestomyceticus* samples were collected and washed using autoclaved water at the end of third day of fermentation and immediately frozen at −80°C for RNA extraction. The RNA samples were prepared using Trizol Reagent (Invitrogen, Waltham, United States) following the manufacturer's instructions. A HiScript III All-in-one RT SuperMix Perfect for qRT-PCR kit (Vazyme, Nanjing, China) was used to remove genomic DNA and synthesize cDNA. The qRT-PCR analysis was performed on Roche Diagnostics LightCycler 480 Instrument II Real-time PCR System using LightCycler® 480 SYBR Green I Master (Roche, Basel, Switzerland). The primers used are listed in **Supplementary Table 1**. The *GAPDH* gene (glyceraldehyde-3-phosphate dehydrogenase) was used as an internal control, and the relative expression levels of *alb22* and *alb45* were normalized to *GAPDH*. Each transcript was performed in triplicate and repeated for three independent biological replicates in qRT-PCR. The relative fold changes in the expression level of each gene were calculated using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001). The *p*-value is computed using Student's *t*-test.

## Albofungin-Based Coating Preparation and Release Rate Measurement

Around 1 g of albofungin compound was isolated, purified, and analyzed using HPLC to guarantee over 95% purity. Albofungin-based hydrolyzable and degradable copolymer coating was prepared as follows: methyl methacrylate (MMA) and tributylsilyl methacrylate (TBSM) copolymer (PMSM0) was synthesized via radical ring-opening polymerization (Zhou et al., 2015). For 5 wt% of albofungin-based coating, 0.95 g of PMSM0 and 0.05 g of albofungin were dissolved in xylene and tetrahydrofuran (v:v = 1:2) and mixed vigorously at room temperature. The solution was then coated on the PVC panels (4 × 7 $cm^2$) and dried at room temperature for 7 days (Ma et al., 2017). Other coatings with different albofungin concentrations (10 and 15 wt%) were prepared using the same procedure. The three concentrations of albofungin-based coatings were optimized based on our previous findings (Wang et al., 2015; Ma et al., 2017). The coating with PMSM0 only was used as a positive control. Each concentration was prepared for three biological replicates. The field test was conducted from March to April 2021 in a fish farm at Yung Shue O, Hong Kong (22°24′N, 114°21′E), which was heavily fouled all year round. The PVC panel was submerged into seawater at a depth of 0.5 m for 2 months, retrieved, washed with seawater, and photographed. The area covered by biofoulers was calculated by Image J (Fiji-2.2.0) (Schindelin et al., 2012). One-way ANOVA was used to compare the albofungin-coated panels and the control panels. The release rate of albofungin under static conditions was determined by measuring the compound concentration with HPLC. Albofungin-based coatings were applied onto a PVC panel (20 × 75 $mm^2$) and submerged in artificial seawater (ASW, NaCl 24.53 g, $MgCl_2{\cdot}6H_2O$ 11.09 g, $Na_2SO_4$ 4.90 g, $CaCl_2$ 1.16 g, KCl 6.95 g, $NaHCO_3$ 0.201 g, KBr 0.101 g, $H_3BO_3$ 0.027 g, $SrCl_2{\cdot}6H_2O$ 0.042 g, NaF 0.003 g per liter of water). After 7 days, the panel was transferred to an individual container with 100 mL of ASW. After 24 h of immersion, 10 mL of the seawater was taken out of the container and extracted with the same volume of ethyl acetate three times. After drying under the SpeedVac vacuum concentrators, the extracts were dissolved in 100 μL of methanol and then subjected to HPLC using a reversed-phase system (Waters 2695) with a Phenomenex Luna C18 column connected to a UV detector at 300 nm. The unique UV absorption of albofungin and retention time were determined, with the amount being calculated from the established standard curves using peak areas plotted against known standard quantities. The logarithmic function was used in curve fitting according to the average of the value.

## Nucleic Acid Extractions, 16S rRNA Amplicon Sequencing and Analysis

The field test for biofilm formation was performed in Pak Sha Wan Pier, Hong Kong (22°38′N, 114°28′E). Glass slides with albofungin/copolymer coatings were submerged at a depth of 0.5 m from sea surface for 12 days and immediately transported into the laboratory for biofilm extraction. The biofilm was scraped using a sterilized cotton swab and was collected in TE buffer

(10 mM Tris-Cl; 0.5 mM EDTA). The samples were centrifuged at 4,000 rpm for 5 min and the supernatant was discarded. Bacteria genomic DNA extraction kit (TIANGEN, Beijing, China) was used to extract the genomic DNA, and the quality was confirmed through BioDrop (Biochrom Ltd., Cambridge, United Kingdom). The 16S rRNA amplicon sequencing of the extracted genomic DNA was performed using an Illumina paired-end platform to generate 250 bp paired-end raw reads (Raw PE) in Novogene (Beijing, China). Sequence data of six samples were subjected to quality control and analyzed using the microbial ecology community software program Mothur (Schloss et al., 2009). Low-quality reads (average quality score < 25) and reads with incorrect length (no shorter than 400 bp and no longer than 430 bp), any ambiguous base, and homopolymers longer than 8 bp were removed. Chimeric sequences were identified and removed by Chimera.uchime in Mothur package. The remaining high-quality sequences were then clustered into the operational taxonomic unit (OTUs) at 97% similarity. Singletons and doubletons were removed before downstream analysis. Taxonomic annotation was performed using Classify. OTU in Mothur with Silver.132 database.

# RESULTS

## Structures of Albofungins

Four compounds (**1**–**4**) in total were isolated: that are, albofungin (**1**), its dimethoxy product, albofungin A (**2**), its monochlorinated derivative, chloroalbofungin (**3**), and its deaminated derivative, chrestoxanthone A (**4**) (**Figure 1A**). All of them were extracted at a high amount from the cultures of *S. chrestomyceticus* BCC 24770 (Bunyapaiboonsri et al., 2016; She et al., 2021). The structure elucidation of albofungins (**1**–**4**) was carried out as described in **Supplementary Figures 1–13**.

## Antibiofilm Activities of Albofungins Against Microfouling Bacteria

The antibiofilm activity of albofungins (**1**–**4**) was evaluated using 6 species of marine bacteria, either isolated from subtidal marine biofilms or described as dominant primary colonizers of submerged surfaces, including *Staphylococcus aureus*, *Micrococcus* sp., *Staphylococcus* sp., *Sulfitobacter pontiacus*, *Pseudomonas pachastrellae*, and *Psychrobacter nivimaris*. The results showed that albofungins (**1**–**4**) strongly prevented the biofilm formation of all selected strains (**Table 1**). The $MBIC_{90}$ values of compounds **1**–**4** against the biofilm formation of Gram-positive bacteria were at a low micromolar range (**Table 2**). More specifically, biofilm formation of *S. aureus*, *Micrococcus* sp., and *Staphylococcus* sp. was almost completely inhibited (>90%) by compounds **1**–**4** at concentrations ranging from 0.03 to 0.5 $\mu g\ mL^{-1}$, 0.06 to 0.5 $\mu g\ mL^{-1}$, and 1.25 $ng\ mL^{-1}$ to 0.2 $\mu g\ mL^{-1}$, respectively (**Figure 2**). As for Gram-negative bacterial strains, compounds **1** and **2** showed strong antibiofilm activities against *S. pontiacus,* and *P. pachastrellae* with $MBIC_{90}$ ranging from 0.02 to 0.50 $\mu g\ mL^{-1}$, whereas compounds **3** and **4** displayed only moderate activities against these bacteria at a concentration of 10–20 $\mu g\ mL^{-1}$ (**Table 3**). All of the compounds, however, showed no apparent biofilm inhibition effects against *P. nivimaris* at concentrations of less than 20 $\mu g\ mL^{-1}$. Additionally, compounds **1**–**4** exhibited antibacterial activities against both Gram-positive bacteria with MIC ranging from 0.8 to 50 $ng\ mL^{-1}$ and Gram-negative bacteria with MIC ranging from 0.008 to 20 $\mu g\ mL^{-1}$ (**Supplementary Figures 18, 19**). These results implied that the biofilm inhibition of albofungins is possibly owing to the inhibition of bacterial growth. Consistent with the antibacterial results, compounds **1** and **2** that contain the hydrazine group but without a chlorinated ring A, exhibited stronger antibiofilm activities than compounds **3** and **4**. According to the preliminary structure-activity relationship analysis, the existence of a hydroxyl group in ring F in compound **2**, instead of a methoxy group in comparison with compound **1**, allows for higher antibiofilm activities against Gram-negative bacteria.

## Antifouling Activity of Albofungins Against the Barnacle *Amphibalanus amphitrite* and Bryozoan *Bugula neritina* Larvae

Then the anti-macrofouling activities of albofungins (**1**–**4**) against the larval settlement of the barnacle *A. amphitrite* and bryozoan *B. neritina* were evaluated. The results showed that the settlement rate of *A. amphitrite* was significantly lower in the treatments of compounds **1** and **2** than that of the control group (0.1% DMSO in FSW) after 48 h of incubation, whereas compounds **3** and **4** did not cause significant difference from the FSW control group (**Figures 3A,C**). The larval settlement rate of *A. amphitrite* in 2.5 $\mu g\ mL^{-1}$ of compound **1** treatment was 33% (± 5.3%), and in 20 $\mu g\ mL^{-1}$ of compound **2** treatment was 17% (± 12.7%), both of which were significantly lower than that in the FSW control group. With respect to *A. amphitrite*, compound **1** exhibited a strong inhibitory effect with an $EC_{50}$ of 1.6 $\mu g\ mL^{-1}$, and compound **2** showed a moderate inhibitory effect with an $EC_{50}$ of 12.2 $\mu g\ mL^{-1}$ (**Figure 3B** and **Supplementary Table 2**). Among these tested compounds, compound **1** showed equivalent antifouling activity against the larval settlement of *A. amphitrite* to that of butenolide, which is a highly potent antifouling compound according to the $EC_{50}$ value (**Supplementary Table 2**). In particular, albofungins up to 40 $\mu g\ mL^{-1}$ had very low lethal effects (**Figure 3A**). Meanwhile, the antifouling activity of compound **1** against the bryozoan *B. neritina* larvae was concentration-dependent and showed low lethal effects at the highest concentration tested as well in the present study (**Figures 3D,E**). Treatment with 2.5 $\mu g\ mL^{-1}$ of compound **1** had 44% (± 9.0%) of larval settlement of *B. neritina*, which was significantly lower than that in the FSW control group (**Figure 3D**).

## Overexpression of the Candidate Activator Genes

To improve the production of albofungin, transcriptional regulators alb22 and alb45 were overexpressed separately in the *S. chrestomyceticus* BCC 24770. The qRT-PCR analysis

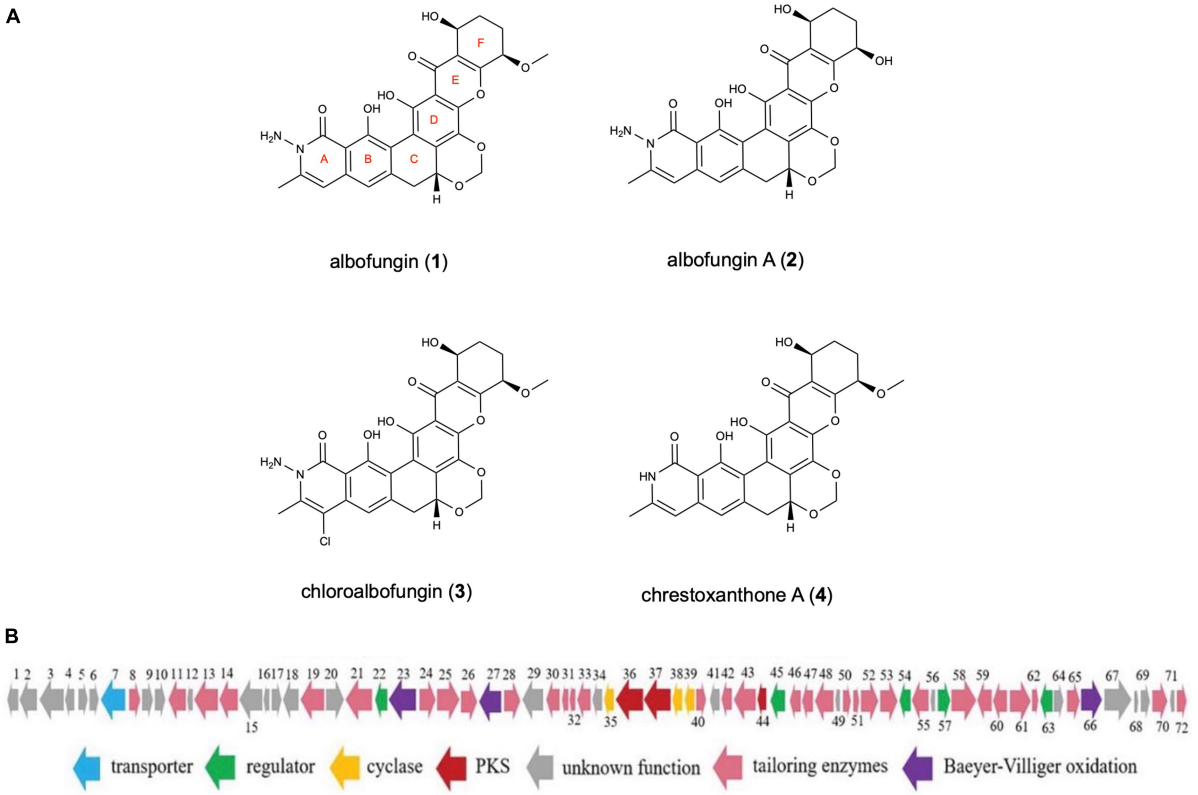



**FIGURE 1 | (A)** Chemical structures of albofungins (**1**–**4**) isolated from *Streptomyces chrestomyceticus* BCC 24770. **(B)** Genetic organization of albofungin biosynthesis gene cluster (She et al., 2021).

**TABLE 2 |** Biofilm inhibition assay of albofungins (**1**–**4**) against gram-positive marine bacteria.

| Gram-positive bacteria | *Micrococcus* sp. Z02 | *Staphylococcus aureus* B04 | *Staphylococcus* sp. Z01 |
|---|---|---|---|
| **Antibiofilm $MBIC_{90}$ (ng $mL^{-1}$)** | | | |
| Albofungin **(1)** | 31.25–62.50 | 15.63–31.25 | 0.63–1.25 |
| Albofungin A **(2)** | 62.5–125 | 15.63–31.25 | 0.63–1.25 |
| Chloroalbofungin **(3)** | 250–500 | 250–500 | 100–200 |
| Chrestoxanthone A **(4)** | 125–250 | 62.5–125 | 5–10 |

*The results are represented as $MBIC_{90}$ in ng $mL^{-1}$.*

**TABLE 3 |** Biofilm inhibition assay of albofungins (**1**–**4**) against gram-negative marine bacteria.

| Gram-negative bacteria | MCCC 1A04899 *Sulfitobacter pontiacus* | MCCC 1A01390 *Pseudomonas pachastrellae* | MCCC 1A11723 *Psychrobacter nivimaris* |
|---|---|---|---|
| **Antibiofilm $MBIC_{90}$ ($\mu$g $mL^{-1}$)** | | | |
| Albofungin **(1)** | 0.13–0.25 | 0.25–0.50 | 10–20 |
| Albofungin A **(2)** | 0.02–0.03 | 0.02–0.03 | >20 |
| Chloroalbofungin **(3)** | 10–20 | >20 | >20 |
| Chrestoxanthone A **(4)** | 10–20 | >20 | >20 |

*The results are represented as $MBIC_{90}$ in $\mu$g $mL^{-1}$.*

revealed that the transcription levels of the two regulatory genes were increased by 0.6- and 3.0-fold in the *alb22* and the *alb45* overexpression strains compared with those in the parental strain (**Figure 4C**). This finding further supported the successful overexpression of these two regulatory genes. Meanwhile, the fermentation results of overexpression strains and parental strain were analyzed using HPLC and the results showed that the overexpression of *alb22* and *alb45* improved

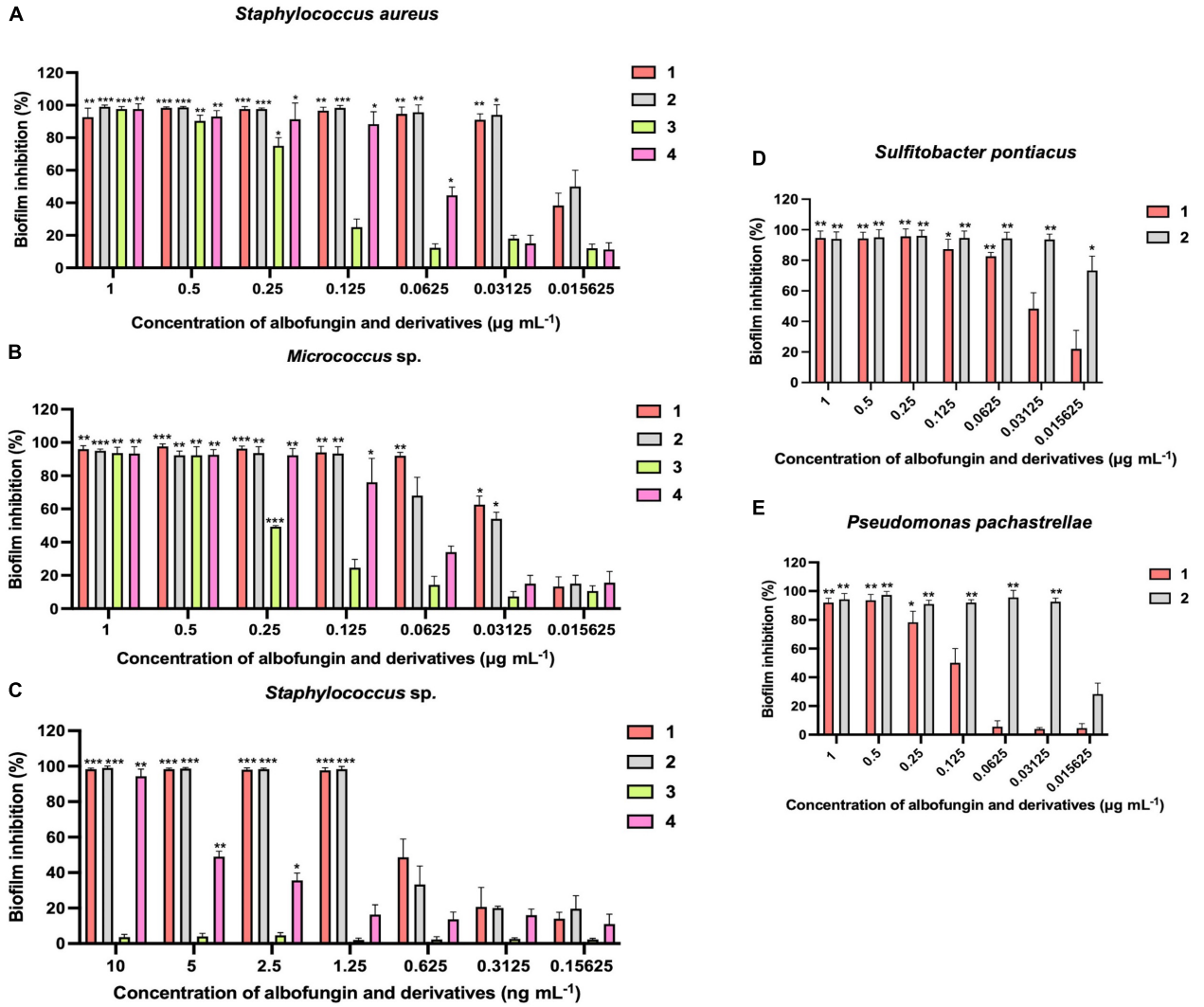



**FIGURE 2 |** Inhibition of biofilm formation of *Staphylococcus aureus* **(A)** *Micrococcus* sp. **(B)** *Staphylococcus* sp. **(C)** *Sulfitobacter pontiacus* **(D)** *Pseudomonas pachastrellae* **(E)** by albofungins (**1**–**4**). Error bars represent SD, $n = 9$ wells from 3 batches of microbial cultures. Significant differences were analyzed by one-way ANOVA compared with the control biofilm, $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$.

albofungin production by 37 and 91% in comparison with the parental strain after 5 days. Noticeably, after 7 days of cultivation, 24770/pPWW-*alb22* strain and 24770/pPWW-*alb45* strain produced 115 ± 9.4 and 153 ± 22.7 mg $L^{-1}$ albofungin in parallel fermentations, which were 0.7- and 1.3-fold higher yields than the parental strain (**Figures 4A,B**), respectively. Because albofungin is yellow powder, its yield can be indicated by the color of the crude extract. Clearly, the color of crude extract from 24770/pPWW-*alb22* and 24770/pPWW-*alb45* dissolved in methanol was darker than that of the parental strain (**Supplementary Figure 16**). Additionally, the overexpression of regulatory genes showed no significant influence on the growth of *S. chrestomyceticus* (**Supplementary Figure 17**). These findings suggested 24770/pPWW-*alb45* as the preferrable genetically modified strain for albofungin production.

