HOST-MICROBE INTERACTION AND COEVOLUTION

EDITED BY: Wei Huang, Kai Fu, Qiang Gao, Rodrigo Pulgar Tejo, Lucile Maria Floeter-Winter and Alejandro P. Gutierrez

PUBLISHED IN: Frontiers in Microbiology and Frontiers in Genetics

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88976-649-9
DOI 10.3389/978-2-88976-649-9

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

HOST-MICROBE INTERACTION AND COEVOLUTION

Topic Editors:
Wei Huang, Johns Hopkins University, United States
Kai Fu, Central South University, China
Qiang Gao, Hunan University, China
Rodrigo Pulgar Tejo, University of Chile, Chile
Lucile Maria Floeter-Winter, University of São Paulo, Brazil
Alejandro P. Gutierrez, University of Stirling, United Kingdom

Citation: Huang, W., Fu, K., Gao, Q., Tejo, R. P., Floeter-Winter, L. M., Gutierrez, A. P., eds. (2022). Host-Microbe Interaction and Coevolution. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88976-649-9

Table of Contents

frontiers | Frontiers in Genetics

TYPE Editorial
PUBLISHED 30 August 2022
DOI 10.3389/fgene.2022.983158

OPEN ACCESS

EDITED AND REVIEWED BY
Ludmila Chistoserdova,
University of Washington, United States

*CORRESPONDENCE
Wei Huang,
whuang38@jhu.edu

SPECIALTY SECTION
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Genetics

RECEIVED 30 June 2022
ACCEPTED 04 August 2022
PUBLISHED 30 August 2022

CITATION
Huang W, Gao Q, Fu K, Tejo RP, Floeter-Winter LM and Gutierrez AP (2022), Editorial: Host-microbe interaction and coevolution. *Front. Genet.* 13:983158. doi: 10.3389/fgene.2022.983158

Editorial: Host-microbe interaction and coevolution

Wei Huang[1]*, Qiang Gao[2], Kai Fu[3], Rodrigo Pulgar Tejo[4], Lucile Maria Floeter-Winter[5] and Alejandro P Gutierrez[6]

[1]Department of Molecular Microbiology and Immunology, Malaria Research Institute, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States, [2]School of Biomedical Science, Hunan University, Changsha, Hunan, China, [3]Institute of Molecular Precision Medicine and Hunan Key Laboratory of Molecular Precision Medicine, Xiangya Hospital, Central South University, Changsha, Hunan, China, [4]Nutrition and Food Technology Institute, University of Chile, Santiago, Chile, [5]Institute of Bioscience, University of São Paulo, São Paulo, Brazil, [6]University of Stirling, Stirling, United Kingdom

KEYWORDS

pathogen, microbiome, agriculture, host-microbe interaction, alkaloid

Editorial on the Research Topic
Host-microbe interaction and coevolution

Over 100 years ago, Robert Koch established that infectious diseases are caused by microbes. Various studies have shown that different symbiotic and pathogenic microbes could impact host health, including causing diseases. As opposed to studying a single species of microbiota or the host independently, more and more researchers have changed their focus to study how the microbes interact with their hosts, including plants, animals, and fungi.

Studies demonstrating the relationship between microbial communities and their hosts have been emphasized since the last century. Various reports have shown that these microbial communities (microbiota) are dynamic, and their composition depends on host genotype, species, and ontogeny, including diet, habitat, geographical location, and anthropogenic disturbance (Kohl, 2020; Gaete et al., 2021). Importantly, it has been demonstrated that diseases such as diabetes and cancers could affect the human microbiota of the gut and other organs (Iida et al., 2013; Grigorescu and Dumitrascu, 2016). During the last 2 decades, high-throughput sequencing technologies have been widely used to characterize the host's microbiome (microbiota gene pool), enhancing our understanding of the microbial networks and their interactions with the hosts.

For this Research Topic, we aim to promote knowledge from recent advances in the field of microbiota-host interaction and coevolution and use this information to develop approaches to promote crop, cattle, and human health in the long run. Ten articles in this issue gathered information encompassing recent studies conducted in various hosts including vertebrates like humans, birds, and chickens; invertebrates like honeybees and fig wasps; and plants like *Macleaya cordata*.

It has been demonstrated that these interactions between the microbe and host could maintain benefits for one or for both sides. However, in other cases, this interaction can also be pathogenic, resulting in severe host diseases and even leading to deaths, such as

diseases caused by SARS-COV2 and monkeypox virus. Some of these pathogens are also constantly evolving to become resistant to a host's defense mechanisms. Understanding pathogenesis and the host defenses in the fight against pathogens is, therefore, an urgent topic. Here, four manuscripts focused on human pathogens. Jiang et al. reported that the non-pathogenic *Mycobacterium smegmatis mc251* strain selected against H2O2 stress showed a mutation in the bacterial fur gene, which is responsible for strong ROS resistance and INH sensitivity. They also showed that the redox-related protein Rv1996 showed distinct phenotypes under different physiological conditions among different species of *Mycobacterium* species. This may somehow explain why despite having similar virulence factors and signaling transduction systems, *Mycobacterium tuberculosis* (TB) and *Mycobacterium smegmatis mc251* strains showed different virulence against their hosts (Jiang et al.). The phenomenon suggests that *M. smegmatis mc251* could be a suitable candidate vector for a vaccine for TB. Shi et al. investigated the polymorphism of PvDBP (PvDBP-II) in the human pathogen *Plasmodium vivax* from the China-Myanmar Border of Yunnan Province, China, which could be targeted to generate vaccines against malaria (Shi et al.). Another important study conducted by Lv et al. found that the two-component system QseBC from Hypervirulent *Klebsiella pneumonia* contributes to the virulence of the human (Lv et al.). This research provided new insights into the functional importance of QseBC in regulating the virulence of hvKP. Wu et al. collected seven *Cyclophyllidea* specimens from rodents in Qinghai-Tibet Plateau (QTP) and found that some of them might affect human health (Wu et al.).

The microbiota is also crucial to the agricultural industry and environmental protection. Three research articles are focused on this area. Yang et al. investigated the microbiota in the asexual hybridization queens (AHQs) of honeybees (Yang et al.). This research revealed that genetic background rather than environmental factors is dominant in shaping the gut microbiota. Li et al. focused on the microbiota in the fig wasp genus *Ceratosolen* (Li et al.). They found that the endosymbiotic *Wolbachia* carried by fig wasps led to a decrease in bacterial diversity. Huang et al. showed that subchronic copper exposure induced metabolic disorders in chickens by disrupting the microbial community, which may induce metabolic disorders in chickens (Huang et al.). This research will guide the chicken farm to optimize the usage of copper.

Two research articles focused on the microbiota diversity in birds from the Tibetan wetland and an endangered bird *Nipponia Nippon* (Zhu et al.; Bo et al.). The authors collected crucial data on microbes correlated with those birds' development and growth.

The final report revealed the consequences of plant niches and alkaloids on the bacterial community, which could help us to better understand microbe-plant interactions. In this study, high-throughput gene sequencing was used to investigate the composition and abundance of bacteria from the rhizospheric soil and different parts of *Macleaya cordata*, which is well known as traditional antibacterial medicine (Kosina et al., 2010; Sai et al., 2015). *Sphingomonas* were observed to have a significant positive correlation with allocryptopine and protopine, while allocryptopine, protopine, chelerythrine, and sanguinarine showed a negative correlation with the microbiota of stems. These results indicated that bacterial abundance was the highest in the rhizospheric soil and there was a higher diversity of endophytic microorganisms with genes related to alkaloid synthesis indicating a mutual benefit for both the host and the microbe (Lei et al.). This is a good example of host-microbe interaction and evolution. The *Macleaya cordata* secretes alkaloid which shapes the microbiota, while, the microbiota also affects the growth and alkaloid production of *Macleaya cordata.*

Overall, this Research Topic highlights the role of microbiota in human pathogens as well as in agricultural animals and beneficial insects, all of which are of great significance to guide clinical applications, agriculture, and wild animal protection. The host-microbe interaction studies contribute to a better understanding of the role of microbiota in organism development and disease pathology. Understanding the interaction between the microbe and the host would guide us to increase the benefits and reduce harm from microbiota. We also noticed six manuscripts in this issue used high-throughput sequencing technologies to study different organisms. There is no doubt the microbiome analysis will provide fundamental data for further study of microbiota networks and host-microbe interaction. In general, high-throughput microbiome sequencing generates a flood of new data, but how to quickly mine valuable data and use the data to guide us to improve human health and environmental protection is still a challenge for future studies.

Author contributions

WH wrote the first draft. QG, KF, and RT provided critical comments and editorial suggestions for revisions.

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

Gaete, A., Pulgar, R., Hodar, C., Maldonado, J., Pavez, L., Zamorano, D., et al. (2021). Tomato cultivars with variable tolerances to water deficit differentially modulate the composition and interaction patterns of their rhizosphere microbial communities. *Front. Plant Sci.* 12, 688533. doi:10.3389/fpls.2021.688533

Grigorescu, I., and Dumitrascu, D. L. (2016). Implication of gut microbiota in diabetes mellitus and obesity. *Acta Endocrinol.* 12 (2), 206–214. doi:10.4183/aeb.2016.206

Iida, N., Dzutsev, A., Stewart, C. A., Smith, L., Bouladoux, N., Weingarten, R. A., et al. (2013). Commensal bacteria control cancer response to therapy by modulating the tumor microenvironment. *Science* 342 (6161), 967–970. doi:10.1126/science.1240527

Kohl, K. D. (2020). Ecological and evolutionary mechanisms underlying patterns of phylosymbiosis in host-associated microbial communities. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 375 (1798), 20190251. doi:10.1098/rstb.2019.0251

Kosina, P., Gregorova, J., Gruz, J., Vacek, J., Kolar, M., Vogel, M., et al. (2010). Phytochemical and antimicrobial characterization of Macleaya cordata herb. *Fitoterapia* 81 (8), 1006–1012. doi:10.1016/j.fitote.2010.06.020

Sai, C. M., Li, D. H., Xue, C. M., Wang, K. B., Hu, P., Pei, Y. H., et al. (2015). Two pairs of enantiomeric alkaloid dimers from Macleaya cordata. *Org. Lett.* 17 (16), 4102–4105. doi:10.1021/acs.orglett.5b02044

frontiers in Microbiology

ORIGINAL RESEARCH
published: 09 June 2021
doi: 10.3389/fmicb.2021.681210

The *Macleaya cordata* Symbiont: Revealing the Effects of Plant Niches and Alkaloids on the Bacterial Community

Fangying Lei[1,2†], Xueduan Liu[1,2†], Haonan Huang[1,2], Shaodong Fu[1,2], Kai Zou[1,2], Shuangfei Zhang[1,2], Li Zhou[3], Jianguo Zeng[3], Hongwei Liu[1,2], Luhua Jiang[1,2], Bo Miao[1,2] and Yili Liang[1,2]*

[1] School of Minerals Processing and Bioengineering, Central South University, Changsha, China, [2] Key Laboratory of Biometallurgy, Ministry of Education, Changsha, China, [3] Hunan Key Laboratory of Traditional Chinese Veterinary Medicine, Hunan Agricultural University, Changsha, China

OPEN ACCESS

Edited by:
Rodrigo Pulgar Tejo,
University of Chile, Chile

Reviewed by:
Debdulal Banerjee,
Vidyasagar University, India
Jincai Ma,
Jilin University, China

***Correspondence:**
Yili Liang
liangyili6@csu.edu.cn

[†] These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: 16 March 2021
Accepted: 07 May 2021
Published: 09 June 2021

Citation:
Lei F, Liu X, Huang H, Fu S, Zou K, Zhang S, Zhou L, Zeng J, Liu H, Jiang L, Miao B and Liang Y (2021) The Macleaya cordata Symbiont: Revealing the Effects of Plant Niches and Alkaloids on the Bacterial Community. Front. Microbiol. 12:681210. doi: 10.3389/fmicb.2021.681210

Endophytes are highly associated with plant growth and health. Exploring the variation of bacterial communities in different plant niches is essential for understanding microbe-plant interactions. In this study, high-throughput gene sequencing was used to analyze the composition and abundance of bacteria from the rhizospheric soil and different parts of the *Macleaya cordata*. The results indicated that the bacterial community structure varied widely among compartments. Bacterial diversity was observed to be the highest in the rhizospheric soil and the lowest in fruits. Proteobacteria, Actinobacteria, and Bacteroidetes were found as the dominant phyla. The genera *Sphingomonas* (∼47.77%) and *Methylobacterium* (∼45.25%) dominated in fruits and leaves, respectively. High-performance liquid chromatography (HPLC) was employed to measure the alkaloid content of different plant parts. Significant correlations were observed between endophytic bacteria and alkaloids. Especially, *Sphingomonas* showed a significant positive correlation with sanguinarine and chelerythrine. All four alkaloids were negatively correlated with the microbiota of stems. The predicted result of PICRUST2 revealed that the synthesis of plant alkaloids might lead to a higher abundance of endophytic microorganisms with genes related to alkaloid synthesis, further demonstrated the correlation between bacterial communities and alkaloids. This study provided the first insight into the bacterial community composition in different parts of *Macleaya cordata* and the correlation between the endophytic bacteria and alkaloids.

Keywords: ***Macleaya cordata*, 16S rRNA, bacterial community structure, niche differentiation, alkaloids**

INTRODUCTION

Endophytes are widely present inside plants during part or all stages of their life cycle. They are able to survive in the root, stem, leaf, fruit, flower, and seed due to the specific colonization conditions provided by plant tissues and organs (Loaces et al., 2010; Bulgarelli et al., 2013; Wellner et al., 2013; Glassner et al., 2015). Plant microbial community composition is influenced by environmental factors, including climate, temperature, geographic location, vegetation density,

and host genotype (Overbeek and Elsas, 2008; Meyer and Leveau, 2012; Ramond et al., 2013; Campisano et al., 2017; Yana et al., 2017; Ting et al., 2019; Sharaby et al., 2020). However, several studies have shown that plant compartment is the main driver of bacterial community composition, while the season, location, and plant species only play a minor role (Miller and Rudgers, 2014; Junker and Keller, 2015; Fonseca-García et al., 2016; Cregger et al., 2018). Soil microorganisms were partially the source of endophytic bacteria (Zarraonaindia et al., 2015; Mangeot-Peter et al., 2020). They migrated and colonized the area around the root of plants under the influence of the roots' secretions (Tahtamouni et al., 2015). Rhizospheric microbial community composition was significantly influenced by the physical and chemical parameters in soil (Winston et al., 2017; Hartman et al., 2018; Ziyuan et al., 2020), which also affected the plant compartments in part (Mighell et al., 2019). However, Two studies on Arabidopsis showed that the endophytic microbial community composition in the Arabidopsis plants growing on four different soils was similar (Davide et al., 2012; Lundberg et al., 2012), indicating the existence of host mediated control mechanisms. Significant plant compartment effects were also observed in the microbiome of plants such as *Populus*, *Cycas panzhihuaensis*, and *Stellera chamaejasme L* (Jin et al., 2014; Cregger et al., 2018; Zheng and Gong, 2019).

Endophytes and plants can be considered as mutualistic symbiosis. They survived and evolved together (Hardoim et al., 2015; Vandenkoornhuyse et al., 2015). Endophytes are an essential part of the plant micro-ecosystem. They were thought to complement the host plant's gene library and could regulate metabolism, enhance stress resistance, and transform host plants' secondary metabolites (Zhou et al., 2015; Prasad et al., 2018; Cordovez et al., 2019; Víctor and Juan, 2019). Some endophytic bacteria were able to synthesize bioactive compounds such as saponins, terpenoids, and alkaloids, which are potential sources of antibacterial, anti-insect, anticancer, and other properties (Yu et al., 2010; Pimentel et al., 2011; Gutierrez et al., 2012). Considering the various effects of plant-associated bacteria on plants, recording the spatial variability of these bacterial communities is critical for further understanding of plant-microbe interactions and the potential value of endophytic bacteria.

Macleaya cordata, a perennial plant mainly distributes in China, Europe, and North America, has been considered as a traditional folk herbal medicine. Its chemical composition and biological activity were a matter of concern because of its detoxifying, analgesic, anti-inflammatory, antimicrobial, and antitumoral properties (Liu et al., 2013; Khadem et al., 2014; Li et al., 2017; Huang et al., 2018). Its extract has been used as a good alternative of antibiotics in feed additives for animal production, and has achieved European Food Safety Certification for its effectiveness in treating inflammation and regulating the intestinal flora of livestock and poultry. Modern pharmacological studies showed that isoquinoline alkaloids are the primary bioactive substances in *Macleaya cordata*. Some of them possess a prominent apoptotic effect on cancer cells. For instance, sanguinarine could inhibit cell invasion and the MMP-9 and COX-2 expression in TPA-induced breast cancer cells by inducing HO-1 expression (Park et al., 2014). The four main alkaloids in *Macleaya cordata* are sanguinarine, chelerythrine, allocrytopine, and protopine (Kosina et al., 2010). Significant differences in the accumulation of these four alkaloids were reported in the root, stem, leaf, and fruit of *Macleaya cordata*. Sanguinarine and chelerythrine were observed to accumulate mainly in the fruits, with allocrytopine and protopine being most abundant in roots and leaves, respectively. The alkaloids content in the stem was very low. Endophytic bacterial communities in different parts might be affected by this difference.

Bacterial communities across different niches covering below-ground and above-ground tissue-level microbial habitats (rhizospheric soil, root, stem, leaf, and fruit) were characterized in this study by using 16S rRNA gene-targeted Illumina MiSeq sequencing. The content of four main alkaloids was also detected to explore the relationship between alkaloids and endophytic bacteria. We hypothesized that: (i) Niche differentiation among different sample types influenced the composition and diversity of the *Macleaya cordata* associated microbial communities (ii) Alkaloids might contribute to the variation of bacterial communities in different niches. This experiment can support the subsequent search of functional microorganisms and guide the cultivation, protection, and increase of crucial metabolites and resource utilization of *Macleaya cordata*.

MATERIALS AND METHODS

Study Location and Sampling Methods

Macleaya cordata used in this study were collected from an experimental cultivar trial in Hunan province at a site managed by the Hunan Agricultural University in July 2019. Mature and healthy roots, stems, leaves, fruits, and rhizospheric soil were selected from six individual plants on the experimental blocks. All six trees were growing in the same general area and experienced similar developing conditions. Samples were collected randomly and cleaned with distilled water, followed by cutting three pieces from each leaf (from the bottom, middle, and top) with five leaves from different branches. Roots, stems, and fruits were also collected from the same location from each tree (Roots and stems were cut into 2.5 cm long pieces using a sterilized scalpel). All samples were collected aseptically wearing bioclean gloves and put in marked aseptic bags. They were placed in dry ice foam boxes and taken back to the lab as soon as possible. The part for alkaloid content detection was frozen in liquid nitrogen immediately and stored in a −80°C freezer. The part for microbiological analysis would be surface sterilized firstly. Then samples were crushed and homogenized to powder using a sterilized mortar and pestle, frozen in liquid nitrogen, and stored at −80°C until DNA extraction. The surface disinfection process is as follows: The samples were rinsed in tap water, sterilized by 75% ethanol (1 min), washed with sterile water three times, soaked in 8% sodium hypochlorite (5 min), washed with sterilized water five times, and left to dry with sterile filter paper.

Preparation of Standard Alkaloid Solutions and Sample Solutions

All solvents used were of liquid chromatography (LC) grade. The sanguinarine and chelerythrine standards were purchased from Sigma (Shanghai, China), while protopine and allocryptopine standards were purchased from the National Institute for the Control of Pharmaceutical and Biological Products (Beijing, China). Standards were dissolved in methanol and diluted to obtain a series of working solutions (5.12–409.52, 4.20–84.04, 8.61–550.75, 7.80–499.30 μg/ml, respectively) and establish the standard curves at 280 nm. 5.0 g of the powder (60 mesh) of each sample was placed in 250 ml corked conical flasks and extracted with 100 ml of methanol-1% HCL (50:50) in an ultrasonic bath (60 min, 35°C, 200 W, 40 KHz), the weight loss was compensated with methanol-1% HCL, and the supernatant was collected. The supernatant was filtered through a 0.22 μm membrane before HPLC analysis. Then a 20 μl aliquot of the solution was injected into the HPLC system. For accurate analysis of high content alkaloids, the sample solution was diluted to avoid a concentration that lies outside the linear range.

HPLC Analysis

Analysis was carried out using a Waters alliance 1,260 liquid chromatographic system (Milford, MA, United States). LC separations were accomplished on a 5 μm, 250 mm × 4.6 mm Ultimate XB C18 column (Welch Materials, Ellicott City, United States) at 35°C. Mobile phases consisted of (A) water containing Acetonitrile and (B) 1% phosphoric acid solution. Gradient elution program was as follows: 0–11 min 25% A, 11–27 min 25–60% A, 27–29 min 60–25% A, 29–35 min 25% A. The flow rate was set at 1 ml/min. The injection volume was 20 μl. A wavelength of 280 nm was selected for quantification. The limit of detection (LOD) was 1.35, 0.97, 0.89, 0.93 μg/ml for protopine, allocryptopine, sanguinarine, and chelerythrine.

DNA Extraction and PCR Amplification of the Bacterial 16S rRNA

Total plant DNA and soil microbial DNA were extracted using DNeasy® Plant Mini Kit (250) (Qiagen, Inc., Duesseldorf, Germany) and ALFA-SEQ Advanced Soil DNA Kit (Guangdong Magigene Biotechnology Co., Ltd.), respectively. All steps were performed strictly according to the manufacturer's protocols. The concentration and purity of DNA were determined using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, United States). 1% (w/v) TAE-agarose gel stained with EB was also used to detect the extracted DNA's quality. The V5 and V6 hypervariable regions of the bacteria 16S rRNA genes were amplified with the primers 799F (5′-AAC MGG ATT AGA TAC CCK G-3′); 1115R (5′-AGG GTT GCG CTC GTT G-3′). The extracted DNA was amplified using a thermocycler polymerase chain reaction (PCR) system with the following process: denaturation for 1 min at 94°C; 35 cycles of 20 s at 94°C, annealing for 25 s at 57°C, elongation for 45 s at 68°C; and a final extension step for 10 min at 68°C. The PCR was performed with reaction volumes of 50 μl containing 25 μl 2 × Taq Master Mix plus (Vazyme), 1.5 μl each primer (10 μM), 2 μl template DNA (100 ng), 20 μl PCR water. The PCR products were assessed by 1.5% (w/v) TAE-agarose gel stained with EB and further purified using the Gel Extraction Kit D2500 (OMEGA bio-tek, Norcross, Georgia, United States). High-throughput paired-end sequencing of the purified amplicons was conducted on the Illumina Miseq platform (Miseq PE250).

Sequencing Data Processing

Gene sequencing data of raw 16S rRNA were imported to the QIIME2 platform to generate the OTU table. The cutadapt program was used to remove the forward and reverse primers. The paired-end fastq files were merged, denoised, chimeras removed using the dada2 program. Meanwhile, they were clustered to operational taxonomic units (OTUs) at 97% similarity. Clean data were classified using BLAST with the Greengenes reference databases. Contaminant sequences unclassified at the domain (bacteria/archaea), mitochondria, and chloroplasts were filtered during the taxa filter-table process. All sequences produced from Illumina sequencing had been uploaded to the sequence read archive (SRA) of the NCBI database. The accession number of all samples is PRJNA714300.

Statistical Analysis

Alpha-diversity indexes, including Shannon index and Pielou evenness, and Non-metric Multidimensional Scaling (NMDS) for assessing the beta-diversity differences in community composition were calculated on Institute for Galaxy | Denglab pipeline[1]. Dissimilarity tests of OTUs and predicted function genes (by PICRUSt2) were computed using MRPP, ANOSIM, and PERMANOVA with Bray-Curtis distance matrices to identify whether the habitat had a significant effect on community composition and the abundance of functional genes. Statistical analyses were conducted using Minitab 19.0 software (Minitab Inc., Pennsylvania State University, Pennsylvania, United States). One-way ANOVA analyses, followed by Tukey's test, were carried out to identify the gene abundance difference. Correlation between bacterial community and four alkaloids was carried out based on the Spearman correlation coefficients and Canonical Correspondence Analysis (CCA). PICRUSt2 was employed to predict each sample's function genes based on the 16S rRNA sequencing data, and the KEGG database was used to match the selected reference OTU.

RESULTS

Analysis of Sequencing Data

A total of 4,298,014 raw reads were obtained among 30 samples. After quality control and contaminant sequences removal, 1,218,950 high-quality paired 16S rRNA sequences with an average read length of 300 bp remained, which can be clustered into 6,866 OTUs (**Supplementary Table 1**). CPM standardization was performed on all samples before analysis of microbial communities. Rarefaction curves were constructed for each sample showing the number of observed OTUs. As

[1]http://mem.rcees.ac.cn:8080/

expected, endophytic microorganisms had lower observed OTUs than rhizospheric soil bacterial communities (**Supplementary Figure 1**). Rarefaction curves evaluating the OTU richness per sample generally approached saturation, indicating that the sequencing depth was adequate.

Diversity of the Microbial Communities Associated With *Macleaya cordata*

Bacterial diversity was observed to be highest in soil, while the situation in plants was relatively complex, with bacterial diversity showing no significant differences in roots, stems and leaves, but a significant decrease in fruits. The highest evenness of bacterial communities was found in the stems among plant parts (**Figure 1**). NMDS revealed strong clustering of bacterial communities based on plant compartments (**Figure 2A**). Specifically, Coordinate1 (x-axis) separated the rhizospheric soil, roots, and above-ground (stems, leaves, fruit) samples. Rhizospheric soil exhibited a significant difference in community composition comparing with the other four plant compartments. However, the bacterial communities in the fruit and leaf were highly similar and separated from other endosphere samples in Coordinate2 (y-axis). Dissimilarity tests based on Bray-Curtis distance, including MRPP, ANOSIM, and PERMANOVA, also showed that the bacterial community composition of these five groups was significantly different ($P < 0.05$) (**Table 1**).

To better understand the OTUs distribution within different niches, the number of OTUs shared by all compartments and OTUs uniquely identified in each sample type were calculated (**Supplementary Figure 2**). The bacterial communities in the rhizospheric soil had the highest number of OTUs, followed by roots. 302, 155, 236, 245, and 5,391 bacterial OTUs were uniquely identified for root, stem, leaf, fruit, and rhizospheric soil, respectively, with 19 OTUs shared among the five sample groups. A high overlap (208) of OTUs from rhizospheric soil and root was observed.

The Bacterial Composition Differences Across Habitat at Phylum Level and Genera Level

At the phylum level, the bacterial community composition showed high variabilities among different compartments (**Figure 3A**). Twenty-five dominant bacterial phyla (≥0.1% relative abundance) were detected across this study. Proteobacteria phyla was the dominant group in all niches (soil, root, stem, leaf, fruit), representing 28.61–54.84%, 37.76–88.36%, 60.43–79.27%, 69.53–88.93%, and 90.30–97.10% of all bacteria, respectively. Alphaproteobacteria accounted for the highest proportion in leaves. The phylum Thermi showed a significantly high abundance in the above-ground part, almost absent in the underground part. Firmicutes and Tenericutes were detected in stems but rare in other parts. The highest abundance of Actinobacteria was found in the roots, followed by the rhizospheric soil, with lower abundance in the above-ground parts. Bacteroidetes had greater abundance in soils and roots than in the above-ground parts whereas TM7 had a greater abundance in root and stem habitats than other habitats. Bacterial community composition was different between rhizospheric soil and plant compartments. Acidobacteria, Bacteroidetes, Gemmatimonadetes, Chloroflexi, Verrucomicrobia, AD3, and Nitrospirae were relatively abundant in the soil compared to other compartments. However, some of them were also detected in endosphere samples in very low abundance.

At the genera level, 334 bacteria were identified from all sequences. Dominant bacteria differed among five compartments. Only the top 20 genera were shown in **Figure 3B**. *Sphingomonas* (47.77%) and *Pseudomonas* (25.09%) were abundant in the fruit, whereas *Methylibium* (45.25%) and *Deinococcus* (10.20%) were mainly distributed in the leaves. *Pseudomonas* (10.09%), *Nocardia* (11.42%), and *Rhodanobacter* (8.82%) were the dominant genera in the roots, and their abundance did not differ much. The microorganisms in stems were relatively uniform. The dominant genera in fruit, leaves, and roots were also abundant in stems. Besides, some microorganisms with low abundance in other endophytic tissues like *Herbaspirillum* (9.79%), *Lysobacter* (6.30%), and *Ralstonia* (6.03%) were also present in relatively high abundance in stems. Most bacterial genera detected in soil could not be identified.

Alkaloids Content of *Macleaya cordata*

The contents of four kinds of alkaloids in each tissue (fruit, leaf, stem, and root) of the six mature plants were calculated according to the standard curve regression equation (**Supplementary Table 2**). Alkaloid content showed a significant difference among organs (**Figure 4A**). The highest content of sanguinarine and chelerythrine was observed in fruit (4.04 and 1.65 mg/g). They were low in roots and leaves and almost zero in the stem. The accumulation of alkaloids in leaves and roots was mainly composed of protopine and allocrytopine. Roots contained the highest allocrytopine (11.67 mg/g) than other organs, while the leaf's protopine content (12.81 mg/g) was the highest. The accumulation of four kinds of alkaloids in the stem was much lower than that in the other three organs.

Correlation Between the Endophytic Bacterial Abundances and Four Alkaloids

The Spearman correlation analysis was performed to explore the relationship between the endophytic bacteria at the phylum and genus level (relative abundance > 0.01%) and alkaloids to quantify the environmental influence (**Tables 2**, **3**). Tenericutes showed a significant ($P < 0.05$) and negative correlation with all four alkaloids. Actinobacteria was significantly ($P < 0.05$) negatively correlated with sanguinarine and chelerythrine. However, most of the genera with higher abundance in this phylum showed a significant ($P < 0.05$) and positive correlation with allocryptopine. The Firmicutes, β- proteobacteria, and γ-proteobacteria were significantly ($P < 0.05$) negatively correlated with protopine and allocryptopine. Only the α-proteobacteria was significantly ($P < 0.05$) and positively correlated with protopine, allocryptopine, and sanguinarine. At the genera level, most endophytic bacteria were significantly ($P < 0.05$) correlated with alkaloids. *Sphingomonas* showed a significant ($P < 0.05$) positive correlation with sanguinarine

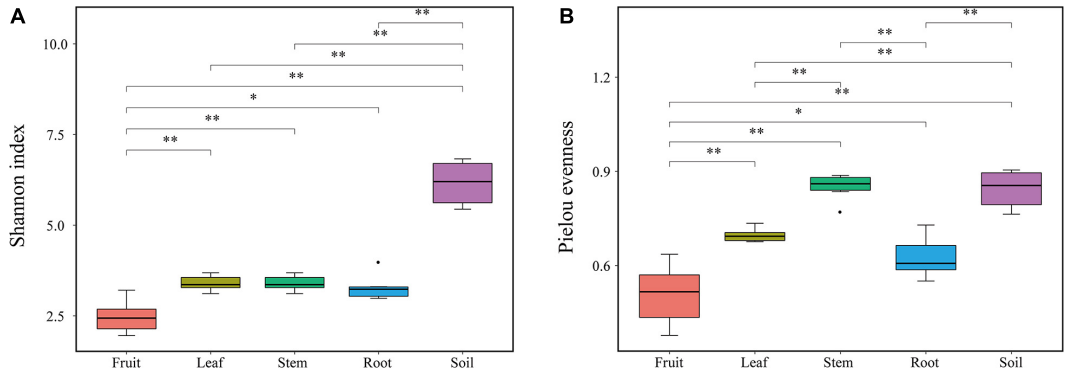

FIGURE 1 | Diversity indexes of endophytic bacterial communities. **(A)** Shannon index **(B)** Pielou's evenness.

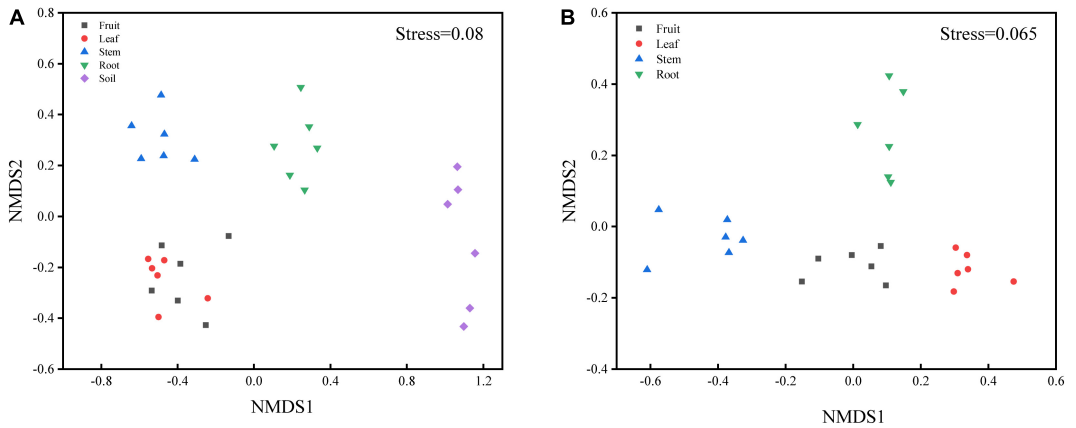

FIGURE 2 | Non-metric multidimensional scaling (NMDS) carried out with **(A)** OTU table based on Bray-Curtis distance (S = 0.08) and **(B)** predicted function gene table based on Bray-Curtis distance (S = 0.065).

TABLE 1 | Dissimilarity tests of OTUs and predicted function genes (by PICRUSt) based on Bray-Curtis distance.

	MRPP		ANOSIM		PERMANOVA	
	r	p	r	p	r	p
OTUs	0.6037	**0.001**	0.9042	**0.001**	8.3007	**0.001**
function genes	0.3827	**0.001**	0.5728	**0.001**	10.5800	**0.001**

Significant differences ($P < 0.05$) are indicated in bold.

and chelerythrine, while *Herbaspirillum*, *Methylibium*, and *Pelomonas* showed a significant ($P < 0.01$) and negative correlation with them. Canonical Correspondence Analysis (CCA) of the top 20 bacterial genera (with significant correlation coefficients) and four critical alkaloids in *Macleaya cordata* was also conducted (**Figure 4B**). The results were consistent with the above analysis. The sanguinarine and chelerythrine were more associated with the fruit bacterial communities, whereas protopine and allocryptopine had greater effects on leaf and root bacterial communities, respectively. However, all four alkaloids were negatively correlated with the microbiota of the stem.

Functional Predictions by PICRUSt2

PICRUSt2 was applied as an exploratory predictive tool for functional annotation analysis. 7,185 KEGG Orthology groups (KOs) were predicted in *Macleaya cordata* endophytic communities. NMDS showed that these KOs were strongly clustered and separated from each other in different niches (**Figure 2B**). Dissimilarity tests of the relative abundance of KOs based on Bray-Curtis distance were also employed, including MRPP, ANOSIM, and PERMANOVA (**Table 1**). These results revealed that the relative abundance of KOs was significantly different among the root, stem, leaf, and fruit samples ($P < 0.05$).

Forty-two gene families were observed at Hierarchical Level 1 (KOs in KEGG). The majority of the KOs belonged to Genetic Information Processing, Environmental Information Processing, and Metabolism. As shown in **Figure 5A**, the frequencies of categories associated with membrane transport gradually increased from 4.57% in fruit to 5.91% in roots. In comparison, the frequencies of categories associated with signal transduction decreased from 4.79% in fruit to 3.23% in roots. The abundance of KOs related to the biosynthesis of other secondary metabolites was the lowest in the stem bacterial community and the highest in the root. To further identify microbiota function

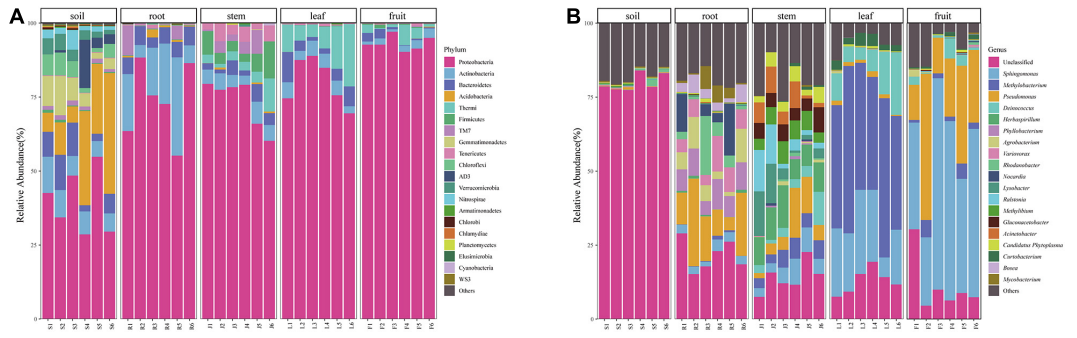

FIGURE 3 | Bacterial community composition of different niches **(A)** bacterial community at the Phylum level and **(B)** bacterial community at the genera level.

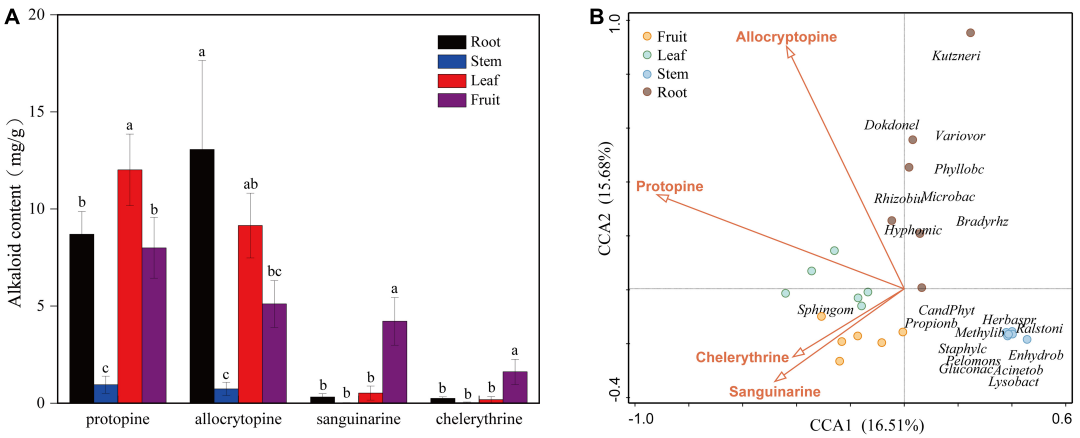

FIGURE 4 | (A) Contents of four kinds of alkaloids in different organs (fruit, leaf, stem, and root) of *Macleaya cordata*. **(B)** Canonical Correspondence Analysis (CCA) of the top 20 bacterial genera (with significant correlation coefficients) and four critical alkaloids in *Macleaya cordata*.

differences in the biosynthesis of secondary metabolites, 17 related KEGG functional categories at the second hierarchy level were analyzed (**Figure 5B**). In comparison between groups with different niches, eight pathways (e.g., Streptomycin biosynthesis, Caffeine metabolism, and Cytochrome P450) were significantly more abundant in the root, while the Staurosporine biosynthesis pathway was significantly more abundant in fruit.

DISCUSSION

The Composition and Diversity of Bacterial Communities Differ in Niches

Bacterial community composition was significantly ($P < 0.01$) different among plant compartments (**Table 1**). The NMDS diagram (**Figure 2A**) also showed that the separation in the bacterial community structure between different plant compartments at the OTU level, indicating differences in the microbes in different niches. Although soil microorganisms were a source of plant endophytes, the migration of soil microorganisms to the root was still mediated by the host (Davide et al., 2012; Lundberg et al., 2012). For above-ground plant tissues, endophytes also derived from horizontal dispersal in the atmosphere and vertical transmission through seeds (Maignien et al., 2014; Fonseca-García et al., 2016), which resulted in bacterial differences between above-ground organs and soil. Similar results have also been reported in maize and wheat (Xiong et al., 2020), poplar (Beckers et al., 2017), *Arabidopsis* thaliana (Davide et al., 2012), and other plant species (Edwards et al., 2015; Lu et al., 2020).

As demonstrated by rarefaction curves and the alpha diversity indices, there were significant differences in species diversity between rhizospheric soil and plant compartments of *Macleaya cordata*. As an ideal habitat for various microorganisms, the bacterial diversity of rhizospheric soil was significantly higher than roots and above-ground niches. This finding is consistent with the general view of microbial colonization (Beckers et al., 2017). No significant difference was observed in the Shannon diversity of bacterial communities in roots, stems, and leaves, while the bacterial diversity in fruits was the lowest. The bacterial diversity within the plant compartments was low, probably because plant tissues were highly variable and complex. Bacterial colonization is limited by the host's immune system, nutritional conditions, and metabolites (Compant et al., 2010). The highest evenness of bacterial communities was found in the stem among plant parts. It should be noticed that there was no significant

TABLE 2 | Spearman Correlation coefficients showing the relationships between endophytic bacterial phyla and four alkaloids.

Phyla	Protopine		Allocryptopine		Sanguinarine		Chelerythrine	
	r	*P*	*R*	*p*	*R*	*p*	*r*	*p*
Acidobacteria	−0.243	0.253	0.171	0.423	−0.241	0.257	−0.151	0.482
Actinobacteria	−0.036	0.867	0.337	0.108	−0.531	**0.008**	−0.424	**0.039**
AD3	0.144	0.503	0.411	**0.046**	0.067	0.756	0.080	0.710
Armatimonadetes	0.163	0.446	0.192	0.368	0.089	0.681	0.145	0.500
Bacteroidetes	0.137	0.525	0.378	0.069	−0.316	0.133	−0.332	0.113
Chlamydiae	−0.010	0.962	0.293	0.165	−0.281	0.184	−0.266	0.209
Chlorobi	0.122	0.571	0.607	**0.002**	−0.151	0.480	0.077	0.722
Chloroflexi	0.348	0.096	0.621	**0.001**	0.198	0.354	0.296	0.160
Cyanobacteria	−0.256	0.227	−0.256	0.227	−0.288	0.173	−0.302	0.152
Euryarchaeota	−0.095	0.658	0.051	0.814	−0.187	0.381	−0.114	0.595
FBP	0.227	0.286	0.437	**0.033**	−0.104	0.628	−0.103	0.632
Firmicutes	−0.763	**0.000**	−0.713	**0.000**	−0.370	0.075	−0.267	0.207
Fusobacteria	0.219	0.303	0.263	0.214	0.126	0.557	0.144	0.502
Gemmatimonadetes	−0.035	0.872	0.180	0.401	0.031	0.886	0.050	0.815
Nitrospirae	0.105	0.624	0.256	0.227	0.030	0.888	0.106	0.624
OD1	0.151	0.482	0.316	0.132	−0.151	0.480	−0.121	0.575
Planctomycetes	−0.111	0.605	0.049	0.820	−0.321	0.126	−0.265	0.211
Proteobacteria	0.051	0.813	−0.225	0.289	0.608	**0.002**	0.577	**0.003**
SR1	0.286	0.175	0.226	0.288	0.166	0.437	0.166	0.439
Tenericutes	−0.481	**0.017**	−0.544	**0.007**	−0.685	**0.000**	−0.659	**0.000**
Thermi	0.283	0.180	−0.173	0.419	0.000	0.998	−0.231	0.277
TM6	−0.083	0.699	0.302	0.151	0.178	0.404	0.321	0.126
TM7	−0.129	0.548	−0.002	0.995	−0.495	**0.014**	−0.418	**0.042**
Verrucomicrobia	−0.035	0.871	0.296	0.160	−0.251	0.237	−0.090	0.675
WPS-2	0.151	0.482	0.316	0.132	−0.151	0.480	−0.121	0.575

Significant differences ($P < 0.05$) are indicated in bold.

difference in bacterial diversity between stems and roots, while many studies have determined that bacterial diversity was highest in roots (Santana et al., 2016; Wang et al., 2019). The variation of endophytic bacterial communities in different compartments was primarily driven by tissue-specific filtering mechanisms within the host (Bonito et al., 2014; Chao et al., 2020). Under the host-mediated control, only a limited number of microorganisms could maintain a symbiotic lifestyle with the host. This pressure sequentially increased from the soil to the plant compartments (Chao et al., 2020; Pasquale et al., 2020), which might be responsible for the lowest diversity of bacterial communities in fruits of *Macleaya cordata*.

Niche Preference Exists for Bacteria of *Macleaya cordata*

We identified several prokaryotic taxa (>0.1%), including microbiota members belonging to Proteobacteria, Actinobacteria, Acidobacteria, Bacteroidetes, TM7, Firmicutes. Bacterial diversity varied among plant-associated habitats, and the dominant phylum found in each habitat was highly comparable to other plant hosts in each habitat. Many studies have reported that plants' bacterial microbiota is generally dominated by three major phyla (Proteobacteria, Actinobacteria, and Bacteroidetes) in both above- and below-ground tissues (Tahtamouni et al., 2015; Wang et al., 2020). This is also consistent with our study. Proteobacteria was the most abundant in the above-ground compartment, while Actinobacteria was widely distributed in the below-ground. Many microorganisms were detected in rhizospheric soil but hardly found in other niches, such as Acidobacteria, Bacteroidetes, and Gemmatimonadetes, which were often discovered in the soil in other studies (Cregger et al., 2018; Wang et al., 2020). Bacteria in the stem were affected by both the above-ground parts (leaf and fruit) and the below-ground part (root) and dominated by Proteobacteria with the enrichment of Actinobacteria, Bacilli, Mollicutes, and TM7. However, some of the microorganisms did not spread from the stem to the leaf and fruit. This might be caused by nonuniform colonization of different compartments, the microbial source difference, or other environmental factors.

As shown in **Figure 3B**, only a few bacterial genera were dominant in fruit and leaves. *Sphingomonas* and *Pseudomonas* were found to be the dominant bacterial genera in the fruit, and *Methylobacterium*, *Sphingomonas*, and *Deinococcus* were detected as predominant groups in the endophytic communities of the *Macleaya cordata* leaf. Different studies have reported that these genera represented a substantial part of various plant species' endophytic microbiota (Mano and Morisaki, 2008; Delmotte et al., 2009). In the previous study, *Sphingomonas* played an essential role in plant stress tolerance, plant

TABLE 3 | Spearman Correlation coefficients showing the relationships between endophytic bacterial genera (relative abundance > 0.01%) and four alkaloids.

Phylum	Genus	Protopine		Allocryptopine		Sanguinarine		Chelerythrine	
		r	*P*	*r*	*p*	*r*	*p*	*r*	*p*
Thermi		0.283	0.180	−0.173	0.419	**0.000**	0.998	−0.231	0.277
	Deinococcus	0.283	0.181	−0.170	0.426	0.012	0.955	−0.219	0.304
Actinobacteria		−0.036	0.867	0.337	0.108	−0.531	**0.008**	−0.424	**0.039**
	Amycolatopsis	0.100	0.642	0.319	0.129	−0.123	0.567	0.010	0.964
	Curtobacterium	0.444	**0.030**	0.004	0.985	0.228	0.284	0.047	0.828
	Kineococcus	0.658	**0.000**	0.262	0.216	0.350	0.094	0.157	0.463
	Kribbella	0.102	0.634	0.493	**0.014**	−0.003	0.987	0.110	0.608
	Kutzneria	0.034	0.876	0.467	**0.022**	−0.032	0.882	0.164	0.445
	Microbacterium	0.387	0.062	0.726	**0.000**	0.125	0.560	0.207	0.332
	Mycobacterium	−0.070	0.745	0.405	**0.050**	−0.128	0.550	0.025	0.908
	Nocardia	0.140	0.515	0.451	**0.027**	0.013	0.951	0.107	0.620
	Propionibacterium	−0.487	**0. 016**	−0.628	**0.001**	−0.067	0.755	−0.058	0.788
	Pseudonocardia	0.326	0.119	0.491	**0.015**	0.194	0.365	0.204	0.338
	Streptomyces	0.344	0.100	0.496	**0.014**	0.117	0.586	0.238	0.262
Bacteroidetes		−0.341	0.103	−0.457	**0.025**	−0.302	0.152	−0.319	0.129
	Dyadobacter	−0.129	0.548	0.339	0.105	−0.008	0.972	0.193	0.365
	Hymenobacter	0.527	**0.008**	0.088	0.684	0.269	0.204	0.081	0.707
	Spirosoma	0.643	**0.001**	0.236	0.268	0.442	**0.031**	0.235	0.269
	Chryseobacterium	−0.054	0.803	−0.176	0.411	0.176	0.409	0.125	0.562
	Pedobacter	−0.312	0.138	−0.005	0.982	0.109	0.613	0.247	0.245
Firmicutes		−0.763	**0.000**	−0.713	**0.000**	−0.37	0.075	−0.267	0.207
	Bacillus	−0.257	0.226	0.081	0.705	−0.469	**0.021**	−0.270	0.202
	Staphylococcus	−0.728	**0.000**	−0.701	**0.000**	−0.328	0.118	−0.270	0.203
α-Proteobacteria		0.804	**0.000**	0.465	**0.023**	0.501	**0.013**	0.368	0.077
	Bradyrhizobium	−0.388	0.061	0.151	0.480	−0.455	**0.026**	−0.220	0.303
	Methylobacterium	0.263	0.215	−0.203	0.339	−0.121	0.572	−0.316	0.132
	Afipia	0.087	0.687	0.614	**0.001**	−0.115	0.592	0.096	0.654
	Agrobacterium	−0.035	0.872	0.261	0.218	−0.103	0.633	0.030	0.891
	Bosea	0.039	0.856	0.317	0.131	0.203	0.340	0.320	0.127
	Caulobacter	−0.253	0.232	0.038	0.860	−0.096	0.656	0.015	0.946
	Hyphomicrobium	−0.230	0.279	0.220	0.302	−0.318	0.129	−0.116	0.589
	Novosphingobium	−0.128	0.551	0.159	0.459	0.088	0.682	0.231	0.276
	Phyllobacterium	−0.145	0.499	0.329	0.117	−0.093	0.666	0.098	0.648
	Rhizobium	−0.132	0.539	0.364	0.080	−0.173	0.419	0.056	0.797
	Rhodoplanes	0.077	0.721	0.522	**0.009**	−0.073	0.733	0.052	0.809
	Sphingomonas	0.257	0.226	−0.196	0.358	0.554	**0.005**	0.428	**0.037**
β-Proteobacteria		−0.718	**0.000**	−0.570	**0.004**	−0.386	0.063	−0.249	0.242
	Delftia	−0.089	0.681	0.275	0.193	0.245	0.249	0.382	0.066
	Herbaspirillum	−0.660	**0.000**	−0.448	**0.028**	−0.751	**0.000**	−0.675	**0.000**
	Janthinobacterium	0.564	**0.004**	0.169	0.429	0.495	**0.014**	0.379	0.068
	Methylibium	−0.726	**0.000**	−0.525	**0.008**	−0.760	**0.000**	−0.649	**0.001**
	Pelomonas	−0.781	**0.000**	−0.691	**0.000**	−0.553	**0.005**	−0.509	**0.011**
	Polaromonas	−0.070	0.744	0.244	0.250	−0.122	0.570	0.014	0.947
	Ralstonia	−0.815	**0.000**	−0.775	**0.000**	−0.318	0.130	−0.264	0.213
	Variovorax	0.105	0.626	0.552	**0.005**	−0.031	0.886	0.098	0.650
γ-Proteobacteria		−0.755	**0.000**	−0.564	**0.005**	−0.068	0.753	0.064	0.766
	Acinetobacter	−0.867	**0.000**	−0.707	**0.000**	−0.395	0.056	−0.305	0.147
	Dokdonella	0.102	0.635	0.641	**0.001**	−0.117	0.585	0.097	0.652
	Enhydrobacter	−0.856	**0.000**	−0.824	**0.000**	−0.324	0.123	−0.233	0.273
	Erwinia	−0.227	0.285	−0.334	0.111	0.231	0.277	0.235	0.269

(Continued)

TABLE 3 | Continued

Phylum	Genus	Protopine		Allocryptopine		Sanguinarine		Chelerythrine	
		r	*P*	*r*	*p*	*r*	*p*	*r*	*p*
	Gluconacetobacter	−0.769	**0.000**	−0.715	**0.000**	−0.397	0.055	−0.397	0.055
	Klebsiella	−0.126	0.558	−0.257	0.225	0.023	0.914	0.012	0.955
	Lysobacter	−0.683	**0.000**	−0.670	**0.000**	−0.332	0.113	−0.352	0.092
	Pantoea	−0.053	0.805	−0.303	0.151	0.373	0.073	0.366	0.079
	Pseudomonas	−0.369	0.076	−0.363	0.082	0.356	0.087	0.447	**0.029**
	Rheinheimera	−0.434	**0.034**	−0.389	0.060	−0.415	**0.044**	−0.317	0.131
	Rhodanobacter	−0.251	0.238	0.297	0.158	−0.127	0.553	0.111	0.606
	Serratia	−0.392	0.058	−0.429	**0.036**	−0.310	0.141	−0.296	0.160
	Stenotrophomonas	−0.067	0.757	−0.172	0.421	0.313	0.137	0.341	0.103
	Steroidobacter	0.075	0.729	0.540	**0.007**	−0.021	0.923	0.189	0.377
δ-Proteobacteria		0.253	0.233	0.086	0.689	0.102	0.636	0.108	0.617
	Cystobacter	0.426	**0.038**	0.068	0.752	0.406	**0.049**	0.244	0.251
Tenericutes		−0.481	**0.017**	−0.544	**0.007**	−0.685	**0.000**	−0.659	**0.000**
	Candidatus Phytoplasma	−0.478	**0.018**	−0.541	**0.007**	−0.685	**0.000**	−0.653	**0.001**

Significant differences (P < 0.05) are indicated in bold.

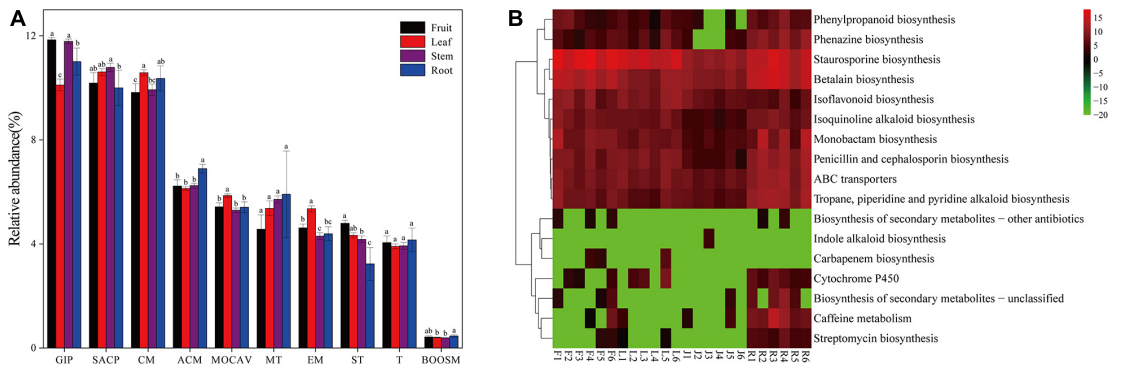

FIGURE 5 | (A) Gene functional predictions of endophytic microbiota between different plant compartments. (GIP: genetic information processing; SACP: signaling and cellular processes; CM: Carbohydrate metabolism; ACM: Amino acid metabolism; MOCAV: Metabolism of cofactors and vitamins; MT: Membrane transport; EM: Energy metabolism; ST: Signal transduction; T: Translation; BOOSM: Biosynthesis of other secondary metabolites). **(B)** Comparisons of the seventeen gene pathways involved in the biosynthesis of secondary metabolites in the bacterial microbiota. Values of each functional gene (row) were log2 transformed. The second level of the KEGG pathway related to the biosynthesis of secondary metabolites was shown in the heatmap (S, Soil; R, Root; J, Stem; L: leaf; F: fruit).

growth promotion, and biodegradation of polycyclic aromatic hydrocarbon (Wilkes et al., 1996; Halo et al., 2015; Asaf et al., 2020). Microbial colonization is related to its ability to adapt to the host's internal environment and its utilization of substrates. Leaves are more often exposed to the vagaries of the environment, including nutrient stress, desiccation, and ultraviolet radiation, providing a special habitat for microorganisms (Hunter et al., 2010). *Methylobacterium*, often isolated from the leaf surface and interior, could specifically colonize the plant by profiting from methanol released by the plant (Galbally and Kirstine, 2002) and has been reported to be drought and radiation-resistant (Yoshida et al., 2017; Jorge et al., 2019; Kim et al., 2019). Therefore, *Methylobacterium* were able to successfully colonize the leave extensively (Delmotte et al., 2009). *Cystobacter* was found only in the fruits and leaves, and *Erwinia* was found only in above-ground tissues. Colonization of these two genera might be due to horizontal transmission (Frank et al., 2017).

Bacterial communities of different ecological niches were assembled under the effect of environmental filtering, ecological drift, and dispersal limitations (Nemergut et al., 2013; Moroenyane et al., 2021). However, there is still a lack of understanding of these processes. Plants are exposed to diverse and highly variable environmental factors, physiological structure (thickness and shape), chemical properties (nutrients contents, water, and secondary metabolites) of each compartment drive the differences in bacterial communities to some extent (Delmotte et al., 2009; Hunter et al., 2010; Arturo et al., 2012). In a study on soybean, the secondary metabolites (ethylamine and betaine) were considered as a robust environmental filter for bacterial communities (Shintaro et al., 2019). Furthermore, the bacterial communities in *Stevia rebaudiana* and *Coptis teeta*, were also demonstrated to be significantly correlated with secondary metabolites (Yu et al., 2015; Liu et al., 2020). On the other hand, the functional capacity of bacterial species is key to

their recruitment by hosts, and Burke et al. (2011) proposed that bacterial community assembly is associated with function rather than the taxonomy.

The Alkaloids May Contribute to the Variation of the Endophytic Bacterial Communities

Protopine and allocrytopine are considered as precursors of sanguinarine and chelerythrine, respectively, and they are abundant in the root. In a previous study on the dynamics of the four alkaloids of *Macleaya cordata,* it was found that after entering the mature fruiting season, the content of protopine in the fruit decreased significantly, while the content of sanguinarine and chelerythrine increased rapidly. This suggested that protopine and allocrytopine were transported into the fruit and converted into sanguinarine and chelerythrine. The transcriptome, proteome, and metabolism data in a previous study conducted by Jianguo Zeng et al. (2013) revealed that the root of *Macleaya cordata* is the primary organ for the biosynthesis of isoquinoline alkaloids. In this study, significant differences were found in alkaloids accumulation in various organs. The content of allocrytopine in the root, leaf, and fruit showed a decreasing trend. Besides, the highest accumulation of protopine was found in leaves, followed by roots and fruit. The sanguinarine and chelerythrine in fruit were much higher than those in other tissues. The accumulation of all four alkaloids in stems was low. Some common endophytes, such as Firmicutes, have been reported to be abundant in the above-ground part, whereas in this study, it was only abundant in the stem. It can be speculated that this may explain to some extent the higher microbial diversity in stems since these four alkaloids have antibacterial effects (Beuria et al., 2005; Li et al., 2009; Razan et al., 2014; Yu et al., 2014). The CCA analysis also showed that all four alkaloids were negatively correlated with the microbiota of the stem of *Macleaya cordata*. The correlation tests between alkaloids and microorganisms also showed that most endophytic bacteria were significantly ($P < 0.05$) correlated with alkaloids.

Endophytes and host plants are closely related, and they adapt to each other and coevolve. Studies have shown that genes and abilities that evolve in one lineage are usually acquired steadily by another lineage (Papke and Gogarten, 2012). Direct gene transfer between species has occurred in all major taxa and seems to occur more frequently in prokaryotes (Moran, 2007). This leads to the possibility that microorganisms could respond to environmental toxins by selecting specific gene sequences that give them a competitive advantage over other organisms. It can be hypothesized that the microbes colonized in the plant might be affected by alkaloids produced by the host. Through the analysis of the functional annotation in Hierarchical level 2, endophytic bacteria in roots was found to contribute most to the gene abundance of the Cytochrome P450, ABC transporters, and secondary metabolite synthesis pathway including the isoquinoline alkaloid synthesis pathway. This was consistent with Jianguo Zeng's reports that all enzymes for protopine and allocrytopine biosynthesis were highly expressed in the host root. Cytochrome P450 is considered as a large family of enzymes involved in many important metabolic pathways, and many enzymes related to the biosynthesis of benzylisoquinoline alkaloids belong to the P450 family (Zeng et al., 2013). Simultaneously, the ABC transporters are involved in transporting secondary metabolites such as alkaloids (Shitan and Yazaki, 2007). The number of genes annotated to the Staurosporine biosynthesis pathway was significantly higher in fruits than that in other parts, which may be due to the high accumulation of sanguinarine in the fruit. Staurosporine, an alkaloid with a diindole chemical structure, has been reported to partially block the accumulation of sanguinarine induced by a fungal activator (Peter et al., 1996). Since the primary function of stems is transport and there is almost no alkaloid stored in the stem, this may contribute to the low gene abundance of the microbiota genes related to the synthesis of secondary metabolites in the stem.

CONCLUSION

This study provided the first insight into the bacterial communities of different plant tissues and rhizospheric soil in *Macleaya cordata*. There were significant differences in bacterial communities among different ecological niches under the influence of plants' vertical stratification structure. A strong correlation between the endophytic bacteria and the alkaloids was found by the Spearman correlation analysis. The predicted results of PICRUST2 further demonstrated that alkaloids might contribute to the variation of bacterial communities in different niches. All in all, this experiment can support the subsequent search of functional microorganisms and guide the cultivation, protection, and increase of crucial metabolites and resource utilization of *Macleaya cordata*.

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

FL, XL, and YL conceived and designed the work. FL, HH, and SF performed the experiments. FL and XL wrote and revised the paper. KZ and SZ helped with the sample collection. FL and XL contributed equally to this study. All authors read and approved the final manuscript.

FUNDING

This work was supported by the following grants: Chinese National Science and Technology Support Program

(2013BAC09B00), National Natural Science Foundation of China (31570113), National Key R&D Program of China (2017YFD0501500), and Graduate Research and Innovation Project (1053320170629) in Central South University, China.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.681210/full#supplementary-material

REFERENCES

Arturo, S.-A., Yumi, O., Fernandes, G. W., Ball, R. A., and Gamon, J. (2012). Relationships between endophyte diversity and leaf optical properties. *Trees Struct. Funct.* 26, 291–299. doi: 10.1007/s00468-011-0591-5

Asaf, S., Numan, M., Khan, A. L., and Al-Harrasi, A. (2020). Sphingomonas: from diversity and genomics to functional role in environmental remediation and plant growth. *Crit. Rev. Biotechnol.* 40, 138–152. doi: 10.1080/07388551.2019.1709793

Beckers, B., Op De Beeck, M., Weyens, N., Boerjan, W., and Vangronsveld, J. (2017). Structural variability and niche differentiation in the rhizosphere and endosphere bacterial microbiome of field-grown poplar trees. *Microbiome* 5:25.

Beuria, T. K., Santra, M. K., and Panda, D. (2005). Sanguinarine blocks cytokinesis in bacteria by inhibiting FtsZ assembly and bundling. *Biochemistry* 44, 16584–16593. doi: 10.1021/bi050767%2B

Bonito, G., Reynolds, H., Robeson, M. S. II, Nelson, J., Hodkinson, B. P., Tuskan, G., et al. (2014). Plant host and soil origin influence fungal and bacterial assemblages in the roots of woody plants. *Mol. Ecol.* 23, 3356–3370. doi: 10.1111/mec.12821

Bulgarelli, D., Schlaeppi, K., Spaepen, S., van Themaat, E., and Schulze-Lefert, P. (2013). Structure and functions of the bacterial microbiota of plants. *Annu. Rev. Plant Biol.* 64, 807–838. doi: 10.1146/annurev-arplant-050312-120106

Burke, C., Steinberg, P., Rusch, D., Kjelleberg, S., and Thomas, T. (2011). Bacterial community assembly based on functional genes rather than species. *Proc. Natl. Acad. Sci. U S A* 108, 14288–14293. doi: 10.1073/pnas.1101591108

Campisano, A., Albanese, D., Yousaf, S., Pancher, M., Donati, C., and Pertot, I. (2017). Temperature drives the assembly of endophytic communities' seasonal succession. *Environ. Microbiol.* 19, 3353–3364. doi: 10.1111/1462-2920.13843

Chao, X., YongGuan, Z., JunTao, W., Brajesh, S., LiLi, H., JuPei, S., et al. (2020). Host selection shapes crop microbiome assembly and network complexity. *New Phytol.* 229, 1091–1104.

Compant, S., Clément, C., and Sessitsch, A. (2010). Plant growth-promoting bacteria in the rhizo- and endosphere of plants: their role, colonization, mechanisms involved and prospects for utilization. *Soil Biol. Biochem.* 42, 669–678. doi: 10.1016/j.soilbio.2009.11.024

Cordovez, V., Dini-Andreote, F., Carrión, V. J., and Raaijmakers, J. M. (2019). Ecology and evolution of plant microbiomes. *Annu. Rev. Microbiol.* 73, 69–88. doi: 10.1146/annurev-micro-090817-062524

Cregger, M. A., Veach, A. M., Yang, Z. K., Crouch, M. J., Vilgalys, R., Tuskan, G. A., et al. (2018). The Populus holobiont: dissecting the effects of plant niches and genotype on the microbiome. *Microbiome* 6:31.

Davide, B., Matthia, R., and Klaus, S. (2012). Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. *Nature* 488, 91–95. doi: 10.1038/nature11336

Delmotte, N., Knief, C., Chaffron, S., Innerebner, G., Roschitzki, B., Schlapbach, R., et al. (2009). Community proteogenomics reveals insights into the physiology of *Phyllosphere bacteria*. *Proc. Natl. Acad. Sci. U S A* 106, 16428–16433. doi: 10.1073/pnas.0905240106

Edwards, J., Johnson, C., Santos-Medellin, C., Lurie, E., Podishetty, N. K., Bhatnagar, S., et al. (2015). Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc. Natl. Acad. Sci. U S A* 112, E911–E920.

Fonseca-García, C., Coleman-Derr, D., Garrido, E., Visel, A., Tringe, S. G., and Partida-Martínez, L. P. (2016). The Cacti Microbiome: interplay between habitat-filtering and host-specificity. *Front. Microbiol.* 7:150. doi: 10.3389/fmicb.2016.00150

Frank, A. C., Saldierna Guzmán, J. P., and Shay, J. E. (2017). Transmission of bacterial endophytes. *Microorganisms* 5:70. doi: 10.3390/microorganisms5040070

Galbally, I. E., and Kirstine, W. (2002). The production of methanol by flowering plants and the global cycle of methanol. *J. Atmospheric chem.* 43, 195–229.

Glassner, H., Zchori-Fein, E., Compant, S., Sessitsch, A., Katzir, N., Portnoy, V., et al. (2015). Characterization of endophytic bacteria from cucurbit fruits with potential benefits to agriculture in melons (Cucumis meloL.). *FEMS Microbiol. Ecol.* 91:fiv074. doi: 10.1093/femsec/fiv074

Gutierrez, R. M. P., Gonzalez, A. M. N., and Ramirez, A. M. (2012). Compounds derived from endophytes : a Review of phytochemistry and. *Curr. Med. Chem.* 19, 2992–3030. doi: 10.2174/092986712800672111

Halo, B. A., Khan, A. L., Waqas, M., Al-Harrasi, A., Hussain, J., Ali, L., et al. (2015). Endophytic bacteria (Sphingomonassp. LK11) and gibberellin can improveSolanum lycopersicumgrowth and oxidative stress under salinity. *J. Plant Interact.* 10, 117–125. doi: 10.1080/17429145.2015.1033659

Hardoim, P. R., van Overbeek, L. S., Berg, G., Pirttila, A. M., Compant, S., Campisano, A., et al. (2015). The hidden world within plants: ecological and evolutionary considerations for defining functioning of microbial endophytes. *Microbiol. Mol. Biol. Rev.* 79, 293–320. doi: 10.1128/mmbr.00050-14

Hartman, K., van der Heijden, M. G. A., Wittwer, R. A., Banerjee, S., Walser, J.-C., and Schlaeppi, K. (2018). Cropping practices manipulate abundance patterns of root and soil microbiome members paving the way to smart farming. *Microbiome* 6:14.

Huang, P., Zhang, Y., Xiao, K., Jiang, F., Wang, H., Tang, D., et al. (2018). The chicken gut metagenome and the modulatory effects of plant-derived benzylisoquinoline alkaloids. *Microbiome* 6:211.

Hunter, P. J., Hand, P., Pink, D., Whipps, J. M., and Bending, G. D. (2010). Both leaf properties and microbe-microbe interactions influence within-species variation in bacterial population diversity and structure in the lettuce (*Lactuca Species*) phyllosphere. *Appl. Environ. Microbiol.* 76, 8117–8125. doi: 10.1128/aem.01321-10

Jin, H., Yang, X.-Y., Yan, Z.-Q., Liu, Q., Li, X.-Z., Chen, J.-X., et al. (2014). Characterization of rhizosphere and endophytic bacterial communities from leaves, stems and roots of medicinal Stellera chamaejasme L. *Syst. Appl. Microbiol.* 37, 376–385. doi: 10.1016/j.syapm.2014.05.001

Jorge, G. L., Kisiala, A., Morrison, E., Aoki, M., Nogueira, A. P. O., and Emery, R. J. N. (2019). Endosymbiotic Methylobacterium oryzae mitigates the impact of limited water availability in lentil (Lens culinaris Medik.) by increasing plant cytokinin levels. *Environ. Exper. Bot.* 162, 525–540. doi: 10.1016/j.envexpbot.2019.03.028

Junker, R. R., and Keller, A. (2015). Microhabitat heterogeneity across leaves and flower organs promotes bacterial diversity. *FEMS Microbiol. Ecol.* 91:fiv097. doi: 10.1093/femsec/fiv097

Khadem, A., Soler, L., Everaert, N., and Niewold, T. A. (2014). Growth promotion in broilers by both oxytetracycline and Macleaya cordata extract is based on their anti-inflammatory properties. *Br. J. Nutr.* 112, 1110–1118. doi: 10.1017/s0007114514001871

Kim, J., Chhetri, G., Kim, I., Kim, H., Kim, M. K., and Seo, T. (2019). Methylobacterium terrae sp. nov., a radiation-resistant bacterium isolated from gamma ray-irradiated soil. *J. Microbiol.* 57, 959–966. doi: 10.1007/s12275-019-9007-9

Kosina, P., Gregorova, J., Gruz, J., Vacek, J., Kolar, M., Vogel, M., et al. (2010). Phytochemical and antimicrobial characterization of Macleaya cordata herb. *Fitoterapia* 81, 1006–1012. doi: 10.1016/j.fitote.2010.06.020

Li, H., Wang, J., Zhao, J., Lu, S., Wang, J., and Jiang, W. (2009). Isoquinoline Alkaloids from Macleaya cordata active against plant microbial pathogens. *Nat. Product Commun.* 4, 1557–1560.

Li, L., YanChun, L., JiaLu, H., XiuBin, L., ZhiXing, Q., JianGuo, Z., et al. (2017). Medicinal plants of the genus Macleaya (*Macleaya cordata, Macleaya microcarpa*): a review of their phytochemistry, pharmacology, and toxicology. *Phytother. Res.* 32, 19–48. doi: 10.1002/ptr.5952

Liu, M., Lin, Y.-L., Chen, X.-R., Liao, C.-C., and Poo, W.-K. (2013). In vitro assessment of Macleaya cordata crude extract bioactivity and anticancer properties in normal and cancerous human lung cells. *Exper. Toxicol. Pathol.* 65, 775–787. doi: 10.1016/j.etp.2012.11.004

Liu, T.-H., Zhang, X.-M., Tian, S.-Z., Chen, L.-G., and Yuan, J.-L. (2020). Bioinformatics analysis of endophytic bacteria related to berberine in the Chinese medicinal plant Coptis teeta Wall. *3 Biotech* 10:96.

Loaces, I., Ferrando, L., and Fernández Scavino, A. (2010). Dynamics, diversity and function of endophytic siderophore-producing bacteria in rice. *Microb. Ecol.* 61, 606–618. doi: 10.1007/s00248-010-9780-9

Lu, Y., Zhang, E., Hong, M., Yin, X., Cai, H., Yuan, L., et al. (2020). Analysis of endophytic and rhizosphere bacterial diversity and function in the endangered plant Paeonia ludlowii. *Arch. Microbiol.* 202, 1717–1728. doi: 10.1007/s00203-020-01882-3

Lundberg, D. S., Lebeis, S. L., Paredes, S. H., Yourstone, S., Jase, G., Stephanie, M., et al. (2012). Defining the core Arabidopsis thaliana root microbiome. *Nature* 488, 86–94. doi: 10.1038/nature11237

Maignien, L., DeForce, E. A., Chafee, M. E., Eren, A. M., and Simmons, S. L. (2014). Ecological succession and stochastic variation in the assembly of Arabidopsis thaliana phyllosphere communities. *mBio* 5:e00682–13.

Mangeot-Peter, L., Tschaplinski, T. J., Engle, N. L., Veneault-Fourrey, C., Martin, F., and Deveau, A. (2020). Impacts of soil microbiome variations on root colonization by fungi and bacteria and on the Metabolome of Populus tremula × alba. *Phytobiomes J.* 4, 142–155. doi: 10.1094/pbiomes-08-19-0042-r

Mano, H., and Morisaki, H. (2008). Endophytic bacteria in the rice plant. *Microb. Environ.* 23, 109–117.

Meyer, K. M., and Leveau, J. H. J. (2012). Microbiology of the phyllosphere: a playground for testing ecological concepts. *Oecologia* 168, 621–629. doi: 10.1007/s00442-011-2138-2

Mighell, K., Saltonstall, K., Turner, B. L., Espinosa-Tasón, J., and Bael, S. A. V. (2019). Abiotic and biotic drivers of endosymbiont community assembly in Jatropha curcas. *Ecosphere* 10:e02941. doi: 10.1002/ecs2.2941

Miller, T. E. X., and Rudgers, J. A. (2014). Niche differentiation in the dynamics of host-symbiont interactions: symbiont prevalence as a coexistence problem. *Am. Nat.* 183, 506–518. doi: 10.1086/675394

Moran, N. A. (2007). Symbiosis as an adaptive process and source of phenotypic complexity. *PNAS* 104, 8627–8633. doi: 10.1073/pnas.0611659104

Moroenyane, I., Mendes, L., Tremblay, J., Tripathi, B., and Yergeau, É (2021). Plant compartments and developmental stages modulate the balance between Niche-Based and neutral processes in Soybean Microbiome. *Microbial. Ecol.* **vol&pg.,

Nemergut, D. R., Schmidt, S. K., Fukami, T., O'Neill, S. P., Bilinski, T. M., Stanish, L. F., et al. (2013). Patterns and processes of microbial community assembly. *Microbiol. Mol. Biol. Rev.* 77, 342–356.

Overbeek, L. V., and Elsas, J. D. V. (2008). Effects of plant genotype and growth stage on the structure of bacterial communities associated with potato (Solanum tuberosum L.). *FEMS Microbiol. Ecol.* 64, 283–296. doi: 10.1111/j.1574-6941.2008.00469.x

Papke, R. T., and Gogarten, J. P. (2012). Ecology. How bacterial lineages emerge. *Science* 336, 45–46. doi: 10.1126/science.1219241

Park, S. Y., Jin, M. L., Kim, Y. H., Lee, S.-J., and Park, G. (2014). Sanguinarine inhibits invasiveness and the MMP-9 and COX-2 expression in TPA-induced breast cancer cells by inducing HO-1 expression. *Oncol. Rep.* 31, 497–504. doi: 10.3892/or.2013.2843

Pasquale, A., Sylvia, S., Silvia, P., and Massimiliano, C. (2020). Diversity and structure of the endophytic bacterial communities associated with Three Terrestrial Orchid Species as revealed by 16S rRNA Gene metabarcoding. *Front. Microbiol.* 11:604964. doi: 10.3389/fmicb.2020.604964

Peter, J. F., Alison, G. J., Julie, P., and Vincenzo, D. L. (1996). Uncoupled defense gene expression and antimicrobial alkaloid accumulation in elicited opium poppy CeII cuitures. *Plant Physiol.* 111, 687–697. doi: 10.1104/pp.111.3.687

Pimentel, M. R., Molina, G., Dionisio, A. P., Marostica Junior, M. R., and Pastore, G. M. (2011). The use of endophytes to obtain bioactive compounds and their application in biotransformation process. *Biotechnol. Res. Int.* 2011:576286.

Prasad, V. S. S. K., Davide, G., and Emilio, S. (2018). Plant growth promoting and biocontrol activity of *Streptomyces spp*. as Endophytes. *Int. J. Mol. Sci.* 19:952. doi: 10.3390/ijms19040952

Ramond, J.-B., Tshabuse, F., Bopda, C. W., Cowan, D. A., and Tuffin, M. I. (2013). Evidence of variability in the structure and recruitment of rhizospheric and endophytic bacterial communities associated with arable sweet sorghum (Sorghum bicolor (L) Moench). *Plant and Soil* 372, 265–278. doi: 10.1007/s11104-013-1737-6

Razan, H., Jurgen, R., and Michael, W. (2014). Synergistic antimicrobial activity of combinations of sanguinarine and EDTA with vancomycin against multidrug resistant bacteria. *Drug Metabol. Lett.* 8:10.

Santana, R. S. M., Fernandes, G. W., Ávila, M. P., Reis, M. P., Araújo, F. M. G. D., Salim, A. C. M., et al. (2016). Endophytic microbiota associated with the root tips and leaves of *Baccharis dracunculifolia*. *Brazilian Arch. Biol. Technol.* 59:e16160287.

Sharaby, Y., Rodríguez-Martínez, S., Lalzar, M., Halpern, M., and Izhaki, I. (2020). Geographic partitioning or environmental selection: what governs the global distribution of bacterial communities inhabiting floral nectar? *Sci. Total Environ.* 749:142305. doi: 10.1016/j.scitotenv.2020.142305

Shintaro, H., Masatoshi, M., and Kiwamu, M. (2019). Growth stage-dependent bacterial communities in soybean plant tissues: Methylorubrum Transiently dominated in the flowering stage of the soybean shoot. *Microb. Environ.* 34, 446–450. doi: 10.1264/jsme2.me19067

Shitan, N., and Yazaki, K. (2007). Accumulation and membrane transport of plant alkaloids. *Curr. Pharm. Biotechnol.* 8, 244–252. doi: 10.2174/138920107781387429

Tahtamouni, M. E., Khresat, S., Lucero, M., Sigala, J., and Unc, A. (2015). Diversity of endophytes across the soil-plant continuum for Atriplex spp. in arid environments. *J. Arid Land* 8, 241–253. doi: 10.1007/s40333-015-0061-9

Ting, O., Wei-Fang, X., Fei, W., Gary, S., Ze-Yang, Z., Zhong-Huai, X., et al. (2019). A Microbiome study reveals seasonal variation in endophytic bacteria among different mulberry cultivars. *Comput. Struct. Biotechnol. J.* 17, 1091–1100. doi: 10.1016/j.csbj.2019.07.018

Vandenkoornhuyse, P., Quaiser, A., Duhamel, M., Le Van, A., and Dufresne, A. (2015). The importance of the microbiome of the plant holobiont. *New Phytol.* 206, 1196–1206. doi: 10.1111/nph.13312

Víctor, J. C., and Juan, P.-J. (2019). Pathogen-induced activation of disease-suppressive functions in the endophytic root microbiome. *Science* 366, 606–612. doi: 10.1126/science.aaw9285

Wang, Y., Wang, C., Gu, Y., Wang, P., Song, W., Ma, J., et al. (2020). The variability of bacterial communities in both the endosphere and ectosphere of different niches in Chinese chives (Allium tuberosum). *PLoS One* 15:e0227671. doi: 10.1371/journal.pone.0227671

Wang, Y., Zhang, W., Zhang, B., Huang, Q., Huang, R., and Su, X. (2019). Endophytic communities of transgenic poplar were determined by the environment and niche rather than by transgenic events. *Front. Microbiol.* 10:588. doi: 10.3389/fmicb.2019.00588

Wellner, S., Lodders, N., Glaeser, S. P., and Kampfer, P. (2013). Methylobacterium trifolii sp. nov. and Methylobacterium thuringiense sp. nov., methanol-utilizing, pink-pigmented bacteria isolated from leaf surfaces. *Int. J. Syst. Evol. Microbiol.* 63, 2690–2699. doi: 10.1099/ijs.0.047787-0

Wilkes, H., Wittich, R., Timmis, K. N., Fortnagel, P., and Francke, W. (1996). Degradation of Chlorinated Dibenzofurans and Dibenzo-p-Dioxins by Sphingomonas sp. Strain RW1. *Appl. Environ. Microbiol.* 62, 367–371. doi: 10.1128/aem.62.2.367-371.1996

Winston, M. E., Hampton-Marcell, J., Zarraonaindia, I., Owens, S. M., Moreau, C. S., Gilbert, J. A., et al. (2017). Understanding cultivar-specificity and soil determinants of the cannabis microbiome. *PLoS One* 9:e99641. doi: 10.1371/journal.pone.0099641

Xiong, C., Yong-Guan, Z., Jun-Tao, W., Singh, B., Li-Li, H., Shen, J. P., et al. (2020). Host selection shapes crop microbiome assembly and network complexity. *New Phytol.* 229, 1091–1104. doi: 10.1111/nph.16890

Yana, A.-G., Ido, I., and Malka, H. (2017). From microhabitat of floral nectar up to Biogeographic Scale: novel insights on neutral and niche bacterial assemblies. *Microbial. Ecol.* 74, 128–139. doi: 10.1007/s00248-017-0935-9

Yoshida, S., Hiradate, S., Koitabashi, M., Kamo, T., and Tsushima, S. (2017). Phyllosphere *Methylobacterium* bacteria contain UVA-absorbing compounds. *J. Photochem. Photobiol. B* 167, 168–175. doi: 10.1016/j.jphotobiol.2016.12.019

Yu, H., Zhang, L., Li, L., Zheng, C., Guo, L., Li, W., et al. (2010). Recent developments and future prospects of antimicrobial metabolites produced by endophytes. *Microbiol. Res.* 165, 437–449. doi: 10.1016/j.micres.2009.11.009

Yu, X., Gao, X., Zhu, Z., Cao, Y., Zhang, Q., Tu, P., et al. (2014). Alkaloids from the tribe *Bocconieae* (Papaveraceae): a chemical and biological review. *Molecules* 19, 13042–13060. doi: 10.3390/molecules190913042

Yu, X., Yang, J., Wang, E., Li, B., and Yuan, H. (2015). Effects of growth stage and fulvic acid on the diversity and dynamics of endophytic bacterial community in Stevia rebaudiana Bertoni leaves. *Front. Microbiol.* 6:867. doi: 10.3389/fmicb.2015.00867

Zarraonaindia, I., Owens, S. M., Weisenhorn, P., West, K., Hampton-Marcell, J., Lax, S., et al. (2015). The soil microbiome influences grapevine-associated microbiota. *mBio* 6:e002527–14.

Zeng, J., Liu, Y., Liu, W., and Liu, X. (2013). Integration of transcriptome, proteome and metabolism data reveals the alkaloids biosynthesis in macleaya cordata and macleaya microcarpa. *PLoS One* 8:e53409. doi: 10.1371/journal.pone.0053409

Zheng, Y., and Gong, X. (2019). Niche differentiation rather than biogeography shapes the diversity and composition of microbiome of Cycas panzhihuaensis. *Microbiome* 7:152.

Zhou, J.-Y., Yuan, J., Li, X., Ning, Y.-F., and Dai, C.-C. (2015). Endophytic bacterium-triggered reactive oxygen species directly increase oxygenous sesquiterpenoid content and diversity in Atractylodes lancea. *Appl. Environ. Microbiol.* 82, 1577–1585. doi: 10.1128/aem.03434-15

Ziyuan, Z., Minghan, Y., Guodong, D., Guanglei, G., and Yingying, H. (2020). Diversity and structural differences of bacterial microbial communities in rhizocompartments of desert leguminous plants. *PLoS One* 15:e0241057. doi: 10.1371/journal.pone.0241057

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.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 09 August 2021
doi: 10.3389/fmicb.2021.723682

Establishment of Gut Microbiome During Early Life and Its Relationship With Growth in Endangered Crested Ibis (*Nipponia nippon*)

Ying Zhu[1], Yudong Li[2], Haiqiong Yang[3], Ke He[4] and Keyi Tang[5]**

[1] Institute of Qinghai-Tibetan Plateau, Southwest Minzu University, Chengdu, China, [2] Sichuan Province Laboratory for Natural Resources Protection and Sustainable Utilization, Sichuan Provincial Academy of Natural Resource Sciences, Chengdu, China, [3] Emei Breeding Center for Crested Ibis, Emei, Chengdu, China, [4] College of Animal Sciences and Technology, Zhejiang A&F University, Hangzhou, China, [5] College of Life Sciences, Sichuan Normal University, Chengdu, China

Gut microbiota during early life could influence host fitness in vertebrates. Studies on how gut microbiota colonize the gut in birds using frequent sampling during early developmental stages and how shifts in microbiota diversity influence host growth are lacking. Here, we examine the microbiome profiles of 151 fecal samples from 14 young crested ibis (*Nipponia nippon*), an endangered bird species, collected longitudinally across 13 time points during the early stages of development and investigated their correlation with host growth. Gut diversity showed a non-linear change during development, which involved multiple colonization and extinction events, mainly associated with Proteobacteria and Firmicutes. Gut microbiota in young crested ibis became more similar with increasing age. In addition, gut microbiota exhibited a strong temporal structure and two specific developmental stages; the beginning of the latter stage coincided with the introduction of fresh loach, with a considerable increase in the relative abundance of Fusobacteria and several Firmicutes, which may be involved in lipid metabolism. Crested ibis chick growth rate was negatively correlated with gut microbiota diversity and negatively associated with the abundance of Halomonadaceae, Streptococci, Corynebacteriaceae, and Dietziaceae. Our findings highlight the importance of frequent sampling when studying microbiome development during early stages of development of vertebrates. The role of microbial diversity in host growth during the early stages of development of birds warrants further investigations.

Keywords: crested ibis (*Nipponia nippon*), host growth, gut microbiota diversity, non-linear change, microbial recruitment patterns, diet change, environment change

OPEN ACCESS

Edited by:
Qiang Gao,
Hunan University, China

Reviewed by:
Monica Rosenblueth,
National Autonomous University of Mexico, Mexico
Xiaolong Cao,
Rutgers, The State University of New Jersey, United States

****Correspondence:***
Ying Zhu
so_zy2003@126.com
Keyi Tang
tangkeyi0214@163.com

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *11 June 2021*
Accepted: *20 July 2021*
Published: *09 August 2021*

Citation:
Zhu Y, Li YD, Yang HQ, He K and Tang KY (2021) Establishment of Gut Microbiome During Early Life and Its Relationship With Growth in Endangered Crested Ibis (Nipponia nippon). Front. Microbiol. 12:723682. doi: 10.3389/fmicb.2021.723682

INTRODUCTION

Gut microbiota participate in host health maintenance, nutrition uptake, digestion, energy release, detoxification, gut development, and the regulation of host physiology and immunity (Waite and Taylor, 2015; Grond et al., 2018), and hence disruptions of normal microbiome could lead to loss of such benefits (Reid et al., 2011; Lewis et al., 2015). In addition, gut microbial communities could influence host fitness, including survival (Kohl et al., 2018), reproductive performance (Hamdi et al., 2011; Rosengaus et al., 2011), and adaption (Blumstein et al., 2017).

In birds, in addition to ecological variables such as diet (Torok et al., 2011; Huang et al., 2018), captivity status (Xenoulis et al., 2010; Wienemann et al., 2011), locality (Hird et al., 2014; Capunitan et al., 2020), and seasons (Lewis et al., 2016; Góngora et al., 2021), biological factors, such as age, are important factors shaping gut microbiota. For example, gut microbiota during early life has been reported to differ markedly from gut microbiota in conspecific adults, as observed in house sparrow (*Passer domesticus*) (Kohl et al., 2019), black-legged kittiwakes (*Rissa tridactyla*) (van Dongen et al., 2013), chinstrap penguins (*Pygoscelis antarctica*) (Barbosa et al., 2016), and folivorous hoatzin (*Opisthocomus hoazin*) (Godoy-Vitorino et al., 2010).

Gut bacterial communities are relatively stable in adults, however, during early life, they are much more transient and dynamic, based on studies conducted on non-avian vertebrates (Grond et al., 2018). Several recent studies have investigated the early establishment of gut microbiota in domestic birds such as chicken (Yin et al., 2010; Oakley et al., 2014; Ballou et al., 2016), turkey (Wilkinson et al., 2017), and wild birds, such as in house sparrows (Kohl et al., 2019), great tits (*Parus major*) (Teyssier et al., 2018), little penguins (*Eudyptula minor*), short-tailed shearwaters (*Ardenna tenuirostris*) (Dewar et al., 2017), black-legged kittiwakes (van Dongen et al., 2013), folivorous hoatzin (Godoy-Vitorino et al., 2010), chinstrap penguins (Barbosa et al., 2016), and dunlin (*Calidris alpina*) (Grond et al., 2017). Such studies on gut microbiota diversity trends with aging during early life have yielded varying results across species. For example, older nestlings have lower microbial diversity than younger nestlings, which was observed in great titis (Teyssier et al., 2018), chicken (Oakley et al., 2014; Ballou et al., 2016), and chinstrap penguins (Barbosa et al., 2016). Opposite trends were observed in black-legged kittiwakes (van Dongen et al., 2013). In addition, some studies have revealed that age does not influence gut microbiota diversity in nestlings in house sparrow (Kohl et al., 2019) and short-tailed shearwaters (Dewar et al., 2017). Furthermore, some species, such as turkey, exhibited much more complex gut microbiota diversity patterns (Wilkinson et al., 2017). These studies have reported strong fluctuations in community composition in the first stages of nestling development. The diverse results across species during early life are thought to be influenced mainly by environmental factors (e.g., chick rearing conditions) (Hird et al., 2014; Grond et al., 2018) and diet (Kohl et al., 2018). Furthermore, infrequent sampling could result in distinct conclusions with regard to gut microbiota colonization processes (de Muinck and Trosvik, 2018). Longitudinal studies with frequent sampling are required to illustrate a comprehensive overview of gut microbiota dynamics in birds (Grond et al., 2018).

Microbial diversity could influence host fitness through its effects on chick growth rate, which is considered a strong predictor of survival later in life (Magrath, 1991). It is unclear whether an increase in diversity could enhance or limit host growth based on the results of the relatively few studies available currently. Most studies have showed that microbial diversity limits growth based on direct evidence come from the comparison of germ-free chickens and conventional chicks, where germ-free chicks grow more rapidly than conventional chicks (Forbes and Park, 1959). Furthermore, antibiotic treatment (assuming antibiotics decrease microbiota diversity) increased growth in chicks (Potti et al., 2002; Dibner and Richards, 2005; Kohl et al., 2018), potentially via the enhancement of food conversion efficiency (Kohl et al., 2018). However, studies performed on ostrich (*Struthio camelus*) have revealed contradicting results at different stages of development: gut microbial diversity was strongly positively associated with growth only during the first week after hatching, and microbial diversity was negatively related with growth after the first week (Videvall et al., 2019). The limited studies and inconsistent results on the association between microbial diversity and host growth in birds highlight the need for further research into the role of microbial diversity in host growth during the early lives of birds.

Crested ibis (*Nipponia nippon*) is an endangered bird species currently found in China, Japan, and South Korea. *Ex situ* conservation, which aims to reintroduce captive individuals to the wild, has been established in China since 1981 (Ding, 2004), where crested ibis are raised in controlled conditions (indoor) after hatching and moved outdoors until reaching approximately 44 days or age. The chicks are raised separately from their parents and are mainly fed on loaches. During the indoor stage, crested ibis undergo a diet change and two environmental changes (Ding, 2004). Gut microbiota composition at hatching is distinguished from other life stages (Ran et al., 2021), however, it is unclear whether community structure is mainly influenced by biological (such as age) or ecological variables (such as diet and environmental changes) during early life. In addition, variation in chick growth rate has been previously observed but whether it is due to gut microbiota remains to be determined.

In the present study, we describe gut microbiota diversity of crested ibis using 16S rRNA amplicon sequencing using frequent sampling strategy. Repeated fecal sampling was conducted under controlled conditions from at hatch until day 44, which corresponds to the developmental stages in this species (Ding, 2004). Our study aims to (1) analyze gut microbiota diversity trends and demonstrate microbial recruitment patterns and colonization processes during growth. (2) Investigate the influence of two ecological variables (diet and environment) and three biological variables (age, sex, and genetic relatedness) on gut microbiota structure and elucidate the functions of the key bacterial groups. (3) Demonstrate the effect of gut microbiota diversity and composition on growth.

MATERIALS AND METHODS

Sample Collection

Fourteen newborn crested ibis from three families were raised at the Emei breeding center of crested ibis and housed in separate incubators (**Supplementary Table 1**). Crested ibis chick were housed in the brood box with the temperature −0.6°C per day from 37°C, fed on loach paste at hatching until day 12 (stage one and diet type 1 in **Figure 1**), and moved to the brood room with consistent temperature of 30°C and where they were fed loach paste until day ∼22 to ∼25 (stage two and diet type 1 in **Figure 1**). Finally, the birds lived in the brood room

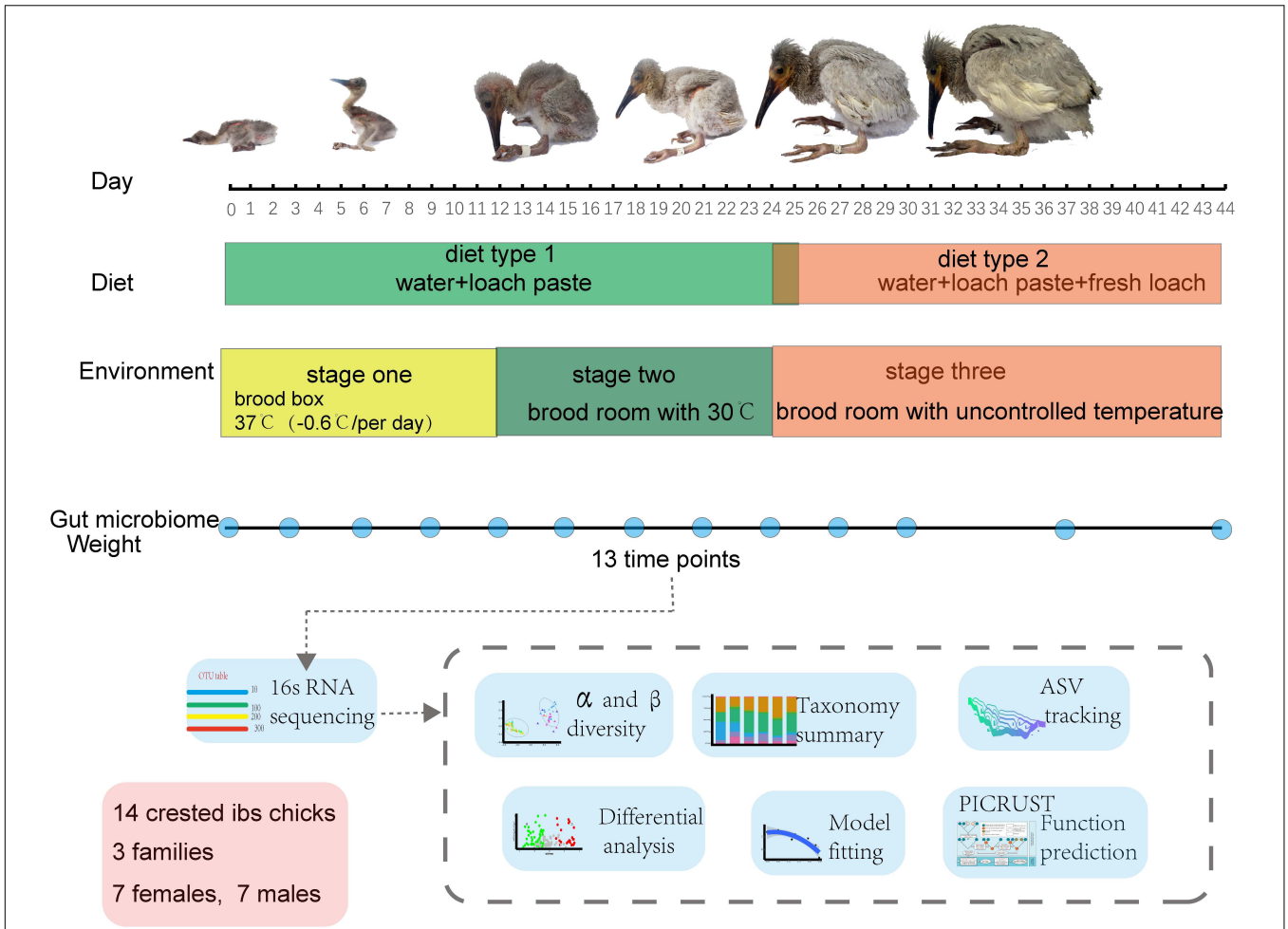

FIGURE 1 | Overview of the study design and sample collection.

without controlled temperature, and fresh loach was added to their formula (stage three and diet type 2 in **Figure 1**). The addition of fresh loach and controlled temperature were both to adapt chick development (Ding, 2004). None of the crested ibis were administered antimicrobial drugs during the sampling period. Fresh feces (n = 182) were obtained from crested ibis, from hatching to day 44 between May and June 2018 (repeatedly sampling at hatching and days 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 37, and 44; **Figure 1** and **Supplementary Table 1**). Crested ibis were weighed every day before the first feeding (**Supplementary Table 1**). Feces were collected using sterile spoons and placed in sterile tubes. The fecal samples were stored in liquid nitrogen until DNA extraction. Sex was identified using the CD1 gene (He et al., 2013).

DNA Extraction, 16S rRNA Gene Sequencing, and Data Processing

DNA extraction from fecal samples was performed using the hexadecyltrimethylammonium bromide method. The V3–V4 regions of the 16S rRNA genes were amplified using primers 341F: 5′-CCTAYGGGRBGCASCAG-3′ and 806R: 5′-GGACTACNNGGGTATCTAAT-3′. Samples with failed PCR amplification were abandoned (n = 31), resulting in 151 samples in the subsequent process. Paired-end sequencing (2 × 250 PE) was conducted at the Nova sequencing company (Novaseq, Tianjin, China) on an Illumina Novaseq 6000 sequencing system (Illumina, San Diego, CA, United States).

Processing of sequence data was conducted using a combination of usearch v10.0.240 and vsearch v2.15.0 (Liu et al., 2021). Dereplication was conducted using the "derep_fulllength" function in vsearch with a minimum unique size of 10, to eliminate artefactual reads. The Divisive Amplicon Denoising Algorithm was performed using the unoise3 function in usearch for correcting amplicon errors (Callahan et al., 2016), yielding 1404 amplicon sequence variations (ASVs). We obtained a total of 9,693,246 high-quality reads from 151 samples (averaging 64193.7 and ranging from 5,0136 to 6,9673 reads per sample). We observed 1404 ASVs after denoising. ASV feature tables were created with the "usearch_global" function in vsearch with a similarity probability of 0.97. Taxonomic profiling was

done against the Ribosomal Database (rdp_16s_v16_sp)[1] using the sintax function in vsearch. Sequencing library sizes were normalized to 50,000 to adjust for sample differences.

Statistical Analysis

All statistical tests were conducted in R (version 4.0.2, 2020-06-22). Sample differences in sequencing library size were normalized to 50,000 using the "Vegan" package (Oksanen et al., 2020). Shannon index were calculated in R package "Vegan". Bray–Curtis distances and weighted UniFrac distances (Lozupone and Knight, 2005) were computed with the "beta_div" function in usearch.

Permutational multivariate analysis of variance (PERMANOVA) was conducted to detect the effects of age, sex, family group (genetic relatedness), individuality, temperature, and diet using the "adonis" function in the "Vegan" package using both Bray–Curtis distances and weighted UniFrac distances with 999 permutations (Oksanen et al., 2020). Principal co-ordinates analysis (PcoA) of Bray–Curtis distance matrices was conducted using the "cmdscale" function.

We used a polynomial linear mixed-effects model to estimate smooth terms in order to fit non-linearity among the Shannon index, Bray–Curtis distance, and age as the predictor variable, with stage used as the covariate factor and individual ID controlled. The relative abundances of taxa of interest were modeled to age, with stage as a covariate factor and individual ID controlled in the linear mixed-effects model.

We modeled young bird growth (weight change per day between time t and $t + 1$) to microbial diversity at time t, including age at time t and diet type at time t, as the covariate and individual ID as a random factor (Videvall et al., 2019) using linear mixed-effect models. To investigate the effect of specific microbial phyla, classes, and families on growth, we also modeled the growth to relative abundances of specific taxa (details in **Supplementary Data 1**). All the linear mixed-effects models were conducted in the NLME package in R (Pinheiro et al., 2021).

The bacterial metagenome was predicted from the 16S rRNA database (Greengenes Database, version gg_13_5) and functional profiles were inferred from the Kyoto Encyclopedia of Gene and Genomes (KEGG) using PICRUSt (phylogenetic investigation of communities by reconstruction of unobserved states) (Douglas et al., 2018).

Differential abundances between two adjacent time points and between two diet types (for microbiota composition and functional profile) were detected in edgeR using a negative binomial generalized linear model with individual ID controlled (Robinson et al., 2010). The model considers sample library size and the dispersion of each ASVs or taxon, which are achieved using the calcNormFactors function. ASVs or taxa with the mean relative abundances $\geq$0.01% were retained. The P values were adjusted for multiple tests with the Benjamini and Hochberg false discovery rate of 0.05 (Benjamini and Hochberg, 1995). ASVs or taxa were significantly abundant or depleted if they had a corrected P value < 0.05 and $|\log FC| > 2.0$ (details in **Supplementary Data 2**).

Spearman's rank correlation coefficients were used to examine the associations between specific genera and pathways (both significantly related to diet) and calculated in the Hmisc package in R (Harrell, 2021).

Sankey plots were produced in imageGP[2] by tracking ASVs with mean relative abundances across all samples greater than 0.1%. Heatmaps were illustrated using the pheatmap package in R (Kolde and Kolde, 2015). Other data visualization procedures were performed using the ggplot2 package (Wickham, 2011).

For the linear mixed-effects models, we removed variables that neither affected the dependent variable nor contributed to the model in order to the keep the model simple. Additionally, we removed variables that could potentially induce serious multicollinearity and reduce the precision of the estimated coefficients.

RESULTS

Gut Microbiota Varied Over Time

Gut microbiota diversity fluctuated notably over time during the first 44 days, decreasing from day 1 after birth to day 12 and then increasing until day 37 (**Figure 2A**). Like alpha diversity, dissimilarity within samples revealed a temporal pattern (**Figure 2B**). A wave trough appeared at day 15, where Bray-Curtis dissimilarity within age group was the smallest. Second peaks emerged for both Shannon diversity and Bray-Curtis distance at day 37. The temporal patterns of alpha diversity and beta diversity both fitted the polynomial age term (linear mixed-effects model; **Supplementary Table 2**).

With an increase in age, gut microbial community similarity within age groups increased to levels greater than in other age groups (days 12, 15, 18, 21, 24, 27, 30, 37, and 44, **Supplementary Figure 1**). The dissimilarity between day 44 (the oldest age we studied) and other time points revealed an overall decrease in dissimilarity with growth (**Supplementary Figure 1**), indicating that gut microbiota structure gradually stabilized after birth.

To identify the main ASV associated with the diversity trend, we calculated the correlation between the relative abundance of the 20 most abundant ASVs (91.7% of total abundance) and the corresponding Bray-Curtis distance. We observed that eight ASVs (ASV10∼12, ASV15∼19, 5 Proteobacteria, 2 Firmicutes, and 1 Fusobacteria, **Supplementary Figure 2**) were significantly correlated with Bray–Curtis distances (**Supplementary Table 3**).

The most prevalent phyla at different ages were Firmicutes and Proteobacteria (Gammaproteobacteria; **Figure 2C** and **Supplementary Figure 3**). The Proteobacteria and Firmicutes phyla were dominated by Enterobacteriaceae, Clostridiaceae_1, Carnobacteriaceae, and Peptostreptococcaceae, respectively, (**Figure 2D** and **Supplementary Figure 4**). Although age did not significantly influence the relative abundance of Firmicutes and Gammaproteobacteria (Firmicutes, $\beta \pm SE = 0.061 \pm 0.044$, $t = 1.369$, $P > 0.05$; Gammaproteobacteria, $\beta \pm SE = -0.536 \pm 0.271$, $t = -1.978$,

[1] www.drive5.com/sintax/

[2] www.ehbio.com

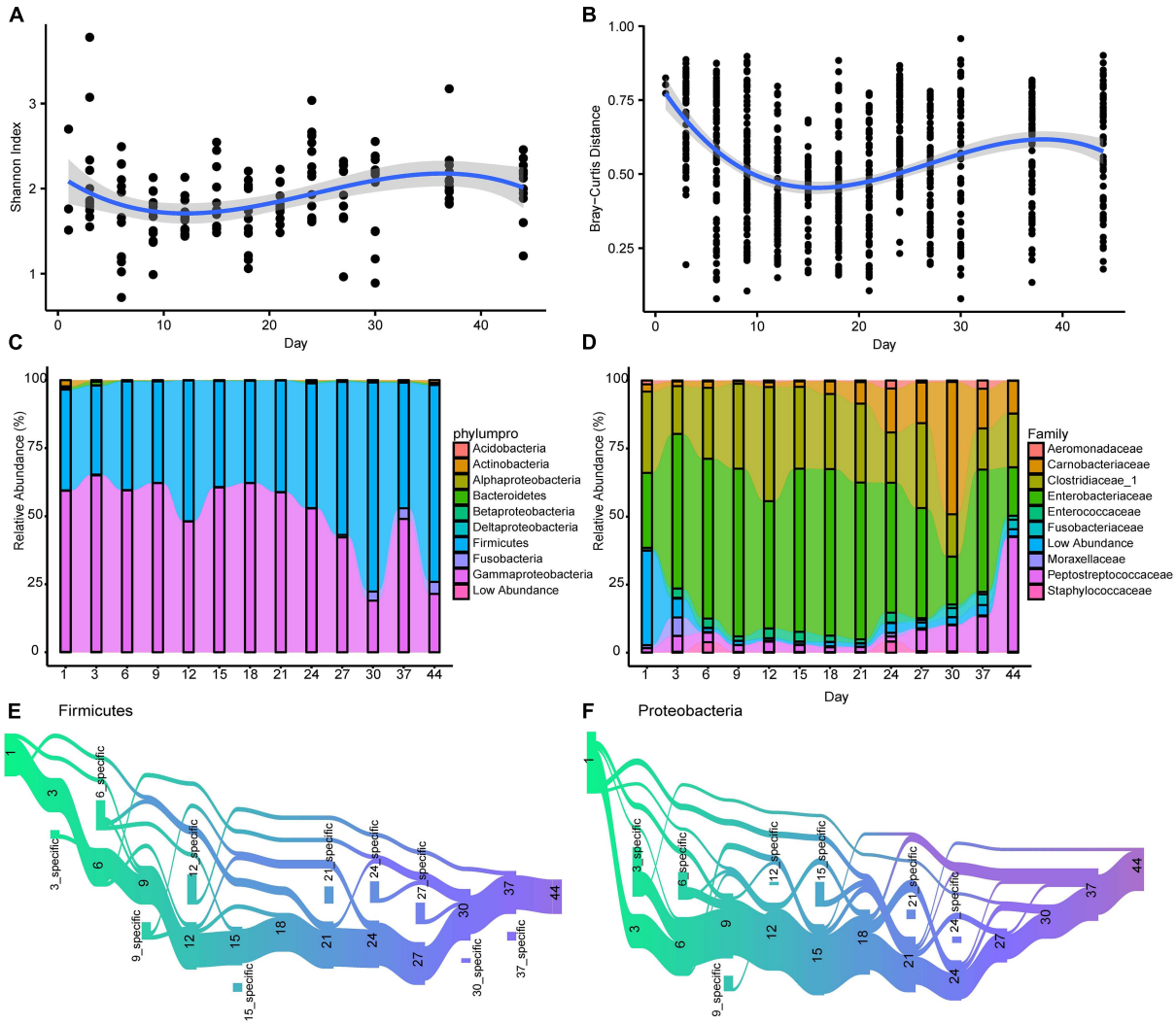

FIGURE 2 | The longitudinal dynamic of gut microbiota in Crested ibis. **(A)** The alpha diversity of gut microbiota over time. The shadow around the lines reveals 95% CI. The alpha diversity was estimated using the Shannon index. **(B)** The beta diversity of gut microbiota over time. The Bray-Curtis distance calculated at the ASV level across microbiota from samples collected at the same time point. Higher values denoted more dissimilar gut microbiota. Samples difference in sequencing library size were normalized to 50,000 reads. **(C)** The temporal changes of bacterial relative abundance at the phylum/class level (top 10). **(D)** The temporal changes of bacterial relative abundance at the family level (top 10). **(E,F)** ASVs that the mean relative abundance across all samples larger than 0.1% are tracked using Sankey plots for both Firmicutes and Proteobacteria. The heights of the rectangles show the relative number of ASVs, and each time point has a distinct color. The lines represent the transfer of ASVs among different time points.

P = 0.050; linear mixed-effects model) when the effect of diet and individual ID was controlled, we observed a positive effect of age on the abundance of one dominant Firmicutes family (Peptostreptococcaceae $\beta \pm SE = 0.478 \pm 0.154$, $t = 3.116$, $P = 0.002$). We did not observe any obvious influence on the other three abundant families (Enterobacteriaceae, Clostridiaceae_1, and Carnobacteriaceae) and the four main genera (*Escherichia*, *Clostridium_sensu_stricto*, *Plesiomonas*, and *Catellicoccus* (**Supplementary Figure 5**). As for the effect of diet, addition of fresh loach decreased Enterobacteriaceae abundance, mainly *Escherichia* ($\beta \pm SE = -16.028 \pm 7.004$, $t = -2.288$, $P = 0.023$), however, it increased Carnobacteriaceae abundance, mainly *Catellicoccus* ($\beta \pm SE = 12.050 \pm 5.136$, $t = 2.346$, $P = 0.020$). Diet change had no effect on the two main phyla or the other two abundant families, Peptostreptococcaceae and Clostridiaceae_1.

Sankey plots revealed distinct temporal dynamics when tracking ASVs within the two dominant phyla, Proteobacteria and Firmicutes (**Figures 2E,F**). A large proportion of the Firmicutes ASVs flew from days 1 to 3 (80%), and then to other time points. Firmicutes ASVs that disappeared reemerged at days 9 and 21 (**Figure 2E**). New Firmicutes ASVs appeared at most of the time points (10 time points). Compared to Firmicutes ASVs, less Proteobacteria ASVs transferred from day 1 to 3 (55%), and nearly a half of the Proteobacteria ASVs disappeared between days 1 and 3 (**Figure 2F**). The Proteobacteria ASVs that disappeared reemerged on days 6, 9,

12, and 21. New Proteobacteria ASVs appeared at few time points (seven time points).

To investigate the colonization and extinction of bacterial groups throughout development in detail, we analyzed the differences between two closest sampling time points (**Figure 3**). Day 3 showed seven depleted ASVs (total: 137) as compared with day 1, with these from Actinobacteria and Gammaproteobacteria. Comparison of day 6 with day 3 showed eight differential ASVs, with disappearance of Gammaproteobacteria, Bacteroidia, and Bacilli.

Moreover, day 9 versus day 6, day 12 versus day 9, day 15 versus day 12, day 18 versus day 15, day 21 versus day 18, and day 37 versus day 30 comparisons revealed the most similarities in overall ASV abundance (no distinct ASVs were detected), showing no obvious colonization or extinction of bacteria groups from day 6 to day 21 or from day 30 to day 37. The fewest similarities were observed in comparisons between day 24 and day 21 (n = 10). Bacilli, Betaproteobacteria, and Gammaproteobacteria were recruited after day 21, and after day 24, more Clostridia than other groups were recruited.

Gene Functional Pathways During the First Six Weeks

The microbiota of crested ibis was mainly associated with metabolism (mean relative abundance, 44.1%), environmental information processing (18.28%), and functions associated with genetic information processing (16.67%). In addition, the microbiota functional compositions of the most abundant KEGG pathways were stable during development (**Supplementary Figure 6**).

The alpha diversity of functional profiles decreased during development and fitted a linear and a quadratic age term (**Figure 4A** and **Supplementary Table 2**). The dissimilarity in functional profiles within age groups revealed decreasing trends, although they exhibited fluctuation, and fitted a linear and a cubic age term (**Figure 4B** and **Supplementary Table 2**).

Age and the Diet Shaped Gut Microbiota Diversity and Gene Functional Pathways in Crested Ibis

The microbiota were highly dynamic during development. With an increase in age, samples within age groups were increased in similarity (**Supplementary Figure 1**) and tended to cluster by age based on PcoA of both Bray-Curtis and weighted Unifrac distances (**Figures 5A,B**, PERMANOVA using BC distances: time: R^2 = 0.112, F_{12} = 2.355, P = 0.001; Unifrac distance: R^2 = 0.142, F_{10} = 2.696, P = 0.001). Samples from stage one to stage two (diet type 1) were separated from samples from stage three (diet type 2) in the first coordinate axis, which suggests diet change was another major factor influencing gut microbiota diversity during development (**Figures 5C,D**, PERMANOVA using BC distances: R^2 = 0.130, F_2 = 12.576, P = 0.001; Unifrac distance: stage: R^2 = 0.130, F_1 = 12.343, P = 0.001).

Similar temporal structures were observed in the gene functional profiles. Age and diet were also major factors influencing the functional profiles (age: R^2 = 0.101, F_{10} = 1.958, P = 0.01; stage: R^2 = 0.188, F_1 = 18.247, P = 0.001; PERMANOVA using BC distances, **Supplementary Figure 7**).

Sex and genetic relatedness had no significant influence of on either bacteria community structure or functional profiles ($P > 0.05$).

Specific Genera Participate in Metabolic Pathways

To reveal diet-related changes in gut microbiota during the growth of crested ibis, we compared raw abundances between diet type 2 and diet type 1 at the ASV and genus levels, respectively, (**Figures 6A,B**). We detected 29 enriched ASVs and 31 depleted ASVs in diet type 2, with most of the different ASVs coming from Firmicutes and Proteobacteria (**Figure 6A**). At the genus level, we detected 21 significantly-different bacterial taxa, with 14 enriched in diet type 2 and seven enriched in diet type 1 and belonging to five phyla (**Figure 6B**). Additionally, the bacterial taxa exhibited obvious temporal variation. The abundances of two Fusobacteria genera, two Bacteroidetes genera, and two Actinobacteria genera were relatively stable across diet type 1 but increased during diet type 2. The other four phyla groups varied across the two diet types.

Among 328 KEGG pathways tested, 10 pathways differed in abundance between diet type 2 and diet type 1 (**Supplementary Figure 8**), with these including pathways associated with metabolism (8) and organismal systems (2). Four pathways that were all associated with metabolism (one lipid metabolism, one metabolism of terpenoids, and polyketides, and two biosynthesis of other secondary metabolites) were enriched in diet type 2, and the other six pathways were enriched in diet type 1. The abundances of metabolism-associated genes were stable in diet type 2. For diet type1, we observed a decreasing trend for two pathways (Chlorocyclohexane and chlorobenzene degradation, and Fluorobenzoate degradation). At 3 days, three pathways, including one involving immune regulation (the RIG-I-like receptor signaling pathway), one involving the digestive system (Carbohydrate digestion and absorption), and one involving lipid metabolism (Steroid hormone biosynthesis), were abundant.

The enriched pathway (lipid metabolism) was highly significant with enriched Firmicutes genera in diet type 2 (**Figure 7**). The most of depleted pathways in diet type 2 were highly positively significant with depleted Proteobacteria and negatively associated with enriched Firmicutes. For example, carbohydrate digestion and absorption, which was depleted in diet type 2, was positively associated with depleted *Enterobacter* (Proteobacteria) and negatively associated with enriched *Catellicoccus* (Firmicutes).