## Antifouling Efficacy of Albofungin and Degradable Copolymer Coatings in Marine Field Test and Release Rate Determination

Albofungin (**1**) showed remarkable anti-microfouling and anti-macrofouling activities with low toxicity toward the target organisms, meanwhile, large-scale fermentation of albofungin high-yield strain could easily provide a gram-scale amount of the compound under laboratory conditions at a low cost. Thus, the antifouling efficacy of albofungin was assessed in a field study. 1 g of pure albofungin (**1**) was obtained from 7 L of *S. chrestomyceticus* 24770/pPWW-*alb45* bacterial culture and then incorporated into different antifouling coatings. These coatings were then applied onto PVC panels that were submerged in marine environments for 60 days. Noticeably, almost no

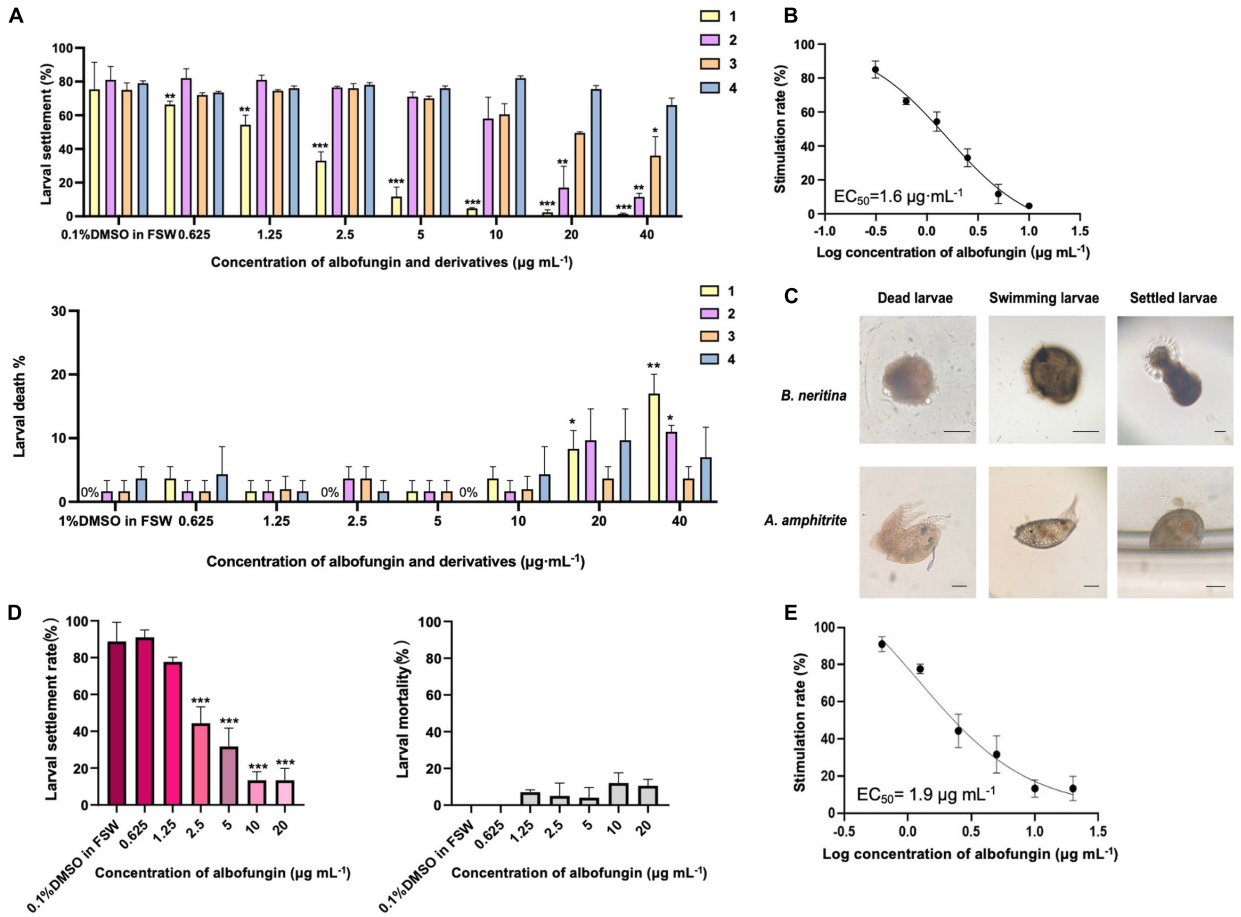



**FIGURE 3 |** Antifouling activity of albofungins (**1**–**4**) against barnacle *Amphibalanus amphitrite* and bryozoan *Bugula neritina*. **(A)** Larval settlement rate and larval mortality rate of *A. amphitrite* under the treatment of albofungins (**1**–**4**) with concentrations of 0.625–40 μg $mL^{-1}$ for 48 h. Error bars represent SD, $n = 9$ wells from 3 batches of larval cultures. Significant differences were analyzed by one-way ANOVA, $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$. **(B)** Stimulation rate curve of albofungin (**1**) treatment on *A. amphitrite* larvae. Error bars represent SD, $n = 9$ wells from 3 batches of larval cultures. **(C)** Different conditions of *A. amphitrite* and *B. neritina* larvae during the bioassay. Scale bars = 100 μm. **(D)** Larval settlement rate and larval mortality rate of *B. neritina* under the treatment of albofungin (**1**) with concentrations of 0.625–20 μg $mL^{-1}$ for 3 h. Error bars represent SD, $n = 9$ wells from 3 batches of larval cultures. Significant differences were analyzed by one-way ANOVA, $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$. **(E)** Stimulation rate curve of albofungin (**1**) treatment on *B. neritina* larvae. Error bars represent SD, $n = 9$ wells from 3 batches of larval cultures.

macrofoulers settled onto the surfaces of the albofungin-coated PVC panels in all three concentrations (5, 10, and 15 wt%) after 1 month (**Figure 5A**), whereas the surface of the negative control group was fouled by the bryozoan *B. neritina* and the polychaete *H. elegans*, which is the most widespread tubebuilding worm in tropical and subtropic regions. After 2 months of submerging in the sea, 96% of the PVC panel surface area was covered by macrofoulers in the control group, whereas the regions fouled in albofungin-coated groups were significantly smaller (**Figures 5A,B**).

The release rate of albofungin from coatings at different concentrations into artificial seawater was measured within 35 days. Overall, the release rate was low during the whole observation period and was positively correlated with the albofungin concentration. The highest release rate of 15, 10, and 5 wt% of albofungin reached 1.4, 1.1, and 0.86 μg $day^{-1}$ $cm^{-2}$ on the first day, respectively, and decreased to approximately 0.08 μg $day^{-1}$ $cm^{-2}$ in a time-dependent manner (**Figure 5C**).

## Changes in Microbial Community Structure Caused by Albofungin and Degradable Copolymer Coatings

Biofilm development on albofungin and degradable copolymer coatings in the field was further examined. During 12 days of observations, the biofilm quickly grew on the panel surfaces in the control group and consisted of diverse microorganisms whereas the diversity of microorganisms on the panel surface covered with 5 wt% of albofungin-based coating reduced significantly. A total of 1,687,617 sequencing reads were analyzed from 6 biofilm samples and microbes in those biofilms were classified into 31 phyla. The 12-day biofilm, which was hereby referred to as "old biofilms" (12–20 days

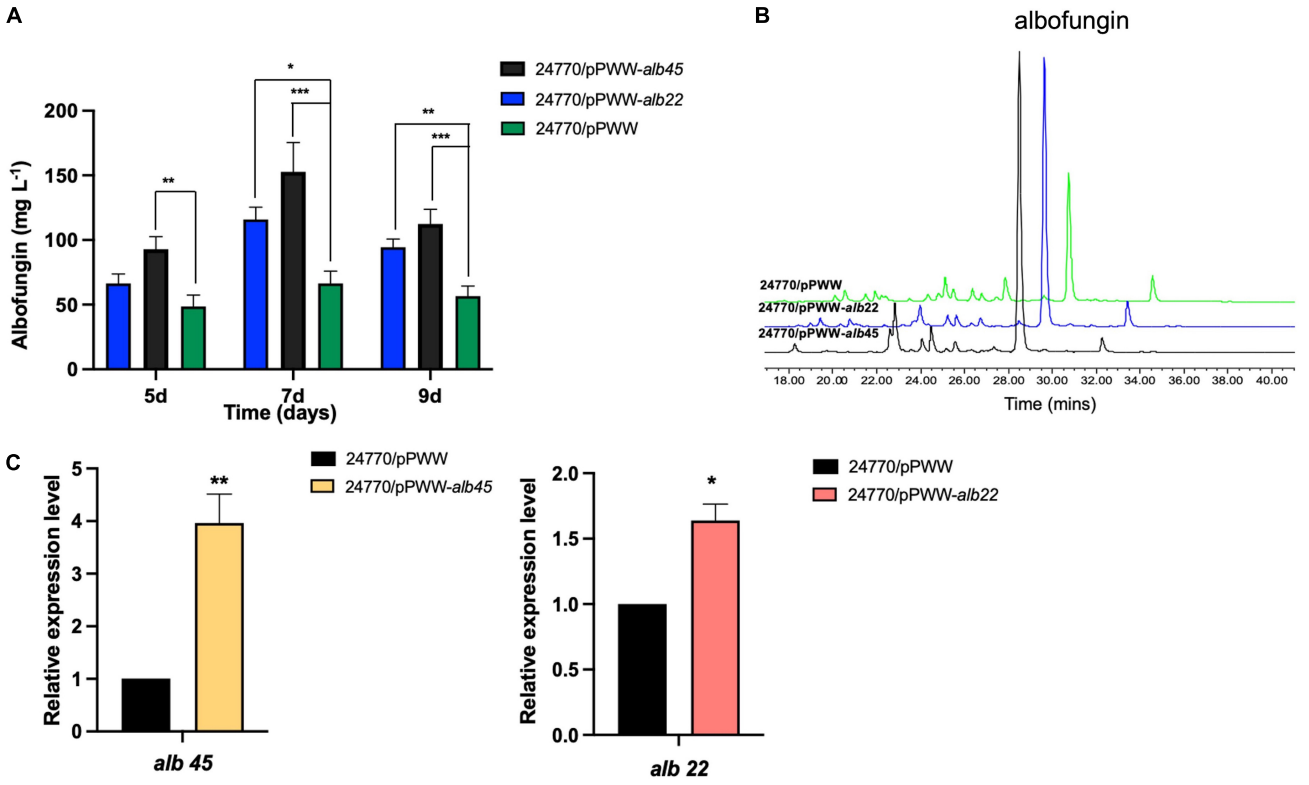



**FIGURE 4 |** Effects of overexpressed regulators on albofungin production. **(A)** Albofungin production (mg $L^{-1}$) by 24770/pPWW (parental strain), 24770/pPWW-*alb22* and 24770/pPWW-*alb45* overexpression strains. Error bars represent SD, $n = 3$ independent cultures. Significant differences were analyzed by one-way ANOVA, $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$. **(B)** HPLC analysis (300 nm) of the crude extracts of 24770/pPWW, 24770/pPWW-*alb45* and 24770/pPWW-*alb22* at 7 days. **(C)** Relative expression level of *alb45* and *alb22* in 24770/pPWW, 24770/pPWW-*alb45* and 24770/pPWW-*alb22* strains. GAPDH was used as a reference gene, and 24770/pPWW served as a control. Error bars represent SD, $n = 3$ independent cultures. Significant differences were analyzed by Student's *t*-test, $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$.

biofilms) (Chung et al., 2010), was dominated mainly by Proteobacteria (Gammaproteobacteria, Alphaproteobacteria, Deltaproteobacteria), Bacteroidetes, Verrucomicrobia, and Actinobacteria (**Figure 6A**). Alpha diversity was used to indicate microbial community diversity in the albofungin-based coating group and control group. The results of the Shannon-Weiner diversity index and observed OTUs were significantly reduced in the albofungin-based coating group, indicating an altered microbial community structure (**Figures 6B,C**). The richness and diversity of microbial communities were also lower than those of the control group.

## DISCUSSION

The albofungin and its derivatives isolated from *S. chrestomyceticus* in this study are belong to polycyclic xanthones. In the family of polyketides, polycyclic xanthones are prominent members, but little research has been done on their antifouling activity (Wang et al., 2017). As previously reported, albofungins showed various biological activities such as potential antibiotics against "ESKAPE pathogens," antifungal activities against pathogenic fungi, and antitumor activities against different cancer cell lines (Bunyapaiboonsri et al., 2016; She et al., 2021). The strong antibiotic effect of albofungins against Gram-positive bacteria was linked to the presence of a hydrazine group (She et al., 2021). In addition, albofungin A inhibited cancer cell proliferation by inducing cellular apoptosis (She et al., 2021). Owing to their prominent and diverse bioactivities, albofungins might have great potential in biofilm control and thus anti-microfouling activities.

In this study, the antifouling activities of albofungins were firstly evaluated. The antibiofilm activities results suggested compounds **1** and **2** contained the hydrazine group had better activities against both gram-positive and gram-negative bacteria. Actually, a similar result was reported in the assessment of bioactivity of hydrazone derivatives, where hydrazide-hydrazones were found to be less active than hydrazides due to the blockage of the $-NH_2$ group (Rollas and Küçükgüzel, 2007). Given that these compounds displayed potent antibiofilm activity, they could be used in preventing undesired microfouling in natural environments. Similar to what was observed with antibiofilm testing, the hydrazine group without chlorination on ring A found in albofungin could cause anti-larval settlement activity. As a compound with $EC_{50} < 5$ μg $mL^{-1}$ and $LC_{50}/EC_{50} > 50$ can be viewed as a non-toxic and high potential

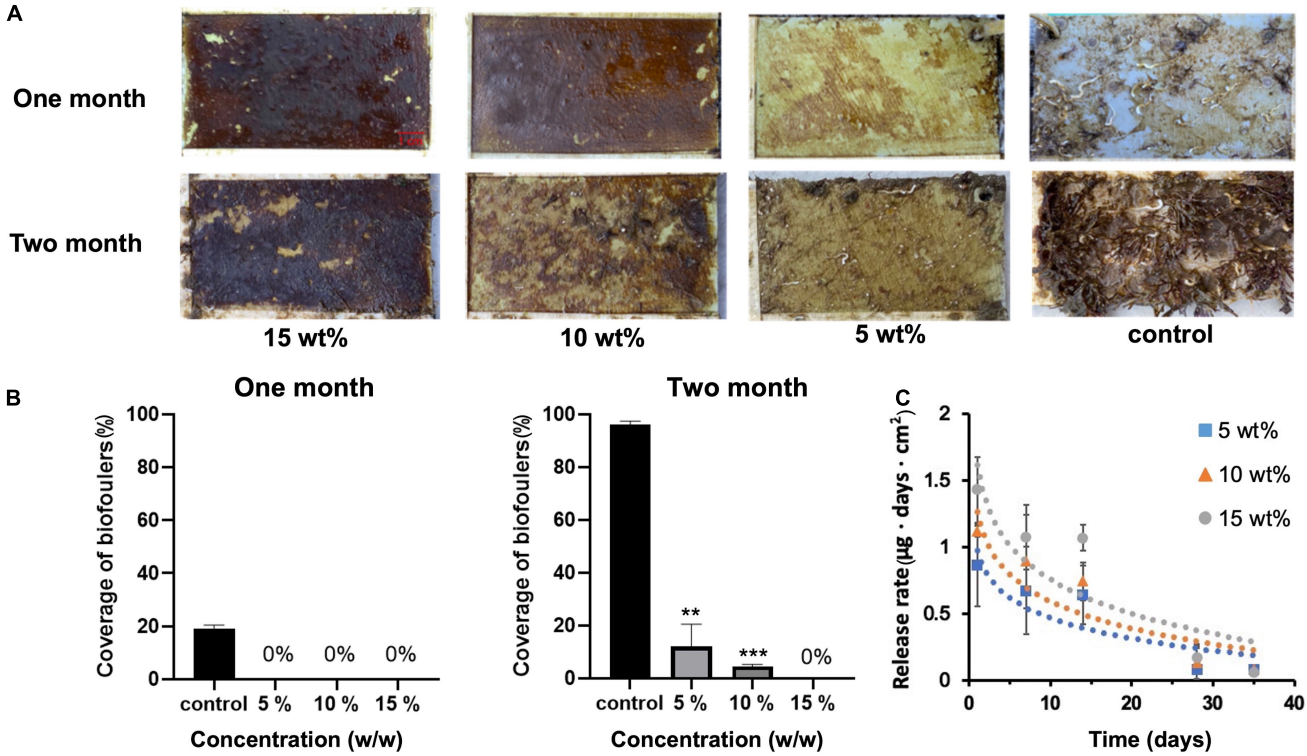



**FIGURE 5 |** Antifouling effect of albofungin-based copolymer coating in the field test after 2 months. **(A)** The copolymer was coated with 5, 10, and 15 wt% albofungin on PVC panels. Control was coated with copolymer only. **(B)** Percentage of coverage of biofoulers after 1 and 2 months. Error bars represent SD ($n = 3$). Significant differences were analyzed by one-way ANOVA, $^{**}p < 0.01$, and $^{***}p < 0.001$. **(C)** Time-dependent release rate of albofungin into artificial seawater. Error bars represent SD ($n = 3$). Significant differences were analyzed by one-way ANOVA and no significant differences were found between 5, 10, and 15 wt% albofungin-based coating groups on the same day.

antifouling compound (Qian et al., 2009), compound **1** with a high $LC_{50}/EC_{50}$ ratio shall be regarded as a low-toxic antifoulant with good antifouling activity. Albofungin has been reported as an inhibitor of bacterial transglycosylase of penicillin-binding protein (Wu et al., 2018), therefore, we proposed albofungins inhibit biofilm formation by cleaving peptidoglycan, which is present in the cell wall of tested bacteria (Roy et al., 2018). The anti-larval settlement assay showed low cytotoxicity against the larvae after 24 h, and according to our observations, the larvae tended to swim inactively. As previously reported that the energy metabolism of larvae is highly needed during the settlement process (Zhang et al., 2010), so we hypothesized that the albofungins blocked the energy metabolism of the larvae during their settlement process, which needs further investigation.