Growth Rate Associated With Gut Microbiota Diversity

Over the 44-day growth period, we identified a negative relationship between growth rate and alpha diversity when diet and individual ID were controlled (Shannon: $\beta \pm SE = -0.020 \pm 0.005$, $t = -3.827$, $P < 0.001$; **Figure 8A**). Further analyses of the relative abundances showed that four families (Halomonadaceae, Streptococcaceae,

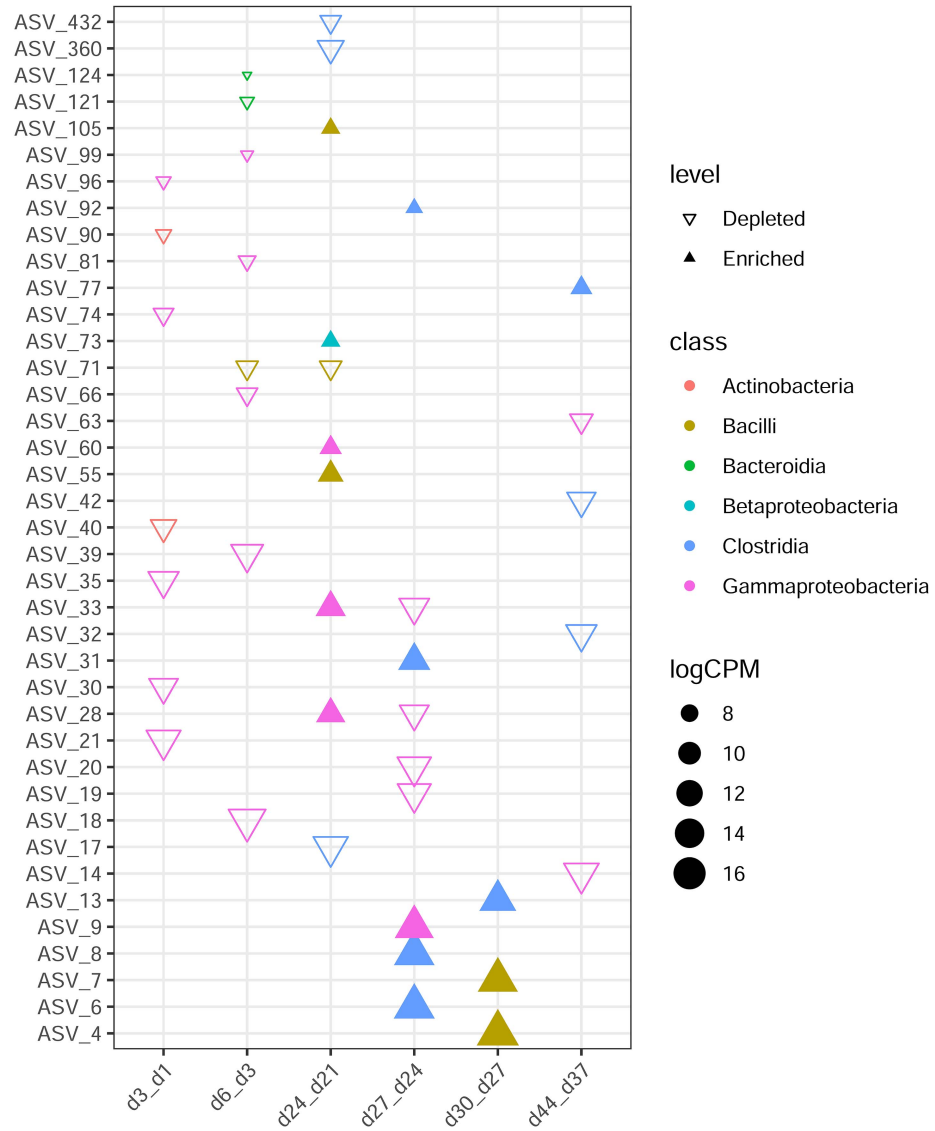

FIGURE 3 | Differences in ASV abundance between two closest ages. The samples of the former time point were taken as control of the latter time point. Each triangle represents a single ASV with adjusted *P* value < 0.05 and | logFC| > 2. Enriched in older age are revealed by filled triangles and hollow triangles shown ASVs enriched in younger age group. ASVs are colored by their class. CPM denotes count per million.

Corynebacteriaceae, and Dietziaceae) were negatively associated with growth during development (Halomonadaceae: $\beta \pm SE = -0.019 \pm 0.008$, $t = -2.530$, $P = 0.013$; Streptococcaceae, $\beta \pm SE = -0.020 \pm 0.007$, $t = -2.855$, $P = 0.005$; Corynebacteriaceae, $\beta \pm SE = -0.060 \pm 0.024$, $t = -2.496$, $P = 0.014$, and Dietziaceae, $\beta \pm SE = -0.035 \pm 0.014$, $t = -2.397$, $P = 0.019$; **Figures 8B–E**).

DISCUSSION

Gut microbiota diversity (both alpha and beta) exhibited non-linear changes (**Figure 2A**) during the sampling period, as expected. Such a pattern was not observed in other age-related studies, including in another study on crested ibis (Ran et al., 2021) and other avian species, due to infrequent sampling. Our findings revealed that bacteria colonization and extinctions occurred throughout the development in crested ibis, and mainly occurred in Proteobacteria and Firmicutes phyla.

The Firmicutes and Proteobacteria phyla dominated the gut microbiota during the first 44 days in crested ibis, with an increase in Firmicutes and a decrease in Proteobacteria. Similar taxonomic changes have been observed in the course of chick development in other birds, such as little penguin (Dewar et al., 2017), arctic shorebirds (Grond et al., 2017), and great tits (Parus major) (Teyssier et al., 2018). Firmicutes produce short-chain fatty acids, which can be absorbed directly by host gut walls as a source of energy

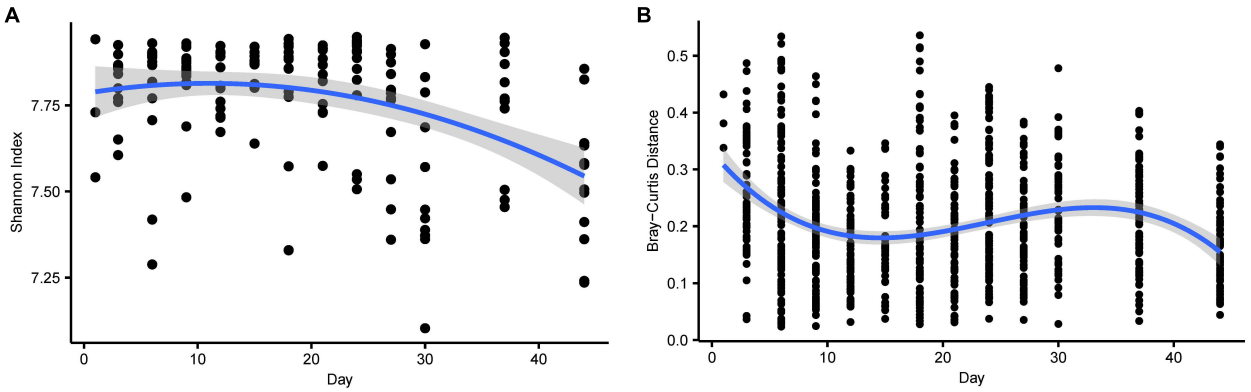

FIGURE 4 | (A) The alpha diversity (Shannon index) of gene function over time. The shadow around the lines reveals 95% CI. The alpha diversity was estimated using the Shannon index. **(B)** The beta (Bray-Curtis distance) diversity of gene function over time. The Bray-Curtis distance calculated at the KO level across microbiota from samples collected at the same time point. Higher values denoted more dissimilar gene function.

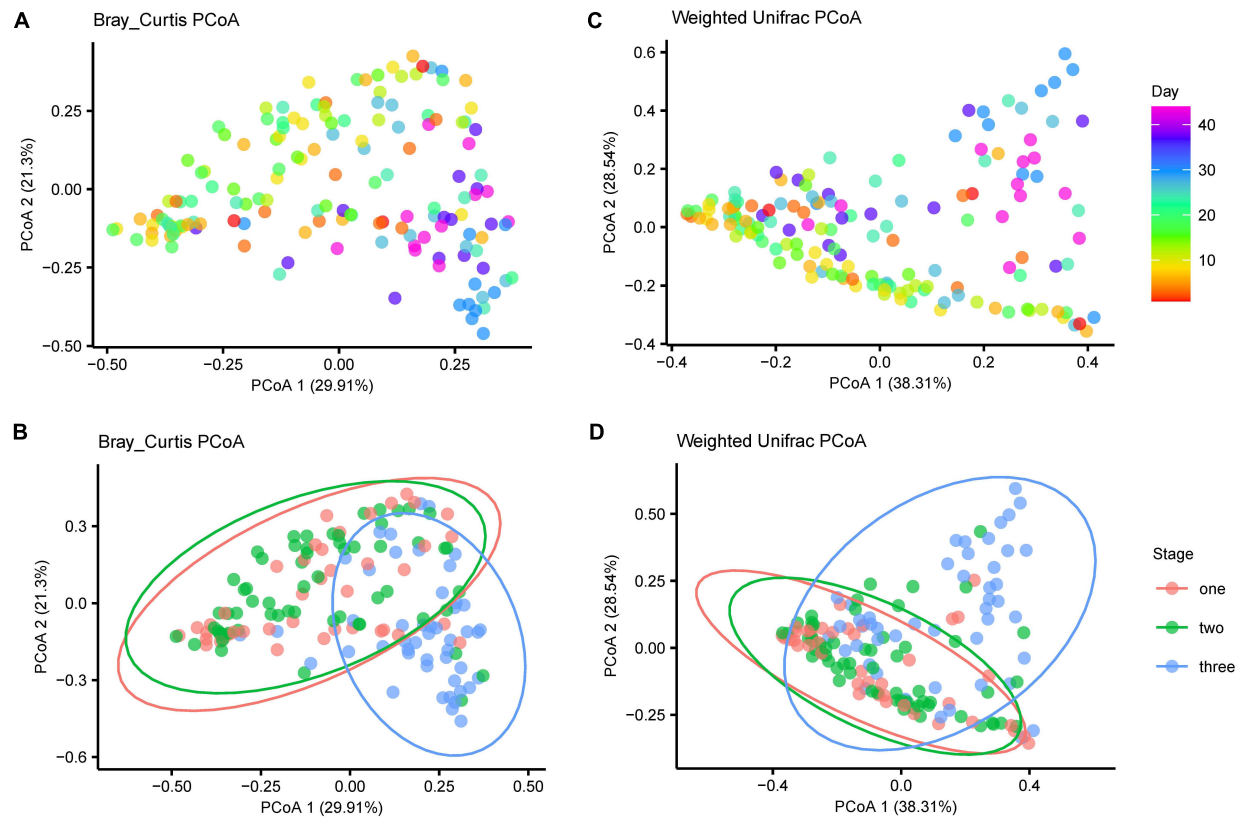

FIGURE 5 | The gut microbiota shift over time. Principle coordinate analysis showing that the gut microbiota with age since birth and stages. **(A)** Bray-Curtis distance for gut microbiota of the age effect **(B)** Bray-Curtis distance for gut microbiota of the three stages **(C)** weighted unifrac distances for gut microbiota of the age effect **(D)** weighted unifrac distances for gut microbiota of the three stages.

(Den Besten et al., 2013) and are positively associated weight gain and immune function in both birds and mammals (Angelakis and Raoult, 2010; Clemente et al., 2012; Liao et al., 2015; John and Mullin, 2016). In particular, an increase in the abundance of Firmicutes was associated with increased Clostridia (mainly Clostridiaceae_1, and Peptostreptococcaceae) and Bacilli (Carnobacteriaceae). Clostridia are also pioneer bacteria in the human gastrointestinal (GI) tract and essential in gut homeostasis (Lopetuso et al., 2013). Bacilli abundance increased considerably at 27 days, which was consistent with the time of fresh loach addition (**Supplementary Figure 1**). Furthermore, Clostridia and Bacilli were significantly

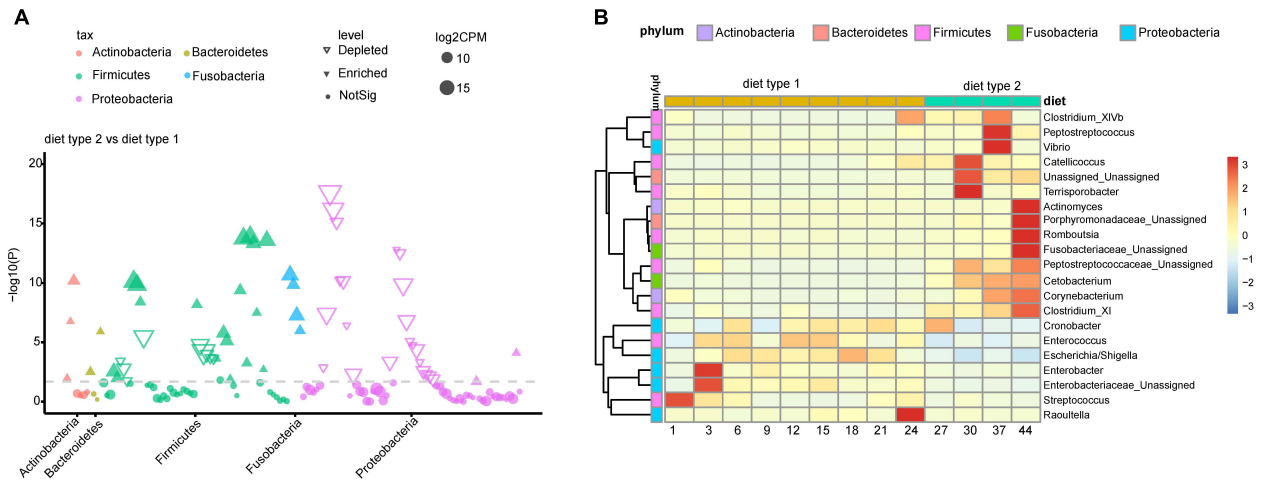

FIGURE 6 | Diet-induced changes in gut microbiota during Crested ibis' growth. **(A)** Enriched and depleted ASVs for comparison between diet type 2 and diet type 1. **(B)** The relative abundance of enriched and depleted bacteria genus for different diet types along with the age. The shown ASVs and genus was obtained from the differential abundance analysis which was conducted by fitting a generalized linear mixed model with a negative binomial distribution in edgeR. Genus counts were normalized for edgeR size factor.

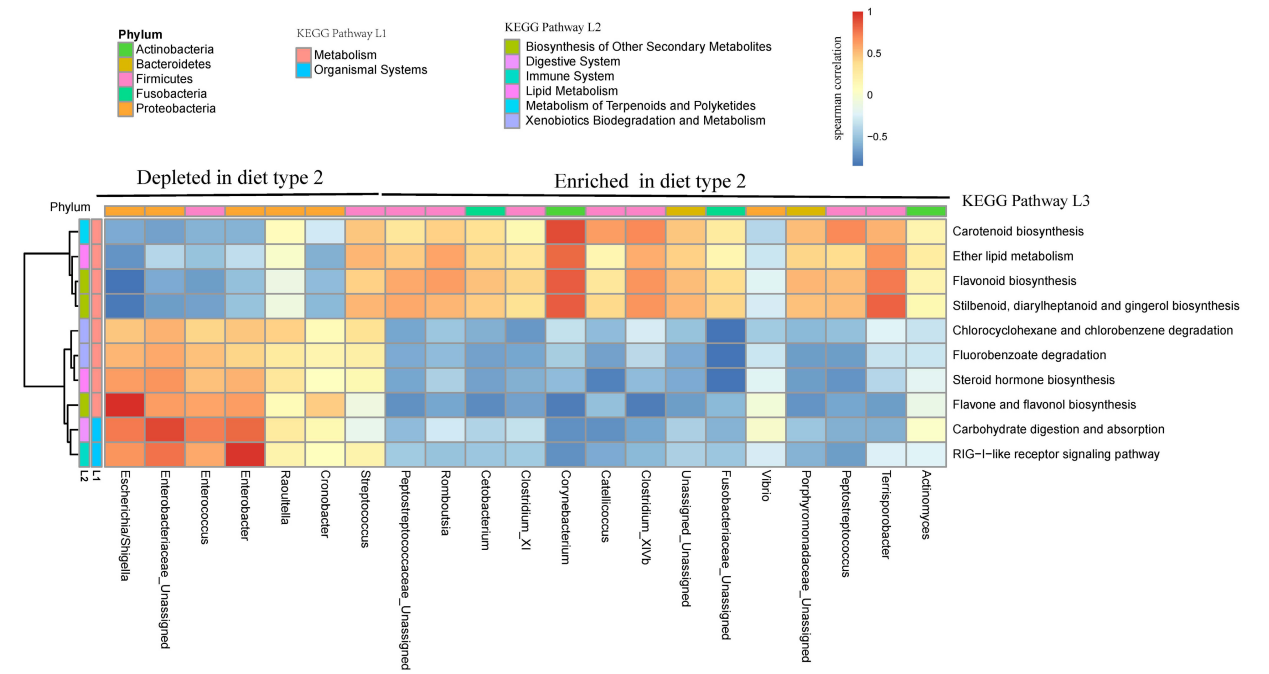

FIGURE 7 | Bacterial taxa from two diet types are related to several KEGG pathways. Gene function of bacteria were predicted using the PICRUST from KEGG annotated databases. Genus and KEGG pathways counts were normalized using edgeR. Spearman's correlation coefficients were estimated for pairwise comparison between genus abundance and KEGG abundance pathways.

associated with lipid metabolism in the present study (**Figure 8**, *Catellicoccus*).

Proteobacteria relative abundance in the present study was higher than those in most wild bird species and domestic chicken (Grond et al., 2018). Furthermore, at the class level, Gammaproteobacteria (mainly Enterobacteriaceae) as opposed to alphaproteobacteria were the dominant Proteobacteria taxa in the course of the growth of the young crested ibis, when compared with the dominant bacteria, alphaproteobacterial, in other wild birds (Grond et al., 2018). In particular, Escherichia or Shigella, which are opportunistic pathogens (Strockbine et al., 2015), were the dominant genera within Gammaproteobacteria. However, their high relative abundances in young crested ibis potentially indicate the presence of unknown non-pathogenic

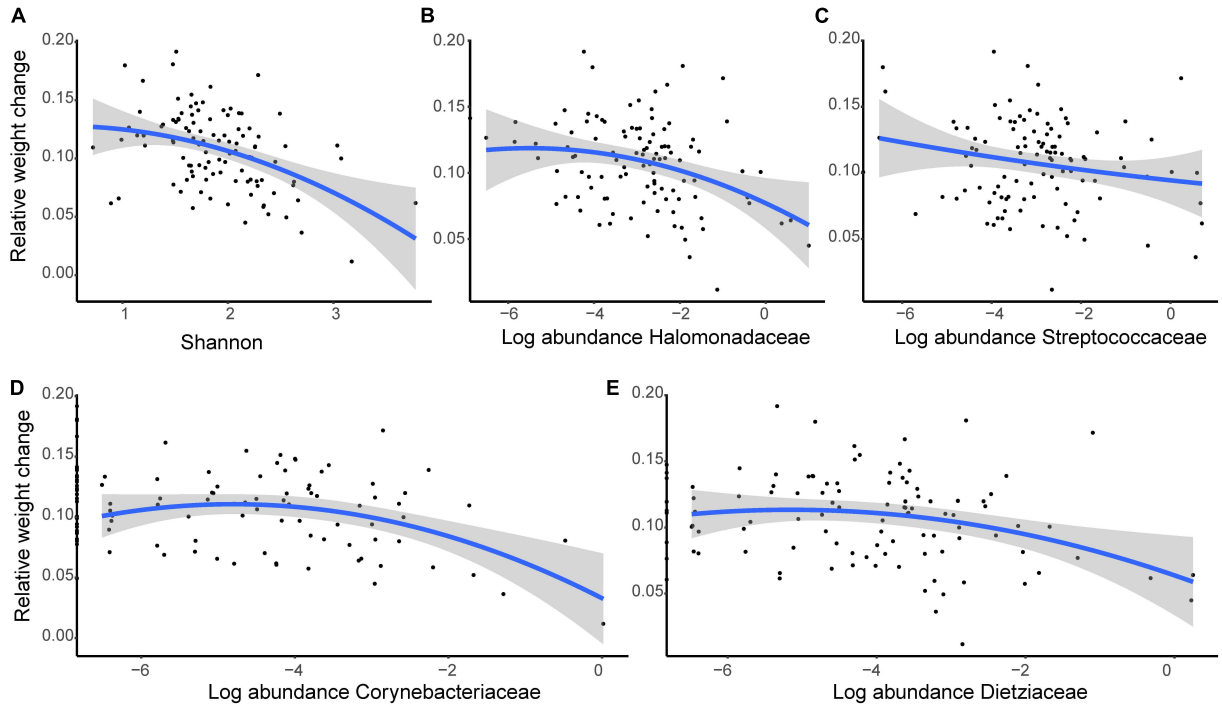

FIGURE 8 | The growth rate of young crested ibis is correlated with gut microbiota. **(A)** Shannon index. The normalized abundance of panel **(B)** Halomonadaceae, **(C)** Streptococcaceae, **(D)** Corynebacteriaceae, and **(E)** Dietziaceae. Relative weight change is measured as the weight change per day (W_{t+1}–W_t) against weight at time *t* (W_t). Age and diet type at time *t* were considered a covariate, and individual ID was controlled as a random factor. The lines represent linear regression lines, while the shaded areas show the 95% confidence interval.

functions, such as a key role of diet as a microbial inoculum or involvement in gut immunity and the digestive system.

Notably, in the present study, Fusobacteria (*Cetobacterium* and other unassigned genera; **Figure 6**) were more dominant from 27 days and enriched in diet type 2 when fresh loach was introduced to chick formula, however, Fusobacteria increased significantly in the Deqing population (another crested ibis's breeding center in China) from 9 days or earlier, as observed by Ran et al. (2021). Fusobacteria colonization was potentially due to the ingestion of fresh loach, since the time of Fusobacteria colonization is consistent with the addition of fresh loach based on the results of two studies on crested ibis (addition of fresh loach began on day 5 in the Deqing population, personal communication with Ran Jian). The prevalence of Fusobacteria in the gut has been observed in other carnivorous birds (Waite and Taylor, 2015), which suggests its involvement in mutualism between microbiota and hosts, excluding its pathogenic role. One of the Fusobacteria genera, *Cetobacterium*, was also predicted to participate in lipid metabolism in the present study (**Figure 8**), however, its functions in other avian species require further research. The taxonomic shifts in Firmicutes, Proteobacteria, and Fusobacteria along with the age observed in the present study suggest selective recruitment of specific gut communities by hosts.

Mature gut microbiota showed strong temporal structure and two specific developmental stages. The beginning of the latter stage coincides with the introduction of fresh loach, an increase in the relative abundance of Fusobacteria, and several groups of Firmicutes that both could be involved in lipid metabolism. Furthermore, samples within similar age groups tended to be much more similar than to other age groups with an increase in age, especially from 27 days, which suggests that diet containing fresh loach makes gut community within age groups to converge. Our findings showed that diet and age (along with morphological modifications and immune maturation, **Figure 1**; Caviedes-Vidal and Karasov, 2001; Killpack et al., 2013) shape gut microbiota during the development of young crested ibis.

When the diet was shifted from diet type 1 to diet type 2 (introduction of fresh loach), more KEGG pathways were identified with more functional including abundant immune, digestive, xenobiotic biodegradation and metabolism, biosynthesis of other secondary metabolites featured in diet type 1 relative to more lip metabolism featured in diet type 2. The higher levels of lipid-metabolism functions in diet type 2 could be attributed to fresh loach requiring more lipid metabolism energy, which may be produced by fresh loach than loach paste (**Figure 8**). Notably, immune function was enhanced as early as day 3, with more abundant RIG-I-like receptor signaling pathway, which are regulated by Proteobacteria genera such as *Enterobacter*, which suggests that the gut immune system at developed at very early stages of development.

Consistent with previous findings (Gaskins et al., 2002; Dibner and Richards, 2005), in the present study, animal-growth rate was

negatively associated with gut bacteria diversity. Four specific taxa (Halomonadaceae, Streptococcaceae, Corynebacteriaceae, and Dietziaceae families from Gammaproteobacteria, Firmicutes, and Actinobacteria phyla) also showed negative effects on the growth rate. A previous studies reported that Streptococcaceae negatively affects ostrich growth (Videvall et al., 2019), however, whether such convergence exists in other wild birds remains to be determined. Additionally, Streptococcaceae and Halomonadaceae are associated with obesity (Garcia-Mantrana et al., 2018) or a high-fat diet (Ijaz et al., 2020), inflammation (Zeng et al., 2016), and other diseases (Chen et al., 2011) in rodents and humans, however, few studies exist on the function of two Actinobacteria families (Corynebacteriaceae and Dietziaceae). These findings suggest conserved interactions between hosts and vertebrate gut microbiota.

Sampling at crested ibis hatching already revealed diverse gut microbiota (**Figure 2A**), which are consistent with previous observations (Ran et al., 2021). According to the results, crested ibis could acquire microbiota when inside the egg and bacterial colonization occurred before hatching. Considering the eggs in the present study were hatched artificially without mother crested ibis indicate that the microbes may enter the GI of crested ibis embryos via penetration through eggshell pores and embryonic membranes after laying (Gantois et al., 2009; Martelli and Davies, 2012). The colonization of gut microbiota before hatching might result in immunological and metabolic advantages for young crested ibis (Grond et al., 2017). Initial gut microbiota colonization took place earlier than the crested ibis could hatch, hence research on crested ibis embryos is warranted.

In conclusion, the present study showed that gut microbiota diversity displayed non-linear changes during the early stages of development of crested ibis, with multiple shifts occurring mainly in Proteobacteria and Firmicutes. The study also provides evidence that both diet and age influence microbiota structure. Microbiota changes observed were correlated with host growth and could influence host fitness over the long term. Our findings and those of studies on humans (de Muinck and Trosvik, 2018) both highlight the importance of frequent sampling strategies, when studying microbiome development during the early stages of development of vertebrates. Gut microbiota diversity could increase after day 44 when young birds move outdoors and are exposed to new diverse environments. Specially, Fusobacteria abundance significantly increases as crested ibis are fed with fresh loach. Further studies on the establishment of gut microbiota after 6 weeks to 1 year would better explain gut microbiota convergence in crested ibis.

DATA AVAILABILITY STATEMENT

The raw sequence data reported in this paper have been deposited in the Genome Sequence Archive (Genomics, Proteomics, and Bioinformatics 2017) in National Genomics Data Center (Nucleic Acids Research 2021), China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences, under accession number CRA004274 that are publicly accessible at https://bigd.big.ac.cn/gsa.

ETHICS STATEMENT

The animal study was reviewed and approved by Decision on Animal Ethics from Sichuan Provincial Academy of Natural Resource Sciences.

AUTHOR CONTRIBUTIONS

HQY and YDL collected the sample. YZ conducted the experiment, data analysis, and drafted the manuscript. YZ and KYT conceived of the idea and designed the study. KYT and KH participated in drafting the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Fundamental Research Funds for the Central Universities of China (3300221466 and 16011211069) and by the Program of the National Natural Science Foundation of China (32001223).

ACKNOWLEDGMENTS

We thank for Qiaoyang Chen and Yanqing Xu for helping the sample collection.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.723682/full#supplementary-material

REFERENCES

Angelakis, E., and Raoult, D. (2010). The increase of Lactobacillus species in the gut flora of newborn broiler chicks and ducks is associated with weight gain. *PLoS One* 5:e10463. doi: 10.1371/journal.pone.0010463

Ballou, A. L., Ali, R. A., Mendoza, M. A., Ellis, J., Hassan, H. M., Croom, W., et al. (2016). Development of the chick microbiome: how early exposure influences future microbial diversity. *Front. Vet. Sci.* 3:2. doi: 10.3389/fvets.2016.00002

Barbosa, A., Balagué, V., Valera, F., Martínez, A., Benzal, J., Motas, M., et al. (2016). Age-related differences in the gastrointestinal microbiota of chinstrap penguins (Pygoscelis antarctica). *PLoS One* 11:e0153215. doi: 10.1371/journal.pone.0153215

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B Stat. Methodol.* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x

Blumstein, D. T., Rangchi, T. N., Briggs, T., De Andrade, F. S., and Natterson-Horowitz, B. (2017). A systematic review of Carrion Eaters' adaptations to avoid sickness. *J. Wildl. Dis.* 53, 577–581. doi: 10.7589/2016-07-162

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016). DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. Methods* 13, 581–583. doi: 10.1038/nmeth.3869

Capunitan, D. C., Johnson, O., Terrill, R. S., and Hird, S. M. (2020). Evolutionary signal in the gut microbiomes of 74 bird species from Equatorial Guinea. *Mol. Ecol.* 29, 829–847. doi: 10.1111/mec.15354

Caviedes-Vidal, E., and Karasov, W. H. (2001). Developmental changes in digestive physiology of nestling house sparrows, *Passer domesticus*. *Physiol. Biochem. Zool.* 74, 769–782. doi: 10.1086/322966

Chen, Y., Yang, F., Lu, H., Wang, B., Chen, Y., Lei, D., et al. (2011). Characterization of fecal microbial communities in patients with liver cirrhosis. *Hepatology* 54, 562–572. doi: 10.1002/hep.24423

Clemente, J. C., Ursell, L. K., Parfrey, L. W., and Knight, R. (2012). The impact of the gut microbiota on human health: an integrative view. *Cell* 148, 1258–1270. doi: 10.1016/j.cell.2012.01.035

de Muinck, E. J., and Trosvik, P. (2018). Individuality and convergence of the infant gut microbiota during the first year of life. *Nat. Commun.* 9, 1–8.

Den Besten, G., Van Eunen, K., Groen, A. K., Venema, K., Reijngoud, D.-J., and Bakker, B. M. (2013). The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. *J. Lipid Res.* 54, 2325–2340. doi: 10.1194/jlr.r036012

Dewar, M. L., Arnould, J. P., Allnutt, T. R., Crowley, T., Krause, L., Reynolds, J., et al. (2017). Microbiota of little penguins and short-tailed shearwaters during development. *PLoS One* 12:e0183117. doi: 10.1371/journal.pone.0183117

Dibner, J., and Richards, J. (2005). Antibiotic growth promoters in agriculture: history and mode of action. *Poult. Sci.* 84, 634–643. doi: 10.1093/ps/84.4.634

Ding, C. Q. (2004). *Research on the Crested Ibis*. Shanghai: Shanghai Science and Technology Education Publishing House.

Douglas, G. M., Beiko, R. G., and Langille, M. G. (2018). "Predicting the functional potential of the microbiome from marker genes using PICRUSt," in *Microbiome Analysis*, Vol. 1849 eds R. Beiko, W. Hsiao, and J. Parkinson. (New York, NY: Humana Press).

Forbes, M., and Park, J. T. (1959). Growth of germ-free and conventional chicks: effect of diet, dietary penicillin and bacterial environment. *J. Nutr.* 67, 69–84. doi: 10.1093/jn/67.1.69

Gantois, I., Ducatelle, R., Pasmans, F., Haesebrouck, F., Gast, R., Humphrey, T. J., et al. (2009). Mechanisms of egg contamination by *Salmonella* Enteritidis. *FEMS Microbiol. Rev.* 33, 718–738.

Garcia-Mantrana, I., Selma-Royo, M., Alcantara, C., and Collado, M. C. (2018). Shifts on gut microbiota associated to mediterranean diet adherence and specific dietary intakes on general adult population. *Front. Microbiol.* 9:890. doi: 10.3389/fmicb.2018.00890

Gaskins, H., Collier, C., and Anderson, D. (2002). Antibiotics as growth promotants: mode of action. *Anim. Biotechnol.* 13, 29–42.

Godoy-Vitorino, F., Goldfarb, K. C., Brodie, E. L., Garcia-Amado, M. A., Michelangeli, F., and Domínguez-Bello, M. G. (2010). Developmental microbial ecology of the crop of the folivorous hoatzin. *ISME J.* 4, 611–620. doi: 10.1038/ismej.2009.147

Góngora, E., Elliott, K. H., and Whyte, L. (2021). Gut microbiome is affected by inter-sexual and inter-seasonal variation in diet for thick-billed murres (Uria lomvia). *Sci. Rep.* 11:1200.

Grond, K., Lanctot, R. B., Jumpponen, A., and Sandercock, B. K. (2017). Recruitment and establishment of the gut microbiome in arctic shorebirds. *FEMS Microbiol. Ecol.* 93:fix142.

Grond, K., Sandercock, B. K., Jumpponen, A., and Zeglin, L. H. (2018). The avian gut microbiota: community, physiology and function in wild birds. *J. Avian Biol.* 49:e01788. doi: 10.1111/jav.01788

Hamdi, C., Balloi, A., Essanaa, J., Crotti, E., Gonella, E., Raddadi, N., et al. (2011). Gut microbiome dysbiosis and honeybee health. *J. Appl. Entomol.* 135, 524–533. doi: 10.1111/j.1439-0418.2010.01609.x

Harrell, F. E. Jr. (2021). *Hmisc: Harrell Miscellaneous*. Available online at: https://CRAN.R-project.org/package=Hmisc

He, X. L., Qing, B. P., Han, J. L., and Ding, C. Q. (2013). Improved molecular assay for sex identification of the endangered Crested Ibis (Nipponia nippon) based on the CHD1 gene and a sex-linked microsatellite locus. *Zool. Sci.* 30, 742–747. doi: 10.2108/zsj.30.742

Hird, S. M., Carstens, B. C., Cardiff, S., Dittmann, D. L., and Brumfield, R. T. (2014). Sampling locality is more detectable than taxonomy or ecology in the gut microbiota of the brood-parasitic Brown-headed Cowbird (Molothrus ater). *Peerj* 2:e321. doi: 10.7717/peerj.321

Huang, T., Gao, B., Chen, W. L., Xiang, R., Yuan, M. G., Xu, Z. H., et al. (2018). Temporal effects of high fishmeal diet on gut microbiota and immune response in Clostridium perfringens-challenged chickens. *Front. Microbiol.* 9:2754. doi: 10.3389/fmicb.2018.02754

Ijaz, M. U., Ahmad, M. I., Hussain, M., Khan, I. A., Zhao, D., and Li, C. (2020). Meat protein in high-fat diet induces adipogensis and dyslipidemia by altering gut microbiota and endocannabinoid dysregulation in the adipose tissue of mice. *J. Agric. Food Chem.* 68, 3933–3946. doi: 10.1021/acs.jafc.0c00017

John, G. K., and Mullin, G. E. (2016). The gut microbiome and obesity. *Curr. Oncol. Rep.* 18, 1–7.

Killpack, T. L., Oguchi, Y., and Karasov, W. H. (2013). Ontogenetic patterns of constitutive immune parameters in altricial house sparrows. *J. Avian Biol.* 44, 513–520. doi: 10.1111/j.1600-048x.2013.00239.x

Kohl, K. D., Brun, A., Bordenstein, S. R., Caviedes-vidal, E., and Karasov, W. H. (2018). Gut microbes limit growth in house sparrow nestlings (*Passer domesticus*) but not through limitations in digestive capacity. *Integr. Zool.* 13, 139–151. doi: 10.1111/1749-4877.12289

Kohl, K. D., Brun, A., Caviedes-Vidal, E., and Karasov, W. H. (2019). Age-related changes in the gut microbiota of wild House Sparrow nestlings. *Ibis* 161, 184–191. doi: 10.1111/ibi.12618

Kolde, R. (2019). *pheatmap: Pretty Heatmaps*. Available online at: https://CRAN.R-project.org/package=pheatmap

Lewis, J. D., Chen, E. Z., Baldassano, R. N., Otley, A. R., Griffiths, A. M., Lee, D., et al. (2015). Inflammation, antibiotics, and diet as environmental stressors of the gut microbiome in pediatric Crohn's disease. *Cell Host Microbe* 18, 489–500. doi: 10.1016/j.chom.2015.09.008

Lewis, W. B., Moore, F. R., and Wang, S. A. (2016). Characterization of the gut microbiota of migratory passerines during stopover along the northern coast of the Gulf of Mexico. *J. Avian Biol.* 47, 659–668. doi: 10.1111/jav.00954

Liao, X., Ma, G., Cai, J., Fu, Y., Yan, X., Wei, X., et al. (2015). Effects of Clostridium butyricum on growth performance, antioxidation, and immune function of broilers. *Poult. Sci.* 94, 662–667. doi: 10.3382/ps/pev038

Liu, Y. X., Qin, Y., Chen, T., Lu, M. P., Qian, X. B., Guo, X. X., et al. (2021). A practical guide to amplicon and metagenomic analysis of microbiome data. *Protein Cell* 12, 315–330. doi: 10.1007/s13238-020-00724-8

Lopetuso, L. R., Scaldaferri, F., Petito, V., and Gasbarrini, A. (2013). Commensal Clostridia: leading players in the maintenance of gut homeostasis. *Gut Pathog.* 5:23. doi: 10.1186/1757-4749-5-23

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

Magrath, R. D. (1991). Nestling weight and juvenile survival in the blackbird, Turdus merula. *J. Anim. Ecol.* 60, 335–351. doi: 10.2307/5464

Martelli, F., and Davies, R. H. (2012). *Salmonella* serovars isolated from table eggs: an overview. *Food Res. Int.* 45, 745–754. doi: 10.1016/j.foodres.2011.03.054

Oakley, B. B., Buhr, R. J., Ritz, C. W., Kiepper, B. H., Berrang, M. E., Seal, B. S., et al. (2014). Successional changes in the chicken cecal microbiome during 42 days of growth are independent of organic acid feed additives. *BMC Vet. Res.* 10:282. doi: 10.1186/s12917-014-0282-8

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2020). *vegan: Community Ecology Package*. Available online at: https://CRAN.R-project.org/package=vegan

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., and R Core Team. (2021). *nlme: Linear and Nonlinear Mixed Effects Models*. Available online at: https://CRAN.R-project.org/package=nlme

Potti, J., Moreno, J., Yorio, P., Briones, V., García-Borboroglu, P., Villar, S., et al. (2002). Bacteria divert resources from growth for magellanic penguin chicks. *Ecol. Lett.* 5, 709–714. doi: 10.1046/j.1461-0248.2002.00375.x

Ran, J., Wan, Q. H., and Fang, S. G. (2021). Gut microbiota of endangered crested ibis: Establishment, diversity, and association with reproductive output. *PLoS One* 16:e0250075. doi: 10.1371/journal.pone.0250075

Reid, G., Younes, J. A., Van der Mei, H. C., Gloor, G. B., Knight, R., and Busscher, H. J. (2011). Microbiota restoration: natural and supplemented recovery of human microbial communities. *Nat. Rev. Microbiol.* 9, 27–38. doi: 10.1038/nrmicro2473

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616

Rosengaus, R. B., Zecher, C. N., Schultheis, K. F., Brucker, R. M., and Bordenstein, S. R. (2011). Disruption of the termite gut microbiota and its prolonged consequences for fitness. *Appl. Environ. Microbiol.* 77, 4303–4312. doi: 10.1128/aem.01886-10

Strockbine, N. A., Bopp, C. A., Fields, P. I., Kaper, J. B., and Nataro, J. P. (2015). "Escherichia, *Shigella*, and *Salmonella*," in *Manual of Clinical Microbiology*, eds J. H. Jorgensen, K. C. Carroll, G. Funke, M. A. Pfaller, M. L. Landry, S. S. Richter et al. (Hoboken, NJ: Wiley), 685–713.

Teyssier, A., Lens, L., Matthysen, E., and White, J. (2018). Dynamics of gut microbiota diversity during the early development of an avian host: evidence from a cross-foster experiment. *Front. Microbiol.* 9:1524. doi: 10.3389/fmicb.2018.01524

Torok, V. A., Hughes, R. J., Mikkelsen, L. L., Perez-Maldonado, R., Balding, K., MacAlpine, R., et al. (2011). Identification and characterization of potential performance-related gut microbiotas in broiler chickens across various feeding trials. *Appl. Environ. Microbiol.* 77, 5868–5878. doi: 10.1128/aem.00165-11

van Dongen, W. F., White, J., Brandl, H. B., Moodley, Y., Merkling, T., Leclaire, S., et al. (2013). Age-related differences in the cloacal microbiota of a wild bird species. *BMC Ecol.* 13:11. doi: 10.1186/1472-6785-13-11

Videvall, E., Song, S. J., Bensch, H. M., Strandh, M., Engelbrecht, A., Serfontein, N., et al. (2019). Major shifts in gut microbiota during development and its relationship to growth in ostriches. *Mol. Ecol.* 28, 2653–2667. doi: 10.1111/mec.15087

Waite, D. W., and Taylor, M. W. (2015). Exploring the avian gut microbiota: current trends and future directions. *Front. Microbiol.* 6:673. doi: 10.3389/fmicb.2015.00673

Wickham, H. (2011). ggplot2. *Wiley Interdiscip. Rev. Comput. Stat.* 3, 180–185. doi: 10.1002/wics.147

Wienemann, T., Schmitt-Wagner, D., Meuser, K., Segelbacher, G., Schink, B., Brune, A., et al. (2011). The bacterial microbiota in the ceca of Capercaillie (Tetrao urogallus) differs between wild and captive birds. *Syst. Appl. Microbiol.* 34, 542–551. doi: 10.1016/j.syapm.2011.06.003

Wilkinson, T. J., Cowan, A., Vallin, H., Onime, L., Oyama, L. B., Cameron, S., et al. (2017). Characterization of the microbiome along the gastrointestinal tract of growing turkeys. *Front. Microbiol.* 8:1089. doi: 10.3389/fmicb.2017.01089

Xenoulis, P. G., Gray, P. L., Brightsmith, D., Palculict, B., Hoppes, S., Steiner, J. M., et al. (2010). Molecular characterization of the cloacal microbiota of wild and captive parrots. *Vet. Microbiol.* 146, 320–325. doi: 10.1016/j.vetmic.2010.05.024

Yin, Y., Lei, F., Zhu, L., Li, S., Wu, Z., Zhang, R., et al. (2010). Exposure of different bacterial inocula to newborn chicken affects gut microbiota development and ileum gene expression. *ISME J.* 4, 367–376. doi: 10.1038/ismej.2009.128

Zeng, H., Ishaq, S. L., Zhao, F.-Q., and Wright, A.-D. G. (2016). Colonic inflammation accompanies an increase of β-catenin signaling and Lachnospiraceae/Streptococcaceae bacteria in the hind gut of high-fat diet-fed mice. *J. Nutr. Biochem.* 35, 30–36. doi: 10.1016/j.jnutbio.2016.05.015

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.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 08 September 2021
doi: 10.3389/fmicb.2021.739577

Effects of Subchronic Copper Poisoning on Cecal Histology and Its Microflora in Chickens

Cheng Huang[1†], Yan Shi[2†], Changming Zhou[1], Lianying Guo[1], Guohui Liu[3], Yu Zhuang[1], Guyue Li[1], Guoliang Hu[1], Ping Liu[1] and Xiaoquan Guo[1]**

[1] Jiangxi Provincial Key Laboratory for Animal Health, College of Animal Science and Technology, Jiangxi Agricultural University, Nanchang, China, [2] School of Computer and Information Engineering, Jiangxi Agricultural University, Nanchang, China, [3] Animal Husbandry and Veterinary Department of Ganzhou, Ganzhou, China

OPEN ACCESS

Edited by:
Wei Huang,
Johns Hopkins University,
United States

Reviewed by:
Mingwei Xing,
Northeast Forestry University, China
Aleksandr G. Bulaev,
Federal Center Research Fundamentals of Biotechnology (RAS), Russia

***Correspondence:**
Ping Liu
Pingliujx@163.com
Xiaoquan Guo
xqguo20720@jxau.edu.cn

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *11 July 2021*
Accepted: *09 August 2021*
Published: *08 September 2021*

Citation:
Huang C, Shi Y, Zhou C, Guo L, Liu G, Zhuang Y, Li G, Hu G, Liu P and Guo X (2021) Effects of Subchronic Copper Poisoning on Cecal Histology and Its Microflora in Chickens. Front. Microbiol. 12:739577. doi: 10.3389/fmicb.2021.739577

Copper (Cu) is an important trace element with a two-sided effect on the growth performance of animals, which depends on the timing and dosage of Cu addition, etc. The purpose of this study was to determine the effects of oral copper sulfate ($CuSO_4$, 350 ppm) on growth performance, cecal morphology, and its microflora of chickens (n = 60) after 30, 60, and 90 days. The results showed that after 90 days of copper exposure, the chickens lost weight, the cecum mucosa was detached, and vacuolation and inflammatory infiltration occurred at the base of the lamina propria. In addition, using the 16S rDNA sequencing method, we observed that copper exposure changed the richness and diversity of intestinal microorganisms. At the phylum level, *Proteobacteria* and *Actinobacteria* both significantly increased, while *Bacteroidetes* significantly decreased in the Cu group compared with control check (CK) group. At the genus level, the relative abundance of *Rikenellaceae_RC9_gut_group* decreased significantly, while *Ruminococcaceae_UCG-014, Lachnoclostridium*, and *[Eubacterium]_coprostanoligenes_group* increased significantly after copper exposure, and the change in microflora was most significant at 90 days. Moreover, the relevance of genus-level bacteria was altered. PICRUST analysis revealed potential metabolic changes associated with copper exposure, such as *Staphylococcus aureus* infection and metabolic disorders of nutrients. To sum up, these data show that subchronic copper exposure not only affects the growth and development of chickens but also causes the imbalance of intestinal microflora, which may further induce metabolic disorders in chickens.

Keywords: Heavy metal pollution, copper, chicken, gut microbiota, 16S rDNA, metabolic disorders

INTRODUCTION

Copper (Cu) is an essential trace element, which plays a key role in maintaining the normal growth and development of animals (Frieden, 1986). Cu can play a catalytic role as an important cofactor of various biological enzymes involved in immune function, protein synthesis, antioxidant defense, and energy metabolism. So adding high copper to diets has become a common practice to improve the production performance of livestock and poultry (Son et al., 2016; Berwanger et al., 2018).

However, it has also been shown that Cu has the ability to inhibit immune response and increase the susceptibility of the host to pathogens (Zhao et al., 2018; Ruan et al., 2019). More importantly, early studies have shown that excessive copper intake may lead to the accumulation of copper in tissues, leading to organ and cytotoxic damage in animals and even humans, as well as intestinal damage (Pal, 2014). However, in order to improve production efficiency, copper is still being abused because it is often used as a promoter to inhibit fungal growth, bacterial and parasitic infections. Hence, people mostly ignore the toxic effect of excessive copper on the body. Additionally, with the development of industry, surface water and soil are also polluted by copper, which increases the risk of subchronic copper poisoning (Shahzad et al., 2012; Bui et al., 2016).

At present, more and more attention has been paid to the research on the effect of excessive copper intake on the body, and previous studies are mainly related to the following aspects. For instance, it was found that the addition of 796 mg/kg Cu to the diet of laying hens for 2 weeks would reduce the food intake, body weight, and egg yield of laying hens (Jackson et al., 1979). A previous study has also found that adding 50 mg/0.1 ml of copper sulfate to chicken embryos by injection can cause liver tissue necrosis and oxidative damage (Oguz et al., 2010). Copper exposure caused pathological damage to the liver and kidney of chickens, as well as changes in organ weight (Shahzad et al., 2012). Recently, some previous studies have reported that excessive copper intake increased the level of autophagy and apoptosis in chicken (Zhao et al., 2015, 2018; Fei et al., 2019). Another study suggested that subchronic copper exposure induced oxidative stress, resulting in mitochondrial disorder and jejunal toxicity increase due to redistribution of trace elements in the jejunum of chicken (Zhao et al., 2018). However, although there is a growing evidence reporting toxicity of heavy metal in chickens, very limited studies have been conducted to directly link toxicity to the changes in the gut microbiota.

The gastrointestinal tract (GIT) of chickens, especially the cecum, has a complex microflora. Intestinal microorganisms are closely related to the metabolism of the host, play a fundamental role in the health of the host, and benefit the host from many aspects, especially nutrition and disease resistance (Jenkins et al., 2019; Macchione et al., 2019). Although the composition of gut microbiota is relatively stable as the animal grows, it could be influenced by various factors, such as food, antibiotics, drugs, and even different environmental chemicals (Patterson and Burkholder, 2003; Gong et al., 2008). The evidence, to date, shows that a disruption of the normal intestinal microbiota in the host has been linked to disruption of physiological, metabolic homeostasis and promoting the development of various diseases to some extent, such as allergies, necrotizing enteritis, diabetes, depression, and cancer (Zhou C. et al., 2019). At present, studies have focused on the effects of copper exposure on intestinal microorganisms in mice (Ruan et al., 2019; Cheng et al., 2020). However, a few studies have explored the interactions between copper and microbiota in chicken, especially in the face of continuous excessive copper uptake. Therefore, it is crucial to research and understand the dynamic changes of intestinal microbial community in chicken during subchronic copper exposure.

Microbial 16S rDNA sequencing technology is an important means to study intestinal flora at the present stage (Jovel et al., 2016). Based on the technical progress of next-generation sequencing (NGS), it provides unparalleled coverage and depth in determining the intestinal dynamics of microorganisms (Mohd Shaufi et al., 2015). In the present study, we used this technique in assessing changes in intestinal bacterial communities induced by Cu exposure in chicken. The body weight and cecal morphology of chickens during subchronic copper poisoning were also observed and compared. We believed that the results obtained here can provide some new information for the toxicity of chicken caused by subchronic copper poisoning. Taken together, the objective of this study was to explore the toxic effects of subchronic copper poisoning on the intestinal tract of chickens.

MATERIALS AND METHODS

Experimental Design and Animal Model

All the experimental schemes and methods used in this study were approved by the Animal Protection Agency and the use Committee (Jiangxi Agricultural University, Nanchang, Jiangxi, China), license No. JXAULL-2016038. The animals involved in the study were treated humanely, and steps were taken to minimize their suffering. One hundred and twenty Hy-line male chickens 1-day old (purchased from China Nanchang Guohua Co., Ltd.) were randomly divided into two groups (60 chickens in each group) after 1 week of breeding, including control group (CK) and the $CuSO_4$-treated group (Cu). According to the test standard of subchronic poisoning in the Organization for Economic Co-operation and Development (OECD) guidelines and referring to the previous studies on the dose of copper poisoning in chickens, we set up the control group to drink tap water without $CuSO_4$, the Cu group only drank $CuSO_4$ working solution containing 350 ppm, and the whole experiment lasted 90 days (Council, 1995; Buschmann, 2013; Zhao et al., 2018). All chickens were free to drink feed and water. They received 12 h of light and 12 h of darkness every day. In addition, chickens were immunized according to routine immunization procedures.

On the 30th, 60th, and 90th day of the experiment, six chickens in each group were randomly weighed and euthanized with pentobarbital sodium. The cecal contents were quickly placed in liquid nitrogen and then stored at −80°C for experiment. At the same time, the cecal tissues were preserved in formalin for histopathological observation.

Histopathological Analysis

The cecal tissues (∼0.3 cm^3) were fixed in 4% paraformaldehyde, dehydrated in graded ethanol (70, 80, 90, and 100%) and xylene, and was embedded in paraffin. The tissue was circumcised to 5-μm thickness and stained with hematoxylin and eosin. Stained sections of feces were examined with a microscope (Olympus DP73, Japan) for morphological and structural changes in the mucosa.

DNA Extraction and Quantitative Polymerase Chain Reaction Amplification

Total microbial genomic DNA from chicken feces was extracted using a combined cetyltrimethylammonium bromide/sodium dodecyl sulfate (CTAB/SDS) method including bacterial and fungal genomic DNA. Briefly, community DNA was extracted from a 0.25-g aliquot of each fecal sample. A 0.25-g aliquot of each fecal sample was extracted for community DNA. This DNA extraction method was employed: Two rounds of beating were performed in the presence of sodium chloride and sodium dodecyl sulfate, followed by sequential ammonium acetate and isopropanol precipitation. The precipitated nucleic acids were then treated with RNase A and Proteinase K, and the DNA was purified using columns of the QIAgen DNA Mini Stool Kit (QIAGEN, MD, United States). The DNA was then tested for purity and concentration by 1% agarose gel electrophoresis. An appropriate amount of DNA was taken in a centrifuge tube and diluted with sterile water to 1 ng/μl. Extracted metagenomic DNA was stored at −20°C.

The hypervariable region V3–V4 of 16S rDNA gene about 470 bp in length was sequenced. Polymerase chain reaction (PCR) was performed using specific primers 338F (5-GCACCTACTCCTACGGGAGCAGCAGCAGCA-3′) and 806R (5-GGACTACHVGGGGTWTCTAAT-3′). Polymerase chain reaction was performed with Q5 high-fidelity DNA polymerase at a final volume of 25 μl, and the template DNA was 20 ng. The thermal cycle included 98°C, 5 min (initial denaturation), 98°C, 10 s (denaturation), 50°C, 30 s (annealing), 72°C, 30 s (25 cycles), and 72°C, 5 min (final prolongation). The samples with the size of about 470 bp were selected by 2% agarose gel electrophoresis in 1.0 × TAE buffer, and then the amplified products were purified by AxyPrep DNA gel extraction kit (Axygen, AP-GX-250). The pyrosequencing of 16S rDNA was carried out on the Illumina HiSeq 2500 PE 250 platform (Illumina, San Diego, CA, United States) of Novogene Bioinformation Technology Co., Ltd. (Beijing, China).

Sequence Processing and Bioinformatics Analysis

We used FLASH software (v1.2.7) to merge into a pair of end read and splice sequences to generate the raw tags. Open-source software system QIIME analysis was used to obtain high-quality clean tags, and UCHIME algorithm was used to remove the chimera sequence to obtain effective tags. The sequences were analyzed by UCLUST software and clustered into operational taxonomic units (OTUs) with 97% similarity. Each OTU was annotated with the Greengenes database. The rarefaction curve and Venn diagram were created with R software (v2.15.3). Microbial alpha diversity was analyzed by QIIME software with Python script. Beta diversity was evaluated by principal component analysis (PCA) to show the difference of bacterial community structure, and R (v2.15.3) was used to separate the bacteria by ANOSIM. Correlation analysis of genus-level bacteria using SparCC network maps. Prediction of the functional potential of bacterial community was done by PICRUST analysis.

Data Presentation

The results were shown as mean and SEM. Significant difference was declared when $p < 0.05$. The 16S rDNA gene amplicon sequencing results were submitted to the Sequence Read Archive of the NCBI (accession number: SRP199017).

RESULTS

Body Weight and the Morphology of the Cecum in Chicken

The weight of chickens were recorded at 30, 60, and 90 days, respectively, in order to evaluate the effect of $CuSO_4$ intake on the growth performance of chickens. The results showed that compared with the CK group, the weight of the copper group decreased at all-time points, with significant difference at 60 days ($p = 0.013$) and 90 days ($p = 0.000$) (**Figure 1A**). The histological changes in the cecum in the CK group and Cu group at 30, 60, and 90 days are shown in **Figure 1B**. There were significant differences in the cecum tissue structure between the CK group and Cu group at different time points. In the whole process, no obvious lesion was found in the CK group, the cecal mucosa was intact, and the structure of the muscular layer was normal. Compared with the CK30 group, there was a small amount of mucosal exfoliation and vacuolation at the bottom of the lamina propria in the Cu30 group. Meanwhile, we found a large number of cecal mucosal exfoliation, lymphocyte proliferation, and vacuolation in the Cu60 group. On the 90th day, severe pathological changes occurred in the Cu group, with massive loss of cecal mucosa, inflammatory infiltration in the lamina propria, and destruction of the tissue structure of the muscular layer.

Quality of Sequencing Analysis

Cecal samples (36) were sequenced by Illumina HiSeq platform and detected by 16S rDNA gene (V3–V4 region). A total of 2,871,406 sequences were obtained from these samples, of which 2,742,829 sequences were effective sequences by quality control filtration. After quality was checked and chimeric sequences were removed, the average number of reads generated from the cecal samples per chicken was 77,863 ± 5,525 [standard deviation (SD)], with an average length of 253 bp, respectively. They were classified using QIIME into different OTUs based on the identity level at 97%, and a total of 1,632 OTUs were identified from all samples. Shannon value and rarefaction curves for each sample reached the saturation plateau (**Supplementary Table 1** and **Figure 2**), indicating that the samples had sufficient sequence coverage to accurately describe the bacterial composition of each group.

Statistical Analysis of Alpha and Beta Diversity

Both alpha and beta diversity metrics were used to estimate microbial community diversity. The alpha diversity reflected the

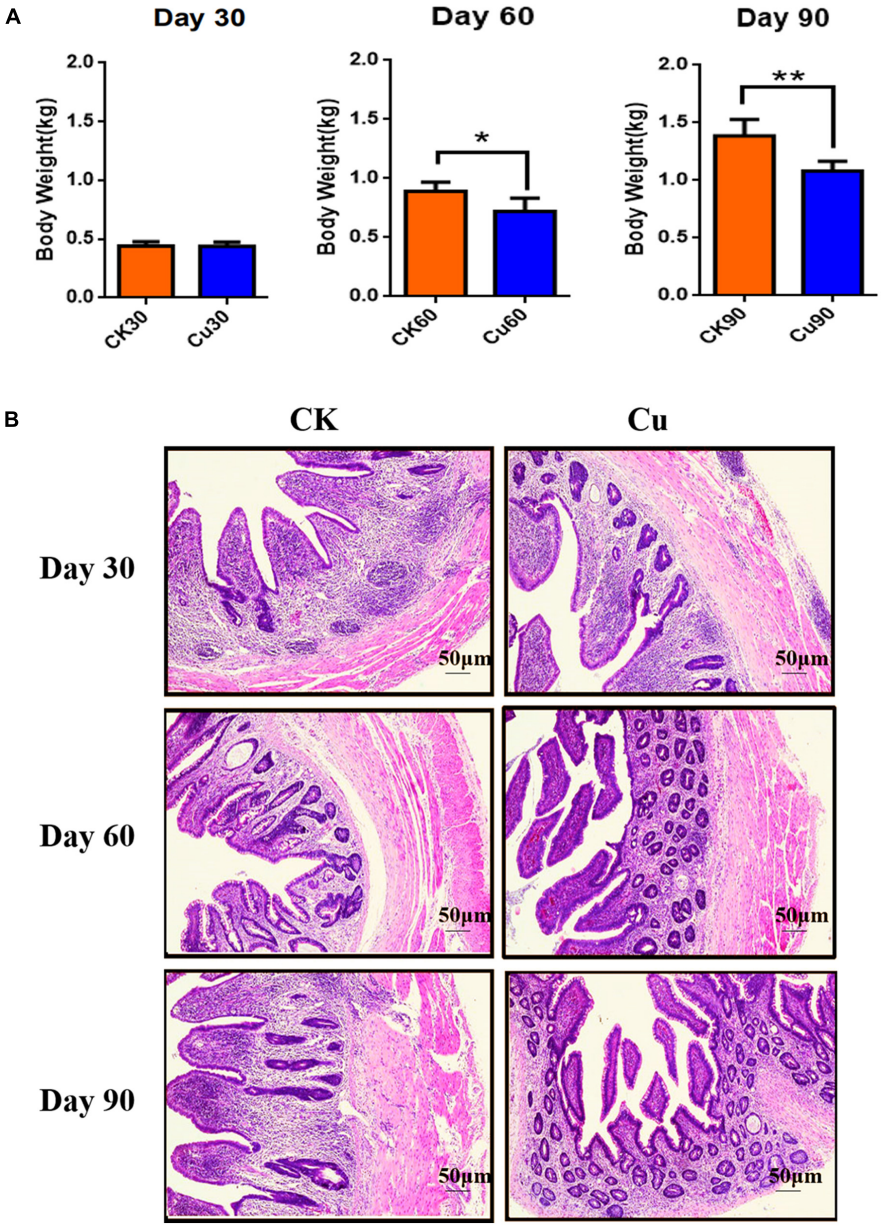

FIGURE 1 | Defines weight and tissue samples at a specified time. Control check (CK): chickens in the control group were fed with copper-free distilled water. Oral copper sulfate ($CuSO_4$)-treated group (Cu group): chickens drank $CuSO_4$ water every day for 90 days. Body weight changes of chickens in **(A)** CK group and Cu group. **(B)** Representative images of hematoxylin and eosin (H&E) staining (original magnification, ×100) of chickens cecal tissue at different time points. All results are expressed as mean ± SD of six chickens in each group. * and ** represent statistically significant at $p < 0.05$ and $p < 0.01$ levels.

comprehensive index of richness and evenness of cecal flora in the CK and Cu groups. Beta diversity reflects whether there are significant differences in microbial communities between groups. In alpha diversity, the Chao1 index and ACE index represent community richness index, and community diversity index includes the Shannon index and Simpson index. **Figure 3A** shows the community richness index of the CK group and Cu group at three sampling times by the Chao1 index. The results showed that compared with the CK group, the community richness of the Cu group increased significantly after adding $CuSO_4$ to drinking water for 30 days (Wilcox test, $p < 0.001$). After that, the community richness of the Cu group showed a downward trend, and at 60 days, compared with the CK group, the community richness of the Cu group decreased significantly ($p < 0.01$). In addition, compared with the CK group, the community richness of the Cu group increased significantly at 90 days ($p < 0.01$). Similarly, the Shannon index proved that the community diversity of the two groups had the same trend over time. This also confirms the results of previous studies, that is, heavy metal treatment can change the diversity

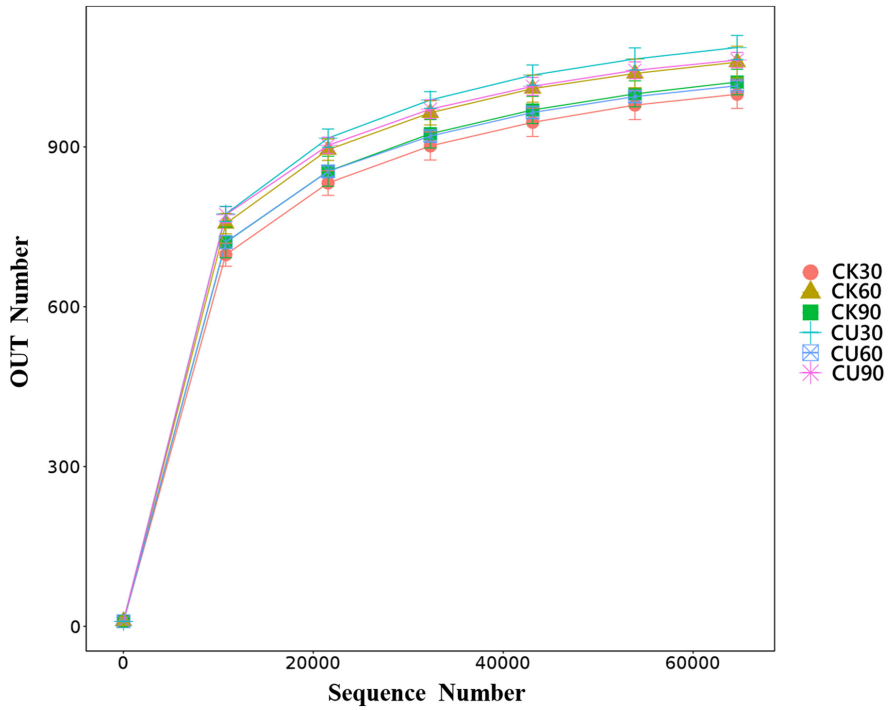

FIGURE 2 | Rarefaction curve. The reads with similarity of 97% is clustered into the same service taxon operation taxonomic unit (OTU). The number of sequences sampled represents the number of sequencing reads.

of intestinal microbiota (Liu et al., 2014; Li et al., 2019). Changes in the intestinal microbial diversity may indicate that microbial diversity is disturbed, and intestinal health status is also changed (Wu et al., 2014).