Given that the total synthesis of albofungin remains challenging, albofungin high-producing strains are needed to satisfy the demand for field test and further application of this great potential compound. Heterologous expression and genetic manipulation are important tools for constructing high-yield strains (Li et al., 2015; Bian et al., 2017). Our previous study revealed that a bacterial artificial chromosome plasmid containing an albofungin biosynthesis gene cluster (72 kb) was successfully introduced into *Streptomyces coelicolor* to yield the albofungin (**Figure 1B**). However, the production of albofungin in *S. coelicolor* M1146 4L19 was lower than that in *S. chrestomyceticus* strain. A similar phenomenon has also been reported in other studies. The heterologous expression of potential anticancer drugs epothilones in *S. colicolor* leads to a low yield, possibly owing to its cytotoxic effect toward heterologous host (Julien and Shah, 2002). It implied that albofungin could inhibit the growth of spores from *S. coelicolor* and *S. chrestomyceticus*. And the results indicated that albofungin showed stronger inhibitory effects towards the heterologous host than the wild-type strain. Thus, genetic manipulation in wild-type *S. chrestomyceticus* was the ideal approach to improve the yield of albofungin.

It is known that the positive regulators activate the transcription of targeted biosynthesis gene clusters, and overexpressing the positive regulators can efficiently optimize the production of important secondary metabolites (Baral et al., 2018; Wei et al., 2018; Xia et al., 2020). *Alb22* and *alb45* encoded for transcriptional enhancer A (TenA) family regulator and *Streptomyces* antibiotic regulatory protein (SARP) family regulator, which were identified as positive regulators for albofungin biosynthesis (She et al., 2021). They are located closer to the polyketide synthase genes than other regulators. The TenA family transcriptional regulator was reported to stimulate the production of extracellular proteases in *Bacillus subtilis*

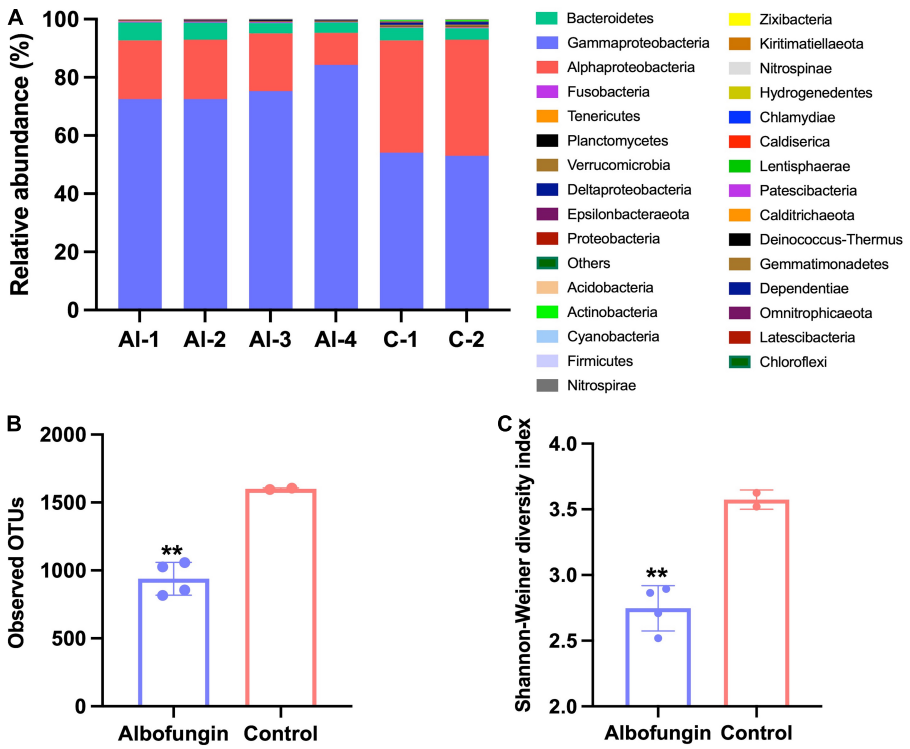



**FIGURE 6 |** Microbial community structure alteration by albofungin-based copolymer coatings based on 16S rRNA amplicon analysis. **(A)** Relative abundance of the major bacterial phylum of biofilms developed in 5 wt% albofungin-based copolymer coatings and control with only copolymer coatings (Proteobacteria are further classified into the classes). **(B)** Observed OTUs of albofungin-based copolymer coatings and control with only copolymer coatings. Error bars represent SD. Significant differences were analyzed by Student's *t*-test, $^{**}p < 0.01$. **(C)** Shannon–Weiner diversity index of albofungin-based copolymer coatings and control with only copolymer coatings (al-1, al-2, al-3, al-4, and control-1, control-2). Error bars represent SD. Significant differences were analyzed by Student's *t*-test, $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$.

and exhibits an enzymatic function in thiamine metabolism in *Pyrococcus furiosus* (Benach et al., 2005; Lee et al., 2005). Moreover, a crystal structure of TenA regulator revealed that it activates gene expression indirectly (Benach et al., 2005), which may imply indirect transcriptional mechanisms of *alb22* in the albofungin biosynthesis gene cluster. However, further study is required to confirm this notion. Meanwhile, many SARPs positively regulate secondary metabolites in *Streptomyces* (Xia et al., 2020). For example, the overexpression of SARP family regulatory gene *pieR* in *Streptomyces piomogeues* var. *hangzhouwanensis* enhances the yield of piericidin A1 by 2.3-fold (Li et al., 2019); while overexpression of *otcR* considerably enhanced oxytetracycline production in *Streptomyces rimosus* by 6.49 times compared to that in the parental strain (Yin et al., 2015). Therefore, in the present study, overexpression of these two positive regulators alb22 and alb45 offers promising genetically modified strains to produce new antifoulants and further promotes the use of albofungins for antifouling application. Moreover, the albofungin yield can be further improved via the co-overexpression of alb22 and alb45 activators, as well as the overexpression of transporters to facilitate the export of the self-produced secondary metabolites.

Subsequently, fermentation of high-yield albofungin strain could easily produce gram-scale albofungin for the preparation of albofungin-based coatings. The low and constant release rate of albofungin-based coatings suggested the slow degradation rate of the copolymer that maintained the relatively constant release of the albofungin over a long period, which could hold an effective concentration longer and play an important role in the antifouling performance (Xie et al., 2019). Thus, the albofungin and PMSM0 copolymer formed a suitable coating to release albofungin in the marine field test. After 2 months of observation, albofungin inhibited macrofoulers attachment, and only a few fouling organisms have settled on the 5 and 10 wt% of albofungin-coated panels because the amount of albofungin being released from the panel into seawater decreased in the later testing stage. The result in the present study was comparable with the field performance of butenolide (Ma et al., 2017) and even better antifouling efficacy than that of DIM and DCOIT (Wang et al., 2015), which was probably due to the broader inhibitory spectrum against biofoulers of albofungin. The field test results suggested that albofungin could prevent the settlement of multiple fouling organisms in marine environments at a low concentration.

Furthermore, albofungin-based coatings changed the microbial community structure, particularly for the abundance of Proteobacteria. More specifically, the comparison of microbial diversity between the control group and the albofungin treated

group demonstrated that the abundance of Alphaproteobacteria was significantly decreased, whereas the abundance of Gammaproteobacteria was increased. It was reported that the class of Alphaproteobacteria plays a crucial role in biomass accumulation and biofilm maintenance (Sheluď ko et al., 2019), which is consistent with the findings in this study. Moreover, Gammaproteobacteria was also demonstrated to be responsible for the metabolism of hydrocarbon, which is important for the defensive response to a chemical stimulus (Dyksma et al., 2016). Accordingly, the increase of Gammaproteobateria after albofungin treatment may suggest the crucial role of Gammaproteobacteria in response to albofungin treatment, implying the potential antifouling mechanism of albofungin, which deserves further investigation. Given that the biofilm can mediate the larval settlement of *H. elegans* (Chung et al., 2010), the changes in microbial community structure and biofilm density may alter the larval settlement response and could cause the inhibitive effect of the albofungin-based copolymer coatings.

The advantages of antifouling compounds isolated from marine microorganisms, such as chemical diversity and uniqueness, make them potential sources for the discovery of new antifouling compounds. These newly discovered antifoulants, however, have always been limited by low yields for commercialization (Long et al., 2021). Also, it is critical to evaluate the highly potent compounds in the field test to provide information for further marine industrial applications (Wang et al., 2017). In our study, albofungins were isolated from bacterial secondary metabolites and we constructed the high-yield strains for albofungin production in bacterial fermentation ($153 \pm 22.7$ mg $L^{-1}$). Our field test results suggested that albofungin had a high potential in preventing the settlement of multiple fouling organisms in marine environments at a low concentration. In terms of limitations, potent antifouling agents must be examined for their acute and chronic toxicity to non-target organisms as well as their degradation kinetics in the marine environment before commercialization (Wang et al., 2017). In the future, more research needs to be conducted on degradation kinetics and potential ecological risks investigations of albofungins.

## CONCLUSION

Novel antifouling agents are needed to address the issues caused by biofouling in the marine industry. In the present study, we identified potent antibiofilm and antifouling compounds albofungins isolated from the metabolites of *S. chrestomyceticus* BCC 24770 with low cytotoxicity. We also constructed high-yield strains for albofungin production and developed albofungin-based antifouling coatings that prevented colonization of macrofoulers in marine environments. Our findings revealed that broad-spectrum and strong antifouling activities, relatively low cytotoxicity, and high yield in bacterial fermentation make albofungins be promising antifouling candidates and indicated their applications in the antifouling area.

## DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the open-source online data repositories hosted at National Center for Biotechnology Information, and the accession numbers are ON014504, ON014498, and PRJNA817672.

## AUTHOR CONTRIBUTIONS

WS, WY, and JC: conceptualization. WS: formal analysis and writing—original draft. AC, P-YQ, WY, JL, and FC: writing—review and editing. WS, RW, WKY, XL, YY, SC, and YL: investigation. CM: resources and methodology. FC, P-YQ, and AC: project administration. P-YQ: funding acquisition. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was supported by the National Key R&D Program of China (2018YFA0903200), Key Special Project for Introduced Talents Team of Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou) (GML2019ZD409), the Hong Kong Branch of Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou) (SMSEGL20SC01), and a CRF grant from the HKSAR Government (C6026-19G-A).

## ACKNOWLEDGMENTS

We thank Mr. Jiansen Pan and Chiang Ho Yin for their help with the field test experiment.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.906345/full#supplementary-material

## REFERENCES

Baral, B., Akhgari, A., and Metsä-Ketelä, M. (2018). Activation of microbial secondary metabolic pathways: avenues and challenges. *Synth. Syst. Biotechnol.* 3, 163–178. doi: 10.1016/j.synbio.2018.09.001

Benach, J., Edstrom, W. C., Lee, I., Das, K., Cooper, B., Xiao, R., et al. (2005). The 2.35 Å structure of the TenA homolog from *Pyrococcus furiosus* supports an enzymatic function in thiamine metabolism. *Acta Crystallogr. D Biol. Crystallogr.* 61, 589–598. doi: 10.1107/S0907444905005147

Bian, X., Tang, B., Yu, Y., Tu, Q., Gross, F., Wang, H., et al. (2017). Heterologous production and yield improvement of epothilones in *Burkholderiales* strain DSM 7029. *ACS Chem. Biol.* 12, 1805–1812. doi: 10.1021/acschembio.7b00097

Bixler, G. D., and Bhushan, B. (2012). Biofouling: lessons from nature. *Philos. Trans. A Math. Phys. Eng. Sci.* 370, 2381–2417. doi: 10.1098/rsta.2011.0502

Bressy, C., NGuyen, M. N., Tanguy, B., and Margaillan, A. (2010). Poly (trialkylsilyl methacrylate) s A family of hydrolysable polymers with tuneable erosion profiles. *Polym. Degrad. Stab.* 95, 1260–1268. doi: 10.1016/j.polymdegradstab.2010.03.017

Bunyapaiboonsri, T., Lapanun, S., Supothina, S., Rachtawee, P., Chunhametha, S., Suriyachadkun, C., et al. (2016). Polycyclic tetrahydroxanthones from *Streptomyces chrestomyceticus* BCC 24770. *Tetrahedron* 72, 775–778. doi: 10.1016/j.tet.2015.12.045

Callow, M. E., and Callow, J. A. (2002). Marine biofouling: a sticky problem. *Biologist* 4, 1–5.

Champ, M. A. (2000). A review of organotin regulatory strategies, pending actions, related costs and benefits. *Sci. Total Environ.* 258, 21–71. doi: 10.1016/s0048-9697(00)00506-4

Chung, H. C., Lee, O. O., Huang, Y. L., Mok, S. Y., Kolter, R., and Qian, P. Y. (2010). Bacterial community succession and chemical profiles of subtidal biofilms in relation to larval settlement of the polychaete *Hydroides elegans*. *ISME J.* 4, 817–828. doi: 10.1038/ismej.2009.157

Dyksma, S., Bischof, K., Fuchs, B. M., Hoffmann, K., Meier, D., Meyerdierks, A., et al. (2016). Ubiquitous Gammaproteobacteria dominate dark carbon fixation in coastal sediments. *ISME J.* 10, 1939–1953. doi: 10.1038/ismej.2015.257

Guan, D., and Pettis, G. S. (2009). Intergeneric conjugal gene transfer from *Escherichia coli* to the sweet potato pathogen *Streptomyces ipomoeae*. *Lett. Appl. Microbiol.* 49, 67–72. doi: 10.1111/j.1472-765X.2009.02619.x

Gust, B., Challis, G. L., Fowler, K., Kieser, T., and Chater, K. F. (2003). PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proc. Natl. Acad. Sci. U.S.A.* 100, 1541–1546. doi: 10.1073/pnas.0337542100

Heidarian, S., Mohammadipanah, F., Maghsoudlou, A., Dashti, Y., and Challis, G. L. (2019). Anti-microfouling activity of *Glycomyces sediminimaris* UTMC 2460 on dominant fouling bacteria of Iran marine habitats. *Front. Microbiol.* 9:3148. doi: 10.3389/fmicb.2018.03148

Hellio, C., and Yebra, D. (eds) (2009). *Advances in Marine Antifouling Coatings and Technologies*. Burlington: Elsevier.

Julien, B., and Shah, S. (2002). Heterologous expression of epothilone biosynthetic genes in *Myxococcus xanthus*. *Antimicrob. Agents Chemother.* 46, 2772–2778. doi: 10.1128/AAC.46.9.2772-2778.2002

Lee, I., Das, K., Xiao, R., Benach, J., Edstrom, W. C., Liu, J., et al. (2005). The 2.35 Å structure of the TenA homolog from *Pyrococcus furiosus* supports an enzymatic function in thiamine metabolism. *Acta Crystallogr. D* 61, 589–598. doi: 10.1107/S0907444905005147

Li, L., Zhao, Y., Ruan, L., Yang, S., Ge, M., Jiang, W., et al. (2015). A stepwise increase in pristinamycin II biosynthesis by *Streptomyces pristinaespiralis* through combinatorial metabolic engineering. *Metab. Eng.* 29, 12–25. doi: 10.1016/j.ymben.2015.02.001

Li, Y., Kong, L., Shen, J., Wang, Q., Liu, Q., Yang, W., et al. (2019). Characterization of the positive SARP family regulator PieR for improving piericidin A1 production in *Streptomyces piomogeues* var. Hangzhouwanensis. *Synth. Syst. Biotechnol.* 4, 16–24. doi: 10.1016/j.synbio.2018.12.002

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Long, L., Wang, R., Chiang, H. Y., Ding, W., Li, Y. X., Chen, F., et al. (2021). Discovery of antibiofilm activity of elasnin against marine biofilms and its application in the marine antifouling coatings. *Mar. Drugs* 19:19. doi: 10.3390/md19010019

Ma, C., Zhang, W., Zhang, G., and Qian, P. Y. (2017). Environmentally friendly antifouling coatings based on biodegradable polymer and natural antifoulant. *ACS Sustain. Chem. Eng.* 5, 6304–6309. doi: 10.1021/acssuschemeng.7b01385

Malit, J. J. L., Liu, W., Cheng, A., Saha, S., Liu, L. L., and Qian, P. Y. (2021). Global Genome Mining Reveals a Cytochrome P450-Catalyzed Cyclization of Crownlike Cyclodipeptides with Neuroprotective Activity. *Org. Lett.* 23, 6601–6605. doi: 10.1021/acs.orglett.1c01022

O'Sullivan, L. M., Patel, S., Ward, J. M., Woodley, J. M., and Doig, S. D. (2001). Large scale production of cyclohexanone monooxygenase from *Escherichia coli* TOP10 pQR239. *Enzyme Microb. Technol.* 28, 265–274. doi: 10.1016/s0141-0229(00)00320-3

Qi, S. H., and Ma, X. (2017). Antifouling compounds from marine invertebrates. *Mar. Drugs* 15:263. doi: 10.3390/md15090263

Qian, P. Y., Chen, L., and Xu, Y. (2013). Mini-review: molecular mechanisms of antifouling compounds. *Biofouling* 29, 381–400. doi: 10.1080/08927014.2013.776546

Qian, P. Y., Xu, Y., and Fusetani, N. (2009). Natural products as antifouling compounds: recent progress and future perspectives. *Biofouling* 26, 223–234. doi: 10.1080/08927010903470815

Ralston, E., and Swain, G. (2009). Bioinspiration—the solution for biofouling control? *Bioinspir. Biomims.* 4:015007. doi: 10.1088/1748-3182/4/1/015007

Rollas, S., and Küçükgüzel, S. G. (2007). Biological activities of hydrazone derivatives. *Molecules* 12, 1910–1939. doi: 10.3390/12081910

Roy, R., Tiwari, M., Donelli, G., and Tiwari, V. (2018). Strategies for combating bacterial biofilms: a focus on anti-biofilm agents and their mechanisms of action. *Virulence* 9, 522–554. doi: 10.1080/21505594.2017.1313372

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682. doi: 10.1038/nmeth.2019

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., et al. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* 75, 7537–7541. doi: 10.1128/AEM.01541-09

She, W., Ye, W., Cheng, A., Liu, X., Tang, J., Lan, Y., et al. (2021). Discovery, Bioactivity Evaluation, Biosynthetic Gene Cluster Identification, and Heterologous Expression of Novel Albofungin Derivatives. *Front. Microbiol.* 12:635268. doi: 10.3389/fmicb.2021.635268

She, W., Ye, W., Shi, Y., Zhou, L., Zhang, Z., Chen, F., et al. (2020). A novel chresdihydrochalcone from *Streptomyces chrestomyceticus* exhibiting activity against Gram-positive bacteria. *J. Antibiot.* 73, 429–434. doi: 10.1038/s41429-020-0298-1

Shelud'ko, A. V., Filip'echeva, Y. A., Telesheva, E. M., Yevstigneeva, S. S., Petrova, L. P., and Katsy, E. I. (2019). Polar flagellum of the alphaproteobacterium *Azospirillum brasilense* Sp245 plays a role in biofilm biomass accumulation and in biofilm maintenance under stationary and dynamic conditions. *World J. Microbiol. Biotechnol.* 35, 1–9. doi: 10.1007/s11274-019-2594-0

Sisson, A. L., Ekinci, D., and Lendlein, A. (2013). The contemporary role of ε-caprolactone chemistry to create advanced polymer architectures. *Polymer* 54, 4333–4350. doi: 10.1016/j.polymer.2013.04.045

Srinivasan, R., Santhakumari, S., Poonguzhali, P., Geetha, M., Dyavaiah, M., and Xiangmin, L. (2021). Bacterial biofilm inhibition: a focused review on recent therapeutic strategies for combating the biofilm mediated infections. *Front. Microbiol.* 12:1106. doi: 10.3389/fmicb.2021.676458

Stewart, P. S., and Costerton, J. W. (2001). Antibiotic resistance of bacteria in biofilms. *Lancet* 358, 135–138. doi: 10.1016/s0140-6736(01)05321-1

Wang, K. L., Wu, Z. H., Wang, Y., Wang, C. Y., and Xu, Y. (2017). Mini-review: antifouling natural products from marine microorganisms and their synthetic analogs. *Mar. Drugs* 15:266. doi: 10.3390/md15090266

Wang, K. L., Xu, Y., Lu, L., Li, Y., Han, Z., and Zhang, J. (2015). Low-toxicity diindol-3-ylmethanes as potent antifouling compounds. *Mar. Biotechnol.* 17, 624–632. doi: 10.1007/s10126-015-9656-6