Beta diversity uses the evolutionary relationship and abundance information between groups to calculate the distance matrix to reflect the differences between groups. Next, non-metric multidimensional calibration method (NMDS) based on weighted Unifrac distances was used to further determine the role of Cu intake in drinking water in changing the distribution of intestinal flora profile. As shown in **Figure 3C**, the NMDS ordination stress value is 0.102, and the microbial community of the Cu group was clearly separated from the CK group. In addition, the UPGMA method based on weighted Uniface is used to cluster the CK and Cu groups at different time points. **Figure 3D** shows that the individual samples of the CK group and the members of the Cu group gather in their respective groups and are significantly closer to each other at three time points. According to the results of the current beta diversity index, we speculated that the composition of microbial community also changed due to the change in copper exposure and time.

Operational Taxonomic Unit Distribution From a Venn Plot

In order to find out the difference between intestinal microbial community during the experiment, Venn diagrams were generated for the CK and Cu groups according to the results of OTUs cluster analysis. The Venn diagrams showed that, in the CK group, there were 1,139 OTUs common to all time point samples, and only 41, 78, and 46 OTUs were identified in the CK30, CK60, and CK90 samples, respectively (**Figure 4A**). In the Cu group, there were 1,149 OTUs that were common to all time points samples, and only 78, 32, and 70 OTUs were identified in the Cu30, Cu60, and Cu90 samples, respectively (**Figure 4B**). During the whole 90-day experimental period, the number of unique OTUS in the control group increased gradually in the first 60 days, but decreased during the last 30 days. In contrast, the number of unique OTUS in the exposed group decreased gradually in the first 60 days, but increased during the last 30 days. These results are consistent with the findings of the Chao1 index. Therefore, through these evidences, we have concluded that Cu exposure might change the cecal microflora of chickens.

Composition Analysis of the Gut Microbiota at Various Taxonomic Levels

According to the results of species annotation, each group of bacterial communities ranked at the level of phylum and genus were selected to form a column accumulation map of relative abundance of species, in order to observe the effect of copper exposure on cecal microorganisms during the experiment. At the phylum level, 10 taxa with high abundance of each sample were counted (**Figure 5A**) (mean relative abundance $>$ 0.5%). *Bacteroidetes* and *Firmicutes* were found to be the most abundant phylum from each group in this study, with relative abundance of 44.75 and 42.40%, respectively (**Figure 5A**). While found in lower relative abundance, the phyla *Proteobacteria* (4.16%), *Tenericutes*

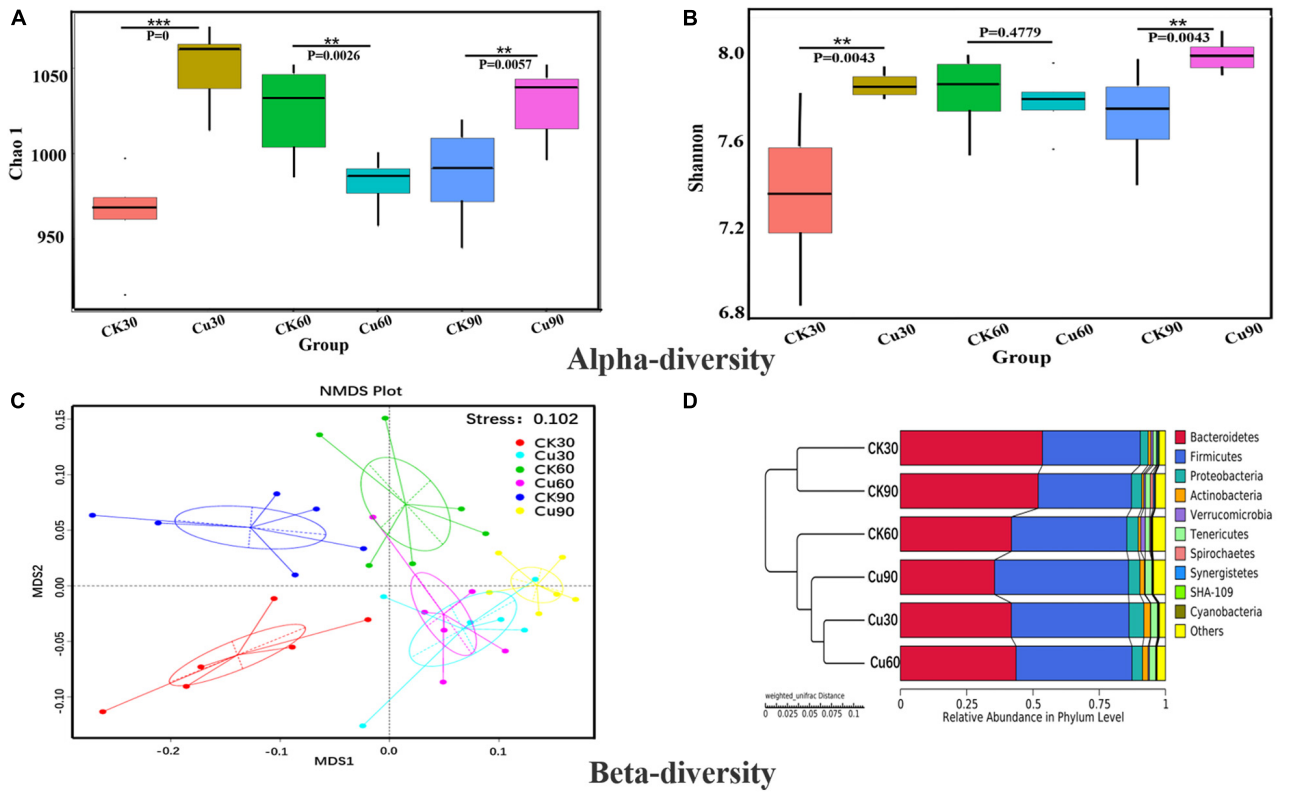

FIGURE 3 | Alpha and beta diversity of cecal bacterial communities between control and Cu treated chickens. **(A)** Difference in bacterial diversity between CK and Cu groups was estimated by the Chao1 estimator at each time-point. The results were compared using the Wilcox test, $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$. **(B)** Difference in bacterial richness between CK and Cu groups was estimated by the Shannon value at each time-point. The results were compared using the Wilcox test, $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$. **(C)** Non-metric multidimensional calibration method (NMDS) generated using the weighted Uniface distance. **(D)** Cluster analysis performed on weighted Uniface distance matrices using the unweighted-pair group method with arithmetic mean (UPGMA).

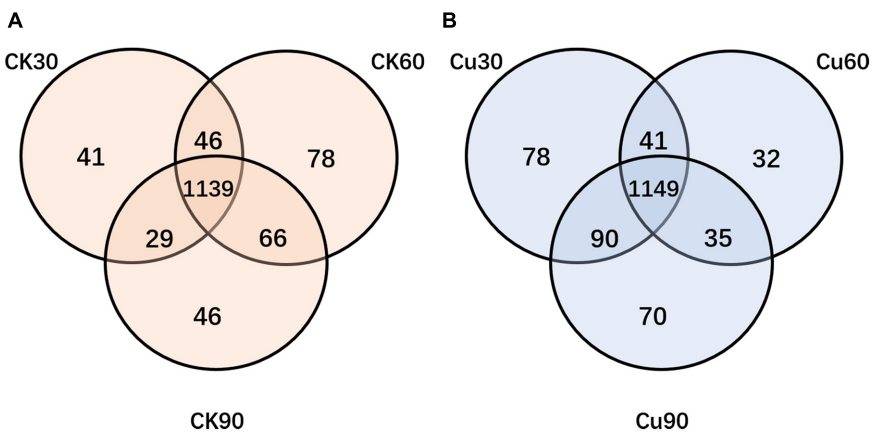

FIGURE 4 | Venn diagram showing the number of unique and shared OTUs (at 97% similarity) among control and Cu-treated cecal microbial communities of chickens. **(A)** The number of unique and shared OTUs was observed in control group among three sampling times. **(B)** The number of unique and shared OTUs was observed in the Cu-treated group three sampling times.

(2.05%), and *Actinobacteria* (1.52%) were overall well represented across the experimental groups. Other minor phyla, such as *Verrucomicrobia* (0.67%), *Spirochaetes* (0.36%), *Synergistetes* (0.27%), *Cyanobacteria* (0.20%), and *SHA-109* (0.16%), were also identified. In addition, it was also found that the specific changes associated with $CuSO_4$ exposure included a significant decrease in the relative abundance of *Bacteroides* in the Cu group on the 30th and 90th day compared with the CK group. Besides, phylum

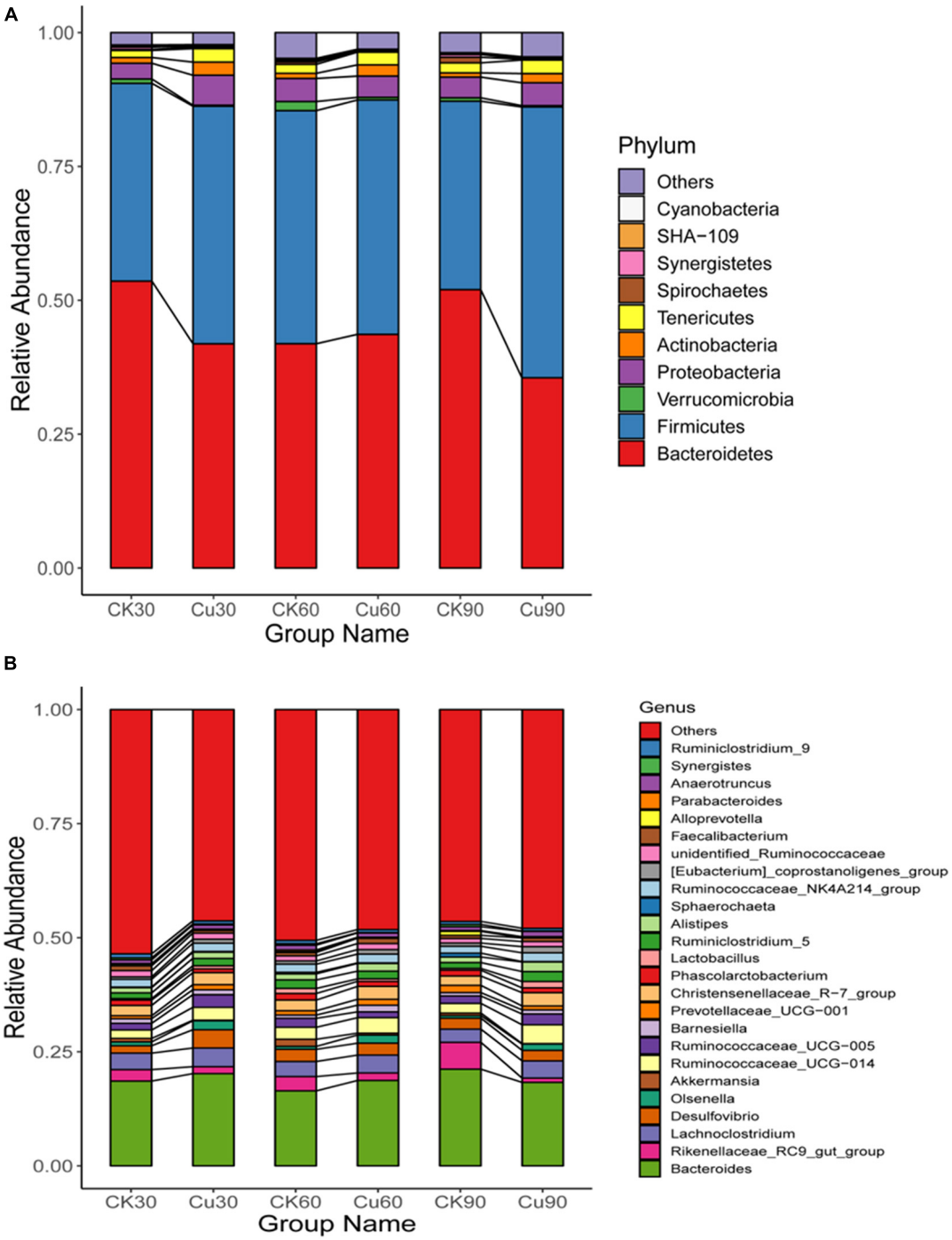

FIGURE 5 | (A) Difference in the relative abundance of the bacterial major phyla between the CK and Cu groups during the experimental period (at days 30, 60, and 90). **(B)** Difference in the relative abundance of the major bacterial genera between the two groups during the experimental period (at days 30, 60, and 90).

Firmicutes increased significantly on the 30th and 90th day, and had an increasing trend on the 60th day. Phylum *Proteobacteria* increased significantly on the 30th day, and *Tenericutes* increased significantly during the first 60-day period, and *Actinobacteria* increased significantly in the latter 60 days (**Figure 5A** and **Supplementary Table 2**). Consistent with the above results, compared with the middle 30-day period, the microflora of chickens exposed to $CuSO_4$ changed more dramatically during the first 30-day period and the latter 30-day period.

At the genus level, 200 genera were identified from the two groups of chickens, and *Bacteroidetes* and *Rikenellaceae_RC9_gut_group* are the dominant genera

of both the CK and Cu groups (**Figure 5B**). In addition, the relative abundance of *Ruminococcaceae_UCG-014*, *Desulfovibrio*, *[Eubacterium]_coprostanoligenes_group*, and *Lactobacillus* increased significantly after $CuSO_4$ exposure for 30 days (**Figure 5B** and **Supplementary Table 3**). Compared with the CK group, the relative abundance of *[Eubacterium]_coprostanoligenes_group* and *Olsenella* increased significantly, and the relative abundance of the *Ruminococcaceae_UCG-005* decreased significantly after 60 days of $CuSO_4$ treatment. It is noticeable that at 90 days, the cecal microorganisms in the Cu group changed dramatically, and the relative abundance of the genus from *Bacteroidetes*, *Firmicutes*, and *Actinobacteria* phylum, such as *Ruminococcaceae_UCG-014*, *Lachnoclostridium*, *Christensenellaceae_R-7_group*, increased significantly.

Co-occurrence Network of Genera

To explore potential interactions between members of the intestinal bacterial community, SparCC network maps were drawn based on genus composition and were considered valuable only if the association between the different genera was strong and the Spearman correlation coefficient was greater than 0.5 in absolute value and $q < 0.05$ (**Figure 6**). Overall, the co-occurrence network in the Cu group exhibited greater complexity with different components and correlations compared with the control group. Specifically, the co-occurrence network of the control group contains 106 edges; the co-occurrence network of the copper group contains 114 edges. We then calculated the closeness of the different bacteria at the genus level in the co-occurrence network. In terms of the tight centrality of nodes in the bacterial network, the three most correlated genera in both groups were *g__Lactobacillus*, *g__Ruminococcaceae_UCG-005*, *g__Desulfovibrio* in the normal group, *g__Synergistes*, *g__Sphaerochaeta*, and *g__Lactobacillus* in the copper group.

Predictive Functional Profiling of Microbial Communities

To further explore the functional changes in cecal microflora after $CuSO_4$ exposure, the function of cecal microbial community was predicted and analyzed by PICRUSt. According to the annotated KEGG results, significant differences were observed between the control group and the $CuSO_4$-treated group in the functional category of level 3, in which a strong separation was noted (**Figure 7**). On the 30th day after $CuSO_4$ exposure, there were only six KEGG pathways, which showed significant changes between the CK30 and Cu30 groups. **Figure 7A** shows that compared with the CK30 group, the Cu30 group predicted higher gene abundance in the RIG-I-like receptor signaling pathway, basal transcription factors, bile secretion, and benzoate degradation, but lower predicted genes associated with pantothenate and CoA biosynthesis, and Phenylalanine, tyrosine, and tryptophan biosynthesis. At day 60, there were significant differences in more KEGG pathways between the two groups, a total of 22. The abundance of eight pathways such as nucleotide metabolism, RNA transport, bacterial motility proteins, and so on, in the CK60 group were upregulated, and the abundance of 14 pathways, such as polycyclic aromatic hydrocarbon degradation, biosynthesis of vancomycin group antibiotics, and ascorbate and aldarate metabolism were downregulated compared with those in the Cu60 group (**Figure 7B**). Finally, 90 days after $CuSO_4$ exposure, the largest number of genomes met the inclusion criteria. A total of 71 KEGG pathways showed significant changes between the CK90 and Cu90 groups. Compared with the CK90 group, 36 pathways were upregulated, and 35 pathways were

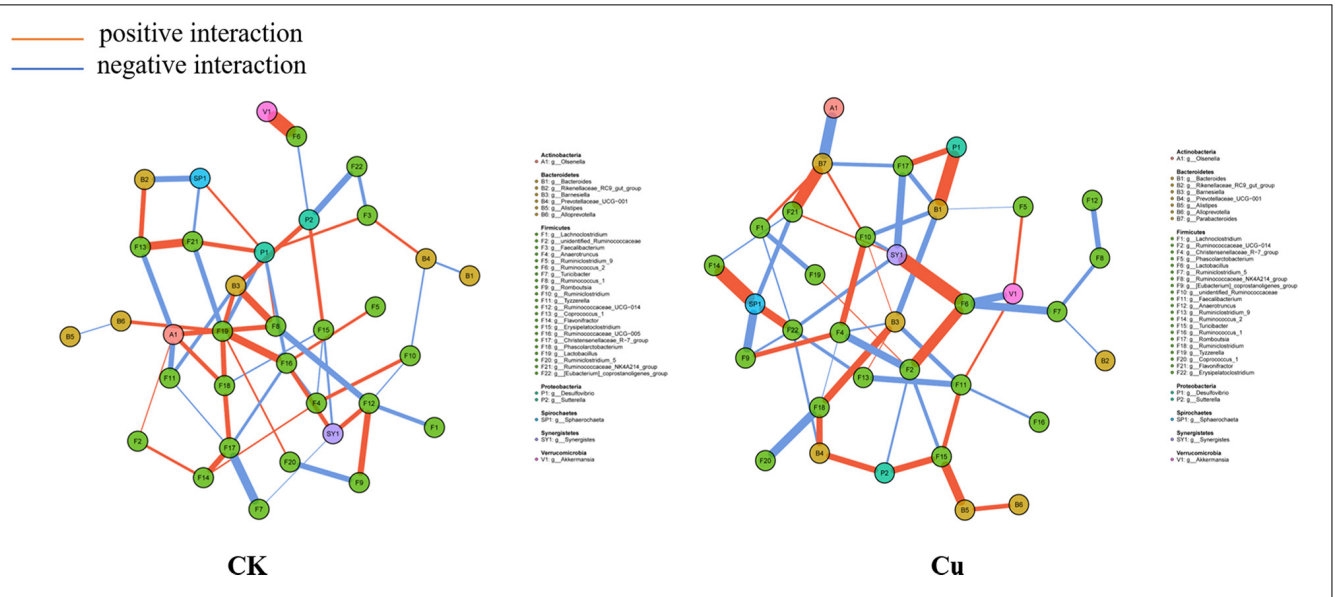

FIGURE 6 | Co-occurrence network analysis of bacterial genera. Each network node represents a genus of bacteria. The connection width represents the relevant values that support this connection. The connection color shows correlation, red indicates positive interaction, and blue indicates negative interaction.

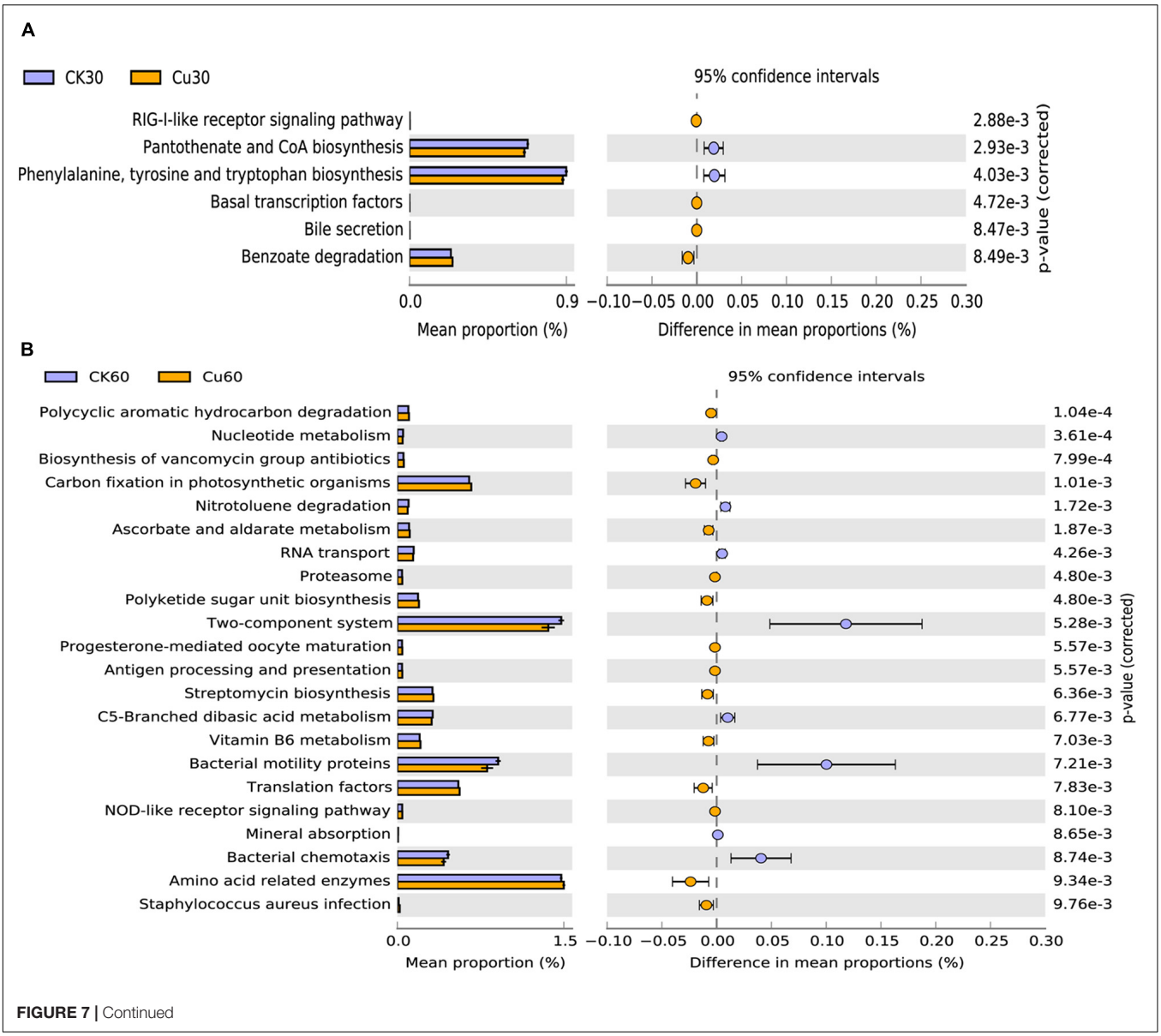

FIGURE 7 | Continued

downregulated in the Cu90 group (**Figure 7C**). These pathway changes may be related to the prolonged exposure time to Cu.

DISCUSSION

Earlier, related studies have shown that high copper may cause damage to the liver, kidney, glandular stomach, and reproductive system of chickens, but there is still a lack of studies on the effects of copper exposure on intestinal damage and intestinal microorganisms in chickens (Jackson et al., 1979; Oguz et al., 2010; Zhao et al., 2018). In this study, we focused on the change in microflora in the cecum of chicken under the subchronic Cu exposure, and further revealed how the structure and composition of intestinal microorganisms evolved. Our results showed that subchronic copper poisoning in chickens caused weight loss, intestinal morphological change, and intestinal flora balance disorder.

The findings of this study suggest that compared with the control groups, the body weight of $CuSO_4$-exposed chickens were decreased in a time-dependent manner. Previous studies have found that copper use of more than 300 mg/kg will lead to growth inhibition, showing an adverse effect on body weight (Pesti and Bakalli, 1996; Shahzad et al., 2012; Hamdi et al., 2018). A large number of histological injuries were also found in our histological observations. The results showed that copper exposure destroyed the normal structure of chicken cecum, mucosa fell off, vacuoles appeared in the lamina propria, and inflammatory infiltration occurred in a time-dependent manner. Interestingly, Ruan et al. (2019) and Cheng et al. (2020) demonstrated that copper exposure could cause intestinal mucosal damage and metabolic disorders, which reduce the

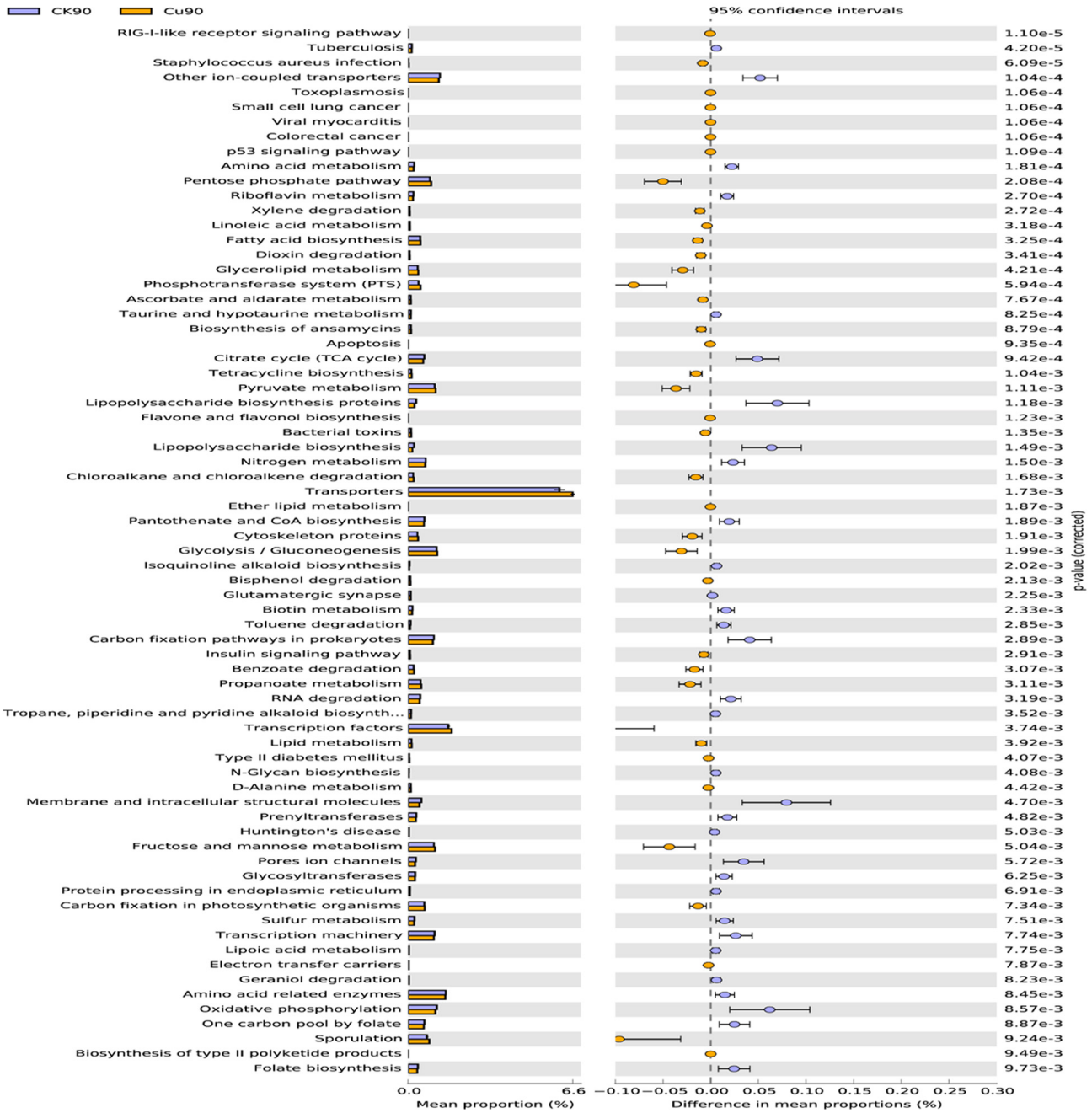

FIGURE 7 | Functional analysis of microbial communities. KEGG categories predicted by PICRUSt (level 3) differentially represented between CK and Cu groups at day 30 **(A)**, day 60 **(B)**, and day 90 **(C)** during the experimental period. The statistical differences between the two groups at the same time point were conducted using a Welch's *t*-test between the two groups; *p*-value < 0.01 was considered as statistical significant difference.

decomposition and absorption of nutrients, and eventually lead to weight loss. Therefore, the pathological damage of cecum may be one of the reasons for the weight loss of copper-exposed chickens.

The chicken intestinal tract is inhabited by a complex and dynamic population of microbial species. Once there is long-term excessive intake of heavy metals, it will not only lead to morphological damage but also cause the disorder of the structure and composition of intestinal flora, and even lead to host metabolic diseases (Kou et al., 2019). According to the α diversity analysis of 16S rDNA sequencing results, copper exposure increased the richness and diversity of cecal

microflora at 30 and 90 days and decreased at 60 days (**Figures 3A,B**), and, this is supported by Venn diagram analysis (**Figure 4**). These results showed that copper exposure causes significant changes in microbial diversity in the chicken cecum. The previous study suggests that it may be due to the fact that some bacteria can tolerate higher levels of copper and grow in copper-contaminated environments. Meanwhile, it was found by NMDS that the change in intestinal microflora was obvious between the two groups, and it depended on the time of exposure to $CuSO_4$ (**Figure 3C**). The UPGMA analysis showed that the OTUs of the CK group was clustered into one group according to phylogeny, while that of the Cu group was clustered into another group (**Figure 3D**). The above results suggest that copper sulfate can disturb the diversity and composition of cecal flora in a time-dependent manner, which is consistent with previous studies suggesting that heavy metals can change intestinal flora (Cheng et al., 2020; Zhou et al., 2020). In addition, it has been reported that disorders of intestinal microflora is related to a variety of diseases such as intestinal barrier permeability and inflammation (Karlsson et al., 2013; Thevaranjan et al., 2017; Zhang et al., 2018). Therefore, we speculate that copper exposure may cause intestinal flora imbalance, which may have adverse effects on the health of chickens.

Previous studies suggest that *Bacteroidetes* and *Firmicutes* are considered to be the most and second common phyla in the intestinal microflora of birds, which is similar to our results (**Figure 5A** and **Supplementary Table 2**; Wei et al., 2013; Zhou et al., 2020). The results also showed that compared with the CK group, the abundance of *Proteobacteria* was increased in the Cu group after $CuSO_4$ exposure. Several studies have reported that the abundance of *Proteobacteria* increased significantly when the intestinal mucosa was damaged (Michail et al., 2012; Tyler et al., 2013). In addition, we also noticed that copper exposure significantly increased the amount of *Tenericutes*. It has been reported that the amount of *Tenericuts* in colonic contents increased after 60 days of fluoride treatment, which may be related to the impairment of intestinal barrier function (Fu et al., 2019). Therefore, the above results showed that after Cu exposure, the balance of intestinal microorganisms at the phylum level is disrupted and may have high correlation with the intestinal infection in chicken.

Moreover, this study also reveals the potential effects of copper exposure on the structure of intestinal bacteria at the genus level. According to the results of species annotation, we found that the dominant bacteria in chicken cecum were *Bacteroides*, *Rikenellaceae_RC9_gut_group*, *Lachnoclostridium*, *Desulfovibrio*, and *olsenella* (**Figure 5B**). This was slightly different from the results of previous studies (Wei et al., 2013). One possible explanation for these results is that different intestinal environments may affect the abundance and composition of intestinal microflora. In addition, according to the *t*-test analysis, we noticed that *Desulfovibrio*, *Alistipes*, and *Anaerotruncus* increased significantly after copper exposure in 13 significantly changed genera, which is noteworthy (**Supplementary Table 3**). *Desulfovibrio* is a common genus in animals and humans. Its abundance increase is often accompanied by diarrhea, weight loss, and loss of appetite, which may further lead to intestinal epithelial hyperplasia, abscess, and inflammatory infiltration (Williamson et al., 2018). Moreover, Lau et al. (2006) demonstrated that *Anaerotruncus* may cause intestinal infection and bacteremia in the body. In addition, previous studies have found that *Alistipes* is considered to be a potentially harmful bacteria that may have an inflammatory effect (Kong et al., 2019). Based on the above observations, the results suggested that these genera with increased abundance in the copper group may play a partial role in intestinal injury and further cause intestinal barrier damage and intestinal inflammation. Therefore, this may explain why the body weight of the Cu group was lower than that of the CK group at all-time points.

The biological and physiological functions of intestinal microflora can be defined in many aspects, such as diversity and taxonomic composition, in order to decipher the potential role of intestinal bacterial community (Liu et al., 2019). In the current experiment, the PICRUST algorithm is used to map the bacterial genetic path to the KEGG database for functional prediction. A deeper analysis discovered that copper sulfate may change the microbiome and metabolic spectrum of chickens in a time-dependent manner. Compared with the CK group, a wide range of pathways were significantly changed under the influence of subchronic copper poisoning in the Cu group, including multiple gene functional families, such as energy production and conversion, amino acid transport and metabolism, signal transduction mechanism, etc. The results also showed that compared with the CK group, the relative abundance of genes regarding amino acid metabolism is decreased, and carbohydrates and lipids are increased in the Cu group (**Figure 7**). Interestingly, previous studies had confirmed that bacterial functional pathways in inflammatory tissues of patients with ulcerative colitis were significantly changed, carbohydrate metabolism decreased, and lipid and amino acid metabolism increased (David et al., 2014). Therefore, this may explain the reason for the difference, which may be related to the host and $CuSO_4$ exposure models used in the study. Likewise, it was worth noting that compared with the control group, the cecal microbial community in the copper group showed increased relative abundance of genes regarding bacterial toxin, apoptosis, *Staphylococcus aureus* infection, sporulation, nucleotide metabolism, and so on, and downregulation of lipoic acid and lipopolysaccharide biosynthesis in sulfur metabolism. This is consistent with previous studies and verified the morphological changes in the cecum in chickens (Zhou L. et al., 2019).

CONCLUDING REMARKS

In short, we found that subchronic copper exposure could lead to weight loss, pathological changes in cecal tissue, change the diversity, structure, and function of microbial community, and finally destroy the balance of intestinal microbial community in chickens. To sum up, our results may play a role in the effects of copper exposure on intestinal microorganisms in chickens.

Subchronic copper exposure disrupts the environmental homeostasis of intestinal microbes and alters the structure and balance of the microbial community (Jovel et al., 2016). Furthermore, alterations in intestinal flora may lead to changes in the intestinal environment such as acidity and alkalinity, intestinal infection, intestinal barrier permeability, and inflammation (Gong et al., 2008; Wei et al., 2013). Long-term effects can disrupt the normal structure of the chicken cecum, with loss of intestinal villi and metabolic disturbances, which may ultimately lead to weight loss (Shahzad et al., 2012; Ruan et al., 2019).

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/, SRP199017.

ETHICS STATEMENT

The animal study was reviewed and approved by the Jiangxi Agricultural University, Nanchang, Jiangxi, China, license No. JXAULL-2016038.

AUTHOR CONTRIBUTIONS

CH, YS, CZ, LG, GLiu, YZ, GLi, GH, PL, and XG contributed to the conception of the review. CH, YS, PL, and XG drafted the manuscript. All authors reviewed and approved the manuscript.

FUNDING

This research was funded by the National Natural Science Foundation of China (31460679) and Key Projects of the Natural Science Foundation of Jiangxi (2017ACB20012).

ACKNOWLEDGMENTS

We thank the team members for their help in the clinical veterinary laboratory of the College of Animal Science and Technology of Jiangxi Agricultural University.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.739577/full#supplementary-material

REFERENCES

Berwanger, E., Vieira, S. L., Angel, C. R., Kindlein, L., Mayer, A. N., Ebbing, M. A., et al. (2018). Copper requirements of broiler breeder hens. *Poult. Sci.* 97, 2785–2797.

Bui, T. K., Do-Hong, L. C., Dao, T. S., and Hoang, T. C. (2016). Copper toxicity and the influence of water quality of Dongnai River and Mekong River waters on copper bioavailability and toxicity to three tropical species. *Chemosphere* 144, 872–878. doi: 10.1016/j.chemosphere.2015.09.058

Buschmann, J. (2013). OECD Guidelines for the Testing of Chemicals. *Methods Mol. Biol.* 947, 37–56.

Cheng, S., Mao, H., Ruan, Y., Wu, C., Xu, Z., Hu, G., et al. (2020). Copper Changes Intestinal Microbiota of the Cecum and Rectum in Female Mice by 16S rRNA Gene Sequencing. *Biol. Trace Elem. Res.* 193, 445–455. doi: 10.1007/s12011-019-01718-2

Council, S. N. (1995). *Nutrient Requirements of Laboratory Animals*. Washington: National Academies Press.

David, L. A., Maurice, C. F., Carmody, R. N., Gootenberg, D. B., Button, J. E., Wolfe, B. E., et al. (2014). Diet rapidly and reproducibly alters the human gut microbiome. *Nature* 505, 559–563. doi: 10.1038/nature12820

Fei, D., Zhao, H., Wang, Y., Liu, J., Mu, M., Guo, M., et al. (2019). The disturbance of autophagy and apoptosis in the gizzard caused by copper and/or arsenic are related to mitochondrial kinetics. *Chemosphere* 231, 1–9.

Frieden, E. (1986). Perspectives on copper biochemistry. *Clin. Physiol. Biochem.* 4, 11–19.

Fu, R., Niu, R., Li, R., Yue, B., Zhang, X., Cao, Q., et al. (2019). Fluoride-Induced Alteration in the Diversity and Composition of Bacterial Microbiota in Mice Colon. *Biol. Trace Elem. Res.* 196, 537–544. doi: 10.1007/s12011-019-01942-w

Gong, J., Yu, H., Liu, T., Gill, J. J., Chambers, J. R., Wheatcroft, R., et al. (2008). Effects of zinc bacitracin, bird age and access to range on bacterial microbiota in the ileum and caeca of broiler chickens. *J. Appl. Microbiol.* 104,ak 1372–1382. doi: 10.1111/j.1365-2672.2007.03699.x

Hamdi, M., Solà, D., Franco, R., Durosoy, S., and Pérez, J. F. (2018). Including copper sulphate or dicopper oxide in the diet of broiler chickens affects performance and copper content in the liver. *Anim. Feed Sci. Technol.* 237, 89–97.

Jackson, N., Stevenson, M. H., and Kirkpatrick, G. M. (1979). Effects of the protracted feeding of copper sulphate-supplemented diets to laying, domestic fowl on egg production and on specific tissues, with special reference to mineral content. *Br. J. Nutr.* 42, 253–266. doi: 10.1079/bjn19790112

Jenkins, S. V., Robeson, M. S. II, Griffin, R. J., Quick, C. M., Siegel, E. R., Cannon, M. J., et al. (2019). Gastrointestinal Tract Dysbiosis Enhances Distal Tumor Progression through Suppression of Leukocyte Trafficking. *Cancer Res.* 79, 5999–6009. doi: 10.1158/0008-5472.CAN-18-4108

Jovel, J., Patterson, J., Wang, W., Hotte, N., O'Keefe, S., Mitchel, T., et al. (2016). Characterization of the Gut Microbiome Using 16S or Shotgun Metagenomics. *Front. Microbiol.* 7:459. doi: 10.3389/fmicb.2016.00459

Karlsson, F. H., Tremaroli, V., Nookaew, I., Bergstrom, G., Behre, C. J., Fagerberg, B., et al. (2013). Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature* 498, 99–103. doi: 10.1038/nature12198

Kong, C., Gao, R., Yan, X., Huang, L., and Qin, H. (2019). Probiotics improve gut microbiota dysbiosis in obese mice fed a high-fat or high-sucrose diet. *Nutrition* 60, 175–184.

Kou, H., Fu, Y., He, Y., Jiang, J., Gao, X., and Zhao, H. (2019). Chronic lead exposure induces histopathological damage, microbiota dysbiosis and immune disorder in the cecum of female Japanese quails (Coturnix japonica). *Ecotoxicol. Environ. Saf.* 183:109588. doi: 10.1016/j.ecoenv.2019.109588

Lau, S. K., Woo, P. C., Woo, G. K., Fung, A. M., Ngan, A. H., Song, Y., et al. (2006). Bacteraemia caused by Anaerotruncus colihominis and emended description of the species. *J. Clin. Pathol.* 59, 748–752. doi: 10.1136/jcp.2005.031773

Li, X., Brejnrod, A. D., Ernst, M., Rykaer, M., Herschend, J., Olsen, N. M. C., et al. (2019). Heavy metal exposure causes changes in the metabolic health-associated gut microbiome and metabolites. *Environ. Int.* 126, 454–467. doi: 10.1016/j.envint.2019.02.048

Liu, B., Lin, W., Chen, S., Xiang, T., Yang, Y., Yin, Y., et al. (2019). Gut Microbiota as a Subjective Measurement for Auxiliary Diagnosis of Insomnia Disorder. *Front. Microbiol.* 10:1770. doi: 10.3389/fmicb.2019.01770

Liu, Y., Li, Y., Liu, K., and Shen, J. (2014). Exposing to cadmium stress cause profound toxic effect on microbiota of the mice intestinal tract. *PLoS One* 9:e85323. doi: 10.1371/journal.pone.0085323

Macchione, I. G., Lopetuso, L. R., Ianiro, G., Napoli, M., Gibiino, G., Rizzatti, G., et al. (2019). Akkermansia muciniphila: key player in metabolic and gastrointestinal disorders. *Eur. Rev. Med. Pharmacol. Sci.* 23, 8075–8083. doi: 10.26355/eurrev_201909_19024

Michail, S., Durbin, M., Turner, D., Griffiths, A. M., Mack, D. R., Hyams, J., et al. (2012). Alterations in the gut microbiome of children with severe ulcerative colitis. *Inflamm. Bowel Dis.* 18, 1799–1808.

Mohd Shaufi, M. A., Sieo, C. C., Chong, C. W., Gan, H. M., and Ho, Y. W. (2015). Deciphering chicken gut microbial dynamics based on high-throughput 16S rRNA metagenomics analyses. *Gut Pathog.* 7:4. doi: 10.1186/s13099-015-0051-7

Oguz, E. O., Yuksel, H., Enli, Y., Tufan, A. C., and Turgut, G. (2010). The effects of copper sulfate on liver histology and biochemical parameters of term Ross broiler chicks. *Biol. Trace Elem. Res.* 133, 335–341. doi: 10.1007/s12011-009-8447-1

Pal, A. (2014). Copper toxicity induced hepatocerebral and neurodegenerative diseases: an urgent need for prognostic biomarkers. *Neurotoxicology* 40, 97–101. doi: 10.1016/j.neuro.2013.12.001

Patterson, J. A., and Burkholder, K. M. (2003). Application of prebiotics and probiotics in poultry production. *Poult. Sci.* 82, 627–631.

Pesti, G. M., and Bakalli, R. I. (1996). Studies on the feeding of cupric sulfate pentahydrate and cupric citrate to broiler chickens. *Poult. Sci.* 75, 1086–1091.

Ruan, Y., Wu, C., Guo, X., Xu, Z., Xing, C., Cao, H., et al. (2019). High Doses of Copper and Mercury Changed Cecal Microbiota in Female Mice. *Biol. Trace Elem. Res.* 189, 134–144. doi: 10.1007/s12011-018-1456-1

Shahzad, M. N., Javed, M. T., Shabir, S., Irfan, M., and Hussain, R. (2012). Effects of feeding urea and copper sulphate in different combinations on live body weight, carcass weight, percent weight to body weight of different organs and histopathological tissue changes in broilers. *Exp. Toxicol. Pathol.* 64, 141–147. doi: 10.1016/j.etp.2010.07.009

Son, J., Lee, Y. S., Kim, Y., Shin, K. I., Hyun, S., and Cho, K. (2016). Joint toxic action of binary metal mixtures of copper, manganese and nickel to Paronychiurus kimi (Collembola). *Ecotoxicol. Environ. Saf.* 132, 164–169. doi: 10.1016/j.ecoenv.2016.05.034

Thevaranjan, N., Puchta, A., Schulz, C., Naidoo, A., Szamosi, J. C., Verschoor, C. P., et al. (2017). Age-Associated Microbial Dysbiosis Promotes Intestinal Permeability, Systemic Inflammation, and Macrophage Dysfunction. *Cell Host Microbe* 21, 455–466.e4.

Tyler, A. D., Knox, N., Kabakchiev, B., Milgrom, R., Kirsch, R., Cohen, Z., et al. (2013). Characterization of the gut-associated microbiome in inflammatory pouch complications following ileal pouch-anal anastomosis. *PLoS One* 8:e66934. doi: 10.1371/journal.pone.0066934

Wei, S., Morrison, M., and Yu, Z. (2013). Bacterial census of poultry intestinal microbiome. *Poult. Sci.* 92, 671–683.

Williamson, A. J., Carlson, H. K., Kuehl, J. V., Huang, L. L., Iavarone, A. T., Deutschbauer, A., et al. (2018). Dissimilatory Sulfate Reduction Under High Pressure by Desulfovibrio alaskensis G20. *Front. Microbiol.* 9:1465. doi: 10.3389/fmicb.2018.01465

Wu, B., Cui, H., Peng, X., Pan, K., Fang, J., Zuo, Z., et al. (2014). Toxicological effects of dietary nickel chloride on intestinal microbiota. *Ecotoxicol. Environ. Saf.* 109, 70–76. doi: 10.1016/j.ecoenv.2014.08.002

Zhang, B., Lv, Z., Li, Z., Wang, W., Li, G., and Guo, Y. (2018). Dietary l-arginine Supplementation Alleviates the Intestinal Injury and Modulates the Gut Microbiota in Broiler Chickens Challenged by Clostridium perfringens. *Front. Microbiol.* 9:1716. doi: 10.3389/fmicb.2018.01716

Zhao, H., Wang, Y., Shao, Y., Liu, J., Liu, Y., and Xing, M. (2018). Deciphering the ionic homeostasis, oxidative stress, apoptosis, and autophagy in chicken intestine under copper(II) stress. *Environ. Sci. Pollut. Res. Int.* 25, 33172–33182. doi: 10.1007/s11356-018-3163-z

Zhao, X. J., Li, Z. P., Wang, J. H., Xing, X. M., Wang, Z. Y., Wang, L., et al. (2015). Effects of chelated Zn/Cu/Mn on redox status, immune responses and hoof health in lactating Holstein cows. *J. Vet. Sci.* 16, 439–446. doi: 10.4142/jvs.2015.16.4.439

Zhou, C., Xu, P., Huang, C., Liu, G., Chen, S., Hu, G., et al. (2019). Effects of subchronic exposure of mercuric chloride on intestinal histology and microbiota in the cecum of chicken. *Ecotoxicol. Environ. Saf.* 188:109920. doi: 10.1016/j.ecoenv.2019.109920

Zhou, L., Chen, C., Xie, J., Xu, C., Zhao, Q., Qin, J. G., et al. (2019). Intestinal bacterial signatures of the "cotton shrimp-like" disease explain the change of growth performance and immune responses in Pacific white shrimp (*Litopenaeus vannamei*). *Fish Shellfish Immunol.* 92, 629–636. doi: 10.1016/j.fsi.2019.06.054

Zhou, C., Xu, P., Huang, C., Liu, G., Chen, S., Hu, G., et al. (2020). Effects of subchronic exposure of mercuric chloride on intestinal histology and microbiota in the cecum of chicken. *Ecotoxicol. Environ. Saf.* 188:109920. doi: 10.1016/j.ecoenv.2019.109920

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.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 05 November 2021
doi: 10.3389/fmicb.2021.722901

The Dominating Role of Genetic Background in Shaping Gut Microbiota of Honeybee Queen Over Environmental Factors

Jiandong Yang[1]†, Yun Zhong[1]†, Liqun Xu[1], Bo Zeng[1], Kang Lai[2], Mingxian Yang[1], Diyan Li[1], Ye Zhao[1], Mingwang Zhang[1] and Debing Li[1]*

[1] College of Animal Sciences and Technology, Sichuan Agricultural University, Chengdu, China, [2] Sichuan Province Apiculture Management Station, Chengdu, China

OPEN ACCESS

Edited by:
Wei Huang,
Johns Hopkins University,
United States

Reviewed by:
Xiaoli Bing,
Nanjing Agricultural University, China
Guohua Xiao,
Sanda University, China
Jiannong Xu,
New Mexico State University,
United States

***Correspondence:**
Jiandong Yang
yangjd@sicau.edu.cn

† *These authors have contributed equally to this work*

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *09 June 2021*
Accepted: *29 September 2021*
Published: *05 November 2021*

Citation:
Yang J, Zhong Y, Xu L, Zeng B, Lai K, Yang M, Li D, Zhao Y, Zhang M and Li D (2021) The Dominating Role of Genetic Background in Shaping Gut Microbiota of Honeybee Queen Over Environmental Factors. Front. Microbiol. 12:722901. doi: 10.3389/fmicb.2021.722901

A balanced, diverse gut microbiota is vital for animal health. The microbial population is shaped by multiple factors including genetic background and environment, but other determinants remain controversial. Numerous studies suggest that the dominant factor is genetic background while others emphasize the environmental factors. Here, we bred asexual hybridization queens (AHQs) of honeybees through nutritional crossbreeding (laid in *Apis mellifera* colony but bred in *Apis cerana* colony), sequenced their gut microbiome, and compared it with normally bred sister queens to determine the primary factor shaping the gut microbiota. Our results showed that the dominant genera in the gut microbiota of AHQs were *Brevundimonas*, *Bombella*, and *Lactobacillus*, and its microbial community was more related to *A. mellifera* queens. The AHQs had a moderate number of different bacterial species and diversity, but total bacterial numbers were low. There were more significant taxa identified in the comparison between AHQ and *A. cerana* queen according to LEfSe analysis results. The only genetic-specific taxon we figured out was *Brevundimonas*. The growth of core bacterial abundance showed different characteristics among different queen groups in the first week after emerging. Collectively, this study suggested that the genetic background played a more dominant role than environmental factors in shaping the gut microbiota of honeybee queen and the microbiota of midgut was more sensitive than that of rectum to this impact.

Keywords: honeybee, queen, gut microbiota, amplicon sequencing, asexual hybridization, crossbreed

INTRODUCTION

Diverse microbial communities colonize different host tissues, with the gut harboring the densest and most diverse range of species (Martinson et al., 2012). Researchers delve into the gut microbiota of animal newborns, which underlines the vital role of the gut microbiota for host's health by maintaining intestinal homeostasis and barrier function, stimulating the development of the immune system, contributing to nutrient digestion, and protecting against pathogens (Sekirov et al., 2010; Maynard et al., 2012; Wopereis et al., 2014). Current evidence indicates that the gut microbiota of honeybees is pivotal to their health as it participates in metabolism and immunity, promotes development, and resists invasion by parasites and pathogens

(Guo et al., 2015; Schwarz et al., 2016; Zheng et al., 2017; Wu et al., 2020). As a result, the functional role of the gut microbiota has drawn much attention worldwide. Besides function, research has been focused on identifying the dominant factors determining the diversity and richness of the gut microbiota. Data from a variety of animal subjects concur that there is a complex interaction between the microbial community and the host, but the primary determinants of the animal gut microbiota include the host's genotype (Knowles et al., 2019; Korach-Rechtman et al., 2019), diet (David et al., 2014; Carmody et al., 2015; Sonnenburg et al., 2016; Griffin et al., 2017; Jones et al., 2017), season (Ludvigsen et al., 2015), host age (Martinson et al., 2012; Tarpy et al., 2015; Anderson et al., 2018), caste (Kapheim et al., 2015; Anderson et al., 2018), and environment (Amato et al., 2016; Ludvigsen et al., 2017; Ren et al., 2017). Multiple nature and nurture effects stemming from differences in host species can greatly influence the interactions between host and microorganism. Even within the same species, conclusions about the dominant influences can vary because of individual differences and the type of calculation methods used. We are only just beginning to understand the processes shaping the composition of host-associated microbial communities over evolutionary and ecological timescales (Foster et al., 2017).

Honeybees are necessary and valuable pollinators of most crops and wild plants, and their economic value in this sense far outweighs their usefulness as honey producers (Van der Sluijs and Vaage, 2016). Their intestinal organs are segmented, and the composition of the gut microbial community is relatively simple, making it an ideal social insect model for studying the impact of social behavior on the dynamics of the gut microbiota. The detailed taxonomic information about the gut microbiota composition of the honeybee (*Apis mellifera*) remained unavailable until high-throughput sequencing (16S amplicon sequencing) was developed and widely employed (Jeyaprakash et al., 2003). Numerous studies suggested that there was a conserved evolutionary pattern of the gut communities in all related corbiculate (pollen basket) bee species, which could insure they have a similar and relatively stable gut microbial community which mainly contained five core members and four non-core members (Martinson et al., 2011; Sabree et al., 2012; Kwong et al., 2017). These core members comprise a remarkably stable characteristic as they can be detected in the gut of every adult worker, whether in the same region or the same colony. Although there is great variability in the microbial population between each individual worker, these microbes could rarely be found in honeybee living environments, including pupae, frame, and hive (Engel and Moran, 2013; Powell et al., 2014). We believe that this phenomenon indicates that there may have been a strong mutual selection between the gut microbiota and the host during evolution, and the explanation for this phenomenon can be seen from the biological characteristics of the social lifestyle of honeybees (Kwong et al., 2017).

Social behavior is a prominent feature of social animals, one of which is honeybee, and one of their social behaviors is mutual feeding. Newly emerged queens and workers are sterile (Martinson et al., 2012; Powell et al., 2014); they usually stay in hives for more than 1 week after emerging and were fed royal jelly by other nurse bees. Without contacting with the outside environment, their gut microbe can develop rapidly within 5 days. During these 5 days, the core members colonize rapidly and the microbial community gradually forms (Guo et al., 2015). Royal jelly, as the main food for newborns, may play an important intermediary role in the microbial transfer process when mutual feeding happened. Thus, as a social behavior, mutual feeding provides a stable pathway for the transfer of the gut microbiota between individuals and is of great significance for the early growth of core members to occupy key metabolic niches (Powell et al., 2014). This pattern of social transmission can also be found in other social living animals and humans (Marcobal and Sonnenburg, 2012). The reason why these gut core microbes plays irreplaceable roles in the host's intestinal tract can be explained as that they occupy some vital metabolic niches, such as helping the host to digest pollen and nectar (Zheng et al., 2016), synthesizing hormones (Zheng et al., 2017), and regulating immune responses (Wu et al., 2020). However, recent studies suggest that some of the non-core microbiota may play important roles in caste development because they maintain a high relative abundance in the early developmental stages of the queen (Jeyaprakash et al., 2003; Corby-Harris et al., 2014b; Anderson et al., 2018); the mechanism remains to be characterized.

The Western honeybee, *A. mellifera*, and the eastern honeybee, *Apis cerana*, are the most widely raised honeybee species in China, bringing the most economic income to Chinese beekeepers compared to other bee species. These two bee species have unique biological characteristics and genetic backgrounds and also have excellent individual productive traits. They do have some shortcomings for beekeepers, however. For instance, the eastern bee produces less honey but has strong disease resistance and is easy to manage manually (Li et al., 2012), while the Italian bee is a high-yielder but more likely to be infected by pathogens and parasites (Guo et al., 2015) and requires more keeper management. It would be advantageous to create a bee variety with high yield, high disease resistance, easy feeding, and management to increase the profitability of beekeeping. Chinese researcher Ming Zhuang has created a hybrid bee with the advantages of both parents by transferring (Zhuang, 1985). This method of hybridization, which does not change the genetic background of the offspring, is a form of asexual reproduction called nutritional crossbreeding. Many additional attempts to produce hybrid bees by nutritional crossbreeding have been made by Chinese researchers, and their production and physiological indices have been measured. They saw some changes in morphology (Zeng et al., 2005a), but surprisingly they found that performance parameters such as birth weight (Zeng et al., 2005b) and mite resistance (Xie et al., 2008) of the offspring of the bees produced by nutritional crossbreeding were also improved. The crossbred queen (transferred from colony A to colony B) is an ideal model for identifying the dominant factors shaping the gut microbiota. The genetic background of the crossbred queens is consistent with that of the queens from colony A, while the living conditions and nutritional factors acquired are those of the queens from colony B.

Here, we use the asexual hybridization queen (AHQ) social-animal model to discover the dominant factors influencing the

composition of the gut microbiota in the early development stage of honeybee queens, to explore the interaction between the gut microbiota and the host. This study may help to reveal how social living affects the gut microbiota and allow a deeper exploration of its relationship with the host during coevolution.

MATERIALS AND METHODS

Grouping of Queens and Nomenclature

Three groups of queens were involved in this study: (1) ACQs, usual *A. cerana* queens that were laid in an *A. cerana* colony and bred by an *A. cerana* nurse bee; (2) AMQs, usual *A. mellifera* queens that were laid in an *A. mellifera* colony (AMC) and bred by an *A. mellifera* nurse bee; and (3) AHQs, an asexual hybridization (nutritionally crossbred) of queens that were laid in an AMC, but fostered by an *A. cerana* nurse bee.

Experimental Design and Management

All the colonies used in this study were located at the affiliated apiculture base of the College of Animal Sciences and Technology, Sichuan Agricultural University, Chengdu, Sichuan, China. The experimental design is shown in **Figure 1**. Three robust *A. mellifera* colonies were selected to breed AMQs, 3 robust *A. cerana* queen-less colonies were chosen to breed ACQs, and 10 additional robust *A. cerana* colonies were used to organize the young *A. cerana* colonies (YACCs). We prepared and introduced 60 queen cells (30 for breeding AMQs, 30 for breeding AHQs) along with larvae for each AMC and 30 queen cells (all for breeding ACQs) for each *A. cerana* colony using artificial queen-breeding technology. The YACCs were organized on the sixth day after the queen cells were introduced successfully and completely. To organize the YACCs, the bee frames along with workers were taken from the 10 additional robust *A. cerana* colonies and placed one to one into each new sterilized hive, at a distance of 50 m away from the original colonies. Adult workers of recognition capability returned to their original hive leaving newly emerged nurse bees in the YACCs to foster the *A. mellifera* queens after they emerged from the queen cells. Once the YACCs were organized, half of the queen cells were transferred from AMCs to YACCs so that there were 30 queen cells along with queens in each experimental group. Each queen in a hive was checked every morning and marked to record age. In order to set up the opposite crossbreeding, we tried transferring *A. cerana* queen pupae to a young AMC to breed the other kind of crossbred queen but failed with an almost total lack of acceptance. This may be because *A. mellifera* workers have a greater ability to recognize and exclude different species than *A. cerana*. Thus, this reverse asexual hybridization group was not included in the study.

The core members of the gut microbe population of the workers were colonized by the fifth day after their host emerged; however, there have been only limited reports on the exact timing of the establishment of queens' gut microbes (Guo et al., 2015). Queens usually flew out from the hive and mated with drones at the seventh day after emerging, when their tissues and organs were almost fully developed. Based on this knowledge, gut samples used for high-throughput sequencing were collected when queens were 5 days old while other samples for use in real-time quantitative PCR (qPCR) were collected on the first, fourth, and seventh days to monitor the absolute abundance changes of the core members in the queens' gut microbial community. Due to the high individual diversity of animals gut microbes, we used as many samples as possible for high-throughput sequencing to minimize the impact of such diversity on our conclusions.

DNA Extraction and qPCR

The queens were first euthanized with carbon dioxide and then pinned in a sterile dissecting plate. The gut tissues were collected by clamping the last part of the sternum with sterilized forceps, separating the midgut and rectum from the gut tissues, and placing them into 2.0-ml microfuge tubes. The gut tissues were immediately frozen in liquid nitrogen and transferred to −80°C until DNA extraction. The entire procedure was conducted under aseptic conditions, and all tools were sterilized. Total genomic DNA was extracted from the midgut and rectum using the TIANamp® Stool DNA Kit (Beijing Tiangen Biotech Ltd., Beijing, China) following the manufacturer's instruction under sterile conditions as described by Powell et al. (2014). DNA purity was determined on a 1% agarose gel, and DNA was diluted to 1 ng/μl using sterile water. One rectal DNA sample was discarded due to the poor quality.

The 16S rRNA gene V3–V4 regions were amplified using universal primers 341F (5′-CCTAYGGGRBGCASCAG-3′) and 806R (3′-GGACTACNNGGGTATCTAAT-5′) (Wang and Qian, 2009; Thijs et al., 2017) with the barcode. All PCR reactions were carried out with 15 μl of Phusion® High-Fidelity PCR Master Mix (New England Biolabs, Ipswich, MA, United States), 0.2 μM of forward and reverse primers, and about 10 ng of template DNA. Thermocycling was performed with an initial denaturation at 98°C for 1 min, followed by 30 cycles of denaturation at 98°C for 10 s, annealing at 50°C for 30 s, elongation at 72°C for 30 s, and lastly, 72°C for 5 min. PCR amplification products were verified by electrophoresis on a 1% agarose gel. Equal volumes of 1× loading buffer containing SYBR Green (New England Biolabs) were mixed with PCR products in equidensity ratios and separated by electrophoresis on a 2% agarose gel for detection. Then, the PCR products were purified with the Qiagen gel extraction kit (Qiagen, Hilden, Germany).

Illumina Sequencing and Sequence Analysis

Sequencing libraries were generated using the TruSeq® DNA PCR-Free sample Preparation Kit (Illumina, San Diego, CA, United States) following the manufacturer's recommendations, and index codes were added. The library quality was assessed on the Qubit 2.0 fluorometer (Thermo Scientific, Waltham, MA, United States) and Agilent Bioanalyzer 2100 system. The libraries were sequenced on an Illumina NovaSeq platform, and 250-bp paired-end reads were generated. Paired-end reads were assigned to samples based on their unique barcode and truncated by cutting off the barcode and primer sequence. All paired-end reads were merged using FLASH

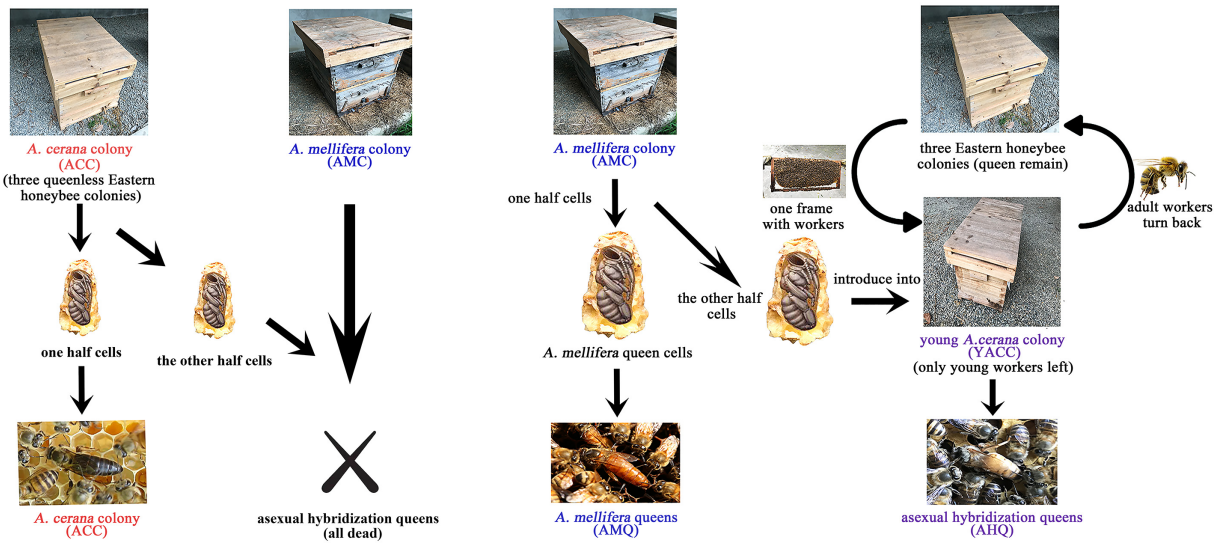

FIGURE 1 | Experimental design. The brief process of ACQs, AHQs, and AMQs included in this study is visualized in this schematic diagram. Experimental details are given in the "Materials and Methods" section. The other opposite crossbreeding (*A. cerana* into *A. mellifera*) was excluded from this study for lack of acceptance.

software (v1.2.7)[1] (Magoč and Salzberg, 2011) and entered into QIIME 2 (v2019.7)[2] (Bolyen et al., 2019) for downstream analysis including demultiplexing, pair joining, de-noising, and clustering. Amplicon sequence variants (ASVs) were defined based on 100% similarity clustering using the deblur (Amir et al., 2017) plugins in QIIME 2. Afterward, the representative sequences of each ASV were aligned and used to generate a phylogenetic tree as a reference for phylogenetic diversity analyses. Lastly, the 16S rRNA gene total length Silva database (v132_99_16S)[3] was specifically retrained for V3–V4 regions and used to classify the representative sequences.

Absolute qPCR

Absolute qPCR was used to determine the variation in abundance of the core members of the queens' gut microbiota. The primers used in this process are listed in **Supplementary Table 1**, and the initial template DNA concentrations were normalized between samples. For absolute qPCR, we first constructed standard samples for each species of bacteria. The corresponding 16S rRNA V3–V4 region sequences obtained by high-throughput sequencing were synthesized by Tsingke Biology Co., Ltd. (Chengdu, China) and cloned into the pMD® 19-T vector (Takara Biotechnology Co., Ltd., Dalian, China). The vectors were transduced into competent *Escherichia coli* DH-5α cells, aliquots were spread on agar plates, and single colonies were selected. After culturing, plasmids were extracted using the TIANprep Mini Plasmid Kit (Tiangen Biotech Co., Ltd., Beijing, China) following the product manual. The concentration of plasmids was measured, and the copy numbers were calculated according to relative plasmid quality. All plasmids containing the target fragments were diluted by 10-fold gradients (at least five gradients) for qPCR to monitor amplification efficiency and to generate standard curves (**Supplementary Table 2**).

Statistical Analysis

For ASV diversity analysis, we used our resampled ASV table at a depth of 12,000 without replacement as a basis. In α-diversity analysis, the richness and evenness of gut microbiota were assessed by calculating the numbers of different species and the Shannon index (Shannon, 1948), respectively. As for β-diversity, both the Bray–Curtis dissimilarity (Beals, 1984) and the unweighted UniFrac distance (Lozupone and Knight, 2005) were used to generate principal component analysis (PCA) plots. Because of the high similarity of the downstream analysis results based on the matrix calculated from the Bray–Curtis dissimilarity and the unweighted UniFrac distance, only the results based on Bray–Curtis dissimilarity were shown in our study (resultant figures based on unweighted UniFrac distance are shown in **Supplementary Materials**). The statistical analysis of both α-diversity and β-diversity between groups was performed in QIIME 2 with pairwise Kruskal–Wallis and permutational multivariate analysis of variance (PERMANOVA) tests (Zapala and Schork, 2006; Chen et al., 2012), respectively. Hierarchical clustering was performed with the UPGMA algorithm using the *hclust* package in R (v3.5.3) (R Core Team, 2015). The phylogenetic tree was constructed from ASVs using the method described by Callahan et al. (2016). Dendrograms were created using the package, *ape* (Paradis et al., 2004). A random forest classifier (RFC)-supervised learning algorithm was implemented in the *randomForest* package (Breiman, 2001) in R. Models were run using CSS-normalized ASV counts with 1,000 trees, and the OOB estimates of error rates were counted. The linear discriminant analysis (LDA) effect size

[1]http://ccb.jhu.edu/software/FLASH/

[2]https://qiime2.org/

[3]https://www.arb-silva.de/

(LEfSe) (Segata et al., 2011) was performed on the Galaxy/Hutab online platform[4] based on the ASV table, and the LDA threshold was set at 3.6 between groups. Other statistical analyses were carried out in SPSS 23. The copy numbers of core microbial members were compared using ANOVA (analysis of variance).

RESULTS

Breeding Model Queens

We succeeded in producing 15 crossbred queens (AHQs) out of 31 attempts, and 16 ACQs and 16 *A. mellifera* queens (AMQs) were bred in the same place during the same period. Considering the high individual diversity of gut microbiota, all the queen samples were used for high-throughput sequencing except the necessary biological duplicates for qPCR. In summary, 20 midgut and 19 rectal gDNAs (one rectal gDNA sample was removed for failing to meet the quality requirements) were extracted from 20 queen samples (seven of ACQs, six of AHQs, and seven of AMQs) and used for sequencing library preparation, while another 27 midgut and 27 rectal gDNA extracted from 27 queen samples participated in qPCR. To verify the early colonization traits of bacteria in queens' gut, the queen samples used for high-throughput sequencing were collected after the queens were fed by nurse bees for 5 days. The samples used for qPCR were collected when the queens were 1, 4, and 7 days of age to monitor the abundance variations of core members of queens' gut microbiota. Notably, in order to get the reciprocal resultant data from the other kind of AHQ (laid in *A. cerana* colony but bred in the AMC), four times efforts have ended and we still failed with 0% acceptance of the reverse crossbreeding experiments.

Composition and Diversity of Queen's Gut Microbiota

Overall, 3,220,523 sequences of the 16S rRNA gene were obtained, forming 1,861,700 ASVs. The number of sequences per sample ranged from 66,247 to 96,700, with an average of 82,578. Our sequencing results showed that the core members of the gut microbiota of queens were different from those of workers. The bacterial community of the three types of queens at the phylum level (**Supplementary Figure 1**) mainly consisted of Proteobacteria (72.0%) and Firmicutes (26.4%), accounting for over 98% of the total microbial composition. At the genus level (**Figure 2A**), only five taxa (relative abundance >1.0%) dominated the midgut and rectum bacterial community: *Brevundimonas* (45.0%), *Lactobacillus* (25.4%), *Bombella* (20.6%), *Klebsiella* (2.0%), and *Escherichia–Shigella* (1.2%). Notably, two of seven ACQ rectal samples were completely dominated by a single taxon, like *Lactobacillus*, while other two rectal samples were mainly dominated by *Brevundimonas*. We found that the composition of the rectum microbiota in ACQs showed two patterns, one dominated by *Lactobacillus* and the other by *Brevundimonas*. The observed ASV numbers (**Figure 2B**) and Shannon index (**Figure 2C**) were used to assess the richness and evenness, relatively. The number of taxa in the midgut or rectum of ACQs was higher than the other two groups, but the difference was only significant for the midgut ($p < 0.01$). As for the microbial evenness, the gut microbiota of the rectum and midgut in the AHQ group was in the middle level, and only the difference between AHQs with AMQs in the rectum was significant ($p < 0.05$).

High Similarity of the Midgut Microbiota Between Asexual Hybridization Queens and *Apis mellifera* Queens

To explore the deeper connections of the gut microbial community of AHQs and other queens, multiple analyses were performed. After that, we found several evidences that could support the gut microbiota composition of AHQs which was more similar to that of AMQs. Firstly, the principal coordinate analysis (PCoA) plots based on Bray–Curtis dissimilarity showed that the AHQ midgut sample clustering had a stronger correlation with AMQs while the AHQ rectal sample clustering fell between ACQs and AMQs (**Figure 3A**). Similar patterns were observed in the PCoA plots based on unweighted UniFrac distances (**Supplementary Figures 2A,B**). Secondly, the topological structure of the UGPMA trees based on Bray–Curtis dissimilarity (**Figures 3C,D**) and unweighted UniFrac distances (**Supplementary Figures 3A,B**) showed that both in the midgut and rectum, most AHQs and AMQs clustered into one branch while ACQs appeared in another single branch; the only exception was the rectum topological structure based on unweighted UniFrac distances in which most AHQs were clustered with ACQs (**Supplementary Figure 3B**). Next, RFC models classified the midgut gut microbial communities from queens' genetic background (AMQs + AHQs vs. ACQs) with great accuracy (95%) while the classification accuracy according to queens' environment (AMQs vs. ACQs + AHQs) was poor (70%). Most definitively, LEfSe tests were employed to identify the taxa of significant differences among AHQs and other queens. Considering the relative simplicity of the composition of queens' gut microbiota and the similarity of core members among different queen types, we initially adjusted the LDA threshold to a relatively high level of 4 to screen out the significant taxa among groups more strictly. As visualized in **Figures 4A–D**, significant differences in taxa between AHQs and the other two groups in the midgut and rectum were detected and varied greatly in number except the comparison between the AHQs vs. AMQs in the midgut, which showed no significant differences in taxon (**Figure 4C**). Even when the threshold was lowered to 3.6 (another commonly used threshold), there were still no significant differences in taxon.

Collectively, these studies consistently revealed the high similarity between AHQs and AMQs, indicating that the queens' genetic background played a more important role than environmental factors in shaping the midgut microbiota. Because AHQs and AMQs were laid by the same queen but bred in different colonies, they had the same genetic background but different nutritional and living environmental conditions. However, attributing all the factors determining the gut microbial composition to genetics would be unrealistic because

[4]http://huttenhower.sph.harvard.edu/galaxy/

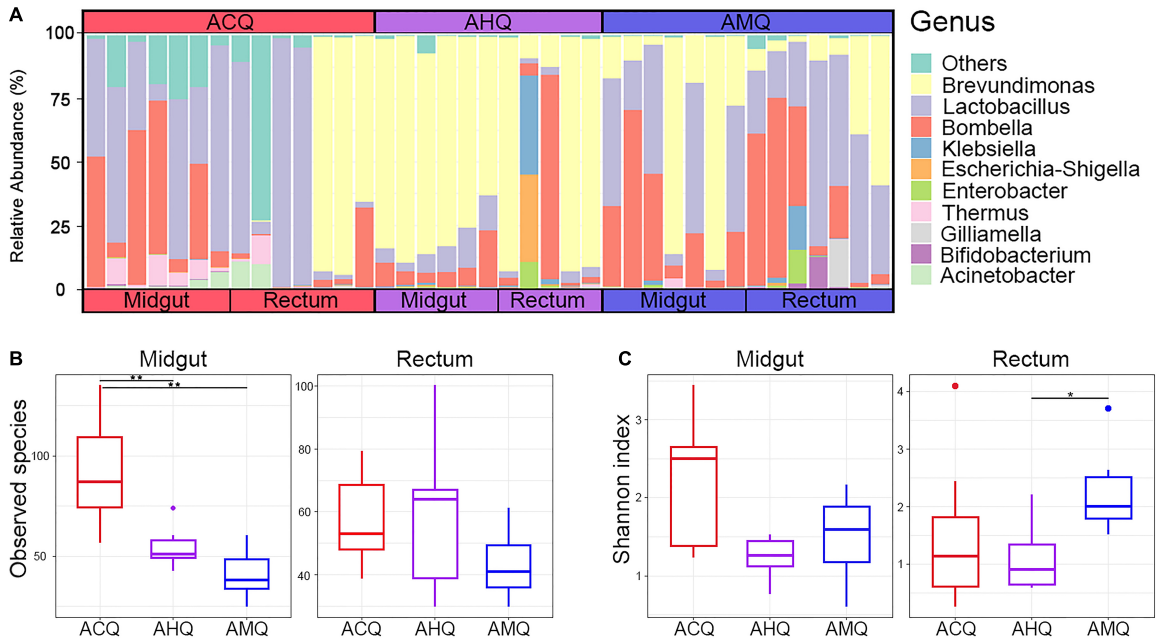

FIGURE 2 | Gut microbiota composition and α-diversity analysis of ACQs, AHQs, and AMQs. **(A)** Bacterial relative abundance of midgut and rectum at the genus level. **(B,C)** The results of observed species and Shannon indexes of α-diversity analysis shown by boxplots colored in red for ACQs, violet for AHQs, and blue for AMQs. $^{*}p < 0.05$ and $^{**}p < 0.01$ represent the significance levels between groups calculated from the Kruskal–Wallis *H* test.

environmental factors did play a part in the determination. For instance, we could not completely separate the AHQ cluster from the ACQ cluster from the midgut or rectum in the vertical axis of the PCoA plots based on Bray–Curtis dissimilarity (**Figures 3A,B**). Early feeding or other environmental factors could contribute to the establishment of the microbial community of AHQs.

Screening for Genetic-Specific Taxa

Based on the results of LEfSe analyses (**Figure 4**) and RFC models (**Supplementary Figure 3**), we tried to seek out the specific taxa of strongly genetic preference according to the following two requirements: firstly, either in midgut or rectum, the absolute abundances of these specific taxa must be significantly different between ACQs and AHQs, but not between AHQs and AMQs. This is because ACQs and AHQs only share the same nurturing conditions, but not the same genetic background. Secondly, to make our results more rigorous, these taxa must be identified simultaneously in both LEfSe analyses and RFC models (**Supplementary Figure 3**). Finally, there were seven taxa were identified from the midgut with high genetic preference. These genera included *Romboutsia*, *Brevundimonas*, *Faecalibacterium*, *Anoxybacillus*, *Thermus*, *Agathobacter*, and *Ruminococcus2*. All of these taxa had a higher abundance in ACQs except *Brevundimonas* in AHQs. As for rectum, only three taxa were figured to have a higher abundance in AHQs, all belonging to Enterobacteriaceae.

Consistent with the results of species composition and α-diversity analysis (**Figures 2A–C**), these findings all indicated that the development of the gut microbiota of AHQs resulted in lower diversity and dominance of several taxa, especially *Brevundimonas*. In our view, the high abundance of *Brevundimonas* can be explained in two ways. Firstly, *Brevundimonas* could have already colonized the advantageous metabolic niches in the early developmental stage of gut microbiota in AHQs, and its rapid proliferation kept it at high levels throughout the establishment of the gut microbiota community (high absolute and high relative abundance). The other explanation could be that it was not the rapid proliferation of *Brevundimonas*, but a slower growth rate of other bacteria, making the dominance of *Brevundimonas*.

Early Developmental Patterns of the Core Members

To explore the developmental patterns of the core members of queens' gut microbiota in host early development, we used qPCR to measure the absolute abundance of total bacteria and of the seven main taxa (including three dominant taxa of queens and four core members of workers) in three groups of 1-, 4-, and 7-day-old queens (parameters of the resultant standard curve are listed in **Supplementary Table 2**). In general, specific bacteria exhibited different proliferation patterns among the different queens (**Figure 5**). The time at which the absolute abundance of specific bacteria reached the highest varied among different queens and in different gut sections. The total bacteria in the midgut of AMQs remained the highest throughout the host's early development stage, followed by ACQs, and the total number of bacteria was the lowest in AHQs ($p < 0.05$). The total bacteria in the other two groups showed a decreased rate of proliferation,

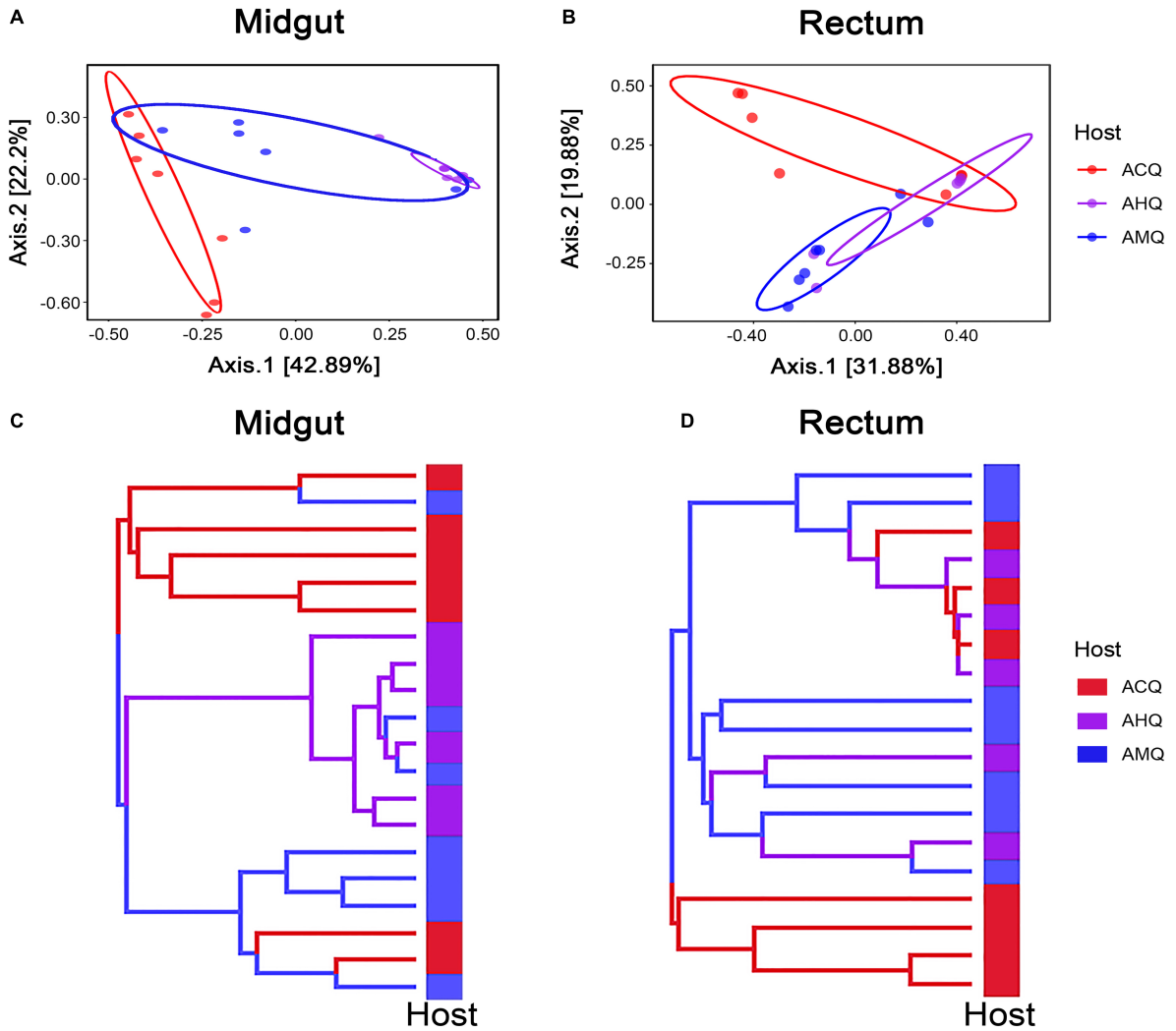

FIGURE 3 | High microbiota similarities between AHQs and AMQs. **(A,B)** β-Diversities of bacterial communities are clustered using principal coordinate analysis (PCoA) based on Bray–Curtis dissimilarity (results of unweighted UniFrac distance are shown in **Supplementary Figures 2A,B**). Each dot on the plot represents the entire microbiota of a gut tissue sample, and the dots were colored correspondingly. **(C,D)** The dendrograms rebuilt using the UPGMA method based on Bray–Curtis dissimilarity. Each topological branch is also colored the same as before.

while bacteria in AMQs maintained a high growth rate. The situation in the rectum was reversed in 7-day-old queens. The bacteria in ACQs proliferated explosively from day 4 to day 7, when it reached the highest, surpassing the numbers in the AMQs ($p < 0.05$). Inexplicably, the number of total bacteria from AHQs remained at a low level during the early developmental stage and showed no growth trend. Changes in absolute abundance of *Brevundimonas* (genetic-specific taxon) in AHQs showed the same pattern of development as in AMQs: high abundance but slow growth rate in the midgut and consistently low abundance in the rectum. However, the other two dominant taxa (*Lactobacillus kunkeei* and *Bombella*) showed unique growth patterns, different from both AMQs and ACQs, and maintained low abundance levels throughout, which were consistent with the results of high-throughput sequencing. Lastly, relatively stable development patterns of the core members of workers were observed in the guts of queens while only a few species of bacteria had extremely high colonization in specific gut sections, like *Bifidobacterium asteroides* and *Lactobacillus Firm4* and *Firm5* in the rectum, indicating differences between core members in the gut microbiota of queens and workers.

DISCUSSION

Overall, our high-throughput sequencing and absolute qPCR results suggest that the rate of core bacterial proliferation of the three groups of queens varied greatly at each time point. The core genera were *Brevundimonas*, *Lactobacillus*, and *Bombella*, and the major phyla were Proteobacteria and Firmicutes. These results differed from the results with adult *A. mellifera* queens in which the core members included *L. Firm5* (51.3%), *Parasaccharibacter*

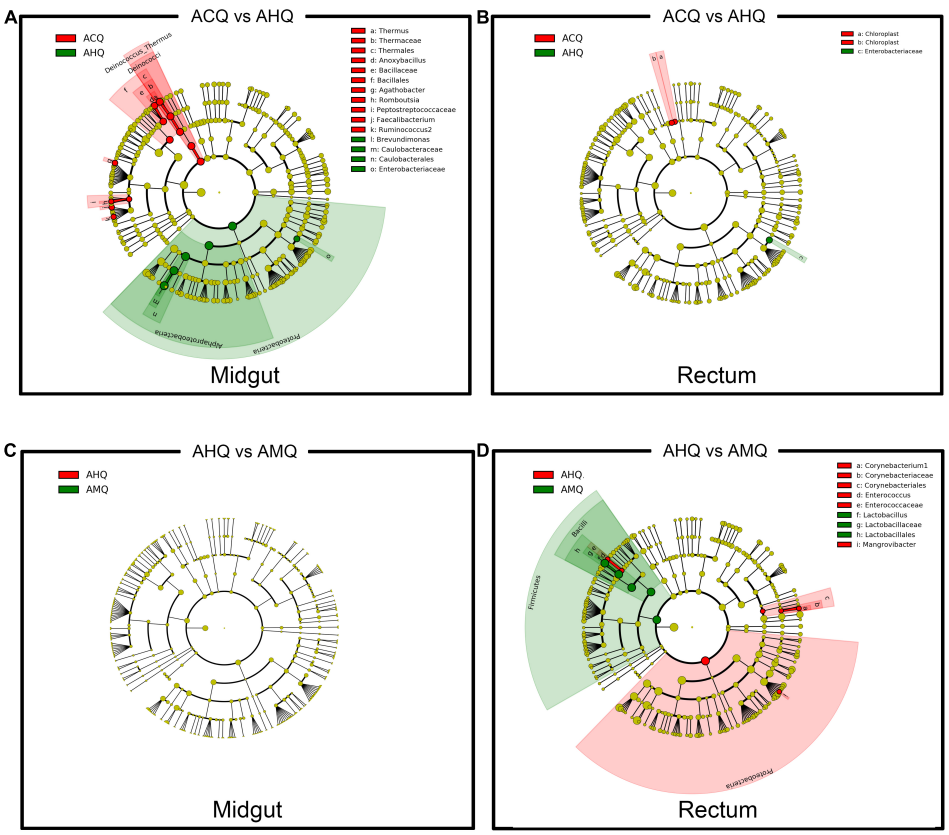

FIGURE 4 | Genetic-specific and environmental-specific taxa filtrating through intergroup comparison. **(A,B)** The differentially abundant taxa of the midgut and rectum comparing ACQs (red) vs. AHQs (green) were identified using LEfSe analysis and displayed in color to detect the genetic-specific taxa. **(C,D)** The other comparison between AHQs (red) and AMQs (green) in the midgut and rectum was also performed to identify the environmental-specific taxa. Each circle diameter is proportional to the taxon abundance.

apium (*Bombella*) (27.1%), and *L. kunkeei* (7.6%) (Anderson et al., 2018; Powell et al., 2018), especially for *Brevundimonas*, which were hardly ever reported. The reason for the difference may be attributed to the age of the queen at sampling as the gut changes with age. The gut microbiota of queens was dynamically changing, and that is why we focused on comparing the early establishment of core members between different queen types. It was worth noting that verifying the gut microbiota of adult and mated queens is equally valuable yet unsuitable for exploring the determinant of the gut microbiota of queens in their early development stage. Therefore, this study was only limited to the perspective of the study on the valuation of the gut microbial community in queens' early stage.

Apis mellifera and *A. cerana* belong to the same genus, but they are quite different in morphology and anatomy (Li et al., 2012) and also in their gut microbiota. In our results, the composition of the gut microbiota of AHQs and AMQs was relatively stable and predictable while the relative abundance of certain taxa varied in each ACQ sample. In most AHQ and AMQ samples, the main microbial members accounted for more than 95% of the total, while the ACQ samples were variable. Interestingly, *Brevundimonas*, a non-core member of the gut microbiota of workers, was almost dominant in the gut microbiota of AHQ and AMQ samples and could also be detected in ACQ rectum samples. This suggests that the initial rectal microbial community of ACQ may have two types, one dominated by *Lactobacillus* and the other by *Brevundimonas*. In terms of microbial composition, the AMQ's gut microbes were more likely to be the further developmental model of AHQ. In this model, the relative abundance of *Brevundimonas* decreased gradually with increasing age of the host while the proportion of other core members increased gradually. The composition of the gut microbiota changed from a single taxon predominating to one in which there were multiple core members dominating the totals. As a result, we speculated, the functional roles of the bacterial community gradually improved and interaction with the host increased. Based on the data available, we suggested that the composition of the gut microbial community was basically determined by the queens' genetic background, while environmental factors influenced the rate of improvement of the gut microbial community. In our study, all non-genetic factors were generally combined as environmental factors. However, the environmental factors were actually composed of a variety of different elements. The weight of each factor should be taken into account when further studying the influence of the environment on the microbiota.

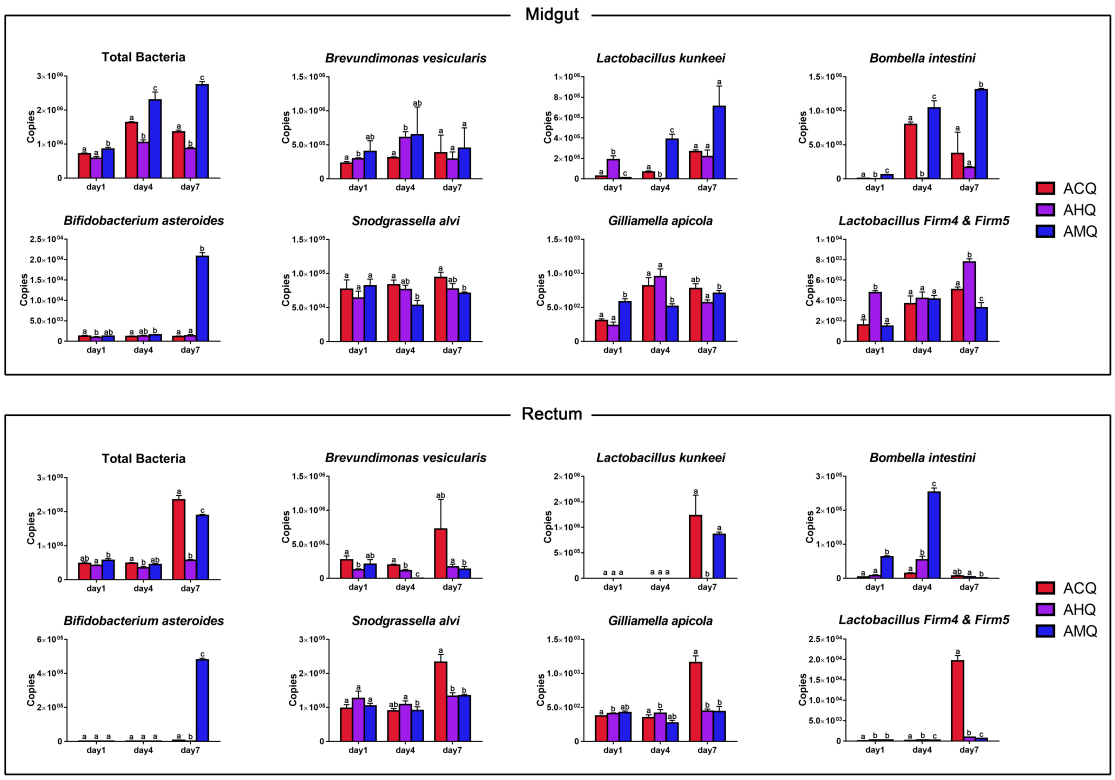

FIGURE 5 | Changes of core members' absolute abundance in queens' early developmental stage (1, 4, and 7 days) measured as copies of the 16S rRNA gene. The columns of host groups were differently colored in red for ACQs, violet for AHQs, and blue for AMQs. Letters above confidence intervals (1 standard deviation) represent significance levels (Tukey's HSD).

Unfortunately, our small sample size may be the major limitation to reach conclusions with high confidence. Although our sample number reached the bottom line of statistical significance, the differences of gut microbiota among organisms were prone to be amplified by a variety of uncertain factors, which further influenced our observations. Moreover, we also tried to transfer the *A. cerana* queen pupae to young AMC to breed the other kind of crossbred queen, which could give the other important reciprocal evidence to support our main conclusion robustly. Unfortunately, four times efforts have ended with 100% failure acceptance (all *A. cerana* fertilized eggs were cleaned out from *A. mellifera* hives), which may be because *A. mellifera* workers had a higher discriminability for other honeybee species than *A. cerana* workers, and these alien individuals were excluded completely as previous studies reported (Rinderer et al., 1985; Robinson, 1987; Ruttner, 1988). The strong rejection of *A. mellifera* to other bee species hinders the applications of crossbreeding, hoping that this hindrance can be removed in the future.

The gut of the honeybee consists of the midgut, ileum, and rectum. The midgut was the main organ for absorbing water and digesting sugars and other nutrients. There were very few bacteria in the midgut of honeybees, which may be because the midgut is unable to provide a stable venue for the establishment and proliferation of bacteria (Martinson et al., 2012; Corby-Harris et al., 2014a). In contrast, the highly anaerobic environment and the stable substrate supply gave the rectum region the greatest concentration of bacteria (Kapheim et al., 2015; Ludvigsen et al., 2015; Jones et al., 2017; Yun et al., 2018). However, our qPCR results showed that the absolute abundance of total bacteria in the midgut was higher overall than that in the rectum, which was contrary to former studies, yet the rate of growth of total rectal bacteria increased sharply between the fourth and seventh days. Together, these results implied that the rectal microbial community was in a rapid development phase and its abundance would soon exceed the microbiota in the midgut at one time point in the future. We found that detecting the significant taxa in the midgut was easier than in the rectum when we tried to screen out genetic-specific taxa, which was also supported by β-diversity analysis results. By comparing α-diversity, we found that both the microbial richness and the variation in evenness of the midgut were greater than in the rectum. In general, these results suggested that the developmental direction of gut microbiota in the midgut and rectum appeared different. The midgut was more likely to show the influence of genetic background during the early life stage of the queens.

In summary, our results from multiple analyses indicated that the genetic background played a more dominant

role than environmental factors in shaping the microbiota of the midgut. Compared to the microbial community of the midgut, the rectal microbiota was more insensitive to the impact of genetic background during the early developmental stage of honeybee queens.

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/ PRJNA751765.

AUTHOR CONTRIBUTIONS

JY and MY designed the study. YuZ and LX collected the samples and data. KL and JY was the project administration. YeZ performed the laboratory analyses while LX and BZ performed the bioinformatics analyses on software. MZ and DiL funded this study. YuZ, MY, and DeL wrote the original draft of manuscript together. All authors contributed to the article and approved the submitted version.

FUNDING

This work was sponsored by grants from the Sichuan Crops and Animals Breeding Special Project in the 14th 5-Year Plan (No. 2021YFYZ0004).

ACKNOWLEDGMENTS

We thank Fanli Kong for her helpful advices about data analysis, and Weidong He and Liuchu Pan for helping in beekeeping.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.722901/full#supplementary-material

Supplementary Figure 1 | The barplot of the composition of queens' gut microbial community in phylum level.

Supplementary Figure 2 | (A) The principal co-ordinates analysis (PCoA) plot based on the unweighted UniFrac matrix of queens' midgut microbiota. **(B)** The principal co-ordinates analysis (PCoA) plot based on the unweighted UniFrac matrix of queens' rectum microbiota.

Supplementary Figure 3 | The dendrograms rebuilt using UPGMA method based on unweighted UniFrac distance.

Supplementary Figure 4 | The top 45 diverse taxa detected by RFC models.

REFERENCES

Amato, K. R., Martinez-Mota, R., Righini, N., Raguet-Schofield, M., Corcione, F. P., Marini, E., et al. (2016). Phylogenetic and ecological factors impact the gut microbiota of two neotropical primate species. *Oecologia* 180, 717–733. doi: 10.1007/s00442-015-3507-z

Amir, A., McDonald, D., Navas-Molina, J. A., Kopylova, E., Morton, J. T., Zech Xu, Z., et al. (2017). Deblur rapidly resolves single-nucleotide community sequence patterns. *mSystems* 2:e00191-16. doi: 10.1128/mSystems.00191-16

Anderson, K. E., Ricigliano, V. A., Mott, B. M., Copeland, D. C., Floyd, A. S., and Maes, P. (2018). The queen's gut refines with age: longevity phenotypes in a social insect model. *Microbiome* 6:108. doi: 10.1186/s40168-018-0489-1

Beals, E. W. (1984). "Bray-curtis ordination: an effective strategy for analysis of multivariate ecological data," in *Advances in Ecological Research*, eds A. MacFadyen and E. D. Ford (Cambridge, MA: Academic Press), 1–55.

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 37, 852–857. doi: 10.1038/s41587-019-0209-9

Breiman, L. (2001). Random forests. *Mach. Learn.* 45, 5–32. doi: 10.1023/A:1010933404324

Callahan, B. J., Sankaran, K., Fukuyama, J. A., McMurdie, P. J., and Holmes, S. P. (2016). Bioconductor workflow for microbiome data analysis: from raw reads to community analyses. *F1000Res* 5:1492. doi: 10.12688/f1000research.8986.2

Carmody, R. N., Gerber, G. K., Luevano, J. M. Jr., Gatti, D. M., Somes, L., Svenson, K. L., et al. (2015). Diet dominates host genotype in shaping the murine gut microbiota. *Cell Host Microbe* 17, 72–84. doi: 10.1016/j.chom.2014.11.010

Chen, J., Bittinger, K., Charlson, E. S., Hoffmann, C., Lewis, J., Wu, G. D., et al. (2012). Associating microbiome composition with environmental covariates using generalized unifrac distances. *Bioinformatics* 28, 2106–2113. doi: 10.1093/bioinformatics/bts342

Corby-Harris, V., Snyder, L. A., Schwan, M. R., Maes, P., McFrederick, Q. S., and Anderson, K. E. (2014b). Origin and effect of alpha 2.2 acetobacteraceae in honey bee larvae and description of parasaccharibacter apium gen. nov., sp. nov. *Appl. Environ. Microbiol.* 80, 7460–7472. doi: 10.1128/aem.02043-14

Corby-Harris, V., Maes, P., and Anderson, K. E. (2014a). The bacterial communities associated with honey bee (*Apis mellifera*) foragers. *PLoS One* 9:e95056. doi: 10.1371/journal.pone.0095056

David, L. A., Maurice, C. F., Carmody, R. N., Gootenberg, D. B., Button, J. E., Wolfe, B. E., et al. (2014). Diet rapidly and reproducibly alters the human gut microbiome. *Nature* 505, 559–563. doi: 10.1038/nature12820

Engel, P., and Moran, N. A. (2013). The gut microbiota of insects - diversity in structure and function. *FEMS Microbiol. Rev.* 37, 699–735. doi: 10.1111/1574-6976.12025

Foster, K. R., Schluter, J., Coyte, K. Z., and Rakoff-Nahoum, S. (2017). The evolution of the host microbiome as an ecosystem on a leash. *Nature* 548, 43–51. doi: 10.1038/nature23292

Griffin, N. W., Ahern, P. P., Cheng, J., Heath, A. C., Ilkayeva, O., Newgard, C. B., et al. (2017). Prior dietary practices and connections to a human gut microbial metacommunity alter responses to diet interventions. *Cell Host Microbe* 21, 84–96. doi: 10.1016/j.chom.2016.12.006

Guo, J., Wu, J., Chen, Y., Evans, J. D., Dai, R., Luo, W., et al. (2015). Characterization of gut bacteria at different developmental stages of Asian honey bees, *Apis cerana*. *J. Invertebr. Pathol.* 127, 110–114. doi: 10.1016/j.jip.2015.03.010

Jeyaprakash, A., Hoy, M. A., and Allsopp, M. H. (2003). Bacterial diversity in worker adults of *Apis mellifera* capensis and *Apis mellifera* scutellata (Insecta: Hymenoptera) assessed using 16S rRNA sequences. *J. Invertebr. Pathol.* 84, 96–103. doi: 10.1016/j.jip.2003.08.007

Jones, J. C., Fruciano, C., Hildebrand, F., Al Toufalilia, H., Balfour, N. J., Bork, P., et al. (2017). Gut microbiota composition is associated with environmental landscape in honey bees. *Ecol. Evol.* 8, 441–451. doi: 10.1002/ece3.3597

Kapheim, K. M., Rao, V. D., Yeoman, C. J., Wilson, B. A., White, B. A., Goldenfeld, N., et al. (2015). Caste-specific differences in hindgut microbial communities of honey bees (*Apis mellifera*). *PLoS One* 10:e0123911. doi: 10.1371/journal.pone.0123911

Knowles, S. C. L., Eccles, R. M., and Baltrūnaitė, L. (2019). Species identity dominates over environment in shaping the microbiota of small mammals. *Ecol. Lett.* 22, 826–837. doi: 10.1111/ele.13240

Korach-Rechtman, H., Freilich, S., Gerassy-Vainberg, S., Buhnik-Rosenblau, K., Danin-Poleg, Y., Bar, H., et al. (2019). Murine genetic background has a stronger impact on the composition of the gut microbiota than maternal inoculation or exposure to unlike exogenous microbiota. *Appl. Environ. Microbiol.* 85:e00826-19. doi: 10.1128/aem.00826-19

Kwong, W. K., Medina, L. A., Koch, H., Sing, K.-W., Soh, E. J. Y., Ascher, J. S., et al. (2017). Dynamic microbiome evolution in social bees. *Sci. Adv.* 3:e1600513. doi: 10.1126/sciadv.1600513

Li, J., Qin, H., Wu, J., Sadd, B. M., Wang, X., Evans, J. D., et al. (2012). The prevalence of parasites and pathogens in Asian honeybees *Apis cerana* in China. *PLoS One* 7:e47955. doi: 10.1371/journal.pone.0047955

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

Ludvigsen, J., Porcellato, D., L'Abée-Lund, T. M., Amdam, G. V., and Rudi, K. (2017). Geographically widespread honeybee-gut symbiont subgroups show locally distinct antibiotic-resistant patterns. *Mol. Ecol.* 26, 6590–6607.

Ludvigsen, J., Rangberg, A., Avershina, E., Sekelja, M., Kreibich, C., Amdam, G., et al. (2015). Shifts in the midgut/pyloric microbiota composition within a honey bee apiary throughout a season. *Microb. Environ.* 30, 235–244. doi: 10.1264/jsme2.ME15019

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

Marcobal, A., and Sonnenburg, J. L. (2012). Human milk oligosaccharide consumption by intestinal microbiota. *Clin. Microbiol. Infect.* 18, 12–15. doi: 10.1111/j.1469-0691.2012.03863.x

Martinson, V. G., Danforth, B. N., Minckley, R. L., Rueppell, O., Tingek, S., and Moran, N. A. (2011). A simple and distinctive microbiota associated with honey bees and bumble bees. *Mol. Ecol.* 20, 619–628. doi: 10.1111/j.1365-294X.2010.04959.x

Martinson, V. G., Moy, J., and Moran, N. A. (2012). Establishment of characteristic gut bacteria during development of the honeybee worker. *Appl. Environ. Microbiol.* 78, 2830–2840. doi: 10.1128/aem.07810-11

Maynard, C. L., Elson, C. O., Hatton, R. D., and Weaver, C. T. (2012). Reciprocal interactions of the intestinal microbiota and immune system. *Nature* 489, 231–241. doi: 10.1038/nature11551

Paradis, E., Claude, J., and Strimmer, K. (2004). APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20, 289–290.

Powell, J. E., Eiri, D., Moran, N. A., and Rangel, J. (2018). Modulation of the honey bee queen microbiota: effects of early social contact. *PLoS One* 13:e0200527. doi: 10.1371/journal.pone.0200527

Powell, J. E., Martinson, V. G., Urban-Mead, K., and Moran, N. A. (2014). Routes of acquisition of the gut microbiota of the honey bee *Apis mellifera*. *Appl. Environ. Microbiol.* 80, 7378–7387. doi: 10.1128/aem.01861-14

R Core Team. (2015). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Ren, T., Boutin, S., Humphries, M. M., Dantzer, B., Gorrell, J. C., Coltman, D. W., et al. (2017). Seasonal, spatial, and maternal effects on gut microbiome in wild red squirrels. *Microbiome* 5:163. doi: 10.1186/s40168-017-0382-3

Rinderer, T. E., Hellmich, R. L. II, Danka, R. G., and Collins, A. M. (1985). Male reproductive parasitism: a factor in the africanization of European honey-bee populations. *Science* 228, 1119–1121. doi: 10.1126/science.228.4703.1119

Robinson, G. E. (1987). Regulation of honey bee age polyethism by juvenile hormone. *Behav. Ecol. Sociobiol.* 20, 329–338. doi: 10.1007/BF00300679

Ruttner, F. (1988). *Biogeography and Taxonomy of Honeybees*. New York, NY: Springer Verlag.

Sabree, Z. L., Hansen, A. K., and Moran, N. A. (2012). Independent studies using deep sequencing resolve the same set of core bacterial species dominating gut communities of honey bees. *PLoS One* 7:e41250. doi: 10.1371/journal.pone.0041250

Schwarz, R. S., Moran, N. A., and Evans, J. D. (2016). Early gut colonizers shape parasite susceptibility and microbiota composition in honey bee workers. *Proc. Natl. Acad. Sci. U. S. A.* 113, 9345–9350. doi: 10.1073/pnas.1606631113

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

Sekirov, I., Russell, S. L., Antunes, L. C., and Finlay, B. B. (2010). Gut microbiota in health and disease. *Physiol. Rev.* 90, 859–904. doi: 10.1152/physrev.00045.2009

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 379–423. doi: 10.1002/j.1538-7305.1948.tb01338.x

Sonnenburg, E. D., Smits, S. A., Tikhonov, M., Higginbottom, S. K., Wingreen, N. S., and Sonnenburg, J. L. (2016). Diet-induced extinctions in the gut microbiota compound over generations. *Nature* 529, 212–215. doi: 10.1038/nature16504

Tarpy, D. R., Mattila, H. R., and Newton, I. L. (2015). Development of the honey bee gut microbiome throughout the queen-rearing process. *Appl. Environ. Microbiol.* 81, 3182–3191. doi: 10.1128/aem.00307-15

Thijs, S., Op De Beeck, M., Beckers, B., Truyens, S., Stevens, V., Van Hamme, J. D., et al. (2017). Comparative evaluation of four bacteria-specific primer pairs for 16S rRNA gene surveys. *Front. Microbiol.* 8:494. doi: 10.3389/fmicb.2017.00494

Van der Sluijs, J. P., and Vaage, N. S. (2016). Pollinators and global food security: the need for holistic global stewardship. *Food Ethics* 1, 75–91. doi: 10.1007/s41055-016-0003-z

Wang, Y., and Qian, P. Y. (2009). Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. *PLoS One* 4:e7401. doi: 10.1371/journal.pone.0007401

Wopereis, H., Oozeer, R., Knipping, K., Belzer, C., and Knol, J. (2014). The first thousand days - intestinal microbiology of early life: establishing a symbiosis. *Pediatr. Allergy Immunol.* 25, 428–438. doi: 10.1111/pai.12232

Wu, Y., Zheng, Y., Chen, Y., Wang, S., Chen, Y., Hu, F., et al. (2020). Honey bee (*Apis mellifera*) gut microbiota promotes host endogenous detoxification capability via regulation of P450 gene expression in the digestive tract. *Microb. Biotechnol.* 13, 1201–1212. doi: 10.1111/1751-7915.13579

Xie, X. b, Peng, W. j, and Zeng, Z. j (2008). Breeding of mite-resistant honeybee by using nutritional crossbreed technology. *Sci. Agric. Sin.* 41, 1530–1535.

Yun, J. H., Jung, M. J., Kim, P. S., and Bae, J. W. (2018). Social status shapes the bacterial and fungal gut communities of the honey bee. *Sci. Rep.* 8:2019. doi: 10.1038/s41598-018-19860-7

Zapala, M. A., and Schork, N. J. (2006). Multivariate regression analysis of distance matrices for testing associations between gene expression patterns and related variables. *Proc. Natl. Acad. Sci. U. S. A.* 103, 19430–19435. doi: 10.1073/pnas.0609333103

Zeng, Z. J, Xie, X. B, and Xue, Y. B (2005a). Effects of nutritional crossbreeding between *Apis cerana* cerana and *Apis mellifera* ligustica on morphological characters of worker bees. *Acta Agric. Univ. Jiangxiensis* 27, 454–457.

Zeng, Z. J, Xie, X. B, and Yan, W. Y (2005b). Effects of nutritional crossbreed on emergent weight of worker bees. *J. Econ. Anim.* 9:149.

Zheng, H., Nishida, A., Kwong, W. K., Koch, H., Engel, P., Steele, M. I., et al. (2016). Metabolism of toxic sugars by strains of the bee gut symbiont *Gilliamella apicola*. *mBio* 7:e01326-16.

Zheng, H., Powell, J. E., Steele, M. I., Dietrich, C., and Moran, N. A. (2017). Honeybee gut microbiota promotes host weight gain via bacterial metabolism and hormonal signaling. *Proc. Natl. Acad. Sci. U. S. A.* 114, 4775–4780. doi: 10.1073/pnas.1701819114

Zhuang, M. (1985). A preliminary study on asexual hybridization of *A.cerana* amd *A.mellifera*. *J. Honeybee* (in Chinese).

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.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 29 November 2021
doi: 10.3389/fmicb.2021.758061

Genetic Diversity and Natural Selection of *Plasmodium vivax* Duffy Binding Protein-II From China-Myanmar Border of Yunnan Province, China

Tian-Qi Shi[1,2,3,4], Hai-Mo Shen[1,2,3,4], Shen-Bo Chen[1,2,3,4], Kokouvi Kassegne[5], Yan-Bing Cui[1,2,3,4], Bin Xu[1,2,3,4], Jun-Hu Chen[1,2,3,4,5], Bin Zheng[1,2,3,4,5]* and Yue Wang[6]**

[1] National Institute of Parasitic Diseases, Chinese Center for Diseases Control and Prevention (Chinese Center for Tropical Diseases Research), Shanghai, China, [2] National Health Commission of the People's Republic of China (NHC) Key Laboratory of Parasite and Vector Biology, Shanghai, China, [3] World Health Organization (WHO) Collaborating Center for Tropical Diseases, Shanghai, China, [4] National Center for International Research on Tropical Diseases, Shanghai, China, [5] School of Global Health, Chinese Center for Tropical Diseases Research, Shanghai Jiao Tong University School of Medicine, Shanghai, China, [6] School of Basic Medical Sciences and Forensic Medicine, Hangzhou Medical College, Institute of Parasitic Diseases, Hangzhou, China

Malaria incidence has declined dramatically over the past decade and China was certified malaria-free in 2021. However, the presence of malaria in border areas and the importation of cases of malaria parasites are major challenges for the consolidation of the achievements made by China. *Plasmodium vivax* Duffy binding protein (PvDBP) performs a significant role in erythrocyte invasion, and is considered a promising *P. vivax* vaccine. However, the highly polymorphic region of PvDBP (PvDBP-II) impedes the development of blood-stage vaccine against *P. vivax*. In this study, we investigated the genetic diversity and natural selection of PvDBP-II among 124 *P. vivax* isolates collected from the China-Myanmar border (CMB) in Yunnan Province, China, during 2009–2011. To compare genetic diversity, natural selection, and population structure with CMB isolates, 85 *pvdbp*-II sequences of eastern Myanmar isolates were obtained from GenBank. In addition, global sequences of *pvdbp*-II were retrieved from GenBank to establish genetic differentiation relationships and networks with the CMB isolates. In total, 22 single nucleotide polymorphisms reflected in 20 non-synonymous and two synonymous mutations were identified. The overall nucleotide diversity of PvDBP-II from the 124 CMB isolates was 0.0059 with 21 haplotypes identified (*Hd* = 0.91). The high ratio of non-synonymous to synonymous mutations suggests that PvDBP-II had evolved under positive selection. Population structure analysis of the CMB and eastern Myanmar isolates were optimally grouped into five sub-populations (*K* = 5). Polymorphisms of PvDBP-II display that CMB isolates were genetically diverse. Mutation, recombination,

OPEN ACCESS

Edited by:
Wei Huang,
Johns Hopkins University,
United States

Reviewed by:
Bhavna Gupta,
Vector Control Research Centre (ICMR), India
Lady Johanna Forero Rodriguez,
National University of Colombia, Colombia

***Correspondence:**
Jun-Hu Chen
chenjh@nipd.chinacdc.cn
Bin Zheng
zhengbin@nipd.chinacdc.cn
Yue Wang
wangyuerr@hotmail.com

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *13 August 2021*
Accepted: *12 October 2021*
Published: *29 November 2021*

Citation:
Shi T-Q, Shen H-M, Chen S-B, Kassegne K, Cui Y-B, Xu B, Chen J-H, Zheng B and Wang Y (2021) Genetic Diversity and Natural Selection of Plasmodium vivax Duffy Binding Protein-II From China-Myanmar Border of Yunnan Province, China. Front. Microbiol. 12:758061. doi: 10.3389/fmicb.2021.758061

and positive selection promote polymorphism of PvDBP-II of *P. vivax* population. Although low-level genetic differentiation in eastern Myanmar was identified along with the more effective malaria control measures, the complexity of population structure in malaria parasites has maintained. In conclusion, findings from this study advance knowledge of the understanding of the dynamic of *P. vivax* population, which will contribute to guiding the rational design of a PvDBP-II based vaccine.

Keywords: *Plasmodium vivax*, Duffy binding protein, genetic diversity, natural selection, China-Myanmar border

INTRODUCTION

Malaria, a common and life-threatening infectious disease, affects 2.29 billion people worldwide. *Plasmodium vivax* infected an estimated 5.3 million people in 2019, with most of the cases concentrated in Southeast Asia (WHO, 2020). The unique biology of *P. vivax* and repeated emergences of antimalarial resistance to vivax infections make this malaria parasite more difficult to control and eliminate than *P. falciparum* (Price et al., 2007). The Greater Mekong Subregion (GMS) in Southeast Asia has committed to achieve the goal of malaria elimination by 2030 (WHO, 2016). The member countries of GMS are Laos, Vietnam, Myanmar, Thailand, Cambodia, and China. Myanmar has the highest malaria burden in the GMS, with a major challenge for disease control and elimination, especially to neighboring countries such as China (Yunnan and Guangxi Provinces) and Thailand (Lo et al., 2017; Brashear et al., 2020). The China-Myanmar border (CMB) was once the area where the burden of malaria was highest in the world. In 2010, China has formulated the "Action plan of China malaria elimination (2010–2020)" and decided to accomplish the target of eliminating malaria across the country by 2020 (Feng et al., 2016). Malaria has substantially decreased over the past 10 years. Although China has been certified malaria-free in 2021, the country is still cautious about imported cases and malaria retransmission. In addition, malaria control in the CMB area is particularly a challenge owing to the high prevalence of malaria in Myanmar, resulting in continuous transmission and infection (Feng et al., 2020). Thus, cases of imported malaria pose a threat for sustainable malaria-free status within the borders of China. The development of an effective malaria vaccine is therefore a research priority for integrated malaria control. One of the main barriers for vaccine design is the antigenic diversity of candidate genes (Takala and Plowe, 2009) and it is essential to identify the determinants of antigenic variation in malaria-endemic areas. Meanwhile, with the progress in the implementation of malaria elimination measures, it is also reasonable to speculate whether there was a change in *P. vivax* population structure.