Wang, R., Ding, W., Long, L., Lan, Y., Tong, H., Saha, S., et al. (2020). Exploring the Influence of Signal Molecules on Marine Biofilms Development. *Front. Microbiol.* 11:571400. doi: 10.3389/fmicb.2020.571400

Wei, J., He, L., and Niu, G. (2018). Regulation of antibiotic biosynthesis in actinomycetes: perspectives and challenges. *Synth. Syst. Biotechnol.* 3, 229–235. doi: 10.1016/j.synbio.2018.10.005

Wu, W. S., Cheng, W. C., Cheng, T. J. R., and Wong, C. H. (2018). Affinity-based screen for inhibitors of bacterial transglycosylase. *J. Am. Chem. Soc.* 140, 2752–2755. doi: 10.1021/jacs.7b13205

Xia, H., Li, X., Li, Z., Zhan, X., Mao, X., and Li, Y. (2020). The application of regulatory cascades in *Streptomyces*: yield enhancement and metabolite mining. *Front. Microbiol.* 11:406. doi: 10.3389/fmicb.2020.00406

Xie, Q., Pan, J., Ma, C., and Zhang, G. (2019). Dynamic surface antifouling: mechanism and systems. *Soft Matter* 15, 1087–1107. doi: 10.1039/c8sm01853g

Xu, Y., He, H., Schulz, S., Liu, X., Fusetani, N., Xiong, H., et al. (2010). Potent antifouling compounds produced by marine *Streptomyces*. *Bioresour. Technol.* 101, 1331–1336. doi: 10.1016/j.biortech.2009.09.046

Ye, W., She, W., Sung, H. Y., Qian, P., and Williams, I. D. (2020). Albofungin and chloroalbofungin: antibiotic crystals with 2D but not 3D isostructurality. *Acta Crystallogr. C Struct. Chem.* 76, 1100–1107. doi: 10.1107/S2053229620015041

Yin, S., Wang, W., Wang, X., Zhu, Y., Jia, X., Li, S., et al. (2015). Identification of a cluster-situated activator of oxytetracycline biosynthesis and manipulation of its expression for improved oxytetracycline production in *Streptomyces rimosus*. *Microb. Cell Fact.* 14, 1–12. doi: 10.1186/s12934-015-0231-7

Zhang, Y., Xu, Y., Arellano, S. M., Xiao, K., and Qian, P. Y. (2010). Comparative proteome and phosphoproteome analyses during cyprid development of the barnacle *Balanus* (= *Amphibalanus*) *amphitrite*. *J. Proteome Res.* 9, 3146–3157. doi: 10.1021/pr1000384

Zhou, X., Xie, Q., Ma, C., Chen, Z., and Zhang, G. (2015). Inhibition of marine biofouling by use of degradable and hydrolysable silyl acrylate copolymer. *Ind. Eng. Chem. Res.* 54, 9559–9565. doi: 10.1021/acs.iecr.5b01819

**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.

*Copyright © 2022 She, Ye, Cheng, Ye, Ma, Wang, Cheng, Liu, Yuan, Chik, Limlingan Malit, Lu, Chen and Qian. 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.*

frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 04 August 2022
DOI 10.3389/fmicb.2022.909493

Check for updates

OPEN ACCESS

EDITED BY
Sailesh Malla,
Chr. Hansen A/S, Denmark

REVIEWED BY
Michelle Meyer,
Boston College, United States
Marcus Lechner,
LOEWE Center for Synthetic Microbiology (SYNMIKRO), Germany

*CORRESPONDENCE
Jan Gorodkin
gorodkin@rth.dk

SPECIALTY SECTION
This article was submitted to Microbiotechnology, a section of the journal Frontiers in Microbiology

RECEIVED 31 March 2022
ACCEPTED 28 June 2022
PUBLISHED 04 August 2022

CITATION
Geissler AS, Poulsen LD, Doncheva NT, Anthon C, Seemann SE, González-Tortuero E, Breüner A, Jensen LJ, Hjort C, Vinther J and Gorodkin J (2022) The impact of PrsA over-expression on the *Bacillus subtilis* transcriptome during fed-batch fermentation of alpha-amylase production. *Front. Microbiol.* 13:909493. doi: 10.3389/fmicb.2022.909493

COPYRIGHT
© 2022 Geissler, Poulsen, Doncheva, Anthon, Seemann, González-Tortuero, Breüner, Jensen, Hjort, Vinther and Gorodkin. 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 impact of PrsA over-expression on the *Bacillus subtilis* transcriptome during fed-batch fermentation of alpha-amylase production

Adrian S. Geissler[1], Line D. Poulsen[2], Nadezhda T. Doncheva[3], Christian Anthon[1], Stefan E. Seemann[1], Enrique González-Tortuero[1], Anne Breüner[4], Lars J. Jensen[3], Carsten Hjort[4], Jeppe Vinther[2] and Jan Gorodkin[1]*

[1]Department of Veterinary and Animal Sciences, Center for non-coding RNA in Technology and Health, University of Copenhagen, Copenhagen, Denmark, [2]Department of Biology, University of Copenhagen, Copenhagen, Denmark, [3]Novo Nordisk Foundation Center for Protein Research, University of Copenhagen, Copenhagen, Denmark, [4]Novozymes A/S, Bagsværd, Denmark

The production of the alpha-amylase (AMY) enzyme in *Bacillus subtilis* at a high rate leads to the accumulation of unfolded AMY, which causes secretion stress. The over-expression of the PrsA chaperone aids enzyme folding and reduces stress. To identify affected pathways and potential mechanisms involved in the reduced growth, we analyzed the transcriptomic differences during fed-batch fermentation between a PrsA over-expressing strain and control in a time-series RNA-seq experiment. We observe transcription in 542 unannotated regions, of which 234 had significant changes in expression levels between the samples. Moreover, 1,791 protein-coding sequences, 80 non-coding genes, and 20 riboswitches overlapping UTR regions of coding genes had significant changes in expression. We identified putatively regulated biological processes *via* gene-set over-representation analysis of the differentially expressed genes; overall, the analysis suggests that the PrsA over-expression affects ATP biosynthesis activity, amino acid metabolism, and cell wall stability. The investigation of the protein interaction network points to a potential impact on cell motility signaling. We discuss the impact of these highlighted mechanisms for reducing secretion stress or detrimental aspects of PrsA over-expression during AMY production.

KEYWORDS

**alpha-amylase, PrsA, *Bacillus subtilis*, RNA sequencing (RNA-seq), enzyme produced by microorganism**

# Introduction

*Bacillus subtilis* is a powerhouse for enzyme production in biotech industries (Schallmey et al., 2004; van Dijl and Hecker, 2013; Hohmann et al., 2016). Amylases are a specific class of enzymes that *B. subtilis* can produce commercially (Schallmey et al., 2004). The amylase enzyme, in particular the alpha-amylase (AMY), is a digestive enzyme (EC 3.2.1.1) that degrades starch molecules. Therefore, AMY is often an active component in laundry detergent for removing sticky stains from clothes. For successful AMY production and subsequent recovery, a host organism needs to both express and secrete AMY proteins in a biologically active form at a high rate (Spinnler, 2021). However, a major issue for commercial production is that the protein folding system of the cell is overwhelmed by the high rate of synthesis unless the strains used for production are genetically modified (Kontinen and Sarvas, 1993). The accumulation of unfolded AMY proteins causes stress that requires a bacterial cell to physiologically adapt to survive (Storz and Hengge, 2010). The Sec secretion pathway secretes AMY co-translationally (Fu et al., 2007). Therefore, unfolded AMY is extracellular, such that the corresponding stress signal triggers the heat shock response (Westers et al., 2004, 2006; Storz and Hengge, 2010; Lim and Gross, 2014; Yan and Wu, 2019). The simplified mechanism of this stress response has two components as follows (Westers et al., 2004, 2006; Storz and Hengge, 2010; Lim and Gross, 2014; Yan and Wu, 2019): First, the membrane-bound CssS receptor transduces the stress signal by phosphorylating CssR. Second, the phosphorylated CssR activates transcription of the two proteases, namely HrtA and HrtB, which degrade unfolded proteins and alleviate the stress condition. Furthermore, stress responses are intertwined with additional regulation in the core energy metabolism (Storz and Hengge, 2010), and such stress responses upregulate flagellar cell motility in order for a cell to physically escape the stress-causing location (Helmann et al., 1988; Márquez-Magaña et al., 1990; Yan and Wu, 2019). For instance, the level of cell motility is boosted by a low level of phosphorylated DegU (Kobayashi, 2007; Verhamme et al., 2007; Gupta and Rao, 2014), which is part of the core stress regulating DegU-DegS two-component system (Storz and Hengge, 2010; Laub, 2014). Nevertheless, these stress alleviating mechanisms can be opposed to the objective of achieving a high AMY yield: (i) the proteolytic degradation of AMY reduces yields, and (ii) a low phosphorylation level of DegU downregulates AMY expression (Gupta and Rao, 2014).

A state-of-the-art approach, which prevents the yield detrimental impact on the secretion of the stress response, is the over-expression of PrsA (Vitikainen et al., 2001; Quesada-Ganuza et al., 2019). Although the over-expression of PrsA reduces secretion stress by aiding AMY folding, it also has detrimental impacts such as hampered cell growth and even cell lysis (Vitikainen et al., 2001; Quesada-Ganuza et al., 2019). These detrimental phenotypes might be caused by protein-protein interactions of specific PrsA protein domains with still unknown partner proteins (Quesada-Ganuza et al., 2019). Another unknown aspect of PrsA over-expression is its impact on the bacterial transcriptome, particularly during industrial fed-batch fermentation. The adaptation to glucose metabolism from maltose metabolism has a global impact on half of all transcriptional regulators even though both carbons are preferred by *B. subtilis* (Buescher et al., 2012). Thus, we would assume a substantially larger global impact on the transcriptome for the extreme secretion stress during PrsA over-expression (Quesada-Ganuza et al., 2019). We consider our assumption to be further supported by a large number of over a hundred proteins that require regulation to adapt bacterial motility (see above concerning stress) (Rajagopala et al., 2007). Furthermore, a pure protein-coding gene focus ignores the essential role regulatory small RNA (sRNA), RNA chaperones, and non-coding RNA (ncRNA) have in facilitating physiological changes impacting the entire cell during stress responses (Storz and Hengge, 2010). General stress regulatory mechanisms have been investigated in public datasets (Arrieta-Ortiz et al., 2015); however, metabolic and stress pathways undergo complex temporal adaptations (Hahne et al., 2010; Otto et al., 2010). Thus, both temporally resolved and condition-specific gene expression levels are needed to study stress pathways. Specifically for secretion stress during *B. subtilis* AMY fed-batch fermentation, no such dataset exists to our knowledge.

Here, we conducted fed-batch fermentation of two commercial *B. subtilis* strains. Both strains produce an AMY and are isogenic, except that one of them over-expresses PrsA. We studied the transcriptome during fermentation at six timepoints with RNA-seq and analyzed the expression levels of both known coding and non-coding annotations, and also of potential novel transcribed, yet unannotated regions. We complemented the differential expression analysis with a network analysis of known protein–protein interactions (PPI). This study found significant changes in gene expression levels between the studied strains for genes in the ATP biosynthesis and cell motility biological processes. Furthermore, the network analysis hints at mechanisms relating to competence transformation and cell motility that might be candidates for further tuning of AMY secretion yields.

# Materials and Methods

## Strains and fed-batch fermentation

The overall experimental setup is as previously described in Geissler et al. (2022). In summary, *B. subtilis* strain 168 Δ*spoIIAC*Δ*amyE*Δ*apr*Δ*nprE*Δ*srfAC* was maintained at 4°C on the LBGG medium. The AMY JE1 [sequence label *je1zyn* in Geissler et al. (2022)] was inserted by splicing by overlapping extension (SOE, inserted sequences are in

**Supplementary Table 2**) linear recombinant transformation, together with the commercial *sigA* promoter sequence P4199 and chloramphenicol marker in the *pel* locus. The PrsA over-expressing strain (referred to as the "+prsA" strain) had the insert by SOE of P4199, *prsA*, and spectinomycin marker in the *amyE* locus. A control strain without the prsA insert was included. After inoculation on SSB4 agar at 37°C, transfer on M-9 medium, sucrose 2M fed-batch fermentations were conducted in proprietary 2L tanks at 38°C. To avoid excessive overflow metabolite formation and to keep the culture in a sucrose metabolizing state, the fermentations were run as fed-batch fermentations without an initial batch phase. The feed medium consisted of sucrose (708 g/L), which was fed at a rate that increased linearly from 0.04 g/min at time = 0 h to 0.125 g/min at time = 8 h. The feed rate after 8 h of cultivation was kept constant at 0.125 g/min. Based on the dissolved oxygen tension data, the cultures entered a carbon-limited state after 9.4 h ± 0.53 of fermentation. Fermentations were run in triplicates for 5 days. The selected replicate size allows detecting significant logFC in the expression of at least ± 0.5 magnitude, as determined in benchmarks (Schurch et al., 2016). Samples were taken at six timepoints: 21, 26, 45, 71, 94, and 118 h after fermentation started. The samples were measured in cell density (OD650), and AMY activities were measured with an in-house assay. The assay (after 1/6,000 dilution) states the enzyme amount that breaks down 5.26 g starch per hour. This activity measure is proportional to the enzyme yield.

## RNA-sequencing dataset

All samples were immediately mixed with 5 ml of 100% ethanol and stored on dry ice. The RNA extraction and purification method is the identical phenol-chloroform protocol of Geissler et al. (2022). RNA libraries and sequencing were conducted by BGI Hong Kong with DNBseq in single-ends of 50 bp length. RNA libraries were prepared with 3′ adapter sequence AAGTCGGAGGCCAAGCGGTCTTAGGAAGACAA and the 5′ adapter AAGTCGGATCGTAGCCATGTCGTTCTGTGAGC CAAGGAGTTG. The 36 samples (triplicates, two strains, and six timepoints) were sequenced in three batches with technical replicates for QC (**Supplementary Table 1**). The computational analyses were conducted in an adapted workflow of Geissler et al. (2022) (doi: 10.5281/zenodo.4534403), which provides a pipeline in a Snakemake framework nested in computational reproducible Anaconda environments (Koster and Rahmann, 2012). In concordance with the read quality assessment of FastQC (version 0.11.8) (Andrews, 2010), any adapter contaminations were removed with Trimmomatic (version 0.39) for up to two seed mismatches at a minimal 10 bp sequence overlap and 30 bp palindromic overlap (Bolger et al., 2014). In a sliding window of 4 bp, reads were clipped for average Phred score quality below 20. From the 3′ of reads, positions with quality below 3 were removed. Finally, a minimal length of 40 bp was required for filtered and cleaned reads. Reads were mapped against the respective +prsA and control genome sequence with Segemehl (version 0.3.4, default settings) (Hoffmann et al., 2009). The mapping and QC filtering statistics are given in **Supplementary Table 2**. Expression levels of coding and non-coding annotations (see below) in the respective strains were quantified for uniquely mapping reads with featureCounts (subread version 1.6.4, ≥50% overlaps). Annotation coordinates in the respective strains were determined by liftOver (version 377) from the reference assembly (NC_000964.3) based on a pairwise alignment with LASTZ (version 1.0.4) (Harris, 2007; Liao et al., 2014; Haeussler et al., 2019).

## Novel potentially transcribed regions

Reference annotations of coding, non-coding RNA (ncRNA), transcripts, untranslated regions (UTRs), and RNA structures were used from the BSGatlas (version 1.1). The BSGatlas uses separate annotation entries to specify which regions of an mRNA transcript are the coding, untranslated, or potential *cis*-regulatory RNA structure parts. Such a distinguishment to the UTR element has advantages since *cis*-regulatory RNA structures can overlap coding regions. The BSGatlas unifies multiple databases and annotation resources, such that it also includes annotations for well-known ncRNA. Additional 141 putative ncRNA annotations from a tiling-array study were used (which are not part of the BSGatlas) (Nicolas et al., 2012; Geissler et al., 2021). Relative to these reference annotations and all transcript and untranslated regions (UTRs) annotated in the BSGatlas, we checked our RNA-seq data for transcription signals in 1,645 unannotated regions. The additional tiling-array annotations and un-annotated regions were determined with the R library plyranges (version 1.6.0) and GenomicRanges (version 1.38.0) combined with an overlap helper script from BSGatlas' analysis code (doi: 10.5281/zenodo.4305872) in R (version 3.6.3) (R Development Core Team, 2008; Lawrence et al., 2013; Lee et al., 2018). Un-annotated regions shorter than 100 bp (the minimum length for >99% of the transcripts in the BSGatlas) were excluded from any further expression analysis. The expression counts for all coding/non-coding sequences and *cis*-regulatory RNA structures were normalized with DESeq2's size-factor estimation (version 1.26.0) (Love et al., 2014). With respect to the downstream analysis of expression signals, we excluded the UTR annotations for improved interpretability, although we still retained all structured RNA *cis*-regulatory annotations. With the possible overlap between *cis*-regulatory RNAs and coding sequences, reads mapping within such overlaps can be counted twice during the quantification of expression. For a total of 542 unannotated regions, we observe

expression signals of normalized read counts relative to gap length of at least 4/50 bp (corresponds to four times average coverage) (**Supplementary Figure 1**). We chose not to narrow down the transcribed regions because we found that a read coverage-based approach (as suggested in the workflow used in the RNA-seq dataset, last section) resulted in fragmented results (see example in **Supplementary Figure 9**). These regions were assumed to be *novel potentially transcribed regions* (NPTRs) (see **Supplementary Table 3**); all other unannotated regions were excluded from the subsequent expression analysis. We used the open reading frame (ORF) predictions of prodigal (version 2.6.3, default settings) to check for potential not-yet annotated coding elements within NPTRs (Hyatt et al., 2010). We also verified the overall quality of the ORF predictions by checking for overlaps with all known coding gene annotations of the BSGatlas. For each overlapping ORF-gene overlap (as detected by plyranges, see above), we computed the Jaccard similarity, which is the ratio of the length in the intersection of both annotations over their union.

## Differential expression analysis

The expression levels of the coding/non-coding sequences, NPTRs, and *cis*-regulatory RNA structures were assessed for biological reproducibility in expression counts with scatter plots (**Supplementary Figure 2**). The scatter plots did not indicate visually striking patterns of batch effects according to the sequencing plan (**Supplementary Table 1**). The principal component analysis (PCA) inspection of the top 100 most variants that expressed annotations (without further diff. expression analysis) confirmed the relevance of the experimental design in the latent structure of the expression data with the principal components corresponding to the strains and time aspect (**Supplementary Figure 3**). Differential expressions for pairwise comparisons between the strains at each of the six time points and within each strain along the time axis (**Figure 1C**) were assessed with the DESeq2's Wald test. Similar to the analysis presented in Geissler et al. (2022), the pairwise tests were weighted in a stage-wise procedure to guarantee an overall false discovery rate relative to the number of annotations: each annotation was screened for dynamic expression with a log-ratio test against a static expression model before confirming which of the pairwise tests had changes in expression. The screening and pairwise tests included a linear factor in the regression models to account for potential batch effects. The stage-wise weighting was conducted with stageR (version 1.8.0) (Van den Berge et al., 2017) and differential expression was called for adjusted $p$-values $< 0.01$. Overall, 2,127 annotations were detected as differentially expressed (**Table 1** and **Supplementary Table 4**). Based on the z-scaled log expected mean expression levels (**Supplementary Table 5**), expression profiles were grouped in 10 k-means clusters (R implementation). The profiles per strain were clustered separately (one gene = two rows in the data matrix). The number of clusters was determined by the "elbow method" over the total within-cluster error curve (**Supplementary Figure 4**; Thorndike, 1953).