Plasmodium vivax invasion depends on a specific receptor-ligand interaction between the parasite and host red blood cells. The *P. vivax* Duffy binding protein (PvDBP) binds to the corresponding receptor Duffy antigen receptor for chemokines (DARC) to form a tight junction that is vital for merozoite invasion. Therefore, PvDBP is increasingly becoming an attractive vaccine molecule against vivax malaria. It could induce strong acquired immune responses in subjects from vivax endemic areas, and more importantly, antibodies against the domain II of PvDBP (PvDBP-II) block the interaction with DARC that prevent merozoite invasion from human erythrocytes (Grimberg et al., 2007; Longley et al., 2017). The PvDBP is a 140-kDa protein characterized by a functionally conserved cysteine-rich region, divided into seven different regions (Wertheimer and Barnwell, 1989). The key erythrocyte binding sites of PvDBP are mapped to a 170-amino acid stretch with cysteine regions (C4–C7) (Chootong et al., 2010). The PvDBP-II is highly polymorphic compared to the rest of the region. Genetic analysis from different geographical *P. vivax* isolates demonstrated that this region is under positive selection, with polymorphic residues varying by geographic region (Ampudia et al., 1996; Xainli et al., 2000; Ju et al., 2013; Chootong et al., 2014; Almeida-de-Oliveira et al., 2020). Despite PvDBP-II presenting a promising vaccine antigenic, diversity may affect the immune recognition and host immune response which may reflect that polymorphisms of PvDBP-II help in immune evasion mechanism (Cole-Tobian and King, 2003). Knowledge of the antigenic variation and natural selection of PvDBP-II is practical significance to fully comprehend the epidemiology of potential vaccine candidates and develop new interventions. Therefore, understanding the nature of genetic polymorphism within PvDBP-II in isolates from different geographic regions is important for developing a broadly reactive DBP-II-based protective vaccine against vivax malaria.

In this study, we collected 124 *P. vivax* samples from the CMB area in Yunnan Province of China at different times to analyze genetic diversity, natural selection, and the dynamic change with the population structure of PvDBP-II in CMB isolates. In addition, the association between CMB and eastern Myanmar *pvdbp*-II sequences were compared.

MATERIALS AND METHODS

Ethics Statement

This study was conducted according to the principles expressed in the Declaration of Helsinki. Before blood collection, the study protocol and potential risks and benefits were explained to the participants, and written informed consent was obtained from each participant. Blood was collected following institutional ethical guidelines reviewed and approved by the Ethics Committee at the National Institute of Parasitic Diseases, Chinese Center for Disease Control and Prevention.

Blood Samples, DNA Extraction, and PCR Amplification

Blood samples were obtained from patients (who met the following inclusion criteria: aged from 18 to 65 years old and diagnosed positive for *P. vivax* infection by microscopic examination of blood smear) from the CMB area in Yunnan Province of China, during 2009–2011 (**Figure 1**). In total, we collected 124 *P. vivax* samples distributed as follows: 35 in 2009, 54 in 2010, and 35 in 2011. The QIAamp DNA mini kit (Qiagen, Germany) was used to extract genomic DNA as reported by previous methods (Chen et al., 2017). The *pvdbp*-II genes of CMB isolates were amplified by PCR using the following specific primers: *pvdbp*-II-F (5′-TGATAGTAAAACTGATAACGG-3′) and *pvdbp*-II-R (5′-TCTGATTTCCATTTTGACCAT-3′). The PCR reaction system was composed of 14 μl ddH_2O, 5 × PrimeSTAR GXL buffer, 2.5 mM dNTP, 10 nM of each primer, 25 U/μl PrimerSTAR GXL Polymerase, and 2.5 μl genomic DNA. The cycling parameters for PCR amplification was performed using the following conditions: 98°C for 5 min, 35 cycles at 98°C for 10 s, 60°C for 15 s, 68°C 1.5 min, followed by extension at 68°C for 5 min. PCR products were examined by 1% agarose and sent to Beijing Genomics Institution (BGI, Shenzhen, China) for sequencing. The sequences have been deposited in the GenBank database under the accession numbers MZ765947-MZ766070.

Genetic Diversity and Natural Selection of Polymorphic Region of PvDBP

To estimate genetic diversity and natural selection of PvDBP-II, the polymorphism between the CMB within-population isolates in 2009–2011 and eastern Myanmar (Laiza, Kachin State) *P. vivax* isolates in 2016 was compared. Sequences of the *pvdbp*-II from eastern Myanmar ($n = 85$, MN233489-MN233573) were obtained from GenBank (Hu et al., 2019). The *pvdbp*-II of Salvador I (DQ156512) was used as a reference sequence. Alignments of all sequences were performed using the CLUSTAL W method in MEGA 7.0 (Kumar et al., 2016). Genetic diversity of *pvdbp*-II was analyzed by using DnaSP 5.0 program (Rozas et al., 2003). Nucleotide diversity (π), number of haplotypes (H), haplotype diversity (Hd), number of segregating sites (S), and mean value of nucleotide differences (K) were calculated to analyze the genetic diversity of *pvdbp*-II. Neutrality of evolution was then detected by applying DnaSP 5.0 to calculate the rates of non-synonymous (dn) to synonymous (ds) mutations. If selective pressure is neutral, the expected value of the dn/ds is 1. However, if non-synonymous mutation is detrimental, the studied population is under purifying selection and the dn/ds is <1. If non-synonymous mutation is beneficial, the dn/ds > 1, and it is under positive selection. Significance was assessed by applying the Z-test using MEGA 7.0. In addition, Tajima's D test was also performed to test neutrality with DnaSP. Tajima's D test estimates the genetic variation within-population; namely, it calculates the number of segregating sites (θ) and the average number of nucleotide (π) differences estimated from pairwise comparison to test departures from the neutral theory of evolution, such as directional selection or balancing selection (Tajima, 1989).

Recombination and Linkage Disequilibrium of Polymorphic Region of PvDBP

The recombination parameter (R: which includes the effective population size and probability of recombination between adjacent nucleotides per generation) and the minimum number of recombination events (Rm) were determined using DnaSP.5.0. Recombination events were evaluated by using ZZ statistic and Rm. Linkage disequilibrium (LD) between different polymorphic sites was computed in terms of the R^2 index using DnaSP for the PvDBP-II of CMB isolates. R^2 for each pair of genetic polymorphisms was plotted on heatmap graphics using the "LD heatmap" package (Rozas et al., 2001; Shin et al., 2006). LD was evaluated using the Zns statistic which represents the average of R^2 overall pairwise comparisons.

Genetic Differentiation, Haplotype Network, and Population Genetic Structure Analysis of Polymorphic Region of PvDBP

To estimate the degree of genetic differentiation of the *pvdbp*-II in global isolates, we download additional *pvdbp-II* sequences from GenBank, including Asia: India (FJ491142.1-FJ491241.1), Iran (EU860429.1-EU860435.1, KF791931.1-KF791925.1), and Thailand (EF379127.1-EF379132.1, EF368159.1-EF368164.1); Oceania: Papua New Guinea (AF289480-AF289483, AF289636-AF289647); Africa: Sudan (MG805621.1-MG805629.1); and South America: Brazil (EU812841.1-EU812845.1; EU812866.1-EU812873.1) (Cole-Tobian et al., 2002; Gosi et al., 2008; Babaeekhou et al., 2009; Sousa et al., 2010; Valizadeh et al., 2014; Hoque et al., 2018). All sequences were aligned and cut to 671 bp by MEGA7.0. Arlequin3.5 with analysis of molecular variance (AMOVA) was used to evaluate fixation (F_{ST}) (Excoffier and Lischer, 2010). To establish a genetic relationship among the CMB, eastern Myanmar, and other countries' haplotypes of PvDBP-II, the haplotype network was analyzed using NETWORK ver. 102200 with the Median-Joining method (Bandelt et al., 1999). The software STRUCTURE (ver. 2.3.4) was used to determine the genetic structure of *P. vivax* population in the CMB and eastern Myanmar (Pritchard et al., 2000). The Bayesian method was used to identify the optimum number of clusters (K). The admixture model for values $K = 3$–9 was run, each with a burn-in period of 10,000 steps and then 100,000 Markov Chain Monte Carlo (MCMC) iterations. The best-fit number of grouping was evaluated using delta K (ΔK) in the STRUCTURE HARVESTER.

RESULTS

Haplotype Variation and Nucleotide Diversity of Polymorphic Region of PvDBP From Different *P. vivax* Isolates

A total of 124 samples from the CMB were successfully amplified and sequenced for the *pvdbp*-II (1077 bp) which

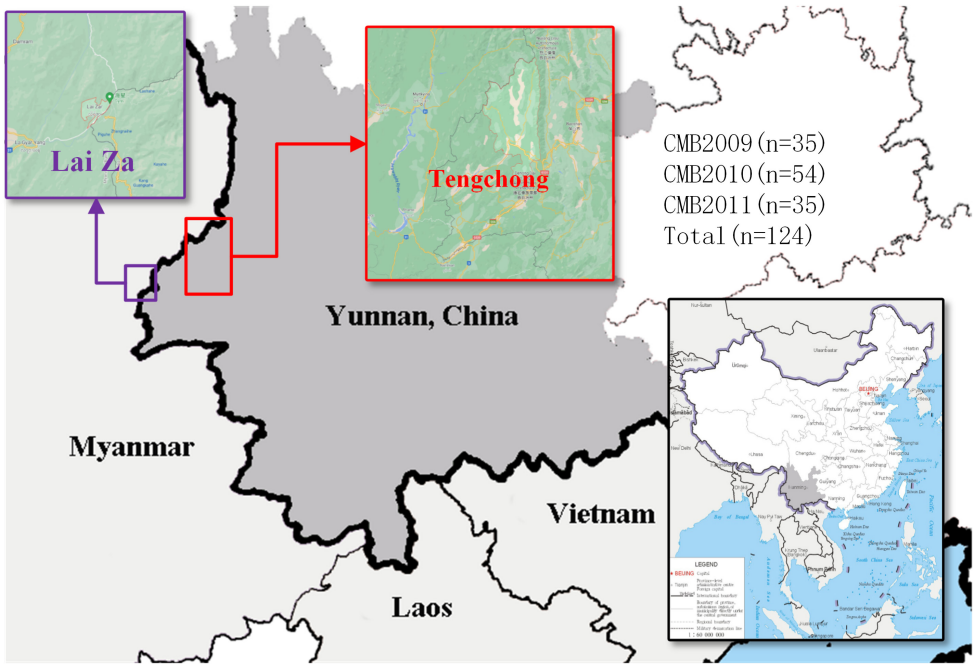

FIGURE 1 | The map of *P. vivax* samples collection. Samples collection was made in Tengchong, marked in red rectangle (China-Myanmar border in Yunnan Province of China).

corresponds to amino acid positions at 222–580 of the Salvador I reference (**Figure 2A**). The 124 CMB samples had 22 single nucleotide polymorphisms (SNPs) of which two and 20 were synonymous and non-synonymous (**Table 1**) polymorphic codons, respectively (nine occurred at the first base of the codon, five at the second, and eight at the third base of the codon), causing significant amino acid changes in protein level. Analysis of the deduced amino acid sequences classified them into 21 haplotypes (Haplotypes 1–21) with amino acid changes at 20 positions, in which all were dimorphic. Haplotype 4 was predominant (n = 27, 21.8%) among CMB isolates (**Figure 2B**).

Nineteen of the 20 non-synonymous changes were previously identified, whereas the remaining one change (Q562H, 1.6%) was unique in CMB isolates, which was not hitherto reported. D834G (83.1%), L424I (71.8%), and W437R (68.6%) SNPs were the most frequent. Comparison of the amino variants observed in PvDBP-II revealed that CMB isolates had a similar pattern compared to eastern Myanmar isolates.

Although eastern Myanmar isolates showed similar amino changes compared to CMB isolates, six variants (F306L, S398T, T404R, P475A, Q486E, and Q562H) detected in the CMB isolates were not found in eastern Myanmar isolates (**Figure 2C**).

Analysis of polymorphism within the 124 *pvdbp*-II CMB isolates revealed lower nucleotide diversity (π = 0.0059) and higher haplotype diversity (*Hd* = 0.91) (**Table 1**). A sliding window plot of π with a window of 100 bp and step size of 25 bp revealed values for all CMB isolates ranging from 0.000 to 0.028. The highest peak of nucleotide diversity within the *pvdbp*-II of CMB isolates was between nucleotide positions 360 and 480. Nucleotide diversity of CMB isolates in 2009, 2010, and 2011 was 0.0060, 0.0056, and 0.0058, respectively, which collectively was a little higher than that of eastern Myanmar isolates (**Figure 3A**).

Natural Selection of Polymorphic Region of PvDBP From Different *P. vivax* Isolates

To determine whether natural selection contributed to the generation of diversity in *pvdbp*-II within the China-Myanmar *P. vivax* population, we calculated the rate of non-synonymous (dn) to synonymous (ds) mutations. The rate of non-synonymous to synonymous substitution (dn/ds) for PvDBP-II for all the 124 CMB isolates was 3.238 (with rates of 2.962, 3.4736, and 3.227 for isolates of 2009, 2010, and 2011, respectively), suggesting that a positive natural selection might occur in the PvDBP-II of CMB isolates. Furthermore, the overall Tajima's D value for PvDBP-II was 1.2158 for all the 124 CMB isolates (with values of 0.9943, 1.5326, and 0.8230 for isolates of 2009, 2010, and 2011, respectively) (**Table 1** and **Figure 3B**). In addition, by sliding the window to detect specific regions under selection, we confirmed that significant positive Tajima's D values (P<0.05) were found in the 1,115–1,185 bp region from the CMB 2009 isolates and the 1,130–1,165 bp region from all the CMB isolates in 2009–2011. Such an informative result suggests a positive balancing selection for PvDBP-II in the CMB population. Moreover, eastern Myanmar isolates were also found under positive balancing selection.

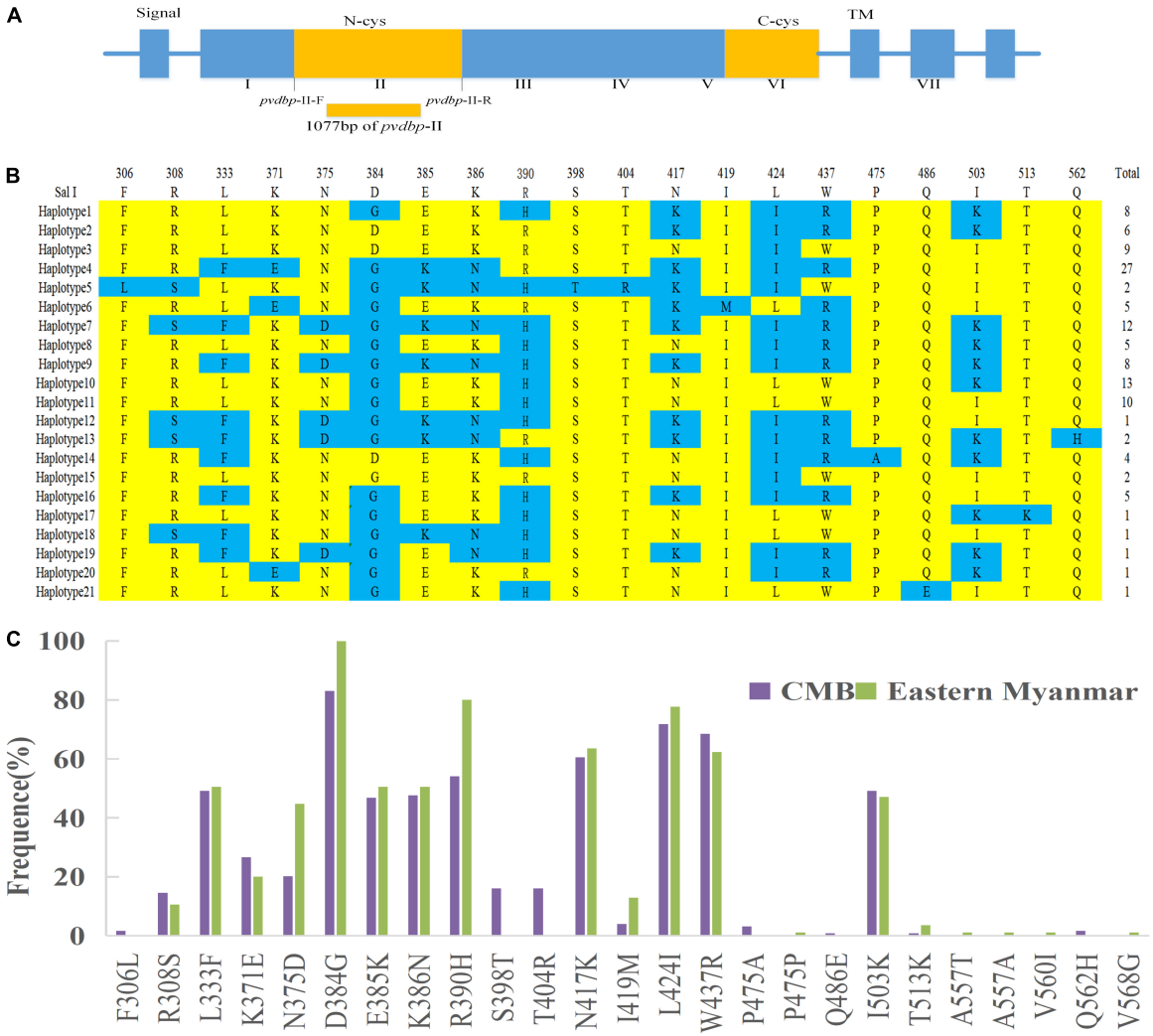

FIGURE 2 | Sequence and polymorphism of PvDBP-II in *P. vivax* CMB isolates. **(A)** The structure of *Plasmodium vivax* Duffy binding protein. **(B)** The changes in amino acid sequences. Polymorphic amino acids are listed for each haplotype. Amino acid residues identical to those of the reference sequence, *Sal* I (DQ156512), are marked in yellow. The dimorphic amino acid changes are marked in blue. Total number of sequences for each haplotype is listed in right panel. **(C)** Frequencies of amino acid changes found in PvDBP-II among CMB and eastern Myanmar isolates.

TABLE 1 | Genetic diversity and natural selection of *pvdbp*-II in China-Myanmar border and eastern Myanmar isolates.

Population	n	S	η	*H*	*Hd*	π	k	Tajima's D	*P*-value	dn	ds	dn/ds	*P*-value
CMB2009	35	19	19	13	0.92	0.006	5.976	0.9943	$P > 0.10$	0.007	0.0024	2.962	$P < 0.05$
CMB2010	54	17	17	14	0.9	0.0056	5.584	1.5326	$P > 0.10$	0.0066	0.0019	3.474	$P < 0.05$
CMB2011	35	20	20	14	0.91	0.006	6.037	0.823	$P > 0.10$	0.0071	0.0022	3.227	$P < 0.05$
CMB2009-2011	124	22	22	21	0.91	0.0059	5.813	1.2158	$P > 0.10$	0.0068	0.0021	3.238	$P < 0.05$
Eastern Myanmar	85	20	21	16	0.85	0.0058	5.784	1.1337	$P > 0.10$	0.0064	0.0038	1.684	$P < 0.05$

CMB, China-Myanmar border; n, number of samples; S, number of segregating sites; η, the total number of mutations; H, number of haplotypes; Hd, haplotype diversity; π, nucleotide diversity; k, the average of nucleotide differences; dn, the rates of non-synonymous substitutions; ds, the rates of synonymous substitutions.

Recombination

The recombination events (Rm) between adjacent polymorphic sites of the 124 CMB isolates was 7 (with values of 6, 5, and 7 for isolates of 2009, 2010, and 2011, respectively) (**Table 2**). In 124 CMB isolates, we found N375D/R378R, E385K/K386N, S398T/T404R, and L424I/W437R on high LD levels. The value of ZZ, Zns, and the decline of R^2 with an increasing distance between the pairs of nucleotide sites suggest that intragenic recombination may also contribute to the diversity of PvDBP-II in CMB isolates (**Figure 4**).

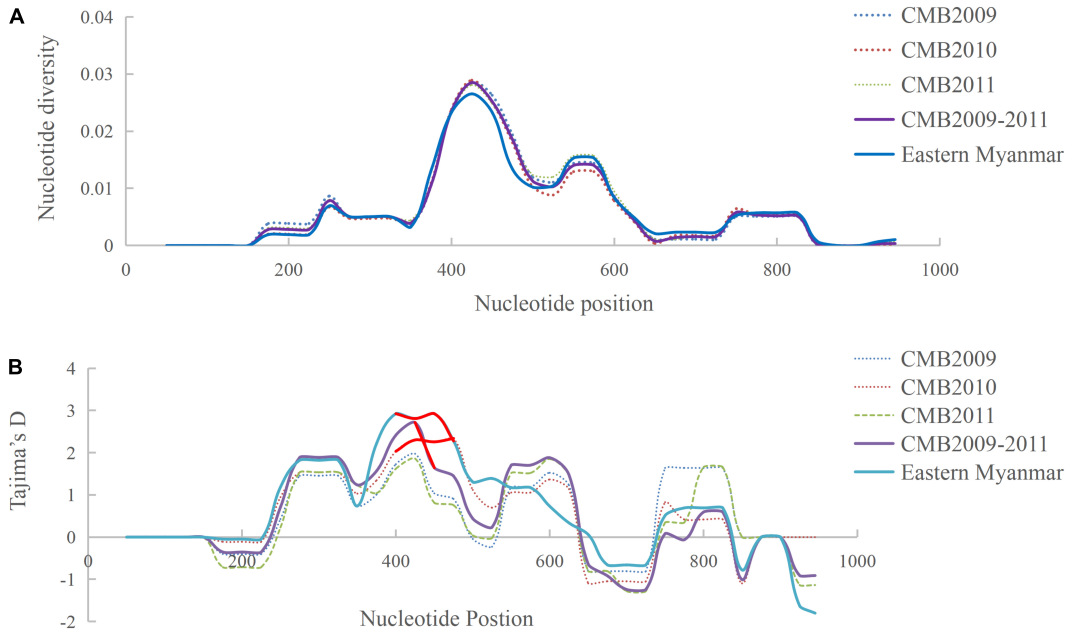

FIGURE 3 | Polymorphism and Tajima's D tests of *pvdbp*-II sequences from CMB and eastern Myanmar isolates. **(A)** Sliding window plots of nucleotide diversity per site (π) comparing the level of genetic diversity at *pvdbp*-II. **(B)** Sliding window plots of Tajima's D for *pvdbp*-II. Statistical significant sites are marked with red lines.

Genetic Differentiation, Population Structure, and Clustering of Polymorphic Region of PvDBP Haplotypes

The level of genetic differentiation of *pvdbp*-II was estimated by F_{ST} values. In general, the value of F_{ST} (0.05–0.15) is poor differentiation, F_{ST} (0.15–0.25) is moderate differentiation, and F_{ST} >0.25 is great differentiation (Balloux and Lugon-Moulin, 2002). The CMB isolates, eastern Myanmar isolates, and Thailand isolates showed a poor genetic differentiation with the values of F_{ST} being 0.0296 and 0.0568, respectively. The high level of genetic differentiation was found between CMB isolates Sudan and Papua New Guinea (PNG) (the values of F_{ST} were 0.2517 and 0.4598, respectively) (**Table 3**).

Collectively, a total of 50 haplotypes were identified from 260 *pvdbp*-II sequences of the CMB, eastern Myanmar, Sudan, Thailand, and PNG isolates, with 26 singleton haplotypes (observed only once). The haplotypes network could be roughly grouped into five major clusters and some small scattered groups (**Figure 5A**). Five major clusters based on geographical distribution were CMB isolates, Myanmar mix isolates, PNG isolates, Brazil isolates, and Sudan isolates. In the network analysis, the haplotype prevalence ranged from 0.38 to 12.69%. Haplotype 4 was shared among CMB, eastern Myanmar, and Thailand. Haplotype 7 and 11 were shared among CMB isolates and eastern Myanmar isolates with relatively high frequencies. H5 and 13 were shared with CMB and Thailand isolates. Whilst H1, H2, H3, H8, H10, H12, H15, H17, H18, H19, and H20 only existed in CMB isolates, H10 was the dominant haplotype.

With regard to the population structure of the haplotype of *pvdbp*-II from the CMB and eastern Myanmar isolates, STRUCTURE analysis showed a clear distribution of haplotypes and demonstrated multiple sub-populations. The haplotypes were optimally grouped into five sub-populations ($K = 5$; **Figure 5B**). Three sub-populations of malaria parasites from CMB in 2009–2011 had admixed haplotypes which showed a similar distribution. Compared with CMB populations, the eastern Myanmar population was composed of four sub-populations and a significant decrease ($P < 0.05$) was noted in the K3 sub-population, while a significant increase ($P < 0.05$) was found in the proportion of the K1 sub-population (**Figure 5B**).

TABLE 2 | Linkage disequilibrium and recombination of *pvdbp*-II in China-Myanmar border and eastern Myanmar isolates.

Population	Rm	Zns	ZZ
CMB2009	6	0.151	0.118
CMB2010	5	0.150	0.077
CMB2011	7	0.138	0.103
CMB2009-2011	7	0.104	0.112
Eastern Myanmar	5	0.166	0.077

CMB, China-Myanmar border; Rm, minimum number of recombination events.

DISCUSSION

Due to vector species, host genetics, and environmental factors, *P. vivax* populations harbor different genetic diversities under different geographies (Arnott et al., 2012). The high level of diversity within *Plasmodium* antigens is a major challenge for effective malaria vaccines. Compared to merozoite surface

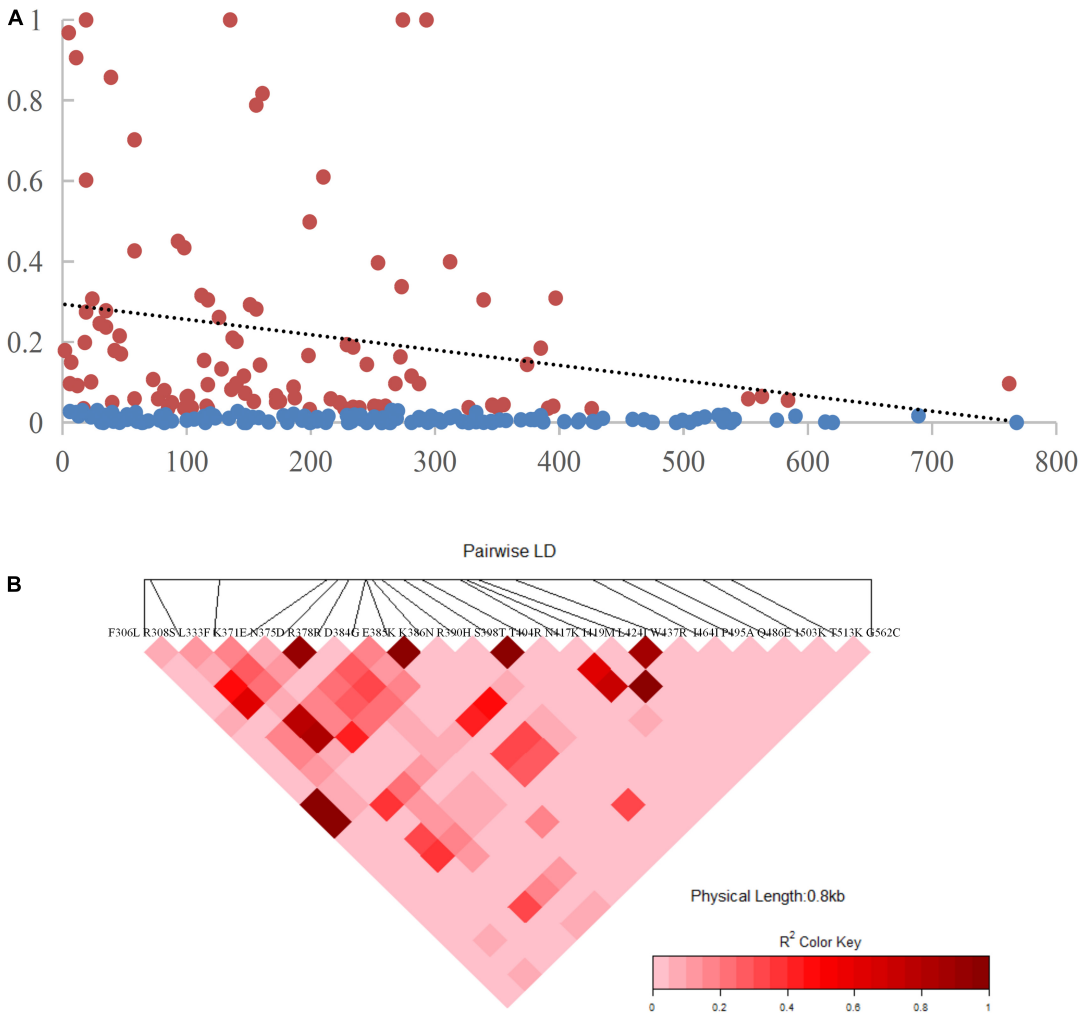

FIGURE 4 | Linkage disequilibrium (LD) plot showing non-random association between nucleotide variants among 124 CMB *P. vivax* isolates at different polymorphic sites. **(A)** The significant LD sites detected by Fisher's exact test are marked red pots and others are marked blue pots. The dashed line represents the regression. **(B)** R^2 for each pair of genetic polymorphisms of CMB isolates.

TABLE 3 | Estimation of genetic differentiation (Fst) of the *pvdbp*-II among other geographical populations.

Population	Brazil	CMB	India	Iran	Eastern Myanmar	PNG	Sudan	Thailand
Brazil (*n* = 13)	0.0000							
CMB (*n* = 124)	0.2764	0.0000						
India (*n* = 100)	0.1566	0.0787	0.0000					
Iran (*n* = 122)	0.1281	0.0816	0.0128	0.0000				
Eastern Myanmar (*n* = 85)	0.3348	0.0296	0.12387	0.1374	0.0000			
PNG (*n* = 16)	0.5709	0.4598	0.4157	0.4338	0.4787	0.0000		
Sudan (*n* = 10)	0.6353	0.2517	0.2775	0.2974	0.3143	0.7044	0.0000	
Thailand (*n* = 12)	0.2498	0.0568	0.0817	0.1083	0.0726	0.4150	0.3421	0.0000

CMB, China-Myanmar border; PNG, Papua New Guinea.

proteins (MSPs) and circumsporozoite protein (CSP), PvDBP-II bears less polymorphism (Shen et al., 2015). Hence, PvDBP-II is considered as the primary fragment of candidate antigen for vaccine against malaria. Investigation on polymorphism and population structure of PvDBP-II may provide deeper criteria for the selection of vaccine candidates. In the present investigation, we analyzed the genetic diversity and molecular evolution of PvDBP-II in the CMB populations in Yunnan Province of China and compared them to that of eastern Myanmar populations. From the 124 CMB *P. vivax* isolates collected during 2009–2011,

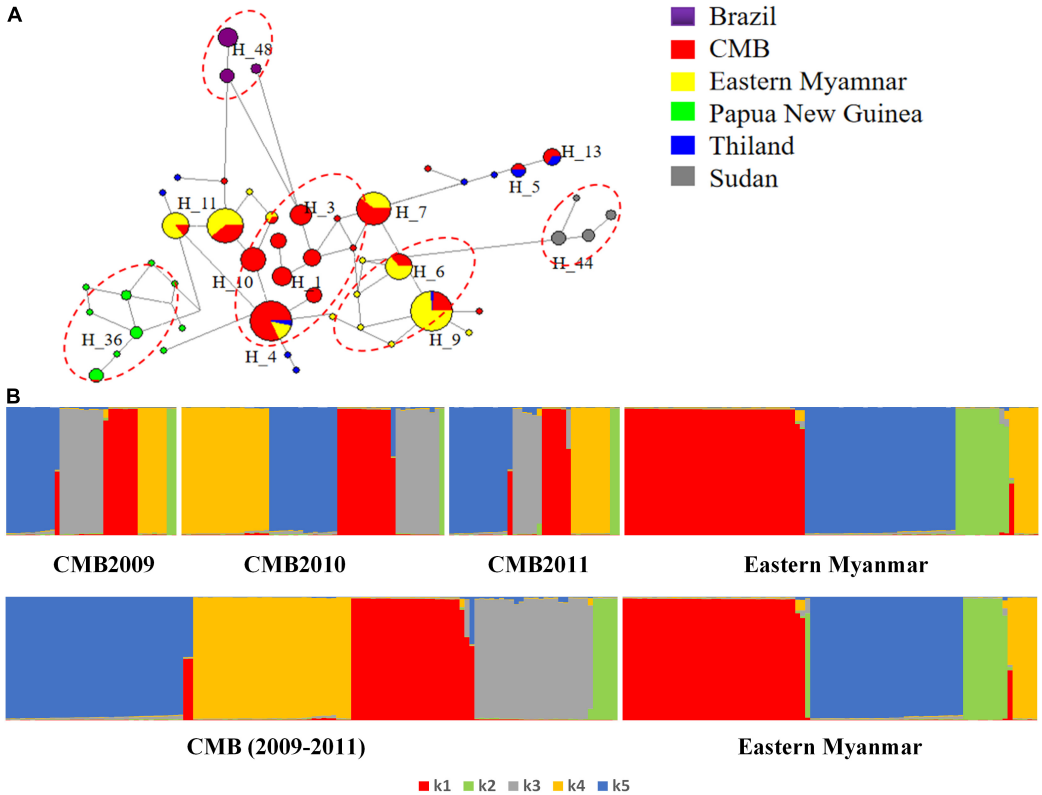

FIGURE 5 | STRUCTURE and Network analysis of PvDBP-II haplotypes. **(A)** The haplotype network shows the relationships among 50 haplotypes present in sequences of 260 isolates. **(B)** STRUCTURE analysis of the full set of variation loci from all isolates. Cluster for each isolate was assessed to an optimized cluster value of $K = 5$.

a total of 22 SNPs were divided into 21 distinct haplotypes. Overall, haplotype diversity (*Hd*) from CMB was 0.91, which showed similar diversity in isolates from other endemic areas such as Brazil, Myanmar, Thailand, and Sudan.

Based on amino acid variants, 22 polymorphic residues were detected in the CMB isolates. Besides one unique mutation (Q562H), the remaining mutations were previously reported from global *P. vivax* isolates. The seven most common mutations (K371E, D384G, R390H, N417K, L424I, W437R, and I503K) of PvDBP-II in global *P. vivax* populations were also found among CMB isolates. D384G (83.1%), L424I (71.8%), and W437R (68.6%) were the most prevalent polymorphic variants in CMB isolates. The frequency of SNPs obtained in this study was similar to that from Myanmar, Sudan, and Thailand, whereas it was smaller than that from PNG and Sri Lanka (Premaratne et al., 2011; Ju et al., 2012). Comparison of SNPs' frequencies identified in this study to those from other areas where vivax malaria is endemic revealed that most observed SNPs in CMB isolates showed different frequencies than what has been previously reported. This information may reflect the intensity of malaria transmission in different areas and the diversity of PvDBP-II may vary by geographic area. Three residues, namely N417K, W437R, and I503K, forming an important discontinuous epitope in PvDBP-II, were described as the main target for binding inhibitory antibodies against erythrocyte binding (Chootong et al., 2010; McHenry et al., 2011). In this study, N417K (60.5%), W437R (68.6%), and I503K (49.2%) were detected and all the three variants were found simultaneously in 18 isolates. Such an informative finding suggests that single or combined sites polymorphisms affect the efficiency of PvDBP-II inhibitory antibodies against invasion and help the parasites evade host immune attacks. All these mutations might generate PvDBP-II-based vaccine-resistant parasite isolates.

The pattern of polymorphism observed in part of PvDBP-II revealed this region under positive balancing selection confirmed by Tajima's D test. The direction of Tajima's D is potentially informative of the population evolution. The positive Tajima's D test of the evolutionary force on the populations indicated CMB population size reduction. Previous studies have found PvDBP-II under a strong selection. The high rates of non-synonymous to synonymous mutations reflected a positive selection promoting greater polymorphism, which may allow the evasion of host immune selection independently of the geographical distribution (Li et al., 2020). In this study, the rate of non-synonymous to synonymous mutations indicated that the PvDBP-II of CMB isolates was also under positive selection. Therefore, the high value of non-synonymous to synonymous mutations and positive Tajima's D support the theory that an increase in polymorphisms within the *pvdbp*-II of CMB isolates results in escape of host immune attacks.

Intragenic recombination is an important factor for genetic diversity of *P. vivax* isolates, which could increase variation in the PvDBP-II. LD decreased as nucleotide distance increased, suggesting that recombination has been taking place in *pvdbp*-II among CMB isolates. Recombination was confirmed by using ZZ tests, which contributed to the diversity. Such an interpretation aligns onto similar findings from previous reports from other regions where *P. vivax* is endemic, such as Myanmar, Korea, and Sri Lanka. Despite *P. vivax* having high diversity, significant LD may reflect multispatial infections within a population. The relatively higher number of recombination events in the CMB than in eastern Myanmar isolates might reflect the existence of *P. vivax* infection in mixture populations. With the progress in control programmes for malaria elimination, the number of recombination events might correlate with a decrease in *P. vivax* transmission.

P. vivax has become the most dominant parasite in the GMS region. Better understanding of transmission dynamics and population structure is essential to develop appropriate interventions for malaria elimination (Li et al., 2020). Myanmar border malaria is a major source of disease transmission and infection in the CMB in Yunnan Province. It was introduced by the highly mobile human populations across the border. The low F_{ST} values were observed between CMB isolates, eastern Myanmar and Thailand isolates, reflecting human mobility which facilitated gene flow between parasite populations of GMS region. STRUCTURE and network analyses further confirmed the relationship between the CMB, Myanmar, and other populations for PvDBP-II. This study demonstrated that the proportion of parasite populations decreased significantly with the time and geographic location difference. Our results displayed a more complex population structure than that of GMS *P. vivax* population which was only composed of four clusters based on PvDBP-II (Hu et al., 2019). Importantly, eastern Myanmar isolates lacked the K3 sub-population that existed in CMB isolates, indicating that no *P. vivax* population was imported from CMB to Myanmar. In 2010, China initiated the "Action plan of China malaria elimination (2010–2020)." Fighting malaria in the CMB has been a remarkable success, with cases having declined significantly, moving from control stage to elimination stage (Hewitt et al., 2013). Hence, in 2016, there was no doubt that CMB isolates or CMB imported cases did not exist in eastern Myanmar. The haplotype network revealed that *P. vivax* populations were highly heterogenetic and dynamics of malaria transmission differed in different areas (Cui et al., 2012). Interestingly, the network analysis of PvDBP-II showed that most of the prevalent haplotypes were shared among CMB and Myanmar isolates. In addition, some haplotypes were unique in CMB isolates which may associate with the mutations and the change of population structure. Such a finding advanced our knowledge of the parasite population dynamics in this region for the rational design of effective interventions to block disease transmission.

Designing an effective vaccine against *P. vivax* requires antigens with limited genetic diversity. Although *pvdbp*-II of CMB isolates bear some diversity, it is noteworthy that the prevalent Haplotype 4 (12.7%) was shared with multiple populations, which may present an attractive point for vaccine development.

CONCLUSION

This investigation provided the description of genetic polymorphism and natural selection of *pvdbp*-II in the CMB of Yunnan Province during 2009–2011. Results indicated that PvDBP-II was genetically diverse in the CMB. Also, findings from this study further confirmed that mutations, natural selection, and recombination might increase and sustain evasion of host immunity. With the remarkable progress made in malaria control at the CMB, this population showed specific structure and temporal differentiation. These findings provide new insights into *P. vivax* population structure and evolution in the CMB, and more importantly, consolidate the basis for rational development of an effective blood-stage malaria vaccine based on antigen variation and dominant haplotype.

DATA AVAILABILITY STATEMENT

All materials and data supporting these findings are contained within the manuscript. The sequences have been deposited in the GenBank database under the accession numbers MZ765947–MZ766070 for the China-Myanmar border isolates in Yunnan Province of China.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee at National Institute of Parasitic Diseases, Chinese Center for Disease Control and Prevention. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

T-QS, J-HC, BZ, and YW conceived, wrote the manuscript, and designed the experiments. T-QS, S-BC, Y-BC, BX, and YW performed the experiments. T-QS, H-MS, KK, and YW analyzed the data. S-BC, KK, Y-BC, BX, and YW contributed the reagents, materials, and analysis tools. All authors contributed to the article and approved the submitted version.

FUNDING

This work was financially supported in part by the Foundation of National Science and Technology Major Program (Grant Nos. 2018ZX10734-404 and 2012ZX10004-220), the Open project of Key Laboratory of Tropical Disease Control (Sun Yat-sen University), the Ministry of Education

(Grant No. 2020kfkt06), the National Sharing Service Platform for Parasite Resources (Grant No. TDRC-2019-194-30), and the National Natural Science Foundation of China (Grant No. 81101266). The funding bodies had no role in the design of the study, collection, analysis, and interpretation of data, or in writing of the manuscript.

ACKNOWLEDGMENTS

We would like to thank the staff from the Community Health Service Centers from Yunnan Province and Yunnan Institute of Parasitic Diseases for the assistance in the collection of blood samples from individuals infected with *P. vivax*.

REFERENCES

Almeida-de-Oliveira, N. K., Lima-Cury, L., de Abreu-Fernandes, R., de Rosa Lavigne, A., de Pina-Costa, A., de Souza Perce-da-Silva, D., et al. (2020). Extensive genetic diversity of *Plasmodium vivax* dbp-II in Rio de Janeiro Atlantic Forest and Brazilian Amazon Basin: evidence of positive selection. *Malar. J.* 19:81. doi: 10.1186/s12936-020-03159-y

Ampudia, E., Patarroyo, M. A., Patarroyo, M. E., and Murillo, L. A. (1996). Genetic polymorphism of the Duffy receptor binding domain of *Plasmodium vivax* in Colombian wild isolates. *Mol. Biochem. Parasitol.* 78, 269–272. doi: 10.1016/s0166-6851(96)02611-4

Arnott, A., Barry, A. E., and Reeder, J. C. (2012). Understanding the population genetics of *Plasmodium vivax* is essential for malaria control and elimination. *Malar. J.* 11:14. doi: 10.1186/1475-2875-11-14

Babaeekhou, L., Zakeri, S., and Djadid, N. D. (2009). Genetic mapping of the duffy binding protein (DBP) ligand domain of *Plasmodium vivax* from unstable malaria region in the Middle East. *Am. J. Trop. Med. Hyg.* 80, 112–118.

Balloux, F., and Lugon-Moulin, N. (2002). The estimation of population differentiation with microsatellite markers. *Mol. Ecol.* 11, 155–165. doi: 10.1046/j.0962-1083.2001.01436.x

Bandelt, H. J., Forster, P., and Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 16, 37–48. doi: 10.1093/oxfordjournals.molbev.a026036

Brashear, A. M., Fan, Q., Hu, Y., Li, Y., Zhao, Y., Wang, Z., et al. (2020). Population genomics identifies a distinct *Plasmodium vivax* population on the China-Myanmar border of Southeast Asia. *PLoS Negl. Trop. Dis.* 14:e0008506. doi: 10.1371/journal.pntd.0008506

Chen, S. B., Wang, Y., Kassegne, K., Xu, B., Shen, H. M., and Chen, J. H. (2017). Whole-genome sequencing of a *Plasmodium vivax* clinical isolate exhibits geographical characteristics and high genetic variation in China-Myanmar border area. *BMC Genomics* 18:131. doi: 10.1186/s12864-017-3523-y

Chootong, P., McHenry, A. M., Ntumngia, F. B., Sattabongkot, J., and Adams, J. H. (2014). The association of Duffy binding protein region II polymorphisms and its antigenicity in *Plasmodium vivax* isolates from Thailand. *Parasitol. Int.* 63, 858–864. doi: 10.1016/j.parint.2014.07.014

Chootong, P., Ntumngia, F. B., VanBuskirk, K. M., Xainli, J., Cole-Tobian, J. L., Campbell, C. O., et al. (2010). Mapping epitopes of the *Plasmodium vivax* Duffy binding protein with naturally acquired inhibitory antibodies. *Infect. Immun.* 78, 1089–1095. doi: 10.1128/iai.01036-09

Cole-Tobian, J., and King, C. L. (2003). Diversity and natural selection in *Plasmodium vivax* Duffy binding protein gene. *Mol. Biochem. Parasitol.* 127, 121–132. doi: 10.1016/s0166-6851(02)00327-4

Cole-Tobian, J. L., Cortés, A., Baisor, M., Kastens, W., Xainli, J., Bockarie, M., et al. (2002). Age-acquired immunity to a *Plasmodium vivax* invasion ligand, the duffy binding protein. *J. Infect. Dis.* 186, 531–539. doi: 10.1086/341776

Cui, L., Yan, G., Sattabongkot, J., Cao, Y., Chen, B., Chen, X., et al. (2012). Malaria in the greater Mekong subregion: heterogeneity and complexity. *Acta Trop.* 121, 227–239. doi: 10.1016/j.actatropica.2011.02.016

Excoffier, L., and Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10, 564–567. doi: 10.1111/j.1755-0998.2010.02847.x

Feng, J., Liu, J., Feng, X., Zhang, L., Xiao, H., and Xia, Z. (2016). Towards malaria elimination: monitoring and evaluation of the "1-3-7" approach at the China-Myanmar border. *Am. J. Trop. Med. Hyg.* 95, 806–810. doi: 10.4269/ajtmh.15-0888

Feng, X., Xia, Z. G., Feng, J., Zhang, L., Yan, H., Tang, L., et al. (2020). The contributions and achievements on malaria control and forthcoming elimination in China over the past 70 years by NIPD-CTDR. *Adv. Parasitol.* 110, 63–105. doi: 10.1016/bs.apar.2020.03.005

Gosi, P., Khusmith, S., Khalambaheti, T., Lanar, D. E., Schaecher, K. E., Fukuda, M. M., et al. (2008). Polymorphism patterns in Duffy-binding protein among Thai *Plasmodium vivax* isolates. *Malar. J.* 7:112. doi: 10.1186/1475-2875-7-112

Grimberg, B. T., Udomsangpetch, R., Xainli, J., McHenry, A., Panichakul, T., Sattabongkot, J., et al. (2007). *Plasmodium vivax* invasion of human erythrocytes inhibited by antibodies directed against the Duffy binding protein. *PLoS Med.* 4:e337. doi: 10.1371/journal.pmed.0040337

Hewitt, S., Delacollette, C., and Chavez, I. (2013). Malaria situation in the greater Mekong Subregion. *Southeast Asian J. Trop. Med. Public Health* 44, 46–72.

Hoque, M. R., Elfaki, M. M. A., Ahmed, M. A., Lee, S. K., Muh, F., Ali Albsheer, M. M., et al. (2018). Diversity pattern of Duffy binding protein sequence among Duffy-negatives and Duffy-positives in Sudan. *Malar. J.* 17:297. doi: 10.1186/s12936-018-2425-z

Hu, Y., Wang, L., Mbenda, H. G. N., Soe, M. T., Yu, C., Feng, H., et al. (2019). Genetic diversity, natural selection and haplotype grouping of *Plasmodium vivax* Duffy-binding protein genes from eastern and western Myanmar borders. *Parasit. Vectors* 12:546. doi: 10.1186/s13071-019-3803-2

Ju, H. L., Kang, J. M., Moon, S. U., Bahk, Y. Y., Cho, P. Y., Sohn, W. M., et al. (2013). Genetic diversity and natural selection of Duffy binding protein of *Plasmodium vivax* Korean isolates. *Acta Trop.* 125, 67–74. doi: 10.1016/j.actatropica.2012.09.016

Ju, H. L., Kang, J. M., Moon, S. U., Kim, J. Y., Lee, H. W., Lin, K., et al. (2012). Genetic polymorphism and natural selection of Duffy binding protein of Plasmodium vivax Myanmar isolates. *Malar. J.* 11:60. doi: 10.1186/1475-2875-11-60

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Li, Y. L., Hu, Y. B., Zhao, Y., Wang, Q. H., Mbenda, H. G. N., Kittichai, V., et al. (2020). Dynamics of *Plasmodium vivax* populations in border areas of the greater Mekong sub-region during malaria elimination. *Malar. J.* 19:145. doi: 10.1186/s12936-020-03221-9

Lo, E., Lam, N., Hemming-Schroeder, E., Nguyen, J., Zhou, G., Lee, M. C., et al. (2017). Frequent spread of *Plasmodium vivax* malaria maintains high genetic diversity at the Myanmar-China border, without distance and landscape barriers. *J. Infect. Dis.* 216, 1254–1263. doi: 10.1093/infdis/jix106

Longley, R. J., White, M. T., Takashima, E., Morita, M., Kanoi, B. N., Li Wai Suen, C. S. N., et al. (2017). Naturally acquired antibody responses to more than 300 *Plasmodium vivax* proteins in three geographic regions. *PLoS Negl. Trop. Dis.* 11:e0005888. doi: 10.1371/journal.pntd.0005888

McHenry, A. M., Barnes, S. J., Ntumngia, F. B., King, C. L., and Adams, J. H. (2011). Determination of the molecular basis for a limited dimorphism, N417K, in the *Plasmodium vivax* Duffy-binding protein. *PLoS One* 6:e20192. doi: 10.1371/journal.pone.0020192

Premaratne, P. H., Aravinda, B. R., Escalante, A. A., and Udagama, P. V. (2011). Genetic diversity of *Plasmodium vivax* Duffy Binding Protein II (PvDBPII) under unstable transmission and low intensity malaria in Sri Lanka. *Infect. Genet. Evol.* 11, 1327–1339. doi: 10.1016/j.meegid.2011.04.023

Price, R. N., Tjitra, E., Guerra, C. A., Yeung, S., White, N. J., and Anstey, N. M. (2007). Vivax malaria: neglected and not benign. *Am. J. Trop. Med. Hyg.* 77, 79–87.

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959. doi: 10.1093/genetics/155.2.945

Rozas, J., Gullaud, M., Blandin, G., and Aguadé, M. (2001). DNA variation at the rp49 gene region of *Drosophila simulans*: evolutionary inferences from an unusual haplotype structure. *Genetics* 158, 1147–1155. doi: 10.1093/genetics/158.3.1147

Rozas, J., Sánchez-DelBarrio, J. C., Messeguer, X., and Rozas, R. (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19, 2496–2497. doi: 10.1093/bioinformatics/btg359

Shen, H. M., Chen, S. B., Wang, Y., and Chen, J. H. (2015). Whole-genome sequencing of a *Plasmodium vivax* isolate from the China-Myanmar border area. *Mem. Inst. Oswaldo Cruz.* 110, 814–816. doi: 10.1590/0074-02760150216

Shin, J.-H., Blay, S., McNeney, B., and Graham, J. (2006). LDheatmap: an R function for graphical display of pairwise linkage disequilibria between single nucleotide polymorphisms. *J. Stat. Softw.* 16, 1–9. doi: 10.18637/jss.v016.c03

Sousa, T. N., Tarazona-Santos, E. M., Wilson, D. J., Madureira, A. P., Falcão, P. R., Fontes, C. J., et al. (2010). Genetic variability and natural selection at the ligand domain of the Duffy binding protein in Brazilian *Plasmodium vivax* populations. *Malar. J.* 9:334. doi: 10.1186/1475-2875-9-334

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585–595. doi: 10.1093/genetics/123.3.585

Takala, S. L., and Plowe, C. V. (2009). Genetic diversity and malaria vaccine design, testing and efficacy: preventing and overcoming 'vaccine resistant malaria'. *Parasite Immunol.* 31, 560–573. doi: 10.1111/j.1365-3024.2009.01138.x

Valizadeh, V., Zakeri, S., Mehrizi, A. A., and Djadid, N. D. (2014). Population genetics and natural selection in the gene encoding the Duffy binding protein II in Iranian *Plasmodium vivax* wild isolates. *Infect. Genet. Evol.* 21, 424–435. doi: 10.1016/j.meegid.2013.12.012

Wertheimer, S. P., and Barnwell, J. W. (1989). *Plasmodium vivax* interaction with the human Duffy blood group glycoprotein: identification of a parasite receptor-like protein. *Exp. Parasitol.* 69, 340–350. doi: 10.1016/0014-4894(89)90083-0

WHO (2016). *Eliminating Malaria in the Greater Mekong Subregion. United to End a Deadly Disease*. Geneva: World Health Organization.

WHO (2020). *World Malaria Report 2020*. Geneva: World Health Organization.

Xainli, J., Adams, J. H., and King, C. L. (2000). The erythrocyte binding motif of *Plasmodium vivax* Duffy binding protein is highly polymorphic and functionally conserved in isolates from Papua New Guinea. *Mol. Biochem. Parasitol.* 111, 253–260. doi: 10.1016/s0166-6851(00)00315-7

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.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 08 February 2022
doi: 10.3389/fmicb.2022.747484

Expansion of Cyclophyllidea Biodiversity in Rodents of Qinghai-Tibet Plateau and the "Out of Qinghai-Tibet Plateau" Hypothesis of Cyclophyllideans

Yao-Dong Wu[1†], Guo-Dong Dai[1†], Li Li[1], D. Timothy J. Littlewood[2,3], John Asekhaen Ohiolei[1], Lin-Sheng Zhang[1], Ai-Min Guo[1], Yan-Tao Wu[1], Xing-Wei Ni[1,4], Nigus Abebe Shumuye[1], Wen-Hui Li[1], Nian-Zhang Zhang[1], Bao-Quan Fu[1], Yong Fu[5], Hong-Bin Yan[1] and Wan-Zhong Jia[1]**

[1] State Key Laboratory of Veterinary Etiological Biology, National Professional Laboratory for Animal Echinococcosis, Key Laboratory of Veterinary Parasitology of Gansu Province, Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Lanzhou, China, [2] Department of Life Sciences, Natural History Museum, London, United Kingdom, [3] London Centre for Neglected Tropical Disease Research, London, United Kingdom, [4] Guizhou Provincial Center for Animal Disease Control and Prevention, Guiyang, China, [5] State Key Laboratory of Plateau Ecology and Agriculture, Qinghai Academy of Animal Science and Veterinary Medicine, Qinghai University, Xining, China

The Cyclophyllidea comprises the most species-rich order of tapeworms (Platyhelminthes, Cestoda) and includes species with some of the most severe health impact on wildlife, livestock, and humans. We collected seven Cyclophyllidea specimens from rodents in Qinghai-Tibet Plateau (QTP) and its surrounding mountain systems, of which four specimens in QTP were unsequenced, representing "putative new species." Their complete mitochondrial (*mt*) genomes were sequenced and annotated. Phylogenetic reconstruction of partial 28S rDNA, *cox*1 and *nad*1 datasets provided high bootstrap frequency support for the categorization of three "putative new species," assigning each, respectively, to the genera *Mesocestoides*, *Paranoplocephala,* and *Mosgovoyia*, and revealing that some species and families in these three datasets, which contain 291 species from nine families, may require taxonomic revision. The partial 18S rDNA phylogeny of 29 species from Taeniidae provided high bootstrap frequency support for the categorization of the "putative new species" in the genus *Hydatigera*. Combined with the current investigation, the other three known Taeniidae species found in this study were *Taenia caixuepengi*, *T. crassiceps,* and *Versteria mustelae* and may be widely distributed in western China. Estimates of divergence time based on *cox*1 + *nad*1 fragment and *mt* protein-coding genes (PCGs) showed that the differentiation rate of Cyclophyllidea species was strongly associated with the rate of change in the biogeographic scenarios, likely caused by the uplift of the QTP; i.e., species differentiation of Cyclophyllidea might be driven by host-parasite co-evolution caused by the uplift of QTP. We propose an "out of QTP" hypothesis for the radiation of these cyclophyllidean tapeworms.

Keywords: Cyclophyllidea, phylogeny, species differentiation, biogeography, Qinghai-Tibet Plateau, rodents

OPEN ACCESS

Edited by:
Rodrigo Pulgar Tejo, University of Chile, Chile

Reviewed by:
Denis Jacob Machado, University of North Carolina at Charlotte, United States
Daxi Wang, Beijing Genomics Institute (BGI), China
Guo-Hua Liu, Hunan Agricultural University, China

***Correspondence:**
Wan-Zhong Jia
jiawanzhong@caas.cn
Hong-Bin Yan
yanhongbin@caas.cn

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *26 July 2021*
Accepted: *10 January 2022*
Published: *08 February 2022*

Citation:
Wu Y-D, Dai G-D, Li L, Littlewood DTJ, Ohiolei JA, Zhang L-S, Guo A-M, Wu Y-T, Ni X-W, Shumuye NA, Li W-H, Zhang N-Z, Fu B-Q, Fu Y, Yan H-B and Jia W-Z (2022) Expansion of Cyclophyllidea Biodiversity in Rodents of Qinghai-Tibet Plateau and the "Out of Qinghai-Tibet Plateau" Hypothesis of Cyclophyllideans. Front. Microbiol. 13:747484. doi: 10.3389/fmicb.2022.747484

INTRODUCTION

Cestoda is a class of parasitic worms in the flatworm phylum Platyhelminthes that parasitize the intestines of all major groups of vertebrates, including fish (Kuchta et al., 2020), reptiles (Yudhana et al., 2019), birds (Okulewicz, 2014), and mammals (Carlson et al., 2020), and cause severe, mild or no symptoms of infection. Their larvae usually achieve development in one or two intermediate hosts (Kuchta et al., 2020) and can cause severe symptoms and even death in animals and humans; species of *Taenia* and *Echinococcus* cause the highest health and economic impact.

To date, there are approximately 4,800 described species in the class Cestoda, belonging to 833 genera and 19 orders, of which the Cyclophyllidea is the most species-rich order in the class, with more than 3,100 valid species, distributed among 16 families (Sharma et al., 2016; Caira and Jensen, 2017; Kuchta et al., 2020). Each free-living metazoan species is considered to harbor at least one protozoan or metazoan parasite species (Poulin and Morand, 2004) with which they interact and usually co-evolve (Ebert and Fields, 2020). Parasites are probably one of the largest groups of undescribed organisms, as most are cryptic whether in their parasitic or free-living forms (Dobson et al., 2008; Larsen et al., 2017; Okamura et al., 2018; Carlson et al., 2020). It is estimated that there are a total of about 100,000–350,000 helminth species in vertebrates around the world, of which 85–95% are undiscovered or recorded to science (Carlson et al., 2020).

Many of the world's biodiversity hotspots are located in large mountain systems, and their role in the evolutionary diversification of organisms is manifold (Hoorn et al., 2018; Rahbek et al., 2019a,b). The Qinghai-Tibet Plateau (QTP) and its surrounding mountain systems of the Eurasian continent, have yielded arguably the biggest and probably the most biologically diverse area of montane species (Päckert et al., 2020). In the past decade, several new Taeniidae species have been described from wild rodents on the QTP (Xiao et al., 2005; Wu et al., 2021). In terms of species diversity and sheer population sizes, rodents are perhaps the most important intermediate and definitive hosts of tapeworms, and are the most widely distributed and diverse group of mammals; about 43% of all species (Singla et al., 2008; Wu et al., 2018). Changes in climate and vegetation caused by the uplift of the QTP may have promoted local adaptations such as the evolution of cold- and hypoxic-tolerant rodent species, which in turn may have led to the co-evolution and radiation of their parasites (Wang et al., 2018; Wu et al., 2021). We hypothesize that there may be many undiscovered tapeworm species parasitizing the rodents in the QTP. Although the number of tapeworm species has been underestimated in general, some practices may lead to the erroneous proposal of new species (Carlson et al., 2020; Kuchta et al., 2020).

Morphological distinctions have been used for the description of many tapeworms, however, the homoplasy in morphology poses quite a challenge to infer their evolutionary lineage (Scholz et al., 2021), where morphological features may be significantly affected by different intermediate host sources (Lymbery, 1998). Many undescribed species may be genetically different but morphologically indistinguishable. The inability to distinguish such cryptic species affects accurate assessment of host range and estimates of total diversity (Carlson et al., 2020). The 28S and 18S nuclear ribosomal RNA genes (rDNA) are relatively conserved within species and are often used to differentiate different species (Wickström et al., 2005; Nakao et al., 2013; Scholz et al., 2021). The mitochondrial (*mt*) genome is largely haploid and uniparentally inherited, so their effective population size is four times smaller than that of the nuclear genome (Toews and Brelsford, 2012). Since the process of lineage sorting is inversely proportional to the effective population size, this means mitochondrial (*mt*) genomes will complete this process faster than nuclear genomes (Toews and Brelsford, 2012). Thus, the genetic nature of a *mt* genome makes it likely more sensitive than any single nuclear marker to distinguish closely related species and study their phylogenetic relationships (Lee et al., 2007).

Therefore, in this study, we sampled parasites in rodents from the QTP and its surrounding areas. In total, seven Cyclophyllidea specimens were collected and characterized using molecular tools. By comparing new (mitochondrial, nuclear 28SrDNA, 18S rDNA) with published homologs (GenBank) we found four isolates to be markedly different in DNA sequence thus likely representing "putative new species." The complete *mt* genomes of these unknown taxa were sequenced and annotated, and their taxonomic status was analyzed and verified through phylogenetic reconstruction of five datasets containing a total of 320 species from 10 families. By combining evolutionary divergence time analyses of *mt* genes of classified cyclophyllideans in NCBI Taxonomy Database[1] and the paleogeography of QTP, we thus speculate an "out of QTP" theory for cyclophyllideans.

MATERIALS AND METHODS

Sample Collection

Rodents were live-trapped in meadows in Tibet, Qinghai, Sichuan, Gansu and Xinjiang province or autonomous region of China in 2013 and from 2018 to 2020. Rodents were euthanized and dissected according to the Ethics Statement mentioned below, cysticerci and host livers were collected from the enterocoelia and thorax, and adults were extracted from the intestines. Detailed sample collection data and host identities are described in **Supplementary Table 1**. After detaching the lesions, parasite specimens and host livers were kept in 75% (v/v) ethanol for molecular identification.

DNA Isolation, Amplification, and Sequencing

DNA samples of hosts and parasites were extracted using Blood and Tissue Kit (Cat. No. 69504, Qiagen, Germany) as instructed by the manufacturer, and were amplified and sequenced for identification by conserved primers of *cyt*b gene of small mammals (Fan et al., 2011) and *cox*1 gene of tapeworm (Bowles et al., 1992), respectively. By means of highly similar BLAST

[1]https://www.ncbi.nlm.nih.gov/taxonomy

search in the nucleotide collection (nr/nt) database,[2] four of the Cyclophyllidea specimens with less than 95% identity of the most similar *cox*1 sequence were identified as putatively unknown species (Xiao et al., 2005). The *mt* genomes of four of the putatively unknown specimens were sequenced and assembled according to the following procedure: firstly, the DNA of the four species was amplified and sequenced using primers published in Wu et al. (2021); missing fragments were amplified with newly designed primers using the program Oligo 6.0 with the method described in Wu et al. (2021), until entire circular *mt*DNAs were amplified and sequenced. A list of primers for each species can be found in **Supplementary Table 2**. The 18S and 28S rDNA fragments of these four species were also amplified and sequenced with conserved primers (Littlewood et al., 2000; Yan et al., 2013) for further species identification and phylogenetic analyses. Primers were synthesized by Tsingke Biotechnology (Xi'an, China). Standard 25 μl PCR protocol was used to amplify the DNA fragments. The PCR products were purified and sequenced according to protocols in Wu et al. (2021). SeqMan software was used to assemble the *cyt*b gene sequences, 18S and 28S rDNA partial sequences and *mt* genomes (see **Supplementary Table 3** for GenBank accession nos.).

Mitochondrial Genome Annotation

The four new *mt*DNAs were annotated preliminarily by Geseq[3] with the reference of the most closely related species (*Mesocestoides corti*, *Anoplocephala magna*, *Hydatigera krepkogorski,* and *Moniezia expansa*) identified by the phylogenetic analyses in **Figures 1**, **2**, whose *mt* genome annotations are available in GenBank. Putative tRNA genes were verified using ARWEN[4] using default parameters (Laslett and Canbäck, 2008). The positions of their open reading frames (ORF) and rRNA genes were further checked and modified using SnapGene (v3.2.1) based on alignments with the reference of the most closely related species mentioned above. SnapGene (v3.2.1) was used to translate the protein-coding genes (PCGs) into their amino acid sequence with echinoderm/flatworm mitochondrial code (NCBI translation table 9) and to illustrate the annotated *mt* genomes.

Phylogenetic Analyses and Sequence Identity

As ingroups for phylogenetic reconstruction, we combined *mt* genes and 18S or 28S rDNA sequenced in this study with those of other classified cyclophyllideans available in GenBank. Three species (*Caryophyllaeus brachycollis*, *Anindobothrium anacolum*, *Spirometra erinaceieuropaei*) of Eucestoda belonging to different orders were chosen as outgroups.

The 28S rDNA fragments of classified cyclophyllideans in NCBI Taxonomy Database are available for most families in GenBank except Taeniidae, while that of the 18S rDNA were only available for Taeniidae, so here the evolutionary trees of 28S and 18S rDNA fragments were constructed separately. Based on the results of the 28S rDNA phylogeny, a simplified 28S evolutionary tree was reconstructed without affecting the topological structure of the tree by removing the species in the same genus and clade. To complement and validate the inferred evolutionary relationship between Cyclophyllidea species, except the species of family Taeniidae, the phylogenies with *cox*1 and *nad*1 fragments were reconstructed.

In summary, 5 datasets containing 320 species in 10 families were assembled for phylogenetic analyses (see **Supplementary Tables 4–8** for GenBank accession nos. and species, genera and families used in each dataset). Since the limited availability of data on different genes of the same species, we performed separate phylogenetic reconstruction for each dataset to cover more species. Datasets were aligned using MAFFT v7.487 with auto option (Katoh and Standley, 2013). The alignments were trimmed by using Trimal v1.2 under the automated1 option (Capella-Gutiérrez et al., 2009) to preserve the same sequence regions and exclude the ambiguously aligned sites. All phylogenetic trees were constructed with maximum likelihood (ML) inference using IQ-TREE v2.1.4 (Nguyen et al., 2015) with ultrafast bootstrap 1,000 replicates in Ubuntu 20.04.2 LTS operating system, where the best-fit models were automatically selected by ModelFinder (**Supplementary Table 9**; Kalyaanamoorthy et al., 2017) and the best number of threads were also selected under AUTO option (using the command line: iqtree -s alignment_file -m -MFP -bb 1000 -nt AUTO). All other parameters were set to their default values. The ML trees were then visualized on the IToL web server (Letunic and Bork, 2016).

The related species of the putatively unknown species were identified based on phylogenetic reconstruction, and the percentage identity of sequences in the 5 datasets between the putatively unknown species and their related species was calculated by BLAST (**Supplementary Table 10**).

Analysis of Divergence Times

Species divergence times were calculated using BEAST v2.6.2 (Bouckaert et al., 2014) based on two datasets without shared species: all available *mt* PCGs dataset of 54 species (**Supplementary Table 11**) and *cox*1 + *nad*1 fragments dataset of other classified 54 species (**Supplementary Table 12**). These two datasets were aligned and trimmed with MAFFT v7.487 and Trimal v1.2 as described above, and were partitioned according to different genes. Model selection of partitions was identified by Partition Finder v2.1.1 (Lanfear et al., 2012) with the set of "linked" branch lengths, "beast" models, "aicc" model selection, and "greedy" search. Partition schemes and substitution models can be found in **Supplementary Table 9**. The Strict Clock model was chosen to ignore the rate differences between the branches in the mode and the gamma category count was set to 4. Other settings, such as substitution rate and shape, in the site model were evaluated in the analysis. The Calibrated Yule model (Heled and Drummond, 2015) was used as the tree prior, as it is a simple model of speciation that is generally appropriate when considering sequences from different species. Time calibration was calibrated with the divergence date between *T. saginata* and *T. asiatica* (∼1.14 Mya) and divergence date between *Schistosoma japonicum* and *S. mansoni* (∼56.10 Mya), which agreed with

[2]https://blast.ncbi.nlm.nih.gov/Blast.cgi

[3]https://chlorobox.mpimp-golm.mpg.de/geseq.html

[4]http://130.235.46.10/ARWEN/

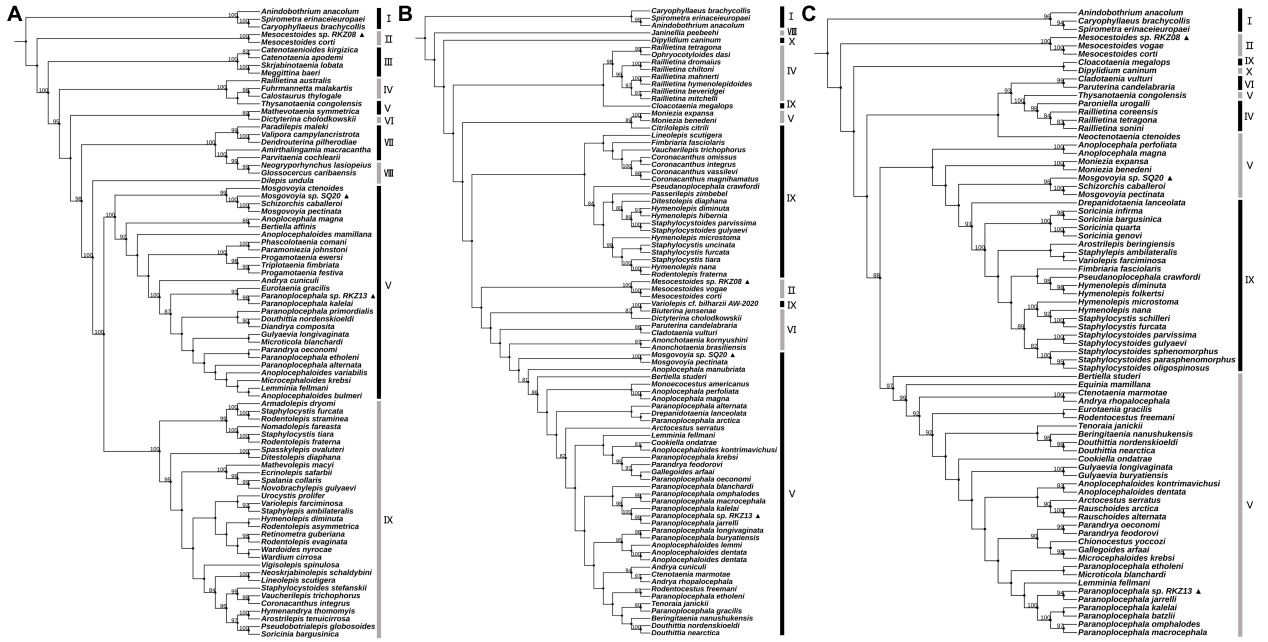

FIGURE 1 | Maximum likelihood analyses of the four "putative new species" with other Cyclophyllidea species, except for Taeniidae, based on simplified 28S rDNA **(A)**, *cox*1 **(B)** and *nad*1 **(C)** fragments. Alternating black and gray bands are used to classify different genera in the Taxonomy of NCBI, The Roman numerals to the right of the bands represent outgroup and different genus names (i.e., I: outgroup, II: Mesocestoididae, III: Catenotaeniidae, IV: Dipylidiidae, V: Anoplocephalidae, VI: Paruterinidae, VII: Gryporhynchidae, VIII: Dilepididae, IX: Hymenolepididae, X: Dipylidiidae). The ▲ after the species name represents "putative new species". Bootstrap frequency support values are stated only for nodes where > 80%.

reported fossil evidence from shark coprolites and previously estimated dates based on *mt* genes (Wang et al., 2016). Posterior probability estimates were sampled every 1,000 iterations over a total of 10,000,000 iterations per MCMC run. Other options were run on their default values. Tracer (v1.7.1) was used to summarize posterior probabilities. Trees were annotated via TreeAnnotator (v2.1.2) using a maximum clade credibility tree and median heights settings with 10% burn-in. The number of divergence nodes in every 2 Mya was summarized based on the evolutionary divergence time trees of two datasets, respectively.

RESULTS

Species Identification and Phylogenetic Relationships

The hosts of parasites were confirmed by BLAST searches (**Supplementary Table 1**; see **Supplementary Table 3** for GenBank accession nos. of host *cyt*b gene). Two of the parasite species were cysticerci and adult worms, each retrieved from the abdominal cavity and intestinal tract of two vole hosts (*Neodon leucurus*) collected from the same pasture location near the Shigatse City of Tibet Autonomous Region (**Supplementary Table 1**). Phylogenetic analyses shown in **Supplementary Figure 1** and **Figures 1A–C** confirmed that these two species likely belong to the genera *Mesocestoides* and *Paranoplocephala*, and were labeled as *Mesocestoides* sp. RKZ08 and *Paranoplocephala* sp. RKZ13, respectively. The cysticerci collected from the liver of plateau pika (*Ochotona curzoniae*) from Xietongmen county of Tibet and zokors (*Eospalax fontanierii*) from Xiahe county of Gansu province were confirmed to be the same species by alignment of *cox*1 and 18S rDNA segments, which was in the monophyletic group of the genus *Hydatigera* in the phylogenetic trees (**Figure 2**), and was marked as *Hydatigera* sp. XHPW10. The final species was an adult tapeworm collected from the intestine of plateau pikas from Shiqu county, Sichuan province, and occurred within the genus *Mosgovoyia* (**Supplementary Figure 1** and **Figures 1A–C**), and was thus named *Mosgovoyia* sp. SQ20. These four species were identified as newly sequenced and "putative new species" as their *cox*1 and *nad*1 sequences having less than 95% identity with available related taxa (**Supplementary Table 10**; Xiao et al., 2005). The degree of differentiation of their *mt* genes was higher than that of the nuclear genes 18S and 28S rDNA (**Supplementary Table 10**), reflecting their differentiation likely results from the so-called deep mitochondrial divergence (DMD, Zhang et al., 2019). In addition, three known cysticerci of Taeniidae, *T. caixuepengi*, *T. crassiceps,* and *Versteria mustelae*, were also identified in the present study (see **Supplementary Table 13** for their *cox*1 fragments).

General Features of the Mitochondrial Genome of "Putative New Species"

The complete *mt* genomes of the four "putative new species" were 13,361 bp (GenBank ID: MW808979), 13,730 bp (GenBank ID: MW808980), 14,148 bp (GenBank ID: MW808981), and

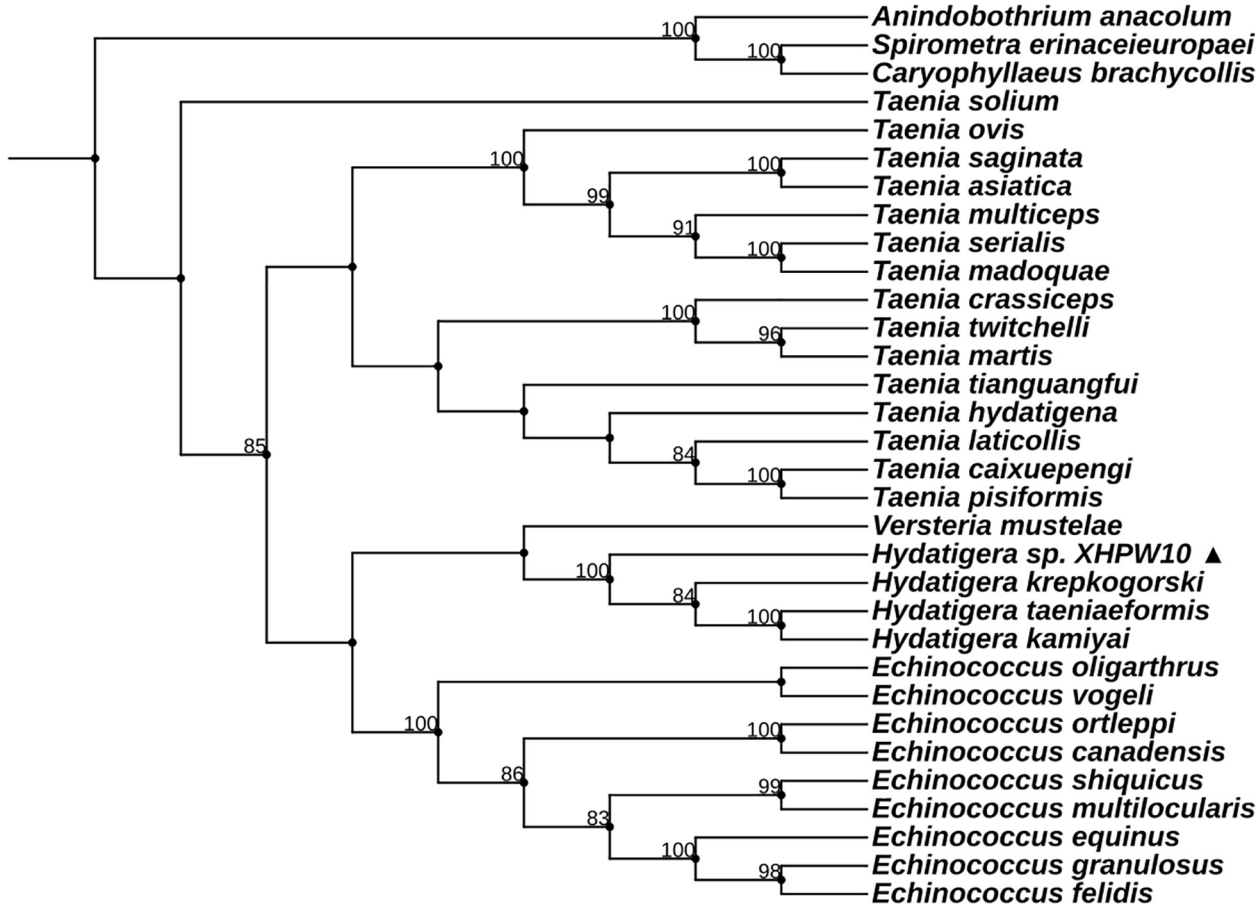

FIGURE 2 | Maximum likelihood analysis of 18S rDNA of family Taeniidae. The ▲ after the species name represents "putative new species." The outgroup is the same as **Figure 1**. Bootstrap frequency support values are stated only for nodes where > 80%.

13,776 bp (GenBank ID: MW808982) in length. Each of them contains two rRNA genes (*rrn*S and *rrn*L) and 12 protein-encoding genes (*atp*6, *cyt*b, *nad*4L, *cox*1-3, and *nad*1-6), which are arrayed in the typical order of *mt* genomes of cestodes. They each contain the 22 typical tRNA genes set of cestodes, and share a common set of anticodons. The order of tRNA genes is roughly the same, except between *nad*6 and *nad*5 genes. Species *Paranoplocephala* sp. RKZ13 had a repeat sequence of *trn*L and *trn*R in the highly variable region between *nad*6 and *nad*5, suggesting that it had an insertion sequence in this region that made its *mt* genome longer than the other three species (**Figure 3** and **Table 1**).

Flatworms use a unique set of *mt* code for protein translation (Nakao et al., 2000; Telford et al., 2000). In addition to ATG, GTG was also used as an initiation codon in a small fraction of coding genes in their *mt* genomes. For the termination codon, all species used only TAA and TAG; TGA was not identified as a termination codon (**Table 1**).

Divergence Times Analysis

The divergence time based on the two datasets used for time calibration is consistent with previous genome-based analysis results (**Figure 4**; Wang et al., 2016). Three of the "putative new species," *Mesocestoides* sp. RKZ08, *Mosgovoyia* sp. SQ20, and *Hydatigera* sp. XHPW10, might have originated from a similar phase in the late Miocene, while *Paranoplocephala* sp. RKZ13 diverged during the Pleistocene (**Figure 4**). By summarizing the number of divergent time nodes in the divergence time trees over time, it was found that the trees of *mt* PCGs and *cox*1 + *nad*1 produced similar differentiation rate trends: there was a slight acceleration of the evolutionary rate in the period 14–24 Mya, and a marked acceleration during the period 4–10 Mya (**Figure 5**). However, compared with *cox*1 + *nad*1, the differentiation rate of *mt* PCGs was faster in the 14–24 Mya but slower in 4–10 Mya, which may have been caused by the species bias used in both datasets.

DISCUSSION

Given the rich diversity and the large rodent population in western China and the few published reports of tapeworms in rodents, except for Taeniidae (Xiao et al., 2005; Zhao et al., 2014; Wang et al., 2018; Zhang et al., 2018; Wu et al., 2021), the present

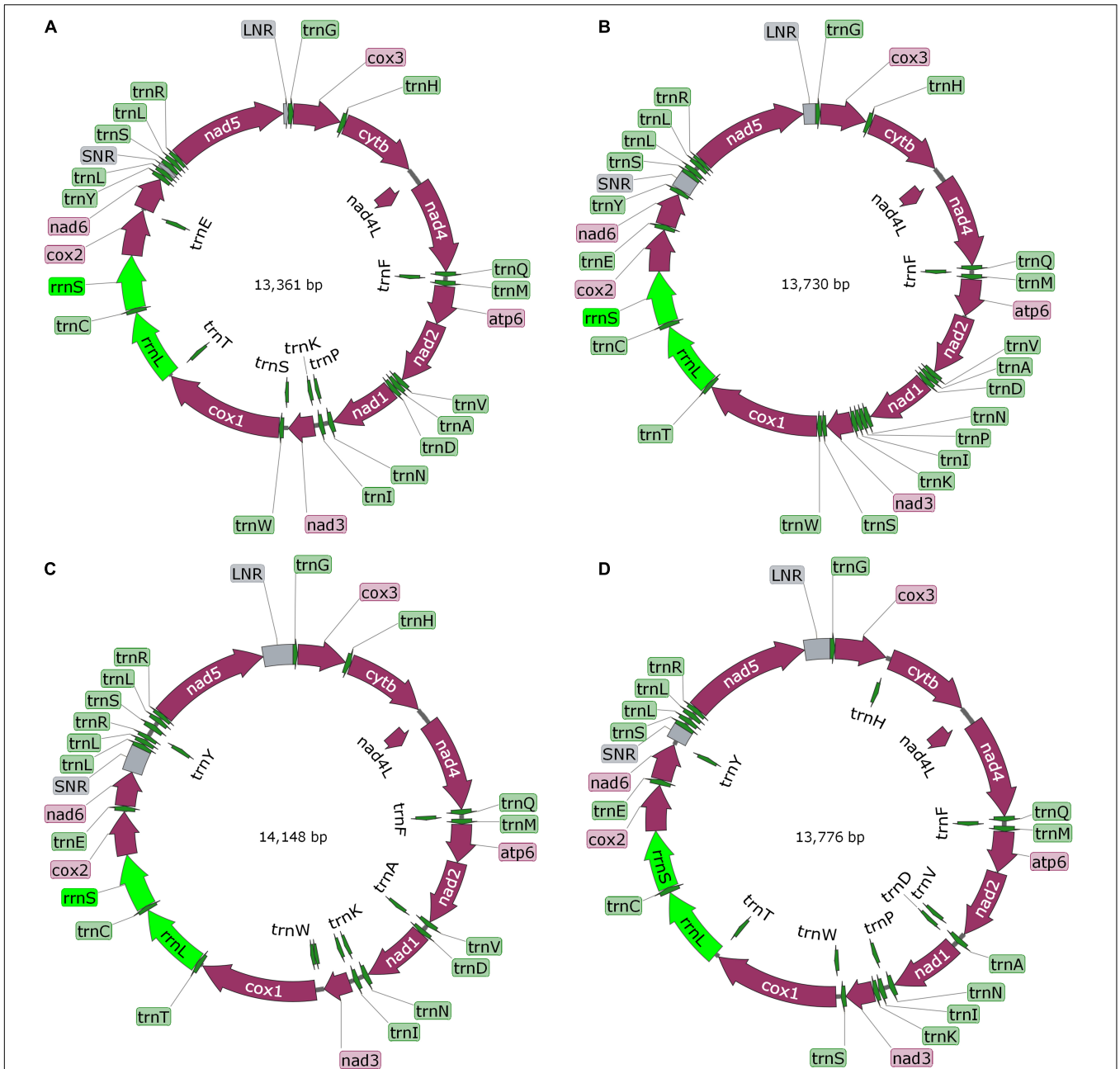

FIGURE 3 | The diagram of complete mitochondrial genome of *Mesocestoides* sp. RKZ08 **(A)**, *Paranoplocephala* sp. RKZ13 **(B)**, *Hydatigera* sp. XHPW10 **(C)**, and *Mosgovoyia* sp. SQ20 **(D)**. The protein-encoding genes are depicted in plum, the tRNAs are depicted in green, the rRNAs are depicted in light green and the non-coding mitochondrial regions (NCRs including LNR and SNR) are depicted in gray. The inferred gene boundaries of them are shown in **Table 1**.

knowledge of tapeworm biodiversity in rodents in western China suggests far greater biodiversity yet to be uncovered.

Here, the *mt* genes and 18S or 28S rDNA fragments of four (two larvae and two adults) unidentified Cyclophyllidea species differed from their related species (**Supplementary Table 10**). However, due to the specimen distortion and insufficient specimen encountered in this study, it is not clear whether they have been described morphologically. All four species showed apparent discordance percentage identity with the related species in their *mt*DNA and 18S or 28S rDNA, which is the common DMD pattern across the animal kingdom, and demonstrated in *Echinococcus granulosus sensu stricto* (Kinkar et al., 2017), possibly due to the parthenogenetic inheritance of mitochondria, gene flow and recombination in the nuclear genome (Toews and Brelsford, 2012).

The larvae *Hydatigera* sp. XHPW10 was found to live in the livers of both zokors from northeast of QTP and plateau pika from southwest of QTP (**Supplementary Table 1**). This adds

TABLE 1 | The list of mitochondrial genome annotation for four "putative new species".

Genes	Positions of nucleotide sequences (bp)				Initiation and termination codons				Anticodons
	RKZ08	RKZ13	XHPW10	SQ20	RKZ08	RKZ13	XHPW10	SQ20	
*trn*G	1–67	1–62	1–67	1–63					TCC
*cox*3	73–717	68–712	70–714	65–715	ATG/TAA	ATG/TAA	ATG/TAA	ATG/TAG	
*trn*H	726–794	713–783	716–784	709–778					GTG
*cyt*b	798–1892	787–1881	787–1,854	782–1876	ATG/TAG	GTG/TAG	ATG/TAG	GTG/TAG	
*nad*4L	1,895–2,155	1,884–2,144	1,877–2,137	1,892–2,152	GTG/TAG	ATG/TAA	GTG/TAG	ATG/TAA	
*nad*4	2,116–3,372	2,111–3,358	2,104–3,354	2,113–3,366	GTG/TAA	GTG/TAG	GTG/TAA	GTG/TAG	
*trn*Q	3,373–3,435	3,359–3,426	3,355–3,417	3,367–3,428					TTG
*trn*F	3,434–3,499	3,425–3,488	3,415–3,479	3,427–3,490					GAA
*trn*M	3,495–3,562	3,485–3,551	3,475–3,542	3,489–3,554					CAT
*atp*6	3,568–4,080	3,557–4,063	3,551–4,069	3,558–4,073	ATG/TAA	ATG/TAA	ATG/TAG	ATG/TAG	
*nad*2	4,103–4,978	4,073–4,948	4,076–4,966	4,083–4,964	ATG/TAG	ATG/TAG	ATG/TAG	ATG/TAA	
*trn*V	4,983–5,047	4,951–5,014	4,967–5,030	4,978–5,041					TAC
*trn*A	5,048–5,115	5,014–5,077	5,036–5,103	5,040–5,106					TGC
*trn*D	5,121–5,187	5,082–5,142	5,108–5,171	5,106–5,166					GTC
*nad*1	5,192–6,079	5,145–6,035	5,176–6,069	5,170–6,060	ATG/TAA	ATG/TAG	ATG/TAG	ATG/TAG	
*trn*N	6,095–6,160	6,041–6,109	6,070–6,137	6,066–6,131					GTT
*trn*P	6,163–6,226	6,117–6,180	6,145–6,207	6,137–6,200					TGG
*trn*I	6,227–6,290	6,180–6,245	6,206–6,269	6,200–6,264					GAT
*trn*K	6,294–6,358	6,244–6,308	6,274–6,336	6,275–6,337					CTT
*nad*3	6,362–6,709	6,310–6,657	6,334–6,681	6,342–6,689	ATG/TAA	ATG/TAG	GTG/TAG	ATG/TAA	
*trn*S	6,718–6,776	6,656–6,714	6,680–6,738	6,692–6,751					GCT
*trn*W	6,780–6,845	6,715–6,776	6,746–6,809	6,756–6,818					TCA
*cox*1	6,851–8,449	6,774–8,342	6,810–8,462	6,816–8,408	ATG/TAA	ATG/TAG	ATG/TAG	ATG/TAA	
*trn*T	8,456–8,518	8,365–8,429	8,427–8,490	8,395–8,458					TGT
*rrn*L	8,519–9,488	8,430–9,393	8,491–9,452	8,459–9,431					
*trn*C	9,489–9,554	9,394–9,450	9,453–9,511	9,432–9,496					GCA
*rrn*S	9,555–10,282	9,451–10,187	9,512–10,232	9,497–10,232					
*cox*2	10,283–10,861	10,188–10,760	10,233–10,859	10,233–10,808	ATG/TAA	ATG/TAA	ATG/TAA	ATG/TAA	
*trn*E	10,864–10,932	10,771–10,837	10,816–10,883	10,813–10,877					TTC
*nad*6	10,933–11,394	10,842–11,300	10,884–11,336	10,881–11,348	ATG/TAG	GTG/TAG	GTG/TAG	ATG/TAA	
*trn*Y	11,403–11,467	11,862–11,928	11,339–11,402	11,340–11,403					GTA
SNR	11,468–11,689	11,301–11,632	11,472–11,526	11,404–11,585					
*trn*L	11,757–11,824	11,633–11,698	11,410–11,471	11,667–11,728					TAG
*trn*L*		11,705–11,768							CAA
*trn*R*		11,782–11,838							ACG
*trn*S	11,690–11,755	11,927–11,999	11,527–11,593	11,586–11,647					TGA
*trn*L	11,863–11,925	12,046–12,111	11,602–11,663	11,751–11,813					TAA
*trn*R	11,926–11,979	12,124–12,180	11,674–11,731	11,830–11,888					ACG
*nad*5	11,984–13,564	12,181–13,752	11,735–13,297	11,889–13,457	ATG/TAA	ATG/TAA	ATG/TAG	ATG/TAA	
LNR	13,565–13,730	13,753–14,148	13,298–13,361	13,458–13,776					

**Stands for the gene in repeat region.*

to the recently described cysticerci species in plateau pika from Qinghai province (Wu et al., 2021). Here the *T. caixuepengi* larva was also found to be parasitic in plateau pikas from Xietongmen, Saga and Sa'gya county of Tibet and Qilian county of Qinghai (**Supplementary Table 1**), and so far, not found in other sympatric rodent species. Due to the absence of comparative studies, it is currently not clear if this species has preference only for plateau pikas as its intermediate host. The larvae of *T. crassiceps* and *V. mustelae* were, respectively, found in Jeminay and Xinyuan county of Xinjiang autonomous region (**Supplementary Table 1**), which were also distributed on the northeast QTP (Li et al., 2013; Zhao et al., 2014). Wide geographic distributions of identical species suggest that their endemic geographic range should be far beyond the available survey data. Except for *Mesocestoides* sp. RKZ08 identified in this study, the *Me. litteratus* is another species of the genus *Mesocestoides* reported in Qinghai and Heilongjiang provinces of China (Wang et al., 2006; Li et al., 2013). However, there is no

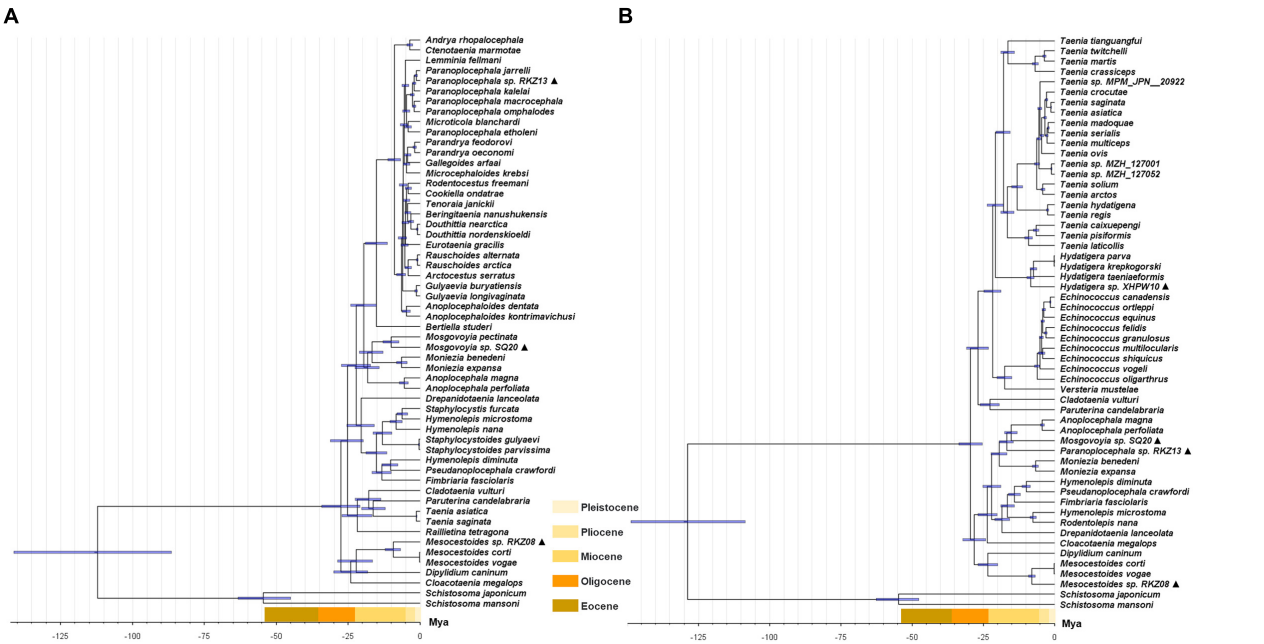

FIGURE 4 | Divergence time construction of concatenated *cox*1 + *nad*1 gene **(A)** and 12 *mt* PCGs **(B)** of Cyclophyllidea species. The ▲ after the species name represents "putative new species". The blue bar represents interval of 95% highest probability density. The time scale bars in different colors shows the extent of the Eocene, Oligocene, Miocene, Pliocene, and Pleistocene period.

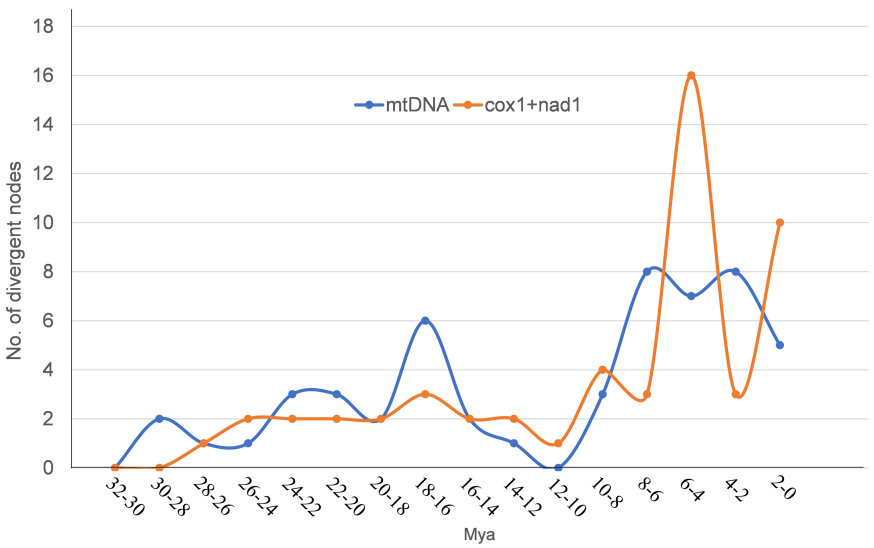

FIGURE 5 | The number of divergent nodes in every 2 Mya over time based on the divergence time trees of **Figure 4**. The orange curve represents the change of divergence nodes number based on **Figure 4A**. The blue curve represents the change of divergence nodes number based on **Figure 4B**.

previous record of *Paranoplocephala* spp. and *Mosgovoyia* spp. available in China except for *Paranoplocephala* sp. RKZ13 and *Mosgovoyia* sp. SQ20 found in this study.

Complete *mt* genomes of the four "putative new species" were sequenced and annotated, and the sequences were clearly different from all available *mt* genomes sequences; however, they were similar in length, gene order and composition as other cyclophyllideans with respect to rRNA, tRNA, and protein-encoding genes (Le et al., 2000; Jeon et al., 2005, 2007; Wu et al., 2021). Different arrangement of genes occurred in tRNA genes between *nad*5 and *nad*6 genes, but was consistent with their respective most relative species, matching the expectation that the arrangement of *mt* genes partly determines the genetic relationship of parasites (Gazi et al., 2016). Moreover, repeat

copies of tRNA gene between the *nad*5 and *nad*6 genes in the *mt* genome of *E. granulosus* sequenced by third-generation sequencing have been reported (Kinkar et al., 2019). We also observed a repeat sequence of *trn*L and *trn*R genes between *nad*5 and *nad*6 genes of the *Paranoplocephala* sp. RKZ13 *mt* genome (**Figure 3** and **Table 1**), which indicates that there may be hidden tRNA gene repeats in the *mt* genome of tapeworms that are hard to identify due to errors in PCR amplification and Sanger sequencing, techniques which often fail to recover repeat regions (Kinkar et al., 2019).

Genetic drift and adaptive differentiation between allopatric populations is responsible for most speciation amongst plants and animals (Turelli et al., 2001). For parasites, however, host association is a key driver in their evolution. Host switching among sympatric populations may lead to ecological isolation, so sympatric speciation of parasites is common (de Meeûs et al., 1998; Paul, 2002; Huyse et al., 2005; Wu et al., 2021). These two models of evolution are not in conflict: the adaptation of parasite to specific host is like the adaptation of animal to specific living environment; so, the evolution of the host, especially its immune system, may be viewed as equivalent to the change of the living environment for the parasite. The environmental and climatic changes caused by the uplift of the QTP are a major driving force for the evolution of associated biotas (Favre et al., 2015). Paleobotanical data suggest that the southeastern margin of the QTP was dominated by a warm and humid subtropical or tropical climate during the Miocene due to the influence of South Asian and East Asian monsoons (Sun and Wang, 2005; Jacques et al., 2011). Since the mid-Miocene, the significant rise of the Himalayas and the Tianshan Mountains, together with worldwide cooling, incurred dramatic changes in air circulation, leading to gradual aridification of the QTP and Central Asia (Miao et al., 2012; Favre et al., 2015). Finally, in the Late Miocene and Early Pliocene, QTP uplift resulted in the accumulation of global ice and the eventual disappearance of the Tethys Sea, which also contributed to the drying of Central Asia (Lu and Guo, 2013).

These timelines of climatic and environmental changes caused by the uplift of the QTP are highly consistent with the timelines of the differentiation rate of Cyclophyllidea species analyzed in this study (**Figures 4**, **5**) and that of plateau pika analyzed in Wang et al. (2020). We speculated in this study that the differentiation of cyclophyllideans may have been driven by host evolution caused by the uplift of the QTP. During the tropical period of QTP, the optimum living environment created the biological diversity (Cai et al., 2020; Päckert et al., 2020), and the species of Cyclophyllidea gradually differentiated. With the rapid uplift of the QTP, the environment changed into a dry and cold climate (Lu and Guo, 2013), and species differentiation of Cyclophyllidea was accelerated by the rapid adaptive evolution of their hosts and geographical isolation caused by the radiation of hosts to the Palaearctic (Favre et al., 2015; Xing and Ree, 2017; Päckert et al., 2020). Finally, in the last 2 million years, Cyclophyllidea differentiation demonstrated an accelerated diversification based on *cox*1 + *nad*1 divergence tree (**Figure 5**), which may be related to the evolution and broad distribution of mammals in Eurasia and the associated population expansion and migration of hominids from Africa to Asia (Dennell, 2004; Rohland et al., 2005; Brugal and Croitor, 2007; Turner et al., 2008; Klein, 2009; Terefe et al., 2014; Wang et al., 2016). The Host-parasite Database of Natural History Museum (HPDNHM) found that most tapeworm were prevalent mainly in the Palaearctic,[5] which is also consistent with the viewpoint that the order Cyclophyllidea originated from the QTP. The Nearctic is another major endemic area where Cyclophyllidea species may have spread over land Bridges across the Bering Strait. However, some species parasitizing birds and other economic and companion animals tend to show a global epidemic, which may be due to the long-distance migration of birds and the spread of human trade and activities.

Phylogenetic reconstruction reveals that some classifications of Cyclophyllidea species may need to be redefined. Closely related species of tapeworm parasites often have similar host specificities and life history (Scholz et al., 2021), a pattern common in the HPDNHM and in our evolutionary analyses, and may provide a basis for revising the classification of Cyclophyllidea species; for example, the family Hymenolepididae and Anoplocephalidae can be divided into multiple families (**Figures 1A–C**), and *Thysanotaenia congolensis* should be reclassified into family Davaineidae (**Figures 1A,C**). However, considering that there is still a large number of undiscovered species, which may provide better support for classification, the current classification status is likely to remain for some time.

CONCLUSION

In conclusion, this study expands the biodiversity of Cyclophyllidea in rodents in QTP and its surrounding mountain systems, and suggests an "out of QTP" hypothesis for the Cyclophyllidea, wherein species differentiation was driven by the uplift of the QTP. Although beyond the scope of this study to consider the evolutionary relationships and history of the whole cyclophyllideans, the species analyzed represent 10 of all 16 families, making this the most extensive study of the evolution of Cyclophyllidea order to date. Verifying the taxonomic revision and the "out of QTP" hypothesis requires more sampling and investigation, including data on a wider geographic and host range, and molecular studies uncovering patterns of host-parasite co-evolution.

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/genbank/, MZ476188–MZ476193; https://www.ncbi.nlm.nih.gov/genbank/, MW808979–MW808982; https://www.ncbi.nlm.nih.gov/genbank/, OM140661–OM140665.

[5]https://www.nhm.ac.uk/research-curation/scientific-resources/taxonomy-systematics/host-parasites/database/index.jsp

ETHICS STATEMENT

The animal study was reviewed and approved by the Animal Ethics Procedures and Guidelines of the People's Republic of China. Animal Ethics Committee of Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Sciences (No. LVRIAEC2012-007).

AUTHOR CONTRIBUTIONS

Y-DW, LL, H-BY, and W-ZJ conceived and designed the experiments. Y-DW, G-DD, L-SZ, A-MG, Y-TW, and YF conducted the sample collection. Y-DW, G-DD, X-WN, and NS performed the experiments. Y-DW and G-DD performed the data analyses. Y-DW prepared the figures and wrote the manuscript. DTJL, JO, W-HL, N-ZZ, and B-QF provided very constructive suggestions for revisions. All authors read and approved the final manuscript.

FUNDING

This work was funded by the National Key Research and Development Plan (2021YFE0191600), Cultivation of Achievements (SKLVEB2020CGPY01), and Open Project (SKLVEB2020KFKT008) of State Key Laboratory of Veterinary Etiological Biology, and the Applied Basic Research of Qinghai Province (2021-ZJ-724).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.747484/full#supplementary-material

Supplementary Figure 1 | Maximum likelihood analysis based on the 28S rDNA fragments of classified cyclophyllideans in NCBI Taxonomy Database. The ▲ after the species name represents "putative new species". The outgroup is the same as **Figure 1**. Bootstrap frequency support values are stated only for nodes where > 80%.

REFERENCES

Bouckaert, R., Heled, J., Kuhnert, D., Vaughan, T., Wu, C. H., Xie, D., et al. (2014). BEAST 2: a software platform for Bayesia evolutionary analysis. *PLoS Comput. Biol.* 10:e1003537. doi: 10.1371/journal.pcbi.1003537

Bowles, J., Blair, D., and McManus, D. P. (1992). Genetic variants within the genus Echinococcus identified by mitochondrial DNA sequencing. *Mol. Biochem. Parasitol.* 54, 165–173. doi: 10.1016/0166-6851(92)90109-w

Brugal, J. P., and Croitor, R. (2007). Evolution, ecology and biochronology of herbivore associations in Europe during the last 3 million years. *Quaternaire* 18, 129–152.

Cai, T., Shao, S., Kennedy, J. D., Alström, P., Moyle, R. G., Qu, Y., et al. (2020). The role of evolutionary time, diversification rates and dispersal in determining the global diversity of a large radiation of passerine birds. *J. Biogeogr.* 47, 1612–1625.

Caira, J. N., and Jensen, K. (eds) (2017). *Planetary Biodiversity Inventory (2008–2017): Tapeworms from Vertebrate Bowels of the Earth*. USA: University of Kansas. 463.

Capella-Gutiérrez, S., Silla-Martínez, J. M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25, 1972–1973. doi: 10.1093/bioinformatics/btp348

Carlson, C. J., Dallas, T. A., Alexander, L. W., Phelan, A. L., and Phillips, A. J. (2020). What would it take to describe the global diversity of parasites? *Proc. Biol. Sci.* 287:20201841. doi: 10.1098/rspb.2020.1841

de Meeûs, T., Michalakis, Y., and Renaud, F. (1998). Santa Rosalia revisited: or why are there so many kinds of parasites in the garden of earthly delights'? *Parasitol. Today* 14, 10–13. doi: 10.1016/s0169-4758(97)01163-0

Dennell, R. W. (2004). Hominid dispersals and Asian biogeography during lower and early middle Pleistocene, c. 2.0–0.5 Mya. *Asian Perspect.* 43, 205–226.

Dobson, A., Lafferty, K. D., Kuris, A. M., Hechinger, R. F., and Jetz, W. (2008). Colloquium paper: homage to Linnaeus: how many parasites? How many hosts? *Proc. Natl. Acad. Sci. U. S. A.* 105, 11482–11489. doi: 10.1073/pnas.0803232105

Ebert, D., and Fields, P. D. (2020). Host-parasite co-evolution and its genomic signature. *Nat. Rev. Genet.* 21, 754–768. doi: 10.1038/s41576-020-0269-1

Fan, Z., Liu, S., Liu, Y., Zhang, X., and Yue, B. (2011). How Quaternary geologic and climatic events in the southeastern margin of the Tibetan Plateau influence the genetic structure of small mammals: inferences from phylogeography of two rodents, *Neodon irene* and *Apodemus latronum*. *Genetica* 139, 339–351. doi: 10.1007/s10709-011-9553-5

Favre, A., Päckert, M., Pauls, S. U., Jähnig, S. C., Uhl, D., Michalak, I., et al. (2015). The role of the uplift of the Qinghai-Tibetan Plateau for the evolution of Tibetan biotas. *Biol. Rev. Camb. Philos. Soc.* 90, 236–253. doi: 10.1111/brv.12107

Gazi, M., Kim, J., García-Varela, M., Park, C., Littlewood, D. T. J., and Park, J. K. (2016). Mitogenomic phylogeny of Acanthocephala reveals novel Class relationships. *Zoologica Scripta* 45, 437–454.

Heled, J., and Drummond, A. J. (2015). Calibrated birth-death phylogenetic time-tree priors for Bayesian inference. *Syst. Biol.* 64, 369–383. doi: 10.1093/sysbio/syu089

Hoorn, C., Perrigo, A., and Antonelli, A. (eds) (2018). *Mountains, Climate, and Biodiversity (1st ed.)*. Hoboken: Wiley.

Huyse, T., Poulin, R., and Théron, A. (2005). Speciation in parasites: a population genetics approach. *Trends. Parasitol.* 21, 469–475. doi: 10.1016/j.pt.2005.08.009

Jacques, F. M. B., Guo, S. X., Su, T., Xing, Y. W., Huang, Y. J., Liu, Y. S., et al. (2011). Quantitative reconstruction of the Late Miocene monsoon climates of southwest China: a case study of the Lincang flora from Yunnan Province. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 304, 318–327. doi: 10.1016/j.palaeo.2010.04.014

Jeon, H. K., Kim, K. H., and Eom, K. S. (2007). Complete sequence of the mitochondrial genome of *Taenia saginata*: comparison with *T. solium* and *T. asiatica*. *Parasitol. Int.* 56, 243–246. doi: 10.1016/j.parint.2007.04.001

Jeon, H. K., Lee, K. H., Kim, K. H., Hwang, U. W., and Eom, K. S. (2005). Complete sequence and structure of the mitochondrial genome of the human tapeworm, *Taenia asiatica* (Platyhelminthes; Cestoda). *Parasitology* 130, 717–726. doi: 10.1017/s0031182004007164

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., and Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14, 587–589. doi: 10.1038/nmeth.4285

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Kinkar, L., Korhonen, P. K., Cai, H., Gauci, C. G., Lightowlers, M. W., Saarma, U., et al. (2019). Long-read sequencing reveals a 4.4 kb tandem repeat region in the mitogenome of *Echinococcus granulosus (sensu stricto)* genotype G1. *Parasit. Vectors* 12:238. doi: 10.1186/s13071-019-3492-x

Kinkar, L., Laurimäe, T., Sharbatkhori, M., Mirhendi, H., Kia, E. B., Ponce-Gordo, F., et al. (2017). New mitogenome and nuclear evidence on the phylogeny and taxonomy of the highly zoonotic tapeworm *Echinococcus granulosus sensu stricto*. *Infect. Genet. Evol.* 52, 52–58. doi: 10.1016/j.meegid.2017.04.023

Klein, R. G. (2009). Darwin and the recent African origin of modern humans. *Proc Natl. Acad. Sci. U. S. A.* 106, 16007–16009. doi: 10.1073/pnas.0908719106

Kuchta, R., Řehulková, E., Francová, K., Scholz, T., Morand, S., and Šimková, A. (2020). Diversity of monogeneans and tapeworms in cypriniform fishes across two continents. *Int. J. Parasitol.* 50, 771–786. doi: 10.1016/j.ijpara.2020.06.005

Lanfear, R., Calcott, B., Ho, S. Y., and Guindon, S. (2012). PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Mol. Biol. Evol.* 29, 1695–1701. doi: 10.1093/molbev/mss020

Larsen, B. B., Miller, E. C., Rhodes, M. K., and Wiens, J. J. (2017). Inordinate fondness multiplied and redistributed: the number of species on earth and the new pie of life. *Q. Rev. Biol.* 92, 229–265. doi: 10.1086/693564

Laslett, D., and Canbäck, B. (2008). ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. *Bioinformatics* 24, 172–175. doi: 10.1093/bioinformatics/btm573

Le, T. H., Blair, D., Agatsuma, T., Humair, P. F., Campbell, N. J., Iwagami, M., et al. (2000). Phylogenies inferred from mitochondrial gene orders-a cautionary tale from the parasitic flatworms. *Mol. Biol. Evol.* 17, 1123–1125. doi: 10.1093/oxfordjournals.molbev.a026393

Lee, S. U., Chun, H. C., and Huh, S. (2007). Molecular phylogeny of parasitic Platyhelminthes based on sequences of partial 28S rDNA D1 and mitochondrial cytochrome c oxidase subunit I. *Korean J. Parasitol.* 45, 181–189. doi: 10.3347/kjp.2007.45.3.181

Letunic, I., and Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* 44, W242–W245. doi: 10.1093/nar/gkw290

Li, W., Guo, Z., Duo, H., Fu, Y., Peng, M., Shen, X., et al. (2013). Survey on helminths in the small intestine of wild foxes in Qinghai, China. *J. Vet. Med. Sci.* 75, 1329–1333. doi: 10.1292/jvms.13-0187

Littlewood, D. T., Curini-Galletti, M., and Herniou, E. A. (2000). The interrelationships of Proseriata (Platyhelminthes: Seriata) tested with molecules and morphology. *Mol. Phylogenet. Evol.* 16, 449–466. doi: 10.1006/mpev.2000.0802

Lu, H. Y., and Guo, Z. T. (2013). Evolution of the monsoon and dry climate in East Asia during late Cenozoic: a review. *Sci. China Earth Sci.* 57, 70–79. doi: 10.1007/s11430-013-4790-3

Lymbery, A. J. (1998). Combining data from morphological traits and genetic markers to determine transmission cycles in the tape worm, *Echinococcus granulosus*. *Parasitology* 117, 185–192. doi: 10.1017/s0031182098002911

Miao, Y. F., Herrmann, M., Wu, F. L., Yan, X. L., and Yang, S. L. (2012). What controlled Mid-Late Miocene long-term aridification in Central Asia? – Global cooling or Tibetan Plateau uplift: a review. *Earth Sci. Rev.* 112, 155–172. doi: 10.1016/j.earscirev.2012.02.003

Nakao, M., Lavikainen, A., Iwaki, T., Haukisalmi, V., Konyaev, S., Oku, Y., et al. (2013). Molecular phylogeny of the genus *Taenia* (Cestoda: Taeniidae): proposals for the resurrection of *Hydatigera* Lamarck, 1816 and the creation of a new genus *Versteria*. *Int. J. Parasitol.* 43, 427–437. doi: 10.1016/j.ijpara.2012.11.014

Nakao, M., Sako, Y., Yokoyama, N., Fukunaga, M., and Ito, A. (2000). Mitochondrial genetic code in cestodes. *Mol. Biochem. Parasitol.* 111, 415–424. doi: 10.1016/s0166-6851(00)00334-0

Nguyen, L. T., Schmidt, H. A., von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300

Okamura, B., Hartigan, A., and Naldoni, J. (2018). Extensive uncharted biodiversity: the parasite dimension. *Integr. Comp. Biol.* 58, 1132–1145. doi: 10.1093/icb/icy039

Okulewicz, A. (2014). Helminths in migrating and wintering birds recorded in Poland. *Ann. Parasitol.* 60, 19–24.

Päckert, M., Favre, A., Schnitzler, J., Martens, J., Sun, Y. H., Tietze, D. T., et al. (2020). "Into and Out of" the Qinghai-Tibet Plateau and the Himalayas: centers of origin and diversification across five clades of Eurasian montane and alpine passerine birds. *Ecol. Evol.* 10, 9283–9300. doi: 10.1002/ece3.6615

Paul, R. (2002). Species concepts versus species criteria. *Trends Parasitol.* 18, 439–440. doi: 10.1016/s1471-4922(02)02319-x

Poulin, R., and Morand, S. (2004). *Parasite Biodiversity*. Washington: Smithsonian Institution Press.

Rahbek, C., Borregaard, M. K., Antonelli, A., Colwell, R. K., Holt, B. G., Nogues-Bravo, D., et al. (2019a). Building mountain biodiversity: geological and evolutionary processes. *Science* 365, 1114–1119. doi: 10.1126/science.aax0151

Rahbek, C., Borregaard, M. K., Colwell, R. K., Dalsgaard, B., Holt, B. G., Morueta-Holme, N., et al. (2019b). Humboldt's enigma: what causes global patterns of mountain biodiversity? *Science* 365, 1108–1113. doi: 10.1126/science.aax0149

Rohland, N., Pollack, J. L., Nagel, D., Beauval, C., Airvaux, J., Pääbo, S., et al. (2005). The population history of extant and extinct hyenas. *Mol. Biol. Evol.* 22, 2435–2443. doi: 10.1093/molbev/msi244

Scholz, T., Waeschenbach, A., Oros, M., Brabec, J., and Littlewood, D.T.J. (2021). Phylogenetic reconstruction of early diverging tapeworms (Cestoda: Caryophyllidea) reveals ancient radiations in vertebrate hosts and biogeographic regions. *Int. J. Parasitol.* 51, 263–277. doi: 10.1016/j.ijpara.2020.09.009

Sharma, S., Lyngdoh, D., Roy, B., and Tandon, V. (2016). Molecular phylogeny of Cyclophyllidea (Cestoda: Eucestoda): an in-silico analysis based on mtCOI gene. *Parasitol. Res.* 115, 3329–3335. doi: 10.1007/s00436-016-5092-4

Singla, L. D., Singla, N., Parshad, V. R., Juyal, P. D., and Sood, N. K. (2008). Rodents as reservoirs of parasites in India. *Integr. Zool.* 3, 21–26. doi: 10.1111/j.1749-4877.2008.00071.x

Sun, X. J., and Wang, P. X. (2005). How old is the Asian monsoon system? Palaeobotanical records from China. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 222, 181–222. doi: 10.1016/j.palaeo.2005.03.005

Telford, M. J., Herniou, E. A., Russell, R. B., and Littlewood, D. T. (2000). Changes in mitochondrial genetic codes as phylogenetic characters: two examples from the flatworms. *Proc. Natl. Acad. Sci. U. S. A.* 97, 11359–11364. doi: 10.1073/pnas.97.21.11359

Terefe, Y., Hailemariam, Z., Menkir, S., Nakao, M., Lavikainen, A., Haukisalmi, V., et al. (2014). Phylogenetic characterisation of *Taenia* tapeworms in spotted hyenas and reconsideration of the "Out of Africa" hypothesis of *Taenia* in humans. *Int. J. Parasitol.* 44, 533–541. doi: 10.1016/j.ijpara.2014.03.013

Toews, D. P., and Brelsford, A. (2012). The biogeography of mitochondrial and nuclear discordance in animals. *Mol. Ecol.* 21, 3907–3930. doi: 10.1111/j.1365-294X.2012.05664.x

Turelli, M., Barton, N. H., and Coyne, J. A. (2001). Theory and speciation. *Trends Ecol. Evol.* 16, 330–343. doi: 10.1016/s0169-5347(01)02177-2

Turner, A., Antón, M., and Werdelin, L. (2008). Taxonomy and evolutionary patterns in the fossil Hyaenidae of Europe. *Geobios* 41, 677–687.