## Regulated biological processes

We investigated the set of differentially expressed genes and their upregulation and downregulation for over-representation in biological processes as annotated in Gene Ontology (GO) terms, which are readily available for 78.3% of coding genes (Caspi et al., 2014; Geissler et al., 2021). For each pairwise differential expression test (**Figure 1C**), we inspected the set of upregulated genes (those with a positive logFC) and downregulated genes separately. The over-representation analysis was performed with topGO (version 2.37.0) (Alexa and Rahnenführer, 2022). Over-representation for the respective upregulated and downregulated genes was determined with a fisher test for the significance level of 0.01 relative to the background of all expressed genes, which were determined by the DESeq2's independent filtering procedure. This procedure discards the on average lowly expressed genes in order to maximize the number of differentially expressed genes (indicated by NA for $p$-values in **Supplementary Table 4**; Love et al., 2014). The minimal term size was set to 10, and the dependencies due to GO's hierarchy were de-correlated with topGO's "elim" algorithm. After filtering for a minimal observed/expected ratio of magnitude 2 (between the 80 and 85th percentile), $p$-values were adjusted for multiple testing with a false discovery rate (FDR). The over-represented processes and the associated differentially expressed genes are listed in **Supplementary Tables 6, 7** and **Supplementary Figure 6**.

## Protein–protein interaction network analysis

The protein–protein interaction network analysis was conducted in Cytoscape (version 3.8.2) (Shannon, 2003) for the differentially expressed protein-coding genes (both with and without significant logFC between strains). High-confidence protein associations (confidence score $>0.8$) were retrieved from the STRING v11 database using stringApp (version 1.6.0) for the *B. subtilis* strain *168* (Doncheva et al., 2019; Szklarczyk et al., 2019). The resulting network was clustered with the MCL algorithm (inflation value of 2.5, confidence scores as edge weights) implemented in the clusterMaker2 app (version 1.3.1) (Enright, 2002; Morris et al., 2011). The visualization of significant between strain logFCs on the network nodes was added with Omics Visualizer (version 1.3.0) (Legeay et al., 2020).

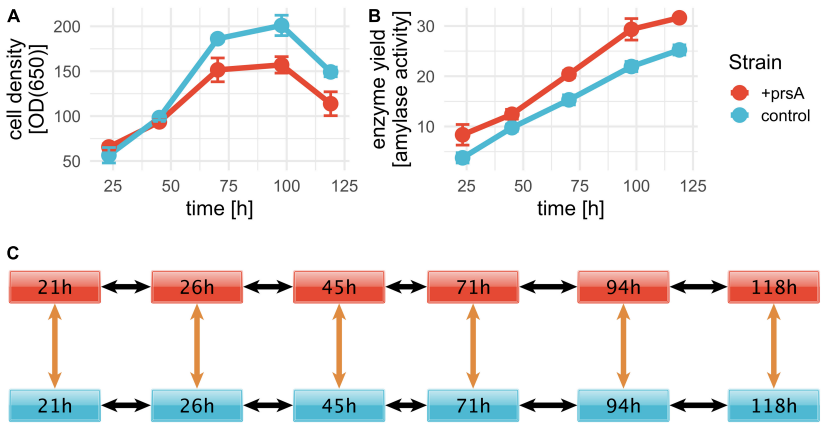



FIGURE 1
AMY fed-batch fermentations. Fed-batch fermentation was conducted in triplicates for a control strain (blue) and +prsA (red). RNA-seq samples were prepared at six timepoints: 21, 26, 45, 71, 94, and 118 h after fermentation start. Cell density and enzyme yield were measured for five timepoints: 23.2, 45, 70.2, 97.8, and 119 h. **(A)** The average cell density per strain over fermentation time was measured in optical density (OD) at 650 nm. The error bars indicate the standard deviation. **(B)** With a progressing fermentation, the yield increases. The shown yield is measured in enzyme activity (see section "Strains and fed-batch fermentation"). **(C)** For the differential expression, we investigated the significance of differential expression between the samples at six pairwise comparisons (orange arrows) and changes in expression over time in either strain for each pair (black arrows).

## Global amino acid composition

In order to interpret the regulated biological processes (see above), we inspected the global amino acid compositions of all *B. subtilis* protein-coding genes. The nucleotide sequences of all coding sequences from the BSGatlas were extracted with BSgenome (version 1.54.0) (Pagès, 2021). The corresponding amino acid sequences were determined according to the bacterial genetic code with Biostrings (version 2.54.0) (Pagès et al., 2019). Here, we used only the 99.3% of the coding genes that were completely relative to their corresponding amino acid sequences; that is, they used all codons encoded in their nucleotide sequences, correctly started with methionine, and ended with a stop codon. The composition in average proportion was determined for these complete sequences (**Table 2**).

# Results

## Novel potentially transcribed regions

### Transcriptome analysis from RNA-seq data

To elucidate potential mechanisms of *B. subtilis* secretion stress during the production of the AMY enzyme JE1 (commercial name Natalase$^{TM}$) with a particular focus on PrsA over-expression, we conducted fed-batch fermentation in triplicates for two isogenic strain conditions: one control strain and one strain with PrsA over-expression (from here on referred to as +prsA). As expected from the reduced growth upon PrsA over-expression (Vitikainen et al., 2001; Quesada-Ganuza et al., 2019), the +prsA strain has a lower cell density (**Figure 1A**) and higher AMY yield (**Figure 1B**). To capture the transcriptome dynamics during fermentation, we took out samples for RNA-seq analysis at six timepoints: 21, 26, 45, 71, 94, and 118 h after fermentation started. These timepoints correspond to sampling every 24 h (within a 3 h window) with one additional sample at the early phase of the fermentation.

### Transcriptional activity for the reference annotations

In order to comprehensively investigate both the coding and non-coding RNA elements, we quantified the RNA-seq expression according to a recently developed transcript atlas for *B. subtilis* (Geissler et al., 2021). We included 141 additional annotations from a tiling-array study that was not included in the atlas due to unclear mechanism of transcription (annotations were ambiguous as to whether they are independent full RNA transcripts or only part thereof) (Nicolas et al., 2012; Geissler et al., 2021). In the following, we refer to these annotations, together with the less well-characterized RNA elements from the atlas, as putative ncRNA. These reference annotations combine gold standard curated information, computational RNA structure biology, and transcriptomic analysis of over 100 experimental conditions (Nicolas et al., 2012; Geissler et al., 2021). Additionally, these experimental conditions suggest that still 5% of remaining unannotated regions have evidence of expression activity (Geissler et al., 2021). Fed-batch fermentations were not part of the above-mentioned experimental conditions, such that there might be a larger potential to discover fed-batch-related regions

TABLE 1 Differentially expressed annotations.

| Annotations | No. Annotations considered for analysis | Differentially expressed | Strain-specific expression |
|---|---|---|---|
| CDS | 2,674 | 1,791 (67.0%) | 1,026 (57.3%) |
| NPTRs | 355 | 234 (65.9%) | 123 (52.6%) |
| Putative ncRNA | 107 | 68 (63.6%) | 38 (55.9%) |
| Riboswitch | 37 | 20 (54.1%) | 10 (50.0%) |
| tRNA | 22 | 9 (40.9%) | 7 (77.8%) |
| sRNA | 9 | 3 (33.3%) | 2 (66.7%) |
| Synthetic PrsA | 1 | 1 (100.0%) | 1 (100.0%) |
| Synthetic AMY | 1 | 1 (100.0%) | 0 |
| asRNA | 1 | 0 | 0 |
| SRP | 1 | 0 | 0 |

For the differential expression analysis, multiple coding and non-coding annotations were considered (first column). The number of genes with minimal expression levels as determined by DESeq2's independent filtering, which were inspected for potential differential expression, is in the second column. The number of detected differentially expressed annotations in any of the pairwise comparisons (Figure 1C) is in the third column. The last column lists the number of annotations detected to have a significant difference in expression between the strains. The percentages provided in parenthesis are relative to the columns to the left (Only 355 of the 542 NPTRs passed the independent filtering).

TABLE 2 Amino acid composition.

| Amino acid | All coding genes | Diff. expressed | Highly expressed | AMY | PrsA |
|---|---|---|---|---|---|
| **Tryptophan** | 1.03% ± 0.99 | 1.06% ± 0.96 | 0.86% ± 0.81 | **4.12% (+3.1 SD)** | 0.35% (−0.7 SD) |
| **Asparagine** | 4.07% ± 2.05 | 3.87% ± 1.75 | 3.86% ± 1.44 | **8.82% (+2.3 SD)** | 2.82% (−0.6 SD) |
| Histidine | 2.30% ± 1.51 | 2.27% ± 1.33 | 2.12% ± 1.10 | 4.31% (+1.3 SD) | 1.06% (−0.8 SD) |
| Tyrosine | 3.57% ± 1.95 | 3.39% ± 1.56 | 3.13% ± 1.32 | 5.49% (+1.0 D) | 3.17% (−0.2 SD) |
| Glycine | 6.67% ± 2.74 | 6.95% ± 2.41 | 7.62% ± 2.08 | 8.82% (+0.8 SD) | 6.34% (−0.1 SD) |
| **Aspartic acid** | 5.11% ± 2.29 | 5.06% ± 2.10 | 5.17% ± 1.83 | 6.86% (+0.8 SD) | **11.27% (+2.7 SD)** |
| Threonine | 5.32% ± 1.87 | 5.44% ± 1.75 | 5.51% ± 1.35 | 5.88% (+0.3 SD) | 4.23% (−0.6 SD) |
| Arginine | 4.19% ± 2.14 | 4.07% ± 1.91 | 4.24% ± 1.92 | 4.31% (+0.1 SD) | 0.70% (−1.6 SD) |
| Proline | 3.48% ± 1.71 | 3.55% ± 1.49 | 3.93% ± 1.24 | 3.53% (+0.0 SD) | 0.35% (−1.8 SD) |
| Glutamine | 3.87% ± 2.04 | 3.85% ± 1.82 | 3.65% ± 1.42 | 3.73% (−0.1 SD) | 6.34% (+1.2 SD) |
| Alanine | 7.36% ± 2.84 | 7.82% ± 2.65 | 8.25% ± 2.29 | 7.06% (−0.1 SD) | 7.39% (+0.0 SD) |
| Phenylalanine | 4.63% ± 2.40 | 4.53% ± 2.20 | 4.21% ± 1.94 | 4.31% (−0.1 SD) | 1.76% (−1.2 SD) |
| Valine | 6.75% ± 2.37 | 6.92% ± 2.11 | 7.38% ± 1.91 | 6.27% (−0.2 SD) | 6.69% (−0.0 SD) |
| Methionine | 2.46% ± 1.43 | 2.52% ± 1.31 | 2.50% ± 1.11 | 2.16% (−0.2 SD) | 1.76% (−0.5 SD) |
| Serine | 6.22% ± 2.26 | 6.28% ± 2.16 | 6.04% ± 2.04 | 4.51% (−0.8 SD) | 4.93% (−0.6 SD) |
| Cysteine | 0.91% ± 1.13 | 0.84% ± 0.91 | 0.70% ± 0.72 | 0.00% (−0.8 SD) | 0.35% (−0.5 SD) |
| **Lysine** | 7.50% ± 3.14 | 7.13% ± 2.78 | 6.79% ± 2.17 | 4.90% (−0.8 SD) | **17.96% (+3.3 SD)** |
| Glutamic acid | 7.37% ± 3.29 | 7.20% ± 3.13 | 7.23% ± 2.77 | 4.31% (−0.9 SD) | 9.15% (+0.5 SD) |
| Leucine | 9.70% ± 3.06 | 9.72% ± 2.86 | 9.42% ± 2.35 | 6.67% (−1.0 SD) | 8.45% (−0.4 SD) |
| Isoleucine | 7.50% ± 2.65 | 7.53% ± 2.48 | 7.37% ± 2.15 | 3.92% (−1.4 SD) | 4.93% (−1.0 SD) |

The average amino acid compositions (rows) are shown for all *B. subtilis* endogenous (thus excl. AMY and PrsA) coding genes (second column), those that were detected as differentially expressed (third column), and the by average expression 10% most highly expressed genes (fourth column). The standard deviations are shown behind the "±" signs. The compositions of amino acids for the AMY enzyme (fifth) and the over-expressed PrsA (sixth) column are shown. The difference in standard deviations relative to the average for all genes are indicated in parenthesis. The bold font highlights amino acids with difference of more than two standard deviations.

from our RNA-seq data. Consequently, we investigated our RNA-seq data for expression in such unannotated regions.

## Novel potentially transcribed regions

There are a total of 1,645 un-annotated contiguous stretches of the genome or gaps (stranded, meaning there can be antisense located annotations) between reference annotations of length >100 bp (minimal length for 99.5% of transcripts in the atlas). We detect novel potentially transcribed regions (NPTRs) by inspecting the average RNA-seq read coverages over the entire unannotated gap region (read counts, DESeq2 size-factor normalized, relative to the lengths). Relative to the 50 bp sequencing lengths (see section "RNA-sequencing dataset"), 70% of atlas annotations were on average expressed by four reads

and 30% by one read. In contrast, only 20% (542) of unannotated regions were on average covered by four reads. This high coverage for these 542 NPTRs (**Supplementary Figure 1**) indicates that the NPTRs may have functional importance and that it would be relevant to include these in subsequent expression analysis (see **Supplementary Table 3**).

## PrsA over-expression changes gene expression regulation of the global transcriptome

### Differential expression

We assessed the impact of PrsA over-expression on the bacterial transcriptome by analyzing the expression levels of coding and non-coding sequences (see section "Transcriptional activity for the reference annotations" above), including the 141 additional annotations and the 542 NPTRs with DESeq2. For each region, we performed 16 pairwise differential expression tests: six tests between the two strains on each timepoint and $2 \times 5$ tests from one timepoint to the next in both strains (**Figure 1C**). Since each pairwise test corresponds to a separate hypothesis test, we used stage-wise testing to adjust for the overall false discovery rate (FDR) per annotation (Love et al., 2014; Van den Berge et al., 2017). Compared to controlling the FDR per hypothesis, the overall FDR increases statistical power and guarantees the FDR relative to the gene/annotation number, independent of the number of hypotheses (Van den Berge et al., 2017). As part of the differential expression analysis, DESeq2's independent filtering detected about half of all coding sequences and 355 of 542 NPTRs as expressed (Love et al., 2014). At an overall FDR $p$-adj. $\leq$0.01, we detected differential expression for 1,793 coding sequences (67% of expressed genes), 234 NPTRs (66%), 68 putative ncRNAs (64%), 20 riboswitches (54%), 9 tRNAs (41%), and 3 sRNAs (33%) (**Table 1** and **Supplementary Table 4**). The differentially expressed coding genes include the AMY enzyme and the over-expressed PrsA. Between 50 and 78% of these biotypes had strain-specific expression patterns (significant difference for at least one of the six between strain tests). PrsA had strain-specific expression (as expected by not being inserted into the control strain's genome). Notably, no strain-specific expression was detected for AMY.

To further assess the coding potential of the differentially expressed NPTRs, we leveraged a set of 4,226 ORF predictions (Hyatt et al., 2010). These ORF recall 4,085 of 4,332 known coding sequences perfectly (overlap with Jaccard similarity >90%), which corresponds to a recall of 94.3% with a precision of 96.7%. Only 141 ORF predictions do not recall coding genes. Furthermore, only 18 overlap unannotated regions (>100 bp); for 3 ORF, the overlap is less than 5% of the ORF length (**Table 3**). Also, only two of the ORFs are fully located within an NPTR that has detected differential expression; in both cases, the ORFs antisense overlaps the 3′ ends of the coding genes: The electron transfer flavoprotein *etfA* and the gene of unknown function *yobB*.

### The regions with the highest expression changes

The strain-specific expression patterns of PrsA and the respective logFC between the two strains on all six timepoints were the most extreme observed in this study with logFC values up to a factor of 20 at each timepoint. Other extreme logFC values were observed for genes from operons encoding a variety of biological functions (**Table 4**). The NAD biosynthesis genes of the *nadABC* operon (Rodionov et al., 2008) also have extreme logFC, but they undergo both extreme upregulation and downregulation in the control strain with *nadA* and *nadB* being downregulated from timepoint 21 to 26 h (both logFCs$<-6$, adj. $p < 0.004$) and subsequently upregulated from 26 to 45 h (both logFCs $\sim+7$, adj. $p <$ 3e-10). Due to the secretion stress, the production strains attempt to sporulate despite being unable to do so (Geissler et al., 2022). Consistently, the two sporulation genes, namely *safA* and *coxA*, were among the most extremely regulated (logFC $> 6$, adj. $p <$ 2.3e-5). Other extreme changes in expression (logFC $< -5$) were observed for the spore killing factors *skfA* and *skfB* (González-Pastor, 2011), the sporulation controlling factor *spoIIGA* (Ramos-Silva et al., 2019), the bacitracin resistance genes *bceA* and *bceB* (Ohki et al., 2003), the for NADH during fermentation essential lactate dehydrogenase *ldh* (Cruz Ramos et al., 2000; Larsson et al., 2005), and an NPTR antisense to the gene of unknown function *ytta* (Asai et al., 2007).

### Biological processes and differentially expressed genes are mutually associated

The investigation of the overall expression profiles from a k-means clustering (**Figure 2A**) on the average expected expression at each timepoint (**Supplementary Table 5**) shows marked differences in the expression dynamics between the strains (**Figure 2C**). Also, all profiles indicate a substantial shift in dynamics between timepoints 45–71 h, during which the cell population increased the most (**Figure 1A**): For instance, profiles 4 and 5 drop in expression levels at that timepoint but recover and even exceed the starting expression level whereas profiles 7 and 8 have drastically downregulated expression at that timepoint and do not recover (**Figure 2B**). Genes and other biotypes with strain-specific expression patterns had predominately different expression profiles between the strains, whereas those without strain-specific expression had the same (**Supplementary Figure 5**). Therefore, *B. subtilis* regulates gene expression both timepoint- and strain-specifically.

We assessed which biological processes [annotated in Gene Ontology (GO), terms (Ashburner et al., 2000)] are over-represented among the differentially expressed genes in each time and strain pairwise comparison (**Figure 1C**). We compared the numbers of respective upregulated or

TABLE 3 ORF overlapping un-annotated regions.

| Overlapping gap | ID | Gap length | ORF start | ORF end | ORF strand | ORF length | % ORF within gap |
|---|---|---|---|---|---|---|---|
| Low expression | gap-865 | 18994 | 164501 | 164923 | – | 18994 | 100 |
| Low expression | gap-964 | 382 | 635149 | 635262 | – | 382 | 100 |
| Low expression | gap-1109 | 277 | 1233396 | 1233599 | – | 277 | 100 |
| Low expression | gap-1244 | 15761 | 1784706 | 1784837 | – | 15761 | 100 |
| Low expression | gap-1402 | 1357 | 2747665 | 2747853 | – | 1357 | 100 |
| Low expression | gap-643 | 8279 | 3341033 | 3341254 | + | 8279 | 100 |
| NPTR† | gap-1322 | 7336 | 2221838 | 2221951 | – | 7336 | 100 |
| **NPTR with diff. expr.** | **gap-1293** | **743** | **2050930** | **2051037** | – | **743** | **100** |
| **NPTR with diff. expr.** | **gap-538** | **7687** | **2915229** | **2915390** | + | **7687** | **100** |
| Low expression | gap-1495 | 1272 | 3419414 | 3419656 | – | 1272 | 96.7 |
| Low expression | gap-1367 | 140 | 2560095 | 2560259 | – | 140 | 84.8 |
| NPTR† | gap-113 | 1795 | 718436 | 718597 | + | 1795 | 83.3 |
| Low expression | gap-1166 | 197 | 1453322 | 1453525 | – | 197 | 79.4 |
| **NPTR with diff. expr.** | **gap-1354** | **1038** | **2460278** | **2460625** | – | **1038** | **26.7** |
| NPTR without diff. expr. | gap-1296 | 581 | 2056650 | 2057114 | – | 581 | 23.2 |
| **NPTR with diff. expr.** | gap-1058 | 2975 | 1033458 | 1034003 | – | 2975 | 4.8 |
| Low expression | gap-449 | 2395 | 2555468 | 2555845 | + | 2395 | 2.1 |
| Low expression | gap-90 | 2286 | 608879 | 609391 | + | 2286 | 1.4 |

Shown are each overlap between ORFs (that did not recall a coding gene) and un-annotated regions (>100 bp). Some of these un-annotated regions with sufficient expression signal were considered as NPTR and further processed for differential expression analysis (column 1–3). Bold texts highlights gaps with detected differential expression. Based on the coordinates of the predicted ORF (column 4–7), the overlap of the ORF with the un-annotated gap is expresses relative to the length of the ORF (column 8).

†Un-annotated region with sufficient expression signal to be considered as NPTR, but which was excluded from further analysis by the independent filtering procedure of the differential expression analysis.

downregulated genes relative to the number of expressed genes (see section "Materials and methods"). A total of 24 processes had significant over-representation (Fisher's exact test, FDR *p*-adj. ≤0.01). We inspected the list of differentially expressed genes per process (**Supplementary Table 7**) in combination with meta-information available in the BSGatlas, particularly KEGG pathway annotations (Kanehisa and Goto, 2000; Geissler et al., 2021). Notably, the detected over-represented processes annotate genes with differentially expressed logFC predominately above the background logFC distribution of genes without detected differential expression (**Supplementary Figure 8**). Furthermore, some of the top 10 most extremely upregulated and downregulated genes (**Table 4**) were annotated by the detected processes (**Supplementary Table 7**), namely cell wall macromolecule catabolic process (*safA* and *skfA*), response to stress (*nadC* and *nadE*), and ATP biosynthetic process (*ldh*). We further inspected the detected biological processes (**Figure 3**) for their relevance with respect to fed-batch fermentation, as described in the sections later.

### Nucleotide biosynthesis

It is well established that an ample supply of nucleotides is needed for efficient AMY protein expression (Hosoda et al., 1959) and thus also the nucleotide precursors, such as UMP and IMP, are of regulatory interest (Peifer et al., 2012; Hohmann et al., 2016). Consistently, the over-representation investigation indicates an upregulation of UMP (GO:0006222) and IMP (GO:0006189) biosynthesis in the +prsA strain from timepoint 26 to 45 h and 95 to 118 h, respectively. The monosaccharide catabolic genes (GO:0046365), especially the genes involved in the ribose synthesis *via* the pentose phosphate pathway (**Supplementary Table 7**), are upregulated in the control strain from timepoint 45 to 71 h. The pteridine-containing compound metabolic process (GO:0042558) was over-represented by genes upregulated from the first to the second timepoint in both strains. These specific genes are also part of the folate biosynthesis pathway, which is essential for both purine and pyrimidine synthesis (Kilstrup et al., 2005) and therefore quintessential for AMY production (Hohmann et al., 2016; Hosseini et al., 2018).

## PrsA over-expression affects genes involved in energy metabolism

### ATP biosynthesis

The ATP biosynthetic process (GO:0006754) was significantly downregulated in +prsA compared to the control strain on the first timepoint of the fermentation. Furthermore, the data suggest that the energy derivation by oxidation of organic compounds (GO:0015980) was further downregulated in +prsA from the first to the second timepoint within the first day of fermentation. The differentially expressed genes associated with both processes comprise a long list

TABLE 4 Most extreme observed logFCs.

| Name | Type | Test | logFC | Adj. *P* | Location |
|---|---|---|---|---|---|
| nadB | Coding | Between strains, 26 h | 7.4 | 1.07E-10 | 2847871<-2849466 |
| nadB | Coding | Control, 26->45 h | 7.2 | 2.65E-10 | 2847871<-2849466 |
| nadC | Coding | Control, 26->45 h | 7.0 | 2.89E-09 | 2847048<-2847917 |
| nadC | Coding | Between strains, 26 h | 6.9 | 5.26E-09 | 2847048<-2847917 |
| nadA | Coding | Control, 26->45 h | 6.8 | 7.82E-11 | 2845955<-2847061 |
| nadA | Coding | Between strains, 26 h | 6.8 | 1.33E-10 | 2845955<-2847061 |
| safA | Coding | Control, 26->45 h | 6.4 | 4.05E-09 | 2844675<-2845838 |
| safA | Coding | Between strains, 26 h | 6.2 | 8.98E-09 | 2844675<-2845838 |
| coxA | Coding | Between strains, 26 h | 6.2 | 1.82E-05 | 2843931<-2844527 |
| coxA | Coding | Control, 26->45 h | 6.1 | 2.28E-05 | 2843931<-2844527 |
| spoIIGA | Coding | Control, 26->45 h | −5.2 | 8.47E-67 | 1603779->1604708 |
| skfB | Coding | Control, 26->45 h | −5.4 | 2.86E-60 | 214175->215407 |
| skfA | Coding | +prsA, 26->45 h | −5.4 | 7.31E-09 | 213941->214108 |
| skfA | Coding | Control, 26->45 h | −5.7 | 4.39E-24 | 213941->214108 |
| bceA | Coding | Control, 26->45 h | −6.2 | 5.76E-107 | 3111327<-3112088 |
| nadA | Coding | Control, 21->26 h | −6.2 | 0.002258927 | 2845955<-2847061 |
| bceB | Coding | Control, 26->45 h | −6.5 | 1.62E-127 | 3109397<-3111337 |
| nadB | Coding | Control, 21->26 h | −6.5 | 0.003566971 | 2847871<-2849466 |
| ldh | Coding | +prsA, 21->26 h | −6.6 | 2.42E-13 | 329774->330739 |
| gap-1449 | NPTR | control, 26->45 h | −6.9 | 1.85E-48 | 3108525<-3109352 |

The table lists the top 10 most extreme upregulated and downregulated genomic elements according to their logFC of differential expression (fourth column). prsA is excluded since it was upregulated with an approximate logFC of 20 between strains at all-time points. For each genomic element, the locations (last column) are relative to the reference genome (see section "Materials and methods"). The pairwise tests (third column) refer to the conducted differential expression analysis (Figure 1C), and the corresponding adjusted *P*-values are listed in the fifth column.

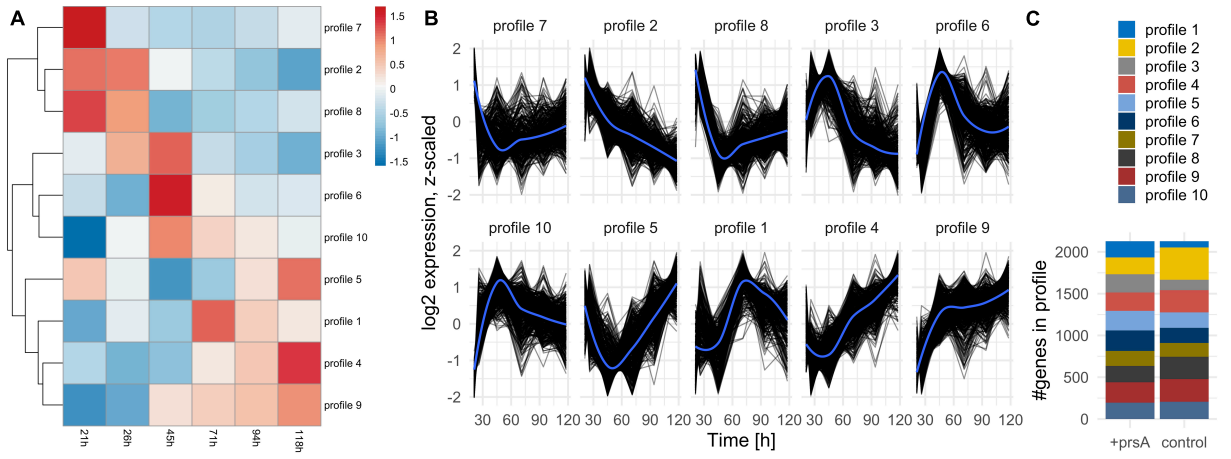



FIGURE 2
Expression profiles. **(A)** Heatmap of the expression profile over time (columns) for all differentially expressed coding and non-coding annotations investigated separately per strain. The resulting profiles were clustered (rows) and re-arranged by a complete linkage tree. **(B)** Profiles of expression per cluster for each annotation (black lines). An overall average expression according to a loess regression is added in blue. **(C)** The number of annotations per profile in either strain. The expression dynamics for each annotation can be in two separate profiles in the strains.

(>50, see Supplementary Table 7) of core energy metabolic enzymes from the citrate cycle, oxidative phosphorylation, and glycolysis. Nevertheless, the list also overlaps with the starch and sucrose metabolism pathway, particularly with glycogen biosynthesis (*glgA*, *glgB*, *glgC*, *glgD*, and *glgP*) (Kiel et al., 1994). Consistent with these observations, the carbohydrate transport (GO:0008643) was also downregulated in +prsA on the first timepoint. In contrast, the cellular ketone metabolic process (GO:0042180) was upregulated in the control strain from the first to the second timepoint. Ketones are essential for the

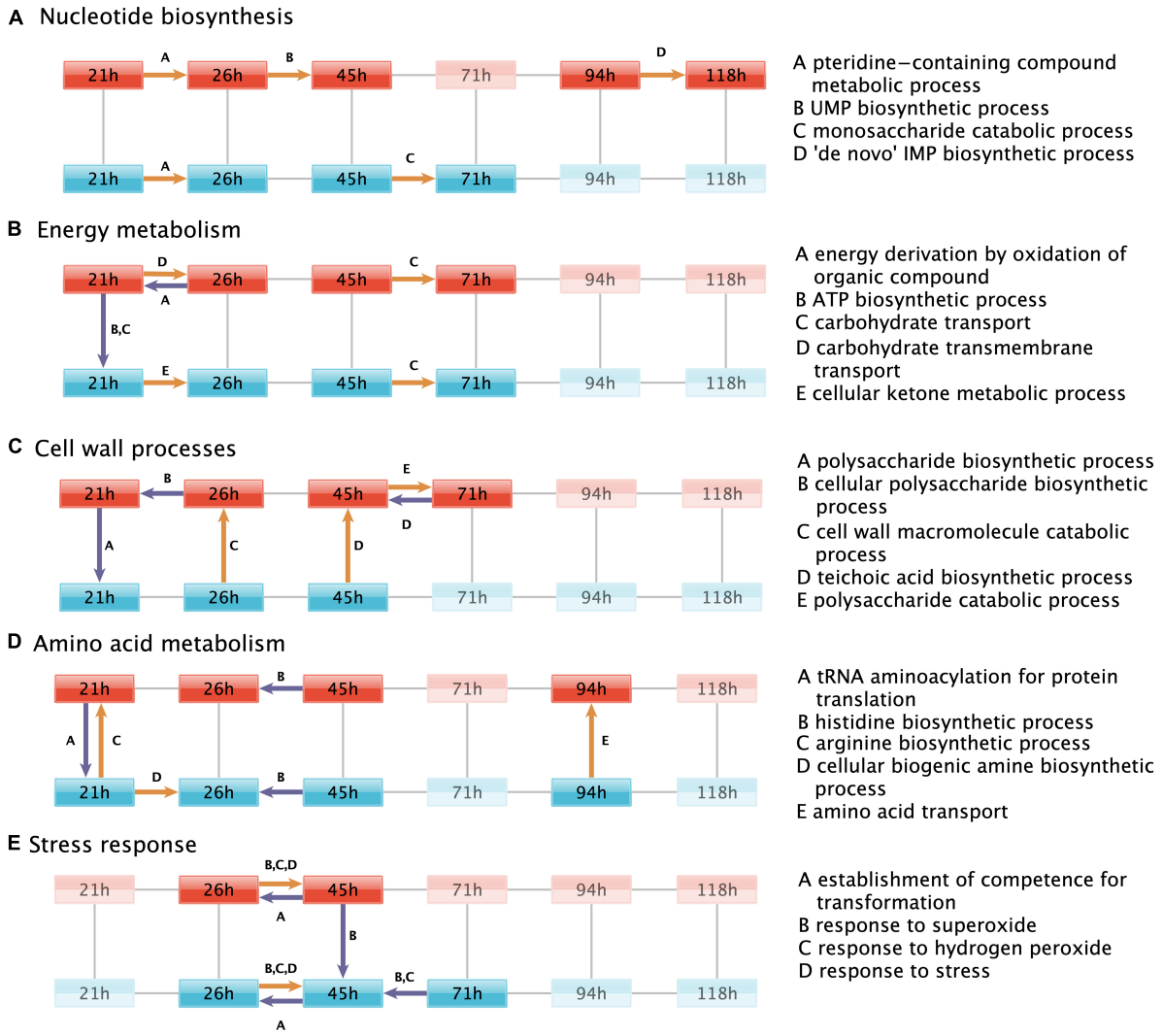



FIGURE 3
Regulated biological processes. Biological processes that are over-represented by the genes differentially expressed in each of the pairwise comparisons (black lines) between the fermentation timepoint in the +prsA (red) and control strain (blue). For simplicity, the regulated processes are grouped in subplots according to the same biological functions discussed in the result sections, which touch upon **(A)** nucleotide biosynthesis, **(B)** energy metabolism, **(C)** cell wall processes, **(D)** amino acid metabolism, and **(E)** stress response. **Supplementary Figure 7** shows the regulated processes without further functional subdivision. Colored arrows indicate a pairwise comparison that was over-represented in a process (see description to the right). The arrows point to the conditions in which expression levels were higher. Upregulation in the +prsA strain or upregulation with time progression of the fermentation is highlighted in orange, whereas downregulation is shown in purple. In each subplot, time-strain conditions not adjacent to an arrow are grayed out.

biosynthesis of menaquinone (Lu et al., 2008). Menaquinone is *B. subtilis*' respiration coenzyme, similar in function to ubiquinone in human mitochondria (Lemma et al., 1990). Nevertheless, the ATP biosynthetic process (GO:0006754) was not detected significantly over-represented by the regulated genes at the other fermentation timepoints.

## Altering carbohydrate transport during fermentation

The over-representation analysis also suggests that both strains have an upregulated carbohydrate transport (GO:0008643, GO:0034219) from 45 to 71 h. The transport might also be upregulated in the +prsA strain from the first to the second timepoint.

## PrsA over-expression affects genes involved in cell wall destabilizing processes

Low PrsA protein abundances and increased concentrations of teichoic acid can reduce cell growth and cell wall disruption

(Driessen et al., 1998; Hyyryläinen et al., 2000). For instance, the inhibition of the dlt operon–which is key to teichoic acid synthesis–increases AMY yields (Hyyryläinen et al., 2000; Yan and Wu, 2017). However, our data suggest that not only *dltB* expression is upregulated in +prsA on timepoint 45 h (logFC = 0.86, adj. $p$ < 2.11e-5) but also the entire teichoic acid biosynthetic process (GO:0019350). Additional processes relating to cell wall molecules and polysaccharide biosynthetic (GO:0033692, GO:0000271) were observed as downregulated in +prsA. Nevertheless, not only does our data suggest that the biosynthesis is downregulated, the corresponding catabolic processes (GO:0016998, GO:0000272) might be upregulated.

## Upregulation of amino acid metabolism during PrsA over-expression

### Regulated amino acid metabolism

Genes of the arginine biosynthetic process (GO:0006526) are over-represented among the genes upregulated in the +prsA strain on the first timepoint and for the amino acid transport (GO:0006865) at timepoint 94 h after fermentation started. The histidine biosynthetic process (GO:0000105) was detected as downregulated from timepoint 26 h to the timepoint 45 h in both strains. The data suggest also that the tRNA aminoacylation for protein translation (GO:0006418) is downregulated in +prsA on the first timepoint, and that the cellular biogenic amine biosynthetic process (GO:0042401) is upregulated in the control strain from the first to the second timepoint.

### Expected changes in amino acid metabolism

Given the observed potential regulation in amino acid metabolism above, we investigated to which extent these might be the result of the peptide sequence of the secreted AMY. The inspection of the codon composition of all coding genes suggests that the AMY and the over-expressed PrsA contain substantially more tryptophan, asparagine, aspartic acid, and lysine (more than 2 standard deviations from the average proportion, **Table 2**). Tryptophan was the strongest over-represented amino acid in AMY (+3.1 standard deviations). In comparison, the subset of neither differentially expressed endogenous (excl. AMY and PrsA) coding genes nor 10% of most highly expressed endogenous genes have changes in the overall composition (within 1 standard deviation). The PrsA was extremely over-expressed in the +prsA strain (logFC > 19, adj. $p \leq$ 5.27e-40). By average expression, AMY was the 5th and PrsA the 34th highest expressed gene (see **Supplementary Table 6**). Thus, the enrichment of these four amino acids in AMY and PrsA should have physiological relevance: given the high energetic cost of tryptophan biosynthesis (Akashi and Gojobori, 2002), the evolutionarily adapted amino acid metabolism will be affected (Smith and Chapman, 2010).

## Protein–protein interactions of stress response and competence transformation

### Stress response turning point

The over-representation investigation reveals that both strains upregulate parts of their stress response concerning the reactive oxygen species (ROS) response (GO:0006950 and the two children terms GO:0042542, GO:0000303) from timepoint 26 h to 45 h. Simultaneously, the strains downregulate the establishment of competence for transformation (GO:0030420). The protein ClpC is the key switch between heat shock (including secretion stress) and competence regulation (Turgay et al., 1997). During stress, a three-protein complex of ClpC, MecA, and ComK is formed (Turgay et al., 1997). The bound central competence regulator ComK can no longer act as a transcription regulator, which prevents the establishment of competence (Turgay et al., 1997). According to our results, *clpC* undergoes significant differential expression during fermentation in both strains, but neither *comK* nor *mecA* had significant expression changes though both were expressed (**Supplementary Table 4**). Given that the molecular mechanism of the ClpC switch (i) is post-translational, (ii) does not directly impact the transcription levels of the involved genes, and (iii) involves a third factor, the analysis by pairwise comparison of expression levels cannot detect that specific interaction. Therefore, we complemented the expression analysis with protein-protein interaction (PPI) network analysis.

### Protein–protein interaction network analysis

We retrieved PPIs from the STRING database for the *B. subtilis* strain 168. STRING provides a list of functional associations from multiple evidence channels, such as curated knowledge from known metabolic pathways and protein complexes, physical PPIs from lab experiments (e.g., pull-down assays), predicted interactions from text mining of the biomedical literature, or associations based on co-expression analysis (Szklarczyk et al., 2019). The resulting network of 4,774 high-confidence associations (confidence score >0.8) among 1,770 of the 1,791 differentially expressed protein-coding sequences was clustered into 201 protein clusters using MCL (Enright, 2002; Morris et al., 2011; Doncheva et al., 2019). In combination with the significant logFCs between the +prsA and control strains (Legeay et al., 2020), we manually inspected four clusters with interesting patterns regarding this study's outset (**Figure 4**). These are described in the following sections below.

### Two-component system

The first PPI cluster consists of the CssRS two-component system, including the involved proteases (see Introduction, **Figure 4A**). However, the cluster contains an additional association between the stress signal transducer CssS and YkoJ of unknown function. The *ykoJ* expression during secretion of

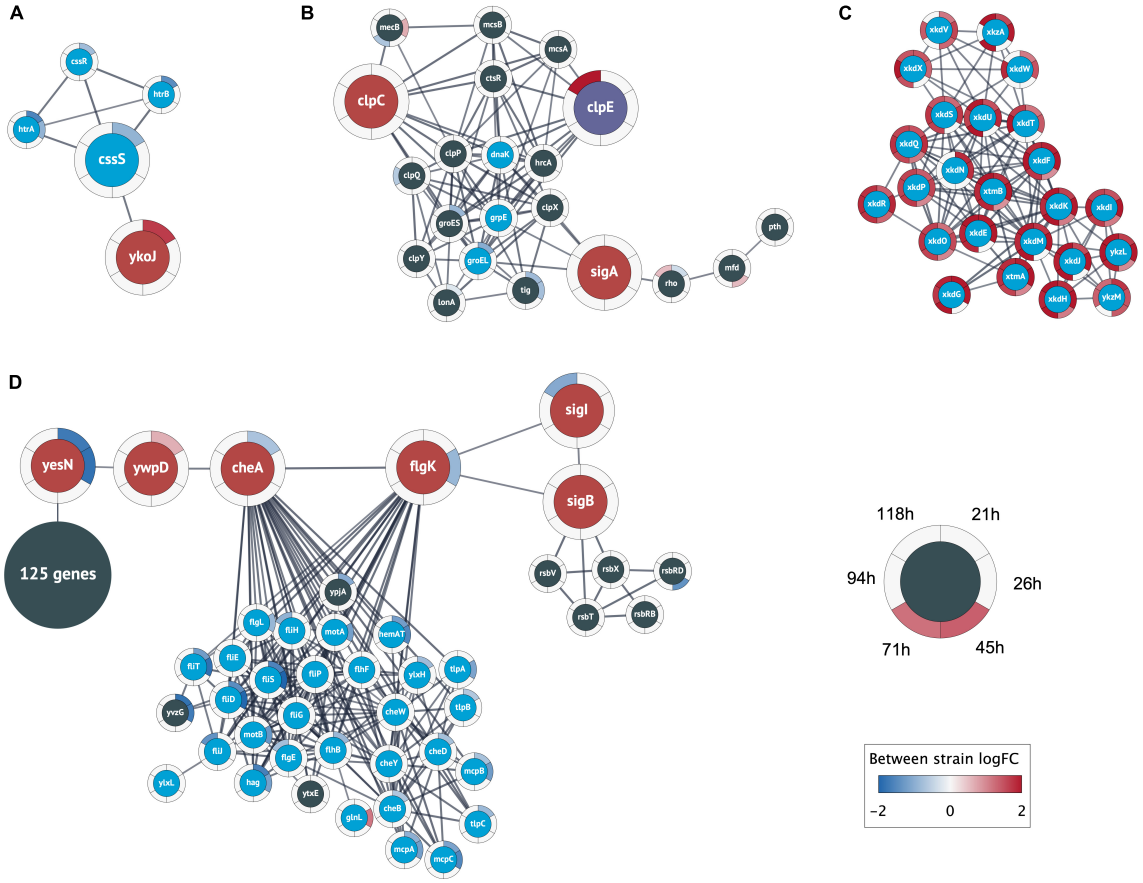



FIGURE 4
Relevant clusters of differentially expressed genes. Nodes represent protein coding genes and edges correspond to high-confidence protein interactions retrieved from STRING. The differential expression between strains is shown as rings around the nodes, where each ring contains the logFC values for each time point comparison in a blue-white-red color gradient (see figure legend). A high positive logFC is colored red and indicates a significantly larger expression in the +prsA strain compared to the control. Non-significant differential expression is shown as 0 logFC (white). The logFC color gradient was truncated at ±2. **(A)** The genes in this cluster include the central heat shock stress two-component system of CssRS and the proteases HtrAB (blue nodes). The cluster also contains the gene ykoJ of unknown function (red node) connected to the stress transducer CssS (large blue node). **(B)** This cluster contains the competence/heat shock switch protein ClpC (leftmost red node) and the universal sigma factor SigA (rightmost red node); SigA and ClpC share interactions with the tree heat shock proteins dnaK, grpE, and groEL (blue nodes). The cluster also contains ClpE (purple node) that had substantially higher expression in +prsA at timepoint 118 h (logFC ∼2.6). **(C)** The analysis found a cluster of 24 prophage or prophage-like genes that were closely interacting and had significantly higher expression in +prsA throughout the fermentation. **(D)** The largest cluster contains a "bottleneck" of high-confidence interactions at two genes of unknown function (yesN and ywqD) between 125 genes of various catalytic function (summarized as one node) and 29 chemotaxis genes (blue nodes) and the central chemotaxis signal protein CheA, the flagellar hook-filament FlgK, the general stress repose sigma factor SigB, and the RNA polymerase sigma factor SigI.

a vaccine compound (beta-toxoid) positively depends on CssS (Nijland et al., 2007). In contrast, the expression during AMY might have a negative dependency with *cssS* being significantly lower expressed in +prsA on timepoint 21 h after fermentation start (logFC = −0.9, adjusted $p$ = 8.5e-10) and *ykoJ* significantly higher (logFC = 1.7, adj. $p$ = 1.2e-7). To our knowledge, the association YkoJ-CssS has not been characterized in the context of AMY production.

## Competence switch

The second cluster (**Figure 4B**) contains the above-described heat shock/competence protein switch ClpC (Turgay et al., 1997). The cluster also contains ClpC' repressor CtsR (Derré et al., 1999) and the universal sigma factor SigA. Furthermore, SigA and ClpC share associations with the three heat shock proteins DnaK, RrpE, and GroEL. Although *mecA* was not detected as differentially expressed, the paralog *mecB* was, and it is part of this second cluster (Persuh et al., 2002). *B. subtilis*' other two Clp-proteins ClpP and ClpE are also part of this cluster. ClpE had a significantly higher expression on timepoint 118 h in +prsA (logFC = 2.6, adj. $p$ = 0.0005), which is relevant because ClpE destabilizes the functionality of the repressor CtsR (Miethke et al., 2006).

### Prophage genes

A third cluster (**Figure 4C**) contains a set of tightly associated 24 PBSX prophage and prophage-like genes that were all significantly higher expressed in +prsA compared to control at various timepoints during the entire duration of the fermentation (logFC ∈ [0.9, 2.4], adj. $p$ ∈ [2.6e-23, 9.2e-3]). PBSX, a defective *B. subtilis* prophage (Wood et al., 1990a), is known to be potentially heat-induced (Wood et al., 1990b), and they have a potential association with the level of lytic stress resistance (Buxton, 1980).

### Potential cell motility regulation

Finally, the fourth cluster has an interesting pattern of associations involving many chemotaxis genes (**Figure 4D**). This cluster is structured into two separate interconnected components: On the one side, there are 29 chemotaxis proteins and on the other 125 protein-coding genes with various catalytic functions [116 of 125 (92.8%) genes are annotated in the general catalytic activity term GO:0003824]; however, both parts are connected by a backbone of associated genes. This backbone includes the central flagella motion frequency regulator CheA, the flagellar hook-filament FlgK, the general stress sigma factor SigB, the heat-shock protein sigma factor SigI, and the two partially characterized signal transducers YesN and YwspD (Fabret et al., 1999; Petersohn et al., 2001; Zuber et al., 2001; Asai et al., 2007; Mukherjee and Kearns, 2014). Interacting with SigB are five stress regulatory proteins induced by SigB (according to STRING annotations). Both YesN and YwsqD are described as histidine kinases, although the corresponding response regulator remains unknown (Fabret et al., 1999; Caspi et al., 2014; Zhu and Stülke, 2018; Geissler et al., 2021). Even if the regulators are unknown, the backbone has an interesting pattern of antagonistic logFC: (i) YesN is significantly lower expressed in +prsA on timepoint 21 h (logFC = −1.7, adj. $p$ = 1.4e-6) and 26 h (logFC = −1.84, adj. $p$ = 8.5e-5), (ii) YwspD is higher expressed in +prsA on 21 h (logFC = 0.6, adj. $p$ = 0.0037), and (iii) CheA lower again on 21 h (logFC = −0.7, adj. $p$ = 0.0025). The bottom-line is that the PPI analysis elucidates the tight associations between heat shock, competence transformation, cell motility, general stress response, and translation.

## Discussion

In this study, we investigated how PrsA over-expression in *B. subtilis* impacts the transcriptome during fed-batch alpha-amylase (AMY) fermentation. We carried out a temporally resolved RNA-seq study to analyze expression levels and regulation of biological processes with respect to secretion stress. We inspected a comprehensive set of coding and non-coding reference annotations and 542 novel potentially transcribed regions (NPTRs). The fermentation process strongly affects gene expression and we observe a large number of differentially expressed genes both between the strange and overtime: a total of 1,793 coding genes (67% of expressed genes), 234 NPTRs (66%), 68 putative ncRNAs (64%), 20 riboswitches (54%), 9 tRNAs (41%), and 3 sRNAs (33%) were differentially expressed. The PrsA over-expressing strain, which is consistent with prior descriptions had increased yield and reduced growth (Quesada-Ganuza et al., 2019), was observed to have a significant strain-specific differential expression for more than half of the transcribed genes. Subsequent in-depth analysis of regulated biological processes (**Figure 3**) and the PPI network of differentially expressed coding genes (**Figure 4**) shed light on the complex intertwined processes of stress pathways, core energy metabolism, and cell motility (Helmann et al., 1988; Márquez-Magaña et al., 1990; Storz and Hengge, 2010; Yan and Wu, 2019).

Concerning the NPTR, we assessed their potential to contain ORF relative to predictions that recalled 94.3% of known genes with high precision of 96.7%. A marginal fraction of these ORF overlap unannotated regions (**Table 3**). Therefore, our data do not suggest the presence of ORF in the NPTR, including those with detected differential expression in this dataset. Future investigation for potential conservation of RNA—let alone assessment of their biological function—requires RNA structure alignments that can have average sequence identities below 40% (Yao et al., 2006, 2007; Weinberg et al., 2010). We predicted the NPTR relative to a reference transcript annotation that integrates a comprehensive set of annotation databases and resources (Nicolas et al., 2012; Caspi et al., 2014; Geissler et al., 2021; Pedreira et al., 2022). Among these resources is SubtiWiki, an active community effort that comprehensively collects previously identified coding and non-coding genes (Zhu and Stülke, 2018). Therefore, we consider the NPTR to extend beyond known transcribed regions.

## Amino acid and energy metabolism

The observation of the potentially downregulated ATP biosynthesis in the +prsA strain surprised us: (i) The AMY hypersecretion is stressful and energy-intensive for the cells (Song et al., 2015). (ii) It has been hypothesized that ATP might be required for PrsA chaperone activity (Yan and Wu, 2017). (iii) The reduction of ATP levels can also increase the general stress response of *B. subtilis* (Haldenwang, 1995; Petersohn et al., 2001; Yan and Wu, 2019). The potential downregulation of ATP biosynthesis in the +prsA strain seems counterintuitive because the strain has both lower stress and higher yield than the control (Quesada-Ganuza et al., 2019). However, the reduced ATP biosynthesis might be due to the impact of the hypersecreted AMY and over-expressed PrsA on amino acid metabolism. Contrary to the evolutionary energetic adaption of the amino acid composition for secreted proteins (Smith and Chapman, 2010), the four amino acids tryptophan, asparagine, aspartic acid, and lysine are over-represented in the AMY and PrsA proteins (**Table 2**). Although the specific metabolism

processes for these four amino acids were not detected as significantly regulated during fermentation (**Figure 3**), more general amino acid processes (e.g., transport) or biosynthetic processes for other amino acids (arginine and histidine) were significantly over-represented by regulated genes. On the one hand, the upregulation of arginine synthesis and related transport mechanisms improves osmotic stress resistance (Du et al., 2011; Zaprasis et al., 2015), which in turn is beneficial to AMY production in *B. subtilis* (Zhao et al., 2018). On the other hand, the over-represented amino acids might explain the reduced ATP biosynthesis. (i) Tryptophan is the amino acid with the highest biosynthetic cost in *B. subtilis,* with a 42.9% higher cost than the second most costly amino acid (phenylalanine) (Akashi and Gojobori, 2002). (ii) The biosynthesis, in particular for costly amino acids, diverges intermediate metabolites from ATP biosynthesis (Akashi and Gojobori, 2002). In the case of tryptophan, the intermediate metabolites have already diverged from glycolysis, which also impacts the downstream citrate cycle (Kanehisa and Goto, 2000; Akashi and Gojobori, 2002). However, a more definite inspection to confirm the regulation of the amino acids and ATP metabolism would require an investigation of concentrations of the individual metabolites with for instance metabolomics.

## Cell wall destabilizing processes

The over-expression of PrsA is known to lead to reduced cell growth and cell lysis (Quesada-Ganuza et al., 2019). It was suggested that protein-protein interactions of specific PrsA protein domains are causal for these phenotypes (Quesada-Ganuza et al., 2019). Our data suggest that, on a transcription regulatory level, the PrsA-over-expressing stain has both increased polysaccharide catabolism and reduced polysaccharide biosynthesis. We hypothesize that this strongly contributes to cell wall breakdown, which leads to detrimental phenotypes. Therefore, investigating the associated differentially expressed genes could potentially be the outset to trace back the causality chain of why their regulation changes, and as a path forward to finding candidates that stabilize cell walls and increase yields. Furthermore, the PPI network analysis highlighted 24 tightly associated PBSX prophage and prophage-like genes (**Figure 4C**) that might be decisive in unraveling the PrsA over-expression lysis phenomena (Buxton, 1980; Quesada-Ganuza et al., 2019), particularly due to the heat-induced (and thus secretion stress-related) expression of the PBSX genes (Wood et al., 1990b).

## Stress and cell motility

The protein-protein interaction network analysis resulted in four clusters of proteins that we found to be relevant to this study's outset (**Figure 4**). These were the genes of the CssRS two-component secretion stress response in one cluster (**Figure 4A**), while the known ClpC regulatory switch and its associations with secretion stress, competence transformation, and associations with the universal sigma factor SigA belong to another cluster (**Figure 3B**; Turgay et al., 1997). Furthermore, the analysis provided a large cluster (**Figure 4D**) of cell motility-related genes, which is consistent with the large number of proteins involved in regulating bacterial motility (Rajagopala et al., 2007). A closer inspection of the latter cluster suggests that the proteins YesN and YwsqD might have a signaling role in balancing between cell motility and 125 genes that are annotated to have various metabolic catalytic functions, e.g., the phosphogluconate dehydrogenase. To our knowledge, the potential relationship between cell motility and AMY fermentation has not been elucidated so far, although a potential hypothesis could be that the signaling facilitates the regulation of flagellar cell motility to escape from the stress region (Helmann et al., 1988; Márquez-Magaña et al., 1990; Yan and Wu, 2019). Nevertheless, a follow-up study is needed to verify cell motility regulation during AMY production.

## Conclusion

In conclusion, our transcriptome study highlights the expression dynamics of secretion stress during fed-batch AMY fermentation. The comparison of expression levels in a PrsA over-expressing strain to a control strain showed differential expression for nearly half of the transcribed genes. A wide variety of upregulated and downregulated biological processes is related to energy and amino acid metabolism. Also, the data shows potential associations of the cell lysis phenomenon of PrsA over-expression with the stress response and cell motility. Overall, these results identify genes and biological processes, which are affected during fermentation and by the overexpression of PrsA and provide a starting point for future genetic modification of *B. subtilis* for improved yield.

# Data availability statement

The genomic sequences and RNA-seq data were deposited in the GEO database (GSE189556). The expression coverages are presented as a browser for interactive investigation (http://rth.dk/resources/bsg/prsa). The annotations of the BSGatlas are accessible at https://rth.dk/resources/bsgatlas/. The additional putative ncRNA annotations are part of the supplementary information of Nicolas et al. (2012). The RNA-seq data were processed with a reproducible pipeline located at doi:10.5281/zenodo.4534403.

## Author contributions

AG conducted the entire computational analysis and wrote the manuscript. LP extracted the RNA. ND contributed to the analysis and methodology design of the PPI network. CA contributed to the discussion of the expression analysis. EG-T contributed to the writing in the early stage. AB prepared the bacterial strains. LJ contributed to the discussion of gene clustering, enrichment analysis, and PPI network analysis. SES, CH, JV, and JG supervised the work. JG and AG made the study design. JG was the main project coordinator. All authors read and approved the manuscript.

## Funding

This work was supported by the Innovation Fund Denmark (5163-00010B) and Novo Nordisk Foundation (NNF14CC0001).

## Acknowledgments

We thank Annaleigh Ohrt Fehler for pivoting the samples for sequencing and both Bjarke Krysel Christensen and Thomas B. Kallehauge for support in conducting the fermentation and sampling.

## Conflict of interest

AB and CH were employed by the Novozymes A/S.

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.2022.909493/full#supplementary-material

## References

Akashi, H., and Gojobori, T. (2002). Metabolic efficiency and amino acid composition in the proteomes of *Escherichia coli* and *Bacillus subtilis*. *PNAS* 99, 3695–3700. doi: 10.1073/pnas.062526999

Alexa, A., and Rahnenführer, J. (2022). *topGO: Enrichment Analysis for Gene Ontology. R package Version 2.48.0*. Available online at: https://git.bioconductor.org/packages/topGO (accessed March 4, 2021).

Andrews, S. (2010). *FastQC: A Quality Control Tool for High Throughput Sequence Data*. Available online at: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (accessed July 18, 2017).

Arrieta-Ortiz, M. L., Hafemeister, C., Bate, A. R., Chu, T., Greenfield, A., Shuster, B., et al. (2015). An experimentally supported model of the *Bacillus subtilis* global transcriptional regulatory network. *Mol. Syst. Biol.* 11:839. doi: 10.15252/msb.20156236

Asai, K., Ootsuji, T., Obata, K., Matsumoto, T., Fujita, Y., and Sadaie, Y. (2007). Regulatory role of RsgI in sigI expression in *Bacillus subtilis*. *Microbiology* 153, 92–101. doi: 10.1099/mic.0.29239-0

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene Ontology: tool for the unification of biology. *Nat. Genet.* 25, 25–29. doi: 10.1038/75556

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Buescher, J. M., Liebermeister, W., Jules, M., Uhr, M., Muntel, J., Botella, E., et al. (2012). Global Network Reorganization During Dynamic Adaptations of *Bacillus subtilis* Metabolism. *Science* 335, 1099–1103. doi: 10.1126/science.1206871

Buxton, R. S. (1980). Selection of *Bacillus subtilis* 168 Mutants with Deletions of the PBSX Prophage. *J. Gene. Virol.* 46, 427–437. doi: 10.1099/0022-1317-46-2-427

Caspi, R., Altman, T., Billington, R., Dreher, K., Foerster, H., Fulcher, C. A., et al. (2014). The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Res.* 42, D459–D471. doi: 10.1093/nar/gkt1103

Cruz Ramos, H., Hoffmann, T., Marino, M., Nedjari, H., Presecan-Siedel, E., Dreesen, O., et al. (2000). Fermentative Metabolism of *Bacillus subtilis*: physiology and Regulation of Gene Expression. *J. Bacteriol.* 182, 3072–3080. doi: 10.1128/JB.182.11.3072-3080.2000

Derré, I., Rapoport, G., and Msadek, T. (1999). CtsR, a novel regulator of stress and heat shock response, controls clp and molecular chaperone gene expression in gram-positive bacteria. *Mol. Microbiol.* 31, 117–131. doi: 10.1046/j.1365-2958.1999.01152.x

Doncheva, N. T., Morris, J. H., Gorodkin, J., and Jensen, L. J. (2019). Cytoscape StringApp: network Analysis and Visualization of Proteomics Data. *J. Proteome Res.* 18, 623–632. doi: 10.1021/acs.jproteome.8b00702

Driessen, A. J., Fekkes, P., and van der Wolk, J. P. (1998). The Sec system. *Curr. Opin. Microbiol.* 1, 216–222. doi: 10.1016/S1369-5274(98)80014-3

Du, Y., Shi, W. W., He, Y. X., Yang, Y. H., Zhou, C. Z., and Chen, Y. (2011). Structures of the substrate-binding protein provide insights into the multiple compatible solute binding specificities of the *Bacillus subtilis* ABC transporter OpuC. *Biochem. J.* 436, 283–289. doi: 10.1042/BJ20102097

Enright, A. J. (2002). An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res.* 30, 1575–1584. doi: 10.1093/nar/30.7.1575

Fabret, C., Feher, V. A., and Hoch, J. A. (1999). Two-component signal transduction in *Bacillus subtilis*: how one organism sees its world. *J. Bacteriol.* 181, 1975–1983. doi: 10.1128/JB.181.7.1975-1983.1999

Fu, L. L., Xu, Z. R., Li, W. F., Shuai, J. B., Lu, P., and Hu, C. X. (2007). Protein secretion pathways in *Bacillus subtilis*: implication for optimization

of heterologous protein secretion. *Biotechnol. Adv.* 25, 1–12. doi: 10.1016/j.biotechadv.2006.08.002

Geissler, A. S., Anthon, C., Alkan, F., González-Tortuero, E., Poulsen, L. D., Kallehauge, T. B., et al. (2021). BSGatlas: a unified *Bacillus subtilis* genome and transcriptome annotation atlas with enhanced information access. *Microb. Genom.* 7:000524. doi: 10.1099/mgen.0.000524

Geissler, A. S., Fehler, A. O., Poulsen, L. D., González-Tortuero, E., Kallehauge, T. B., Alkan, F., et al. (2022). CRISPRi screen for enhancing heterologous α-amylase yield in *Bacillus subtilis*. *bioRxiv* [Preprint]. doi: 10.1101/2022.03.30.486407

González-Pastor, J. E. (2011). Cannibalism: a social behavior in sporulating *Bacillus subtilis*. *FEMS Microbiol. Rev.* 35, 415–424. doi: 10.1111/j.1574-6976.2010.00253.x

Gupta, M., and Rao, K. K. (2014). Phosphorylation of DegU is essential for activation of amyE expression in *Bacillus subtilis*. *J. Biosci.* 39, 747–752. doi: 10.1007/s12038-014-9481-5

Haeussler, M., Zweig, A. S., Tyner, C., Speir, M. L., Rosenbloom, K. R., Raney, B. J., et al. (2019). The UCSC Genome Browser database: 2019 update. *Nucleic Acids Res.* 47, D853–D858. doi: 10.1093/nar/gky1095

Hahne, H., Mäder, U., Otto, A., Bonn, F., Steil, L., Bremer, E., et al. (2010). A Comprehensive Proteomics and Transcriptomics Analysis of *Bacillus subtilis* Salt Stress Adaptation. *J. Bacteriol.* 192, 870–882. doi: 10.1128/JB.01106-09

Haldenwang, W. G. (1995). The sigma factors of *Bacillus subtilis*. *Microbiol. Rev.* 59, 1–30.

Harris, R. S. (2007). *Improved Pairwise Alignment of Genomic DNA*. Ph.D thesis. University Park, PA: Pennsylvania State University.

Helmann, J. D., Marquez, L. M., and Chamberlin, M. J. (1988). Cloning, Sequencing, and Disruption of the *Bacillus subtilis* σ28 Gene. *J. Bacteriol.* 170:7.

Hoffmann, S., Otto, C., Kurtz, S., Sharma, C. M., Khaitovich, P., Vogel, J., et al. (2009). Fast Mapping of Short Sequences with Mismatches, Insertions and Deletions Using Index Structures. *PLoS Comput. Biol.* 5:e1000502. doi: 10.1371/journal.pcbi.1000502

Hohmann, H. P., van Dijl, J. M., Krishnappa, L., and Prágai, Z. (2016). "Host Organisms: *Bacillus subtilis*," in *Industrial Biotechnology*, eds C. Wittmann and J. C. Liao (Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA), 221–297. doi: 10.1002/9783527807796.ch7

Hosoda, J., Kohiyama, M., and Nomura, M. (1959). Studies on amylase formation by *Bacillus subtilis*: vii. Effect of purine, pyrimidine and their analogues on exoenzyme formation by uracil- and adenine-requiring mutants. *J. Biochem.* 46, 857–864. doi: 10.1093/oxfordjournals.jbchem.a126976

Hosseini, S., Curilovs, A., and Cutting, S. M. (2018). Biological Containment of Genetically Modified *Bacillus subtilis*. *Appl. Environ. Microbiol.* 84, e02334–17. doi: 10.1128/AEM.02334-17

Hyatt, D., Chen, G. L., LoCascio, P. F., Land, M. L., Larimer, F. W., and Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. doi: 10.1186/1471-2105-11-119

Hyyryläinen, H.-L., Vitikainen, M., Thwaite, J., Wu, H., Sarvas, M., Harwood, C. R., et al. (2000). d-Alanine Substitution of Teichoic Acids as a Modulator of Protein Folding and Stability at the Cytoplasmic Membrane/Cell Wall Interface of *Bacillus subtilis*. *J. Biol. Chem.* 275, 26696–26703. doi: 10.1016/S0021-9258(19)61432-8

Kanehisa, M., and Goto, S. (2000). KEGG: kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 28, 27–30. doi: 10.1093/nar/28.1.27

Kiel, J. A. K. W., Boels, J. M., Beldman, G., and Venema, G. (1994). Glycogen in *Bacillus subtilis*: molecular characterization of an operon encoding enzymes involved in glycogen biosynthesis and degradation. *Mol. Microbiol.* 11, 203–218. doi: 10.1111/j.1365-2958.1994.tb00301.x

Kilstrup, M., Hammer, K., Ruhdal Jensen, P., and Martinussen, J. (2005). Nucleotide metabolism and its control in lactic acid bacteria. *FEMS Microbiol. Rev.* 29, 555–590. doi: 10.1016/j.fmrre.2005.04.006

Kobayashi, K. (2007). Gradual activation of the response regulator DegU controls serial expression of genes for flagellum formation and biofilm formation in *Bacillus subtilis*. *Mol. Microbiol.* 66, 395–409. doi: 10.1111/j.1365-2958.2007.05923.x

Kontinen, V. P., and Sarvas, M. (1993). The PrsA lipoprotein is essential for protein secretion in *Bacillus subtilis* and sets a limit for high-level secretion. *Mol. Microbiol.* 8, 727–737. doi: 10.1111/j.1365-2958.1993.tb01616.x

Koster, J., and Rahmann, S. (2012). Snakemake–a scalable bioinformatics workflow engine. *Bioinformatics* 28, 2520–2522. doi: 10.1093/bioinformatics/bts480

Larsson, J. T., Rogstam, A., and von Wachenfeldt, C. (2005). Coordinated patterns of cytochrome bd and lactate dehydrogenase expression in *Bacillus subtilis*. *Microbiology* 151, 3323–3335. doi: 10.1099/mic.0.28124-0

Laub, M. T. (2014). "The Role of Two-Component Signal Transduction Systems in Bacterial Stress Responses," in *Bacterial Stress Responses*, eds G. Storz and R. Hengge (Washington, DC, USA: ASM Press), 45–58. doi: 10.1128/9781555816841.ch4

Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., et al. (2013). Software for Computing and Annotating Genomic Ranges. *PLoS Comput. Biol.* 9:e1003118. doi: 10.1371/journal.pcbi.1003118

Lee, S., Lawrence, M., and Cook, D. (2018). *plyranges: A Fluent Interface for Manipulating GenomicRanges*. Available online at: https://github.com/sa-lee/plyranges (accessed July 9, 2017).

Legeay, M., Doncheva, N. T., Morris, J. H., and Jensen, L. J. (2020). Visualize omics data on networks with Omics Visualizer, a Cytoscape App. *F1000 Res.* 9:157. doi: 10.12688/f1000research.22280.2

Lemma, E., Unden, G., and Kröger, A. (1990). Menaquinone is an obligatory component of the chain catalyzing succinate respiration in *Bacillus subtilis*. *Arch. Microbiol.* 155, 62–67. doi: 10.1007/BF00291276

Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. doi: 10.1093/bioinformatics/btt656

Lim, B., and Gross, C. A. (2014). "Cellular Response to Heat Shock and Cold Shock," in *Bacterial Stress Responses*, eds G. Storz and R. Hengge (Washington, DC, USA: ASM Press), 91–114. doi: 10.1128/9781555816841.ch7

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-8

Lu, X., Zhang, H., Tonge, P. J., and Tan, D. S. (2008). Mechanism-based inhibitors of MenE, an acyl-CoA synthetase involved in bacterial menaquinone biosynthesis. *Bioorg. Med. Chem. Lett.* 18, 5963–5966. doi: 10.1016/j.bmcl.2008.07.130

Márquez-Magaña, L. M., Helemann, J. D., Ferrari, E., Helen, Parker HM, Ordal, G. W., and Chamberlin, M. J. (1990). Studies of σ or D-Dependent Functions in *Bacillus subtilis*. *J. Bateriol.* 172, 3435–3443.

Miethke, M., Hecker, M., and Gerth, U. (2006). Involvement of *Bacillus subtilis* ClpE in CtsR Degradation and Protein Quality Control. *J. Bacteriol.* 188, 4610–4619. doi: 10.1128/JB.00287-06

Morris, J. H., Apeltsin, L., Newman, A. M., Baumbach, J., Wittkop, T., Su, G., et al. (2011). clusterMaker: a multi-algorithm clustering plugin for Cytoscape. *BMC Bioinformatics* 12:436. doi: 10.1186/1471-2105-12-436

Mukherjee, S., and Kearns, D. B. (2014). The Structure and Regulation of Flagella in *Bacillus subtilis*. *Annu. Rev. Genet.* 48, 319–340. doi: 10.1146/annurev-genet-120213-092406

Nicolas, P., Mäder, U., Dervyn, E., Rochat, T., Leduc, A., Pigeonneau, N., et al. (2012). Condition-Dependent Transcriptome Reveals High-Level Regulatory Architecture in *Bacillus subtilis*. *Science* 335, 1103–1106. doi: 10.1126/science.1206848

Nijland, R., Heerlien, R., Hamoen, L. W., and Kuipers, O. P. (2007). Changing a Single Amino Acid in Clostridium perfringens β-Toxin Affects the Efficiency of Heterologous Secretion by *Bacillus subtilis*. *AEM* 73, 1586–1593. doi: 10.1128/AEM.02356-06

Ohki, R., Giyanto, N., Tateno, K., Masuyama, W., Moriya, S., Kobayashi, K., et al. (2003). The BceRS two-component regulatory system induces expression of the bacitracin transporter, BceAB, in *Bacillus subtilis*. *Mol. Microbiol.* 49, 1135–1144. doi: 10.1046/j.1365-2958.2003.03653.x

Otto, A., Bernhardt, J., Meyer, H., Schaffer, M., Herbst, F.-A., Siebourg, J., et al. (2010). Systems-wide temporal proteomic profiling in glucose-starved *Bacillus subtilis*. *Nat. Commun.* 1:137. doi: 10.1038/ncomms1137

Pagès, H. (2021). *BSgenome: Software Infrastructure for Efficient Representation of Full Genomes and their SNPsR Package Version 1.64.0*. Available online at: https://bioconductor.org/packages/BSgenome (accessed November 22, 2021).

Pagès, H., Aboyoun, P., Gentleman, R., and DebRoy, S. (2019). *Biostrings: Efficient Manipulation of Biological Strings*. Available online at: https://bioconductor.org/packages/Biostrings (accessed July 9, 2019).

Pedreira, T., Elfmann, C., and Stülke, J. (2022). The current state of Subti Wiki, the database for the model organism *Bacillus subtilis*. *Nucleic Acids Res.* 50, D875–D882. doi: 10.1093/nar/gkab943

Peifer, S., Barduhn, T., Zimmet, S., Volmer, D. A., Heinzle, E., and Schneider, K. (2012). Metabolic engineering of the purine biosynthetic pathway in

*Corynebacterium glutamicum* results in increased intracellular pool sizes of IMP and hypoxanthine. *Microb. Cell Factories* 11:138. doi: 10.1186/1475-2859-11-138

Persuh, M., Mandic-Mulec, I., and Dubnau, D. (2002). A MecA Paralog, YpbH, Binds ClpC, Affecting both Competence and Sporulation. *J. Bacteriol.* 184, 2310–2313. doi: 10.1128/JB.184.8.2310-2313.2002

Petersohn, A., Brigulla, M., Haas, S., Hoheisel, J. D., Völker, U., and Hecker, M. (2001). Global Analysis of the General Stress Response of *Bacillus subtilis*. *J. Bacteriol.* 183, 5617–5631. doi: 10.1128/JB.183.19.5617-5631.2001

Quesada-Ganuza, A., Antelo-Varela, M., Mouritzen, J. C., Bartel, J., Becher, D., Gjermansen, M., et al. (2019). Identification and optimization of PrsA in *Bacillus subtilis* for improved yield of amylase. *Microb. Cell Factories* 18:158. doi: 10.1186/s12934-019-1203-0

R Development Core Team (2008). *R: A Language and Environment for Statistical Computing.R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Rajagopala, S. V., Titz, B., Goll, J., Parrish, J. R., Wohlbold, K., McKevitt, M. T., et al. (2007). The protein network of bacterial motility. *Mol. Syst. Biol.* 3:128. doi: 10.1038/msb4100166

Ramos-Silva, P., Serrano, M., and Henriques, A. O. (2019). From Root to Tips: sporulation Evolution and Specialization in *Bacillus subtilis* and the Intestinal Pathogen Clostridioides difficile. *Mol. Biol. Evol.* 36, 2714–2736. doi: 10.1093/molbev/msz175

Rodionov, D. A., Li, X., Rodionova, I. A., Yang, C., Sorci, L., Dervyn, E., et al. (2008). Transcriptional regulation of NAD metabolism in bacteria: genomic reconstruction of NiaR (YrxA) regulon. *Nucleic Acids Res.* 36, 2032–2046. doi: 10.1093/nar/gkn046

Schallmey, M., Singh, A., and Ward, O. P. (2004). Developments in the use of Bacillus species for industrial production. *Can. J. Microbiol.* 50, 1–17. doi: 10.1139/w03-076

Schurch, N. J., Schofield, P., Gierliński, M., Cole, C., Sherstnev, A., Singh, V., et al. (2016). How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use? *RNA* 22, 839–851. doi: 10.1261/rna.053959.115

Shannon, P. (2003). Cytoscape: a Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303

Smith, D. R., and Chapman, M. R. (2010). Economical Evolution: microbes Reduce the Synthetic Cost of Extracellular Proteins. *mBio* 1, e00131–10. doi: 10.1128/mBio.00131-10

Song, Y., Nikoloff, J. M., and Zhang, D. (2015). Improving Protein Production on the Level of Regulation of both Expression and Secretion Pathways in *Bacillus subtilis*. *J. Microbiol. Biotechnol.* 25, 963–977. doi: 10.4014/jmb.1501.01028

Spinnler, H. E. (2021). "Production of enzymes: Fermentation and genetic engineering," in *Enzymes Novel Biotechnological Approaches for the Food Industry*, eds S. Kermasha and N. A. Michael Eskin (Amsterdam: Elsevier), 45–58. doi: 10.1016/B978-0-12-800217-9.00003-4

Storz, G., and Hengge, R. (eds) (2010). *Bacterial Stress Responses*. Washington, DC, USA: ASM Press, doi: 10.1128/9781555816841

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 47, D607–D613. doi: 10.1093/nar/gky1131

Thorndike, R. L. (1953). Who belongs in the family? *Psychometrika* 18, 267–276. doi: 10.1007/BF02289263

Turgay, K., Hamoen, L. W., Venema, G., and Dubnau, D. (1997). Biochemical characterization of a molecular switch involving the heat shock protein ClpC, which controls the activity of ComK, the competence transcription factor of *Bacillus subtilis*. *Genes Dev.* 11, 119–128. doi: 10.1101/gad.11.1.119

Van den Berge, K., Soneson, C., Robinson, M. D., and Clement, L. (2017). stageR: a general stage-wise method for controlling the gene-level false discovery rate in differential expression and differential transcript usage. *Genome Biol.* 18:151. doi: 10.1186/s13059-017-1277-0

van Dijl, J., and Hecker, M. (2013). *Bacillus subtilis*: from soil bacterium to super-secreting cell factory. *Microb. Cell Factories* 12:3. doi: 10.1186/1475-2859-12-3

Verhamme, D. T., Kiley, T. B., and Stanley-Wall, N. R. (2007). DegU co-ordinates multicellular behaviour exhibited by *Bacillus subtilis*. *Mol. Microbiol.* 65, 554–568. doi: 10.1111/j.1365-2958.2007.05810.x

Vitikainen, M., Pummi, T., Airaksinen, U., Wahlstrom, E., Wu, H., Sarvas, M., et al. (2001). Quantitation of the Capacity of the Secretion Apparatus and Requirement for PrsA in Growth and Secretion of alpha-Amylase in *Bacillus subtilis*. *J. Bacteriol.* 183, 1881–1890. doi: 10.1128/JB.183.6.1881-1890.2001

Weinberg, Z., Wang, J. X., Bogue, J., Yang, J., Corbino, K., Moy, R. H., et al. (2010). Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea, and their metagenomes. *Genome Biol.* 11:R31. doi: 10.1186/gb-2010-11-3-r31

Westers, H., Darmon, E., Zanen, G., and van Dijl, J. M. (2004). The Bacillus secretion stress response is an indicator for a-amylase production levels. *Lett. Appl. Microbiol.* 39, 65–73. doi: 10.1111/j.1472-765X.2004.01539.x

Westers, H., Westers, L., Darmon, E., van Dijl, J. M., Quax, W. J., and Zanen, G. (2006). The CssRS two-component regulatory system controls a general secretion stress response in *Bacillus subtilis*. *FEBS J.* 273, 3816–3827. doi: 10.1111/j.1742-4658.2006.05389.x

Wood, H. E., Dawson, M. T., Devine, K. M., and McConnell, D. J. (1990a). Characterization of PBSX, a defective prophage of *Bacillus subtilis*. *J. Bacteriol.* 172, 2667–2674. doi: 10.1128/jb.172.5.2667-2674.1990

Wood, H. E., Devine, K. M., and McConnell, D. J. (1990b). Characterisation of a repressor gene (xre) and a temperature-sensitive allele from the *Bacillus subtilis* prophage. *PBSX Gene* 96, 83–88. doi: 10.1016/0378-1119(90)90344-Q

Yan, S., and Wu, G. (2017). Bottleneck in secretion of α-amylase in *Bacillus subtilis*. *Microb. Cell Factories* 16:124. doi: 10.1186/s12934-017-0738-1

Yan, S., and Wu, G. (2019). Proteases HtrA and HtrB for α-amylase secreted from *Bacillus subtilis* in secretion stress. *Cell Stress Chaperones* 24, 493–502. doi: 10.1007/s12192-019-00985-1

Yao, Z., Barrick, J., Weinberg, Z., Neph, S., Breaker, R., Tompa, M., et al. (2007). A Computational Pipeline for High- Throughput Discovery of cis-Regulatory Noncoding RNA in Prokaryotes. *PLoS Comput. Biol.* 3:e126. doi: 10.1371/journal.pcbi.0030126

Yao, Z., Weinberg, Z., and Ruzzo, W. L. (2006). CMfinder-a covariance model based RNA motif finding algorithm. *Bioinformatics* 22, 445–452. doi: 10.1093/bioinformatics/btk008

Zaprasis, A., Bleisteiner, M., Kerres, A., Hoffmann, T., and Bremer, E. (2015). Uptake of Amino Acids and Their Metabolic Conversion into the Compatible Solute Proline Confers Osmoprotection to *Bacillus subtilis*. *Appl. Environ. Microbiol.* 81, 250–259. doi: 10.1128/AEM.02797-14

Zhao, L., Ye, J., Fu, J., and Chen, G.-Q. (2018). Engineering peptidoglycan degradation related genes of *Bacillus subtilis* for better fermentation processes. *Bioresour. Technol.* 248, 238–247. doi: 10.1016/j.biortech.2017.05.134

Zhu, B., and Stülke, J. (2018). SubtiWiki in 2018: from genes and proteins to functional network annotation of the model organism *Bacillus subtilis*. *Nucleic Acids Res.* 46, D743–D748. doi: 10.1093/nar/gkx908

Zuber, U., Drzewiecki, K., and Hecker, M. (2001). Putative Sigma Factor SigI (YkoZ) of *Bacillus subtilis* Is Induced by Heat Shock. *J. Bacteriol.* 183, 1472–1475. doi: 10.1128/JB.183.4.1472-1475.2001

# 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

**Contact us**

+41 (0)21 510 17 00
frontiersin.org/about/contact