Wang, C. R., Qiu, J. H., Zhao, J. P., Xu, L. M., Yu, W. C., and Zhu, X. Q. (2006). Prevalence of helminthes in adult dogs in Heilongjiang Province, the People's Republic of China. *Parasitol. Res.* 99, 627–630. doi: 10.1007/s00436-006-0219-7

Wang, S., Wang, S., Luo, Y., Xiao, L., Luo, X., Gao, S., et al. (2016). Comparative genomics reveals adaptive evolution of Asian tapeworm in switching to a new intermediate host. *Nat. Commun.* 7:12845. doi: 10.1038/ncomms12845

Wang, X., Liang, D., Jin, W., Tang, M., Liu, S., and Zhang, P. (2020). Out of Tibet: genomic perspectives on the evolutionary history of extant pikas. *Mol. Biol. Evol.* 37, 1577–1592. doi: 10.1093/molbev/msaa026

Wang, X., Liu, J., Zuo, Q., Mu, Z., Weng, X., Sun, X., et al. (2018). *Echinococcus multilocularis* and *Echinococcus shiquicus* in a small mammal community on the eastern Tibetan Plateau: host species composition, molecular prevalence, and epidemiological implications. *Parasit. Vectors* 11:302. doi: 10.1186/s13071-018-2873-x

Wickström, L. M., Haukisalmi, V., Varis, S., Hantula, J., and Henttonen, H. (2005). Molecular phylogeny and systematics of anoplocephaline cestodes in rodents and lagomorphs. *Syst. Parasitol.* 62, 83–99. doi: 10.1007/s11230-005-5488-5

Wu, Y. D., Li, L., Fan, Y. L., Ni, X. W., Ohiolei, J. A., Li, W. H., et al. (2021). Genetic evolution and implications of the mitochondrial genomes of two newly identified *Taenia* spp. in rodents from Qinghai-Tibet Plateau. *Front. Microbiol.* 12:647119. doi: 10.3389/fmicb.2021.647119

Wu, Z., Lu, L., Du, J., Yang, L., Ren, X., Liu, B., et al. (2018). Comparative analysis of rodent and small mammal viromes to better understand the wildlife origin of emerging infectious diseases. *Microbiome* 6:178. doi: 10.1186/s40168-018-0554-9

Xiao, N., Qiu, J., Nakao, M., Li, T., Yang, W., Chen, X., et al. (2005). *Echinococcus shiquicus* n. sp., a taeniid cestode from Tibetan fox and plateau pika in China. *Int. J. Parasitol.* 35, 693–701. doi: 10.1016/j.ijpara.2005.01.003

Xing, Y., and Ree, R. H. (2017). Uplift-driven diversification in the Hengduan Mountains, a temperate biodiversity hotspot. *Proc. Natl. Acad. Sci. U. S. A.* 114, E3444–E3451. doi: 10.1073/pnas.1616063114

Yan, H., Lou, Z., Li, L., Ni, X., Guo, A., Li, H., et al. (2013). The nuclear 18S ribosomal RNA gene as a source of phylogenetic information in the genus *Taenia*. *Parasitol. Res.* 112, 1343–1347. doi: 10.1007/s00436-012-3199-9

Yudhana, A., Praja, R. N., and Supriyanto, A. (2019). The medical relevance of *Spirometra* tapeworm infection in Indonesian Bronzeback snakes (*Dendrelaphis pictus*): a neglected zoonotic disease. *Vet. World* 12, 844–848. doi: 10.14202/vetworld.2019.844-848

Zhang, D., Tang, L., Cheng, Y., Hao, Y., Xiong, Y., Song, G., et al. (2019). 'Ghost introgression' as a cause of deep mitochondrial divergence in a bird species complex. *Mol. Biol. Evol.* 36, 2375–2386. doi: 10.1093/molbev/msz170

Zhang, X. Y., Jian, Y. N., Ma, L. Q., Li, X. P., and Karanis, P. (2018). A case of coenurosis in a wild rabbit (*Lepus sinensis*) caused by *Taenia serialis* metacestode in Qinghai Tibetan Plateau area, China. *Korean J. Parasitol.* 56, 195–198. doi: 10.3347/kjp.2018.56.2.195

Zhao, F., Ma, J. Y., Cai, H. X., Su, J. P., Hou, Z. B., Zhang, T. Z., et al. (2014). Molecular identification of *Taenia mustelae* cysts in subterranean rodent plateau zokors (*Eospalax baileyi*). *Dongwuxue Yanjiu* 35, 313–318. doi: 10.13918/j.issn.2095-8137.2014.4.313

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.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 14 February 2022
doi: 10.3389/fmicb.2021.800190

The Phylosymbiosis Pattern Between the Fig Wasps of the Same Genus and Their Associated Microbiota

Jiaxing Li, Xianqin Wei, Dawei Huang and Jinhua Xiao**

College of Life Sciences, Nankai University, Tianjin, China

Microbial communities can be critical for many metazoans, which can lead to the observation of phylosymbiosis with phylogenetically related species sharing similar microbial communities. Most of the previous studies on phylosymbiosis were conducted across the host families or genera. However, it is unclear whether the phylosymbiosis signal is still prevalent at lower taxonomic levels. In this study, 54 individuals from six species of the fig wasp genus *Ceratosolen* (Hymenoptera: Agaonidae) collected from nine natural populations and their associated microbiota were investigated. The fig wasp species were morphologically identified and further determined by mitochondrial CO1 gene fragments and nuclear ITS2 sequences, and the V4 region of 16S rRNA gene was sequenced to analyze the bacterial communities. The results suggest a significant positive correlation between host genetic characteristics and microbial diversity characteristics, indicating the phylosymbiosis signal between the phylogeny of insect hosts and the associated microbiota in the lower classification level within a genus. Moreover, we found that the endosymbiotic *Wolbachia* carried by fig wasps led to a decrease in bacterial diversity of host-associated microbial communities. This study contributes to our understanding of the role of host phylogeny, as well as the role of endosymbionts in shaping the host-associated microbial community.

Keywords: insect, host-associated microbiota, holobiont, 16S rDNA, coevolution, *Wolbachia*

OPEN ACCESS

Edited by:
Rodrigo Pulgar Tejo,
University of Chile, Chile

Reviewed by:
Anne Duplouy,
University of Helsinki, Finland
Christian Hodar Quiroga,
University of Chile, Chile

****Correspondence:***
Dawei Huang
huangdw@ioz.ac.cn
Jinhua Xiao
xiaojh@nankai.edu.cn

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *22 October 2021*
Accepted: *22 December 2021*
Published: *14 February 2022*

Citation:
Li J, Wei X, Huang D and Xiao J (2022) The Phylosymbiosis Pattern Between the Fig Wasps of the Same Genus and Their Associated Microbiota. Front. Microbiol. 12:800190. doi: 10.3389/fmicb.2021.800190

INTRODUCTION

Microbes play important roles in hosts' biology. All insects are colonized by microbes, and the microbiota accounts for 1–10% of the insect's biomass (Douglas, 2015). Bacterial symbionts have been accepted by biologists as considerable drivers of insect nutrition, protection, detoxification, behavior, reproduction, communication, and evolution (Douglas, 2015; Salcedo-Porras et al., 2020). Historically, research on symbiotic relationship between insects and microorganisms has focused mainly on insect hosts and their obligate bacterial symbionts, such as aphids and *Buchnera aphidicola* (Chong et al., 2019), whiteflies and *Portiera aleyrodidarum* (Santos-Garcia et al., 2020), psyllids and *Carsonella ruddii* (Thao et al., 2000), and mealybugs and *Tremblaya princeps* (Lopez-Madrigal et al., 2015), or focused on specific endosymbionts known to be widespread in arthropods, such as *Wolbachia* (Hertig and Wolbach, 1924) or *Spiroplasma* (Davis et al., 1972; Cisak et al., 2015). However, symbioses are formed by complex multi-part interactions, including interactions between the hosts and their resident microbiota as well as interactions within the microbial community. Under these circumstances, binary interactions between hosts and endosymbionts

were broadened to multivariate interactions between hosts and all microbes of their associated microbial community, as they are so important to the hosts (Brinker et al., 2019). A host organism and its associated microbial community form an entity that is termed as holobiont, which is considered as a unit of selection in evolution (Rosenberg et al., 2007). This understanding clarifies the holobiont as a complex ecosystem in which the host and associated microbiota are closely linked.

A deluge of data has enabled unprecedented insights into the extensive taxonomic, genetic, and functional compositions of host-associated microbial communities (Lim and Bordenstein, 2020). However, our understanding of the interaction between microbiota and host is still relatively superficial due to quantitative limiting factors, such as incredibly diverse interactions between microbiota and host, and the technical limitations that plenty of microbes *in vivo* are difficult to be isolated and cultured *in vitro* (Lagier et al., 2018). It is generally considered that the host-associated microbiota can be shaped primarily by diets, environment, and host phylogeny (Kartzinel et al., 2019). In addition, interactions within microbiota can also shape the diversity and structure of insect bacterial communities (Guegan et al., 2018), in which the endosymbionts *Wolbachia* and *Spiroplasma* are widely studied. For example, parts of the native mosquito microbiota can impede vertical transmission of *Wolbachia* in *Anopheles* (Hughes et al., 2014). In *Drosophila melanogaster*, *Spiroplasma* reduced *Wolbachia* density but not vice versa (Goto et al., 2006), while in spider mite, *Wolbachia* dominated relative to *Spiroplasma* (Yang et al., 2020). These observations suggest exclusion or competition within microbiota. Unfortunately, data on microbial interactions are still scarce, and the mechanisms involved remain unclear.

In recent years, a new term phylosymbiosis has been proposed to describe the interaction between host and associated microbiota (Brucker and Bordenstein, 2013). According to the study of phylosymbiosis, host-associated microbial community relationships recapitulate the phylogeny of their hosts, indicating that the relationships of microbiota across host species maintain an ancestral signal of the host's evolution (Brucker and Bordenstein, 2013). Therefore, host phylogeny can reflect or be reflected by microbial community structure. In other words, phylosymbiosis may reveal whether there is an interaction between host phylogeny and bacterial community. Coevolution, cospeciation, codiversification, or cocladogenesis may lead to phylosymbiosis, and this pattern may alternatively arise by antagonistic interactions and/or horizontal transmission of the microbiota due to the direct microbial transfer between related individuals (Lim and Bordenstein, 2020). To date, a great deal of phylosymbiosis systems have been explored, such as mammals (Groussin et al., 2017), omnivorous cockroaches (Tinker and Ottesen, 2020), widow spider (Dunaj et al., 2020), freshwater snails (Huot et al., 2019), coral reef fish (Chiarello et al., 2018), and their associated microbiota. However, the host taxa in these studies were mainly across orders, families, or genera. Due to the long-term differentiation and distant genetic relationship between hosts, their microbial communities are distinct. Therefore, the phylosymbiosis patterns are easily detected in the case of the studies described above. In the lower taxa, within a genus for instance, different species are phylogenetically closely related, and according to the studies of phylosymbiosis, the more closely related the taxa are, the more similar the compositions of their associated microbial communities are (Brucker and Bordenstein, 2012a,b; O'Brien et al., 2019). In a previous study, phylosymbiosis signals were observed in salamanders and frogs at the taxonomic levels of order, but were not observed within genera and species (Ellison et al., 2019). There are few studies on phylosymbiosis within genera, and even if there are, the subjects were mostly lab-fed organisms, such as *Nasonia* wasps (Brucker and Bordenstein, 2012a, 2013), *Drosophila* flies, and *Peromyscus* deer mice (Brooks et al., 2016).

In this study, we collected fig pollinators (Hymenoptera: Agaonidae) to explore their microbiota and test whether a phylosymbiosis signal between insect hosts and their associated microbiota is prevalent among closely related species within a genus. Fig-pollinating wasps are the only pollinators of fig trees (*Ficus*, Moraceae) (Corner, 1958; Berg, 1989), and they spend almost their whole lives in fig fruits (syconia), which is a relatively closed and stable system. The larvae of fig pollinators only feed on the galled fig flowers (Janzen, 1979), so their diets are simple. The characteristic life histories and diets make the fig pollinators and their microbial community an excellent model to experimentally investigate evolutionary dynamics of host-microbiota interactions. In addition, endosymbiont *Wolbachia* is highly prevalent in fig pollinators (Chen et al., 2010), which allows us to investigate the effects of the endosymbiont *Wolbachia*, the phylogenetic relationship of fig wasp hosts, and the unique symbiotic environment provided by fig fruit on the host bacterial communities.

In this study, six fig pollinator species of the genus *Ceratosolen* and their microbial community were investigated. The mitochondrial cytochrome c oxidase subunit 1 (CO1) gene fragment and nuclear marker internal transcribed spacer 2 (ITS2) sequences were used to reconstruct the phylogenetic relationship of nine populations of the six species. The 16S ribosomal DNA amplicon sequencing was used to analyze the bacterial communities of these nine populations. We detected significant phylosymbiosis signal between the fig pollinators and their microbial communities. Our results also showed that *Wolbachia* was the dominant bacteria in the infection samples, and the fig wasps infected with *Wolbachia* had a lower bacterial diversity than those not infected. These results revealed the phylosymbiosis relationship between hosts and microbial communities in natural insect populations at a low taxon level, and the effect of the endosymbiotic *Wolbachia* on shaping host microbial communities.

MATERIALS AND METHODS

Sample Collection

Figs of *Ficus semicordata*, *Ficus racemosa*, *Ficus tikoua*, *Ficus fistulosa, Ficus auriculata*, and *Ficus hispida*, which are hosts of *Ceratosolen gravelyi*, *Ceratosolen fusciceps*, *Ceratosolen* sp., *Ceratosolen hewitti*, *Ceratosolen emarginatus*, and *Ceratosolen*

solmsi, respectively, were collected from Yunnan, Hainan, Guangxi, Guizhou, and Guangdong provinces in China (**Supplementary Table 1**). The figs we collected were wiped with alcohol cotton balls before being placed in sterile plastic cups. Each fig was put in a plastic cup and reserved in a climate chamber (humidity 70%, 16:8 h/L: D, 25°C/25°C). We collected adult fig wasps as soon as we found them emerging from figs, and then collected every 5–10 min for about 5 h per fig. The wasps collected were put in ethanol and stored at −20°C refrigerator. We collected dozens or even hundreds of fig pollinators per fig, and then randomly selected the unwounded female individuals that do not lack appendages for the experiment. The fig wasps immersed in ethanol were identified and selected under stereozoom microscope (Motic SMZ-168), and the identification was confirmed by CO1 gene fragment and ITS2 sequences (as described below). Fig wasps of nine populations from six species were used in the following analysis, with six samples from each population.

DNA Extraction and PCR Amplification

Total genomic DNA from each fig wasp was extracted with DNeasy Blood and Tissue Kit (Qiagen, Germany) according to the protocol. Single fig pollinator was washed three times with sterile water before the extraction of genomic DNA. DNA concentration and purity were monitored on 1% agarose gels. According to the concentration, DNA was diluted to 1 ng/μl using sterile water and stored at −20°C refrigerator. The host CO1 gene was amplified by primers LCO1490 (5′-GGTCA ACAAATCATAAAGATATTGG-3′) and HCO2198 (5′-TAAA CTTCAGGGTGACCAAAAAATCA-3′) (Folmer et al., 1994). The nuclear marker ITS2 was amplified by four pairs of primers. Specifically, the ITS2 of *C. gravelyi*, *C. emarginatus*, and *Ceratosolen* sp. was amplified by primers 5.8s-Fc (5′-TGAACATCGACATTTYGAACGCACAT-3′) and 28S-D4-5R (5′-GTTACACACTCCTTAGCGGA-3′) (Cruaud et al., 2010). The ITS2 of *C. solmsi*_2 was amplified by primers Aed5.8F (5′-GTGAACTGCAGGACACATGAAC-3′) and AedAB28 (5′-ATATGCTTAAATTTAGGGGGT-3′) (Kjer et al., 1994; Brust et al., 1998). The ITS2 of *C. solmsi*_1 and *C. fusciceps* was amplified by primers designed based on *C. solmsi* genome (accession number: PRJNA277475), namely, ITS2-L11F (5′-TTTGCGCGTCAACTTGTGAA-3′) and ITS2-L11R (5′-TCG CCGCTACTGAGGAAATC-3′), and ITS2-L9F (5′-GCAGG ACACATGAACATCGAC-3′) and ITS2-L9R (5′-TCTCAA GCAACCCGACTCTG-3′), respectively. The ITS2 sequence of *C. hewitti* failed to be amplified using the primers described above. PCR reaction system included 5 ng DNA template, 2.0 μM each primer, 0.2 mM dNTPs, 1.25 U of *EasyTaq*® DNA Polymerase, and 1 × *EasyTaq*® Buffer (TransGen Biotech, Beijing, China), then sterile water was added to a total volume of 25 μl. The PCR conditions of CO1 consisted of 5 min at 94°C, 35 cycles of 94°C for 30 s, 51°C for 45 s, 72°C for 1 min, and 10 min at 72°C. The amplification of ITS2 followed the protocol outlined for CO1 above, but 40 cycles of amplification, an annealing temperature of 54°C, and an extension time of 2 min and 30 s were employed. Blank controls were set during DNA extraction and PCR amplification to exclude contamination. Finally, the PCR products were purified and sequenced by conventional Sanger sequencing (Sangon Biotech, Shanghai, China).

Detection of *Wolbachia* in Fig Wasps

The presence/absence of *Wolbachia* in a fig wasp was detected by a PCR-based assay with *Wolbachia*-specific primers. Three pairs of *Wolbachia*-specific primers including wsp81F (5′-TGGTCCAATAAGTGATGAAGAAAC-3′) and wsp691R (5′-AAAAATTAAACGCTACTCCA-3′) (Zhou et al., 1998), FtsZ-F (5′-TACTGACTGTTGGAGTTGTAACTAACGCGT-3′) and FtsZ-R (5′-TGCCAGTTGCAAGAACAGAAACTCTAACTC-3′) (Jeyaprakash and Hoy, 2000), and 16SwolF (5′-TTGTAGC CTGCTATGGTATAACT-3′) and 16SwolR (5′-GAATAGGTA TGATTTTCATGT-3′) (O'Neill et al., 1992) were used. Each fig wasp was individually diagnosed with *Wolbachia*. Amplifying conditions were the same as used in the CO1 gene except for the annealing temperature, which were 55, 55, and 47°C for *wsp*, *fts*Z, and 16S rRNA gene, respectively. Blank control was set during PCR amplification, and sterile water of equal volume was used instead of template. Amplified fragments were revealed under UV light after migration on 1% agarose gel electrophoresis. A fig wasp was confirmed to be infected with *Wolbachia* only when all three genes were successfully amplified. Similarly, if none of the three genes could be amplified, the individual was considered not infected with *Wolbachia*.

Phylogenetic Reconstruction of Pollinators and Fig Trees

Phylogenetic reconstruction of insects was performed by combining CO1 fragments and ITS2 sequences. We obtained CO1 sequences for all 54 samples and 40 ITS2 sequences (at least one for each population except *C. hewitti*). The genetic distance over CO1 and ITS2 sequences between populations was estimated with MEGA7 (Tamura and Nei, 1993; Kumar et al., 2016). For phylogenetic trees, a fig pollinator species of *Kradibia gibbosae* was set as the outgroup and its mitochondrial sequence containing CO1 was downloaded from GenBank (accession number: MT947598.2). The ITS2 sequence of *K. gibbosae* was extracted from the genome (accession number: PRJNA641212). We combined CO1 and ITS2 sequences and used two methods to perform phylogenetic analysis in PhyloSuite v1.2.2 (Zhang et al., 2020). First, we made multiple sequence alignments with MAFFT v7.036 (Katoh et al., 2019). According to the Akaike Information Criterion (AICc), PartitionFinder2 (Lanfear et al., 2017) revealed GTR + G as the best evolutionary model for our data (Guindon et al., 2010; Lanfear et al., 2012). Maximum likelihood (ML) tree was constructed using IQ-TREE (Nguyen et al., 2015). Second, we constructed a Bayesian tree (Ronquist et al., 2012) under GTR + G model (Guindon et al., 2010; Lanfear et al., 2012). Phylogenetic trees were visualized with FigTree v1.4.4[1].

In order to elucidate the genetic relationships among various fig trees, a phylogenetic reconstruction based on three nuclear DNA markers, internal transcribed spacer (ITS), external transcribed spacer (ETS), and glyceraldehyde-3-phosphate

[1]http://tree.bio.ed.ac.uk/software/figtree/

dehydrogenase (G3pdh) was conducted. The sequences of three nuclear markers of the six species of fig trees corresponding to the six species of pollinators and the host of *K. gibbosae* (as outgroup) were obtained from NCBI, and the accession numbers were listed in **Supplementary Table 2**. The genetic distance over ITS-ETS-G3pdh sequences between fig trees was estimated with MEGA7 (Kimura, 1980; Kumar et al., 2016). The reconstruction process of ML (Nguyen et al., 2015) and Bayesian phylogenetic tree (Ronquist et al., 2012) was the same as described above, but under GTR + I + G, K81, TRN + I models or GTR + I + G, K80, F81 + I models for ITS, ETS, and G3pdh, respectively (Guindon et al., 2010; Lanfear et al., 2012, 2017). Then, the outgroup was removed and the phylogenetic tree was used for phylosymbiosis analysis.

Molecular Species Delimitation

We carried out four approaches for species delimitation in order to have more robust results, including Automatic Barcode Gap Discovery (ABGD) (Puillandre et al., 2012), which detects the barcode gap based on the user-defined boundaries for intraspecific variability, then sorts the sequences into hypothetical species with *p*-values; a Java program uses an explicit, determinate algorithm to define Molecular Operational Taxonomic Unit (jMOTU) (Jones et al., 2011), and clustering-based approaches, e.g., Generalized Mixed Yule Coalescent (GMYC) (Pons et al., 2006), which uses likelihood to identify species boundaries by detecting the transition point between the speciation process and intraspecific lineage coalescence, and Bayesian implementation of the Poisson Tree Processes model (bPTP) (Zhang et al., 2013), which assumes independent exponential distributions to model the branch lengths for speciation and for coalescence. The first two methods (ABGD and jMOTU) were based on genetic distances, and the latter two methods (GMYC and bPTP) were based on the inferred tree (Pons et al., 2006; Jones et al., 2011; Puillandre et al., 2012; Zhang et al., 2013). The ABGD analyses were performed at the web server[2], with the following settings: relative gap width $X = 1.0$, K2P distance and intraspecific divergence (P) values range from 0.001 to 0.1, and other parameter values employed defaults. The bPTP analyses were conducted on the web server[3] using rooted phylogenetic input tree attached to **Supplementary Figure 1**, and the following settings were employed: 100,000 MCMC generations, thinning interval of 100, and the first 10% were discarded as burn-in. GMYC delimits distinct genetic clusters by optimizing the set of modes defining the transitions between inter- and intraspecific processes; the analysis was conducted using BEAST 1.8.0 under a strict clock model and speciation with Birth-Death process model (Drummond et al., 2012); the runs consisted of 10 million generations sampled every 1,000 cycles, convergence was assessed by ESS values, and a burn-in with 25% was set to obtain an optimal consensus tree; we then used the obtained tree to analyze the data under the GMYC species delimitation approach in the software R v4.0.1 (R Core Team, 2021) with the package *splits* using the single-threshold method (Ezard et al., 2015). The jMOTU uses predefined thresholds to calculate the genetic differences within average sequence length. The results of the species delimitation were visualized *via* iTOL v6[4] (Letunic and Bork, 2021).

16S rDNA Library Preparation

The V4 region of the 16S rRNA gene was PCR-amplified at Novogene Bioinformatics Technology Co., Ltd. (Beijing, China) by specific primers (515F-806R) with the barcode following Kueneman et al. (2014). Each PCR reaction was carried out with 10 ng template DNA, 15 μl of Phusion® High-Fidelity PCR Master Mix (New England Biolabs, Ipswich, MA, United States), and 2 μM of forward and reverse primers. One thermal cycle consisted of initial denaturation at 98°C for 1 min, followed by 30 cycles of denaturation at 98°C for 10 s, annealing at 50°C for 30 s, elongation at 72°C for 30 s, and finally 72°C for 5 min. Blank control was set during PCR amplification, and sterile water of equal volume was used instead of template.

Each sample was amplified, and the PCR products were tested for concentration using Qubit@ 2.0 Fluorometer (Thermo Fisher Scientific, Cleveland, OH, United States). There was only one sequencing library per sample. Equal concentrations of each sample were pooled, and the pooled amplicons were cleaned using Qiagen Gel Extraction Kit (Qiagen, Germany). Sequencing libraries were generated using TruSeq® DNA PCR-Free Preparation Kit (Illumina, Inc., San Diego, CA, United States) following manufacturer's protocols and index codes were added. The library quality was assessed on the Qubit@ 2.0 Fluorometer (Thermo Fisher Scientific, Cleveland, OH, United States) and Agilent Bioanalyzer 2100 system. Finally, the library was sequenced on Illumina NovaSeq6000 platform at the Novogene Bioinformatics Technology Co., Ltd. (Beijing, China) and 250 bp paired-end reads were obtained.

Sequence Processing

The 54 sequencing libraries we obtained ranged in size from 30.4 to 57.1 M, with an average of 44.6 M. Paired-end reads were assigned to samples based on their unique barcode and truncated by cutting off the barcode and primer sequence. Paired-end reads were merged into a single sequence using Fast Length Adjustment of Short Reads (FLASH) v1.2.7, and the splicing sequences were called raw tags (Magoč and Salzberg, 2011). Then, raw tags were filtered and pre-processed in Qualitative Insights into Microbial Ecology (QIIME) v1.9.1 (Caporaso et al., 2010b; Bokulich et al., 2013). During this process, QIIME's default quality-control parameters were used. The tags were compared with the Silva132 database using UCHIME algorithm to detect chimera sequences, and then all chimera sequences were removed (Edgar et al., 2011; Haas et al., 2011).

Sequence analyses were performed by Uparse v7.0.1001. Sequences with $\geq$ 97% similarity were assigned to the same operational taxonomic unit (OTU) (Edgar, 2013). Representative sequence for each OTU was screened for further annotation. The Silva132 database was used based on Mothur algorithm

[2]https://bioinfo.mnhn.fr/abi/public/abgd/abgdweb.html

[3]https://species.h-its.org

[4]https://itol.embl.de

to annotate taxonomic information for each unique OTU (Quast et al., 2013).

Data Analysis

The sequence number of the sample with the least sequences was used as the standard for normalization to obtain OTUs abundance information. Subsequent analyses of alpha diversity and beta diversity were all performed basing on this output normalized data.

The rarefaction curves, Venn diagram, and heat map of the top 20 classes in bacterial abundance were carried out in R v4.0.1 software (R Core Team, 2021). Then, we inferred phylogenies of the top 20 genera in bacterial abundance of fig wasp by FastTree (Price et al., 2009) based on the sequence alignments of Greengenes database (DeSantis et al., 2006) and PyNAST (Caporaso et al., 2010a) software, and the bacterial abundance information was added.

Observed species, Chao1, ACE, Shannon, Simpson, and PD whole tree were calculated for all samples as measures of alpha diversity. All these indices in our samples were obtained by QIIME v1.9.1 and visualized with R v4.0.1 software (Caporaso et al., 2010b; R Core Team, 2021). One-way ANOVA was performed to test the difference of alpha diversity indices among host populations using SPSS v24. Spearman method was used to examine the correlations between the proportion of *Wolbachia* and the Shannon, Simpson indices and performed by *ggpubr* package (Kassambara, 2020) in R v4.0.1 software (R Core Team, 2021) and functions in the package that comes with R itself. The significance of differences in alpha diversity indices between group *Wolbachia*-infected and non-infected was evaluated by *t*-test, and *t*-test was carried out in R v4.0.1 software (R Core Team, 2021). Before the *t*-test, function qqplot in *car* package (Fox and Weisberg, 2019) and function bartlett.test of R v4.0.1 (R Core Team, 2021) software were used for the test of normal distribution and variance homogeneity. The sample sizes of group *W*+ and *W*− were 36 and 18, respectively; the sample size of *C. solmsi*_1 or *C. solmsi*_2 was six. When two variables have equal variances, the two sample *t*-test was used; otherwise, the Welch two sample *t*-test was used (R Core Team, 2021).

Beta diversity was calculated based on weighted and unweighted Unifrac, Bray–Curtis, and binary Jaccard distances in QIIME v1.9.1 (Caporaso et al., 2010b). Non-Metric Multi-Dimensional Scaling (NMDS) analysis was performed to visualize complex, multidimensional data. In this study, NMDS analyses based on Bray–Curtis and weighted Unifrac were displayed by *vegan* package in R v4.0.1 software (Oksanen et al., 2020; R Core Team, 2021). In order to determine whether the inter-population differences were significantly different from those intra-population and to evaluate whether the grouping was meaningful, the significance test of inter-population differences using Analysis of Similarity (ANOSIM) based on the Bray–Curtis distance value was calculated. At the same time, Multi-Response Permutation Procedure (MRPP) parametric test based on Bray–Curtis was used to analyze whether there is significant difference in microbial community structure between populations. Both ANOSIM and MRPP conducted by *vegan* package in R v4.0.1 software (Oksanen et al., 2020; R Core Team, 2021). Unweighted Pair-group Method with Arithmetic Means (UPGMA) Clustering was performed as a type of hierarchical clustering method to interpret the distance matrix using average linkage and was conducted by QIIME v1.9.1 (Caporaso et al., 2010b).

The phylosymbiosis was validated using both matrix and topological comparisons. Mantel test was used to analyze the correlation between the genetic distance of fig pollinators or fig trees and the microbial beta diversity distance matrices with the *ade4* package in R v4.0.1 software (Dray and Dufour, 2007; R Core Team, 2021). In addition to the Mantel test, the Robinson–Foulds (Robinson and Foulds, 1981) and Matching Cluster (Bogdanowicz and Giaro, 2013) congruency analysis between the phylogenetic tree of pollinators or fig trees and microbial UPGMA cluster tree was carried out according to Brooks' Python script (Brooks et al., 2016) and the TreeCmp program (Bogdanowicz et al., 2012). Statistical significance was evaluated by determining the probability of 100,000 randomized bifurcating dendrogram topologies with the same leaf nodes as the phylogenetic tree yielding equivalent or more congruent phylosymbiotic patterns than the microbiome dendrogram (Brooks et al., 2016). Normalized Robinson–Foulds (nRF) and normalized Matching Cluster (nMC) scores range from 0 (complete congruence) to 1.0 (complete incongruence) (Brooks et al., 2016).

RESULTS

The Phylogeny of Insects and *Wolbachia* Infection

After constructing the host phylogenetic tree by ML and Bayesian methods, we found that the same dendrogram was obtained from the two methods, and the phylogenetic relationship of all the species was well defined by CO1 genes combined with ITS2 sequences (**Supplementary Figure 1**). All the 54 samples were separated into nine clades, supported by a confidence value ranging from 56 to 100 (**Figure 1A**). Each of *C. gravelyi*, *C. emarginatus*, and *C. solmsi* had two clades. All conspecific individuals from the same population were clustered into one clade (**Figure 1A** and **Supplementary Table 1**).

Based on the approaches of ABGD, jMOTU, bPTP, and GYMC, all the samples were defined as nine clades, which was consistent with the results obtained from phylogenetic analysis (**Figure 1**). *C. gravelyi*, *C. emarginatus*, and *C. solmsi* were separated into two clades, although we did not find the morphological differences.

All the fig wasp individuals of *C. gravelyi*_1, *C. gravelyi*_2, *C. emarginatus*_1, *C. emarginatus*_2, *Ceratosolen* sp., and *C. solmsi*_1 were positive for *Wolbachia* infection (*W*+ group). No *Wolbachia* was detected using specific primers in the individuals from *C. fusciceps*, *C. hewitti*, and *C. solmsi*_2 (*W*− group) (**Figure 1B**).

Overall Microbial Community Compositions of the Fig Pollinators

Overall, 54 samples were successfully examined. Bacterial communities' compositions were studied using the Illumina

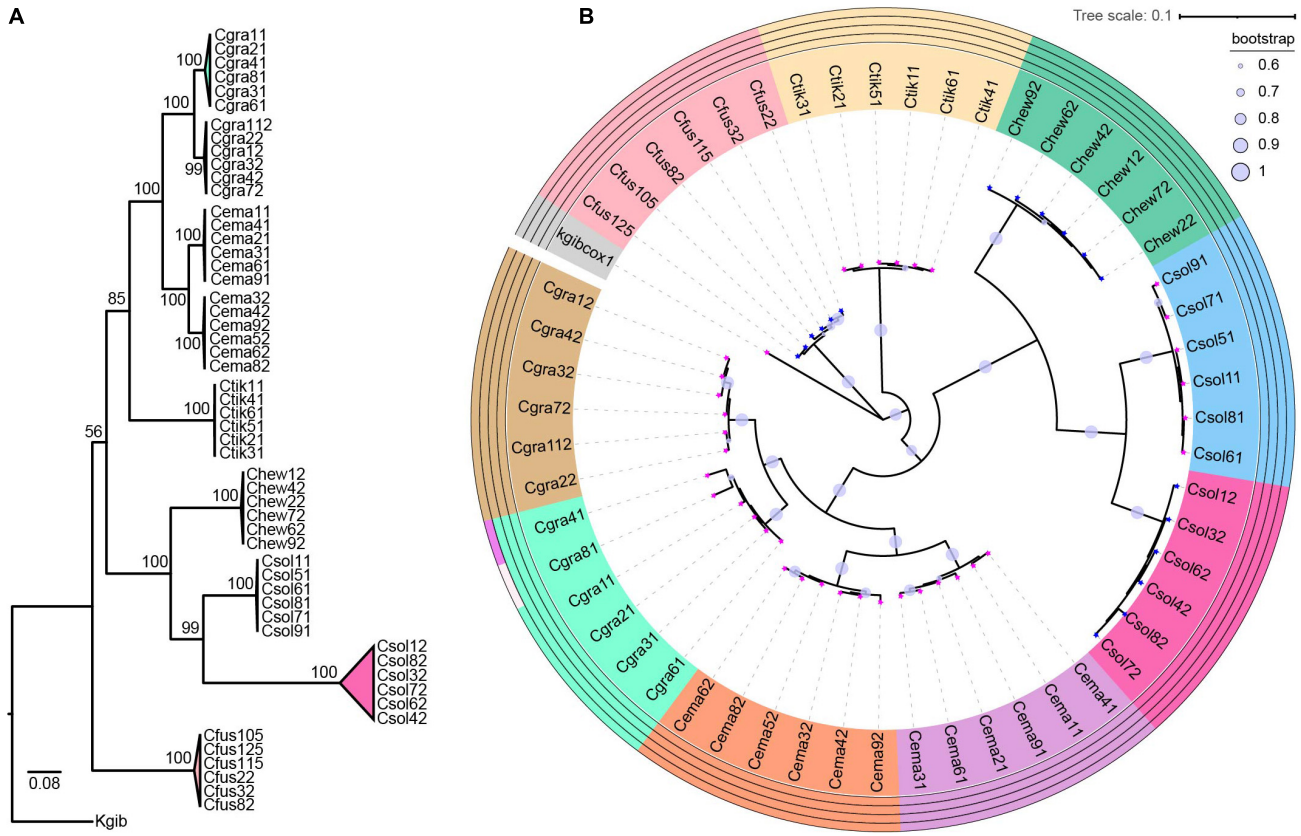

FIGURE 1 | Host phylogeny and molecular species delimitation. **(A)** Phylogenies of *Ceratosolen* by ML method based on CO1 gene fragments and ITS2 sequences. **(B)** According to four methods of ABGD, jMOTU, bPTP, and GYMC, all the samples were defined as nine clades. Species delimitation results are shown in the outer circles. At the tip of the branch, the magenta pentagram indicates *Wolbachia*-infected, and the blue pentagram indicates uninfected. Fig wasp species are abbreviated as follows: Cgra, *Ceratosolen gravelyi*; Cfus, *Ceratosolen fusciceps*; Ctik, *Ceratosolen* sp.; Chew, *Ceratosolen hewitti*; Cema, *Ceratosolen emarginatus*; Csol, *Ceratosolen solmsi*; Kgib, *Kradibia gibbosae*.

NovaSeq6000 platform. The filtered, high-quality sequence database obtained was 3,478,098 sequences, which were classified into 13,331 unique OTUs by a cutoff of 97% sequence similarity. Rarefaction analysis was performed at a threshold of 3% sequences dissimilarity for all samples. The great majority of the samples reached an asymptote level, indicating that our sampling efforts were sufficient to obtain an accurate estimate of OTU richness (**Supplementary Figure 2**).

The Venn diagram showed the unique and shared core OTUs among the populations. Of the total 13,331 OTUs obtained, 259 OTUs were shared by nine populations (**Figure 2A**). The proportions of shared OTUs in populations of *W*− group were lower than that of *W*+ group. Correspondingly, the populations in *W*− group had more specific OTUs than the populations in *W*+ group. In particular, the specific OTUs of *C. solmsi* in the *W*− group were almost 17 times higher than that in the *W*+ group (**Figure 2B**). A similar trend was also shown in the heat map of the top 20 classes in bacterial abundance, with more hot areas in the populations of *W*− group (**Figure 2C**).

In total, 12,763 (95.74%) OTUs were annotated based on Silva132 database. Approximately 95.75, 72.86, 61.20, 51.38, 29.80, and 8.12% of the OTUs were annotated at the level of phylum, class, order, family, genus, and species, respectively. Overall, 79 different phyla were detected. At the phylum level, the composition of bacteria was similar in all *Ceratosolen* samples, with Proteobacteria being the predominant phylum for all the populations except for *C. solmsi*_2 (**Figure 3A**). The subsequent dominant phyla were Firmicutes and Bacteroidetes. In the case of *C. solmsi*_2, Deinococcus-Thermus was dominant and followed by Proteobacteria, Firmicutes, and Bacteroidetes. At the class level, the populations from the *W*+ group were dominated by Alphaproteobacteria, while the populations from the *W*− group except for *C. solmsi*_2 were dominated by Gammaproteobacteria (**Figure 3B**). At the genus level, the dominant genus identified was *Wolbachia* for the *Wolbachia*-infected populations (**Supplementary Figure 3**), at a prevalence of 20.3–96.4%. By contrast, the dominant genus in *C. fuscipes*, *C. hewitti*, or *C. solmosi*_2 was totally different. It was *Acinetobacter*, *Citrobacter*, or *Thermus*, respectively. Surprisingly, almost half of the sequences could not be identified at the level of genera in *C. hewitti*.

In our study, a total of nine unique OTUs were classified as *Wolbachia*. In detail, OTU1 (93.5%) and OTU8 (6.2%)

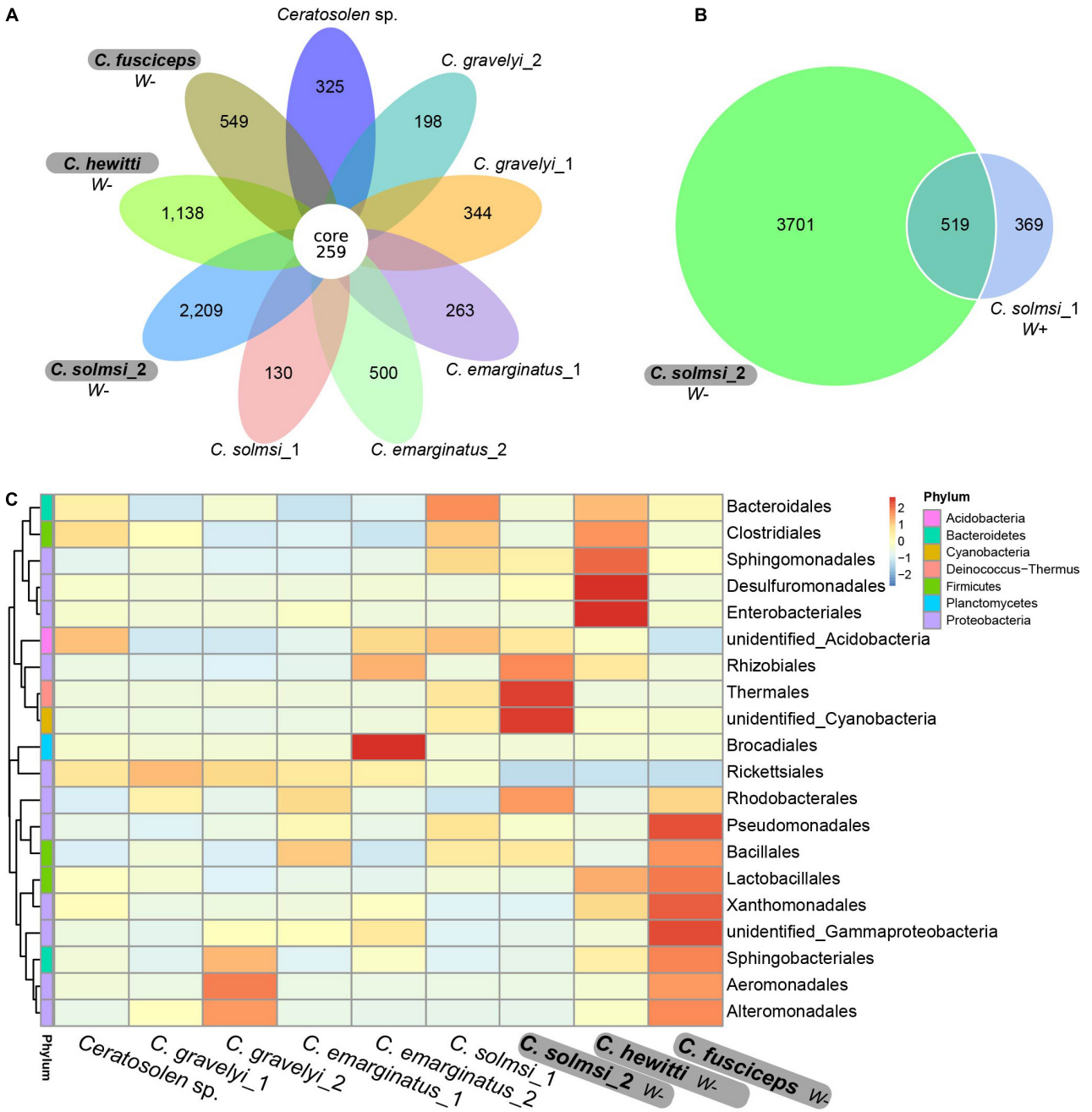

FIGURE 2 | Venn diagram and heat map. **(A)** Venn diagram based on all nine population. **(B)** Venn diagram between *Wolbachia* infected and uninfected group of *C. solmsi*. **(C)** Heat map of the top 20 classes in bacterial abundance. Three populations in bold are free of *Wolbachia* (*W*–), while the other six populations are infected with *Wolbachia* (*W*+).

had higher abundance and totally accounted for 99.7% of OTUs classified as *Wolbachia* in *W*+ group. In *C. gravelyi*_1, OTU1 and OTU8 accounted for 86.8 and 13.0% of OTUs classified as *Wolbachia*, respectively. Similarly, OTU1 and OTU8 accounted for 83.3 and 16.3% of OTUs classified as *Wolbachia* in *C. gravelyi*_2. As for the other four *W*+ populations, OTU1 accounted for more than 98.9% of the OTUs classified as *Wolbachia*.

Similarity of Bacterial Communities Within and Between Fig Pollinator Species

We calculated and compared alpha diversity indices of the fig wasp microbiota composition at the OTU level, and the variations of microbiota between and within various populations were shown in **Table 1**. Among all the

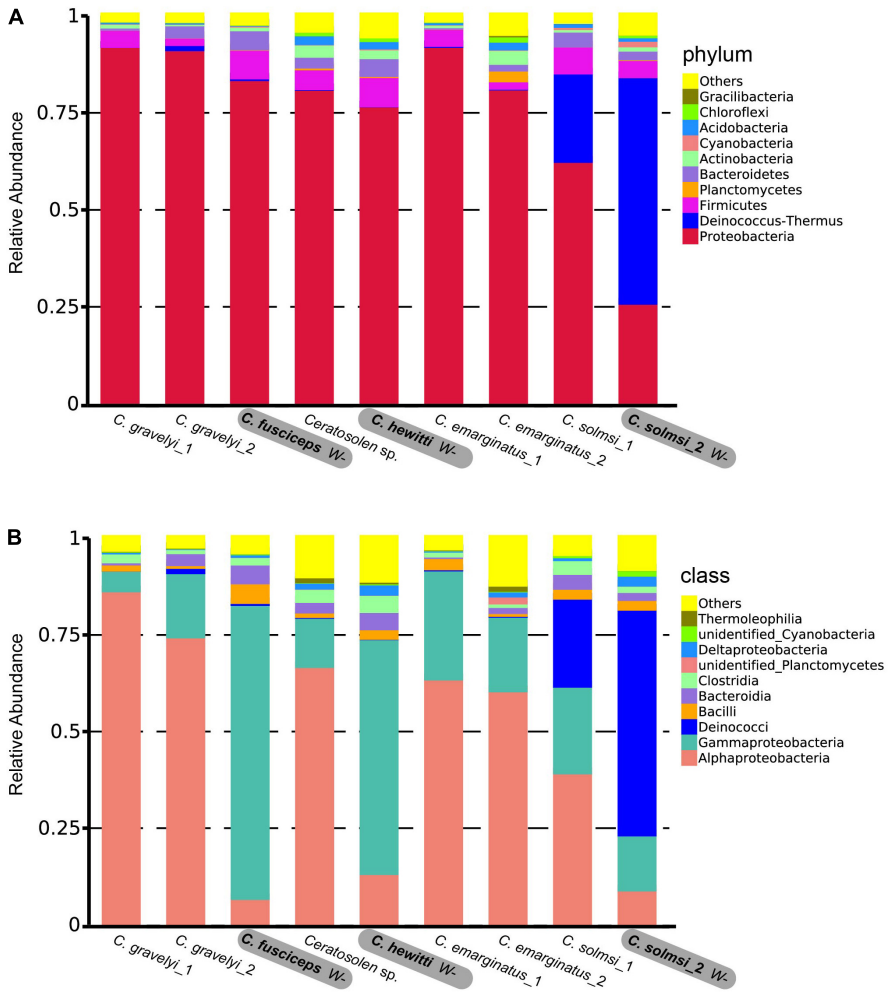

FIGURE 3 | Microbial community composition. Relative abundance of bacteria at the level of phylum **(A)** and class **(B)** in nine populations of fig wasp. Three populations in bold are free of *Wolbachia* (*W*–), while the other six populations are infected with *Wolbachia* (*W*+).

TABLE 1 | Alpha diversity indices and percentages of shared OTUs among populations.

	Richness estimates			Diversity estimates			Core microbiota
Population	**Observed species**	**Chao 1**	**ACE**	**Shannon**	**Simpson**	**PD whole tree**	**Shared OTUs (%)**
C. hewitti	2613 ± 975^{a}	2994 ± 1032^{a}	3236 ± 1061^{a}	5.44 ± 2.64^{a}	0.742 ± 0.292^{ab}	230.8 ± 70.5^{a}	18.54
Ceratosolen sp.	2259 ± 748^{a}	2681 ± 834^{a}	2940 ± 891^{a}	3.94 ± 1.34^{a}	0.567 ± 0.160^{bc}	187.1 ± 36.9^{abc}	44.35
C. fusciceps	1239 ± 325^{a}	1420 ± 349^{a}	1536 ± 364^{a}	4.38 ± 1.25^{a}	0.777 ± 0.119^{a}	195.2 ± 57.4^{abc}	32.05
C. emarginatus_2	1901 ± 769^{ab}	2233 ± 836^{a}	2454 ± 891^{a}	3.91 ± 1.10^{a}	0.640 ± 0.084^{abc}	198.4 ± 35.8^{ab}	34.12
C. emarginatus_1	1017 ± 344^{ab}	1290 ± 464^{ab}	1473 ± 567^{ab}	2.73 ± 0.61^{a}	0.564 ± 0.146^{bc}	142.6 ± 53.0^{cd}	49.62
C. gravelyi_1	1033 ± 427^{ab}	1211 ± 468^{ab}	1340 ± 490^{ab}	2.41 ± 1.10^{a}	0.456 ± 0.137^{c}	95.8 ± 24.8^{d}	42.95
C. gravelyi_2	728 ± 486^{ab}	858 ± 556^{ab}	949 ± 603^{ab}	2.66 ± 0.82^{a}	0.577 ± 0.109^{bc}	92.3 ± 27.6^{d}	56.67
C. solmsi_2	1280 ± 1115^{ab}	1412 ± 1233^{ab}	1503 ± 1314^{ab}	4.30 ± 2.66^{a}	0.722 ± 0.215^{ab}	160.1 ± 69.1^{bc}	10.49
C. solmsi_1	456 ± 47^{b}	497 ± 50^{b}	521 ± 51^{b}	3.93 ± 0.42^{a}	0.809 ± 0.091^{a}	101.2 ± 28.7^{d}	66.58

The values of alpha diversity indices are expressed as mean ± SD. One-way ANOVA was used to compare the differences of alpha diversity indices among populations. Significant differences are indicated by different letters ($P < 0.05$) in the same column.

populations, *C. hewitti* had the highest bacterial richness indices, while *C. solmsi*_1 had the lowest. *C. solmsi*_2 had the greatest variability.

The NMDS analysis was performed on the bacterial compositions of nine populations of fig wasps. The global differences in microbial community compositions were clearly

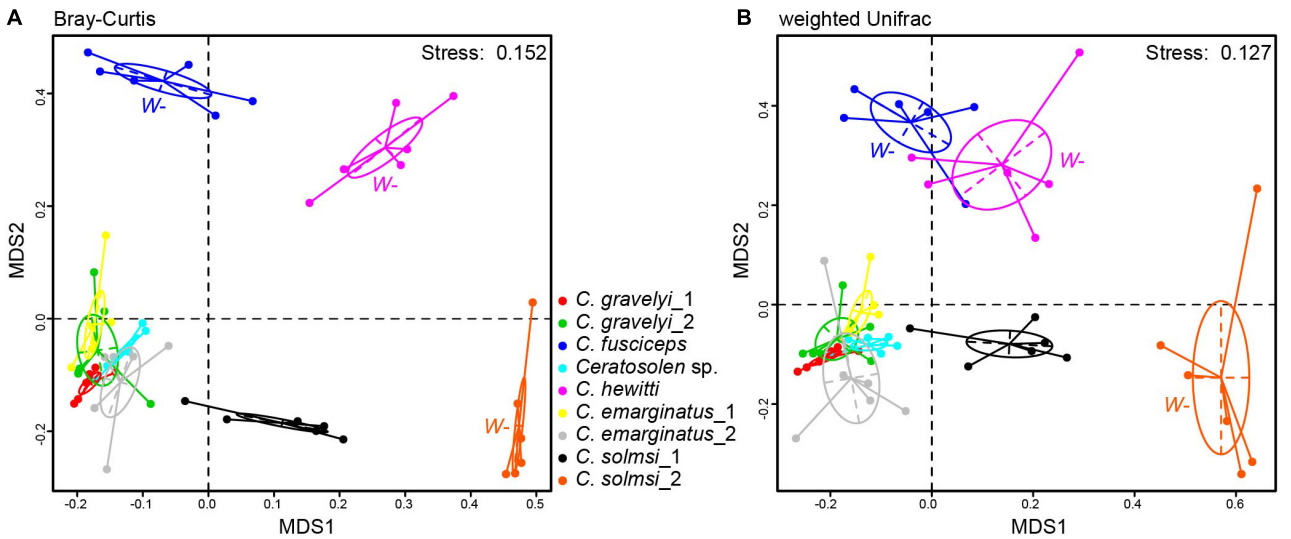

FIGURE 4 | Microbial community structure. NMDS ordinations of fig wasp microbiota based on Bray–Curtis (stress = 0.152) **(A)** and weighted Unifrac (stress = 0.127) dissimilarity matrices **(B)** on OTU level. Each dot represents a sample. Fig wasps from the same populations are colored uniformly. Three populations that are not infected with *Wolbachia* are labeled *W*–, while the other six populations are infected. Ellipses = 80% confidence intervals.

visualized in the NMDS plots based on Bray–Curtis and weighted UniFrac distance metrics (**Figure 4**). The individuals of the same population clustered together. ANOSIM revealed that each population harbored a unique bacterial community composition ($P < 0.05$, $R > 0$) (**Supplementary Table 3**). Consistently, MRPP analysis confirmed the significant differences in bacterial community composition among various populations ($P < 0.05$, $A > 0$) (**Supplementary Table 3**).

In addition, the distribution of *Wolbachia*-infected samples was relatively concentrated in NMDS analysis, which should be caused by the high abundance of *Wolbachia* (**Figure 4**). The populations from *W*+ group were concentrated, and close to each other (**Figure 4**). The populations from *W*− group were scattered, and the bacterial community structures were quite different from each other. On the whole, the *W*− group was deviated from the *W*+ group.

Effects of *Wolbachia* on the Microbial Communities

Spearman's rank correlation analyses showed that the proportion of *Wolbachia* was significantly negatively correlated with the Shannon ($r = -0.72$, $P < 0.001$) and Simpson ($r = -0.94$, $P < 0.001$) indices across all *W*+ samples (**Figures 5A,B**). For samples not infected with *Wolbachia* (*W*− group, the relative abundance of *Wolbachia* was close to zero), the effect of *Wolbachia* on the diversity of microbial communities was not counted. When the analysis was carried out by population, the populations that were not infected with *Wolbachia* tended to have a more diverse microbial community compared with the populations that were infected, and with the presence of *Wolbachia* and the increase in relative abundance, a significant negative correlation between the relative abundance of *Wolbachia* and diversity index was also observed at the population level (**Figures 5C,D**). We also examined whether there were significant differences in the diversity index of the microbiota between the *Wolbachia*-infected and uninfected samples. The Shannon and Simpson indices of group *W*+ and *W*−, *C. solmsi*_1, and *C. solmsi*_2 conformed to the normal distribution; the Simpson indices of group *W*+ and *W*− ($P = 0.1858$) and *C. solmsi*_1 and *C. solmsi*_2 ($P = 0.0849$) had equal variances, while the Shannon indices of group *W*+ and *W*− ($P = 0.0007$) and *C. solmsi*_1 and *C. solmsi*_2 ($P = 0.0011$) did not have equal variances. According to the *t*-test, there were significant differences between *W*+ and *W*− groups in Shannon ($P = 0.016$) and Simpson ($P = 0.006$) indices. However, there was no significant difference between *C. solmsi*_1 and *C. solmsi*_2 in Shannon ($P = 0.75$) and Simpson ($P = 0.38$) indices.

Phylosymbiosis Between Fig Pollinators or Fig Trees and Microbiota

To examine whether the *Ceratosolen* fig wasp phylogeny and their bacterial communities followed patterns of phylosymbiosis, we used both matrix comparisons and topological comparisons to quantify the signal of phylosymbiosis. Based on the genetic matrix of insect hosts and the beta diversity distance matrices, Mantel tests showed significant patterns of phylosymbiosis as measured by Bray–Curtis ($r = 0.61$, $P = 0.008$), weighted Unifrac ($r = 0.73$, $P = 0.003$), and unweighted Unifrac ($r = 0.55$, $P = 0.005$) distance metrics, while no significant pattern as measured by binary Jaccard ($r = 0.42$, $P = 0.066$) distance metric. Both the Robinson–Foulds and matching cluster metrics were used to detect phylosymbiosis between the microbiota dendrogram and their hosts' phylogenetic tree. We observed significant phylosymbiosis signals with Bray–Curtis distances, although the microbiota dendrogram and their hosts' phylogenetic tree were not completely

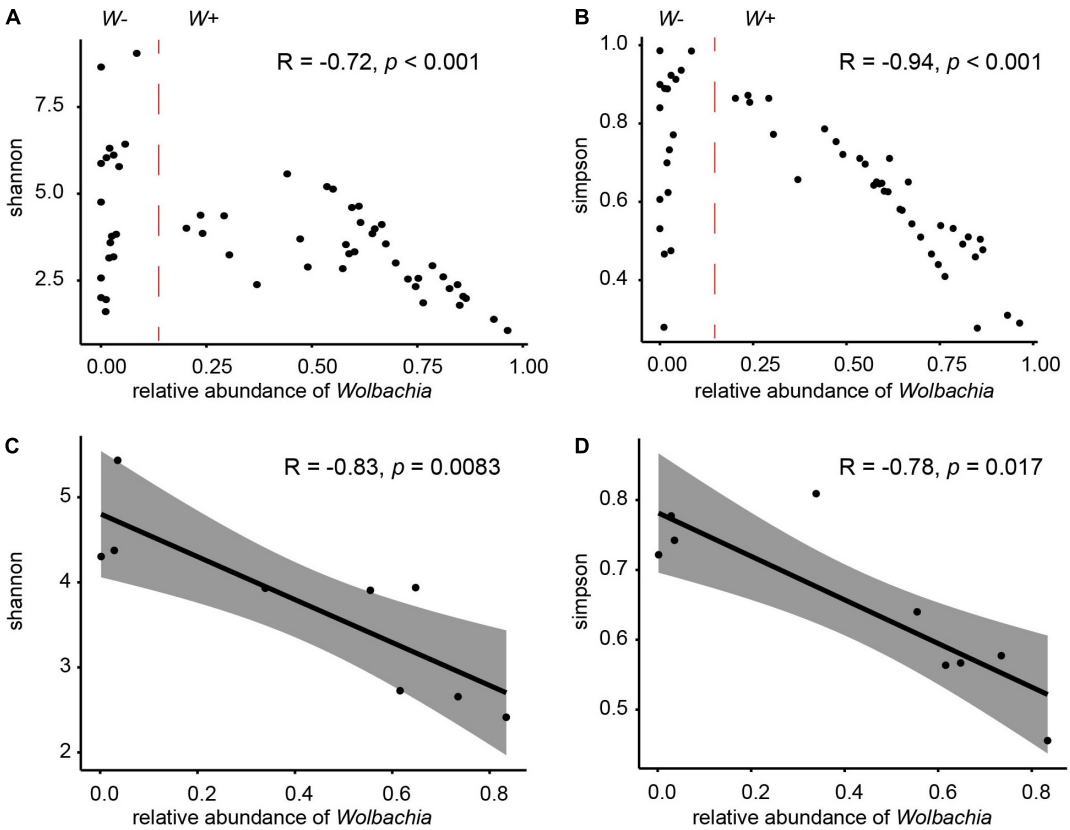

FIGURE 5 | Effects of *Wolbachia* on the microbial communities. Relationships between the proportions of *Wolbachia* and the Shannon indices **(A)** and Simpson indices **(B)** of microbial community in samples from group *W*+. Relationships between the proportions of *Wolbachia* and the Shannon indices **(C)** or Simpson indices **(D)** of microbial community in nine populations. *R* values and *P*-values of Spearman's rank correlation are provided.

congruent. More specifically, we found statistically significant congruence between fig wasp phylogeny and microbiota dendrograms using normalized Robinson–Foulds (nRF = 0.5, P = 0.006) and the matching cluster method (nMC = 0.23, P = 0.0002) (**Figure 6A**).

In order to determine whether the phylogeny of fig tree has an effect on the microbiota of fig pollinators, we examined whether there were phylosymbiotic signals between the fig trees and microbiota of fig pollinators based on both matrix comparisons and topological comparisons. Mantel tests showed no significant patterns of phylosymbiosis as measured by Bray–Curtis (r = 0.22, P = 0.15), weighted Unifrac (r = 0.19, P = 0.18), unweighted Unifrac (r = 0.14, P = 0.32), and binary Jaccard (r = 0.19, P = 0.26) distance metrics. Topological comparisons also showed that there were no significant phylosymbiotic signals between the fig pollinators' microbiota and the fig trees (nRF = 0.67, P = 0.30; nMC = 0.58, P = 0.31), although four of the six species shared the same topological location (**Figure 6B**).

DISCUSSION

In this study, we use the fig pollinators of the genus *Ceratosolen* to explore the effects of the phylogenetic relationship of fig wasp hosts, the endosymbiont *Wolbachia*, and the unique symbiotic environment provided by fig fruit on the host bacterial communities. Our results show significant phylosymbiosis signals between fig pollinators and their microbiota, regardless of matrix comparisons or topological comparisons. In the analysis of phylosymbiosis, we detect statistically significant congruence between fig wasp phylogeny and microbiota dendrogram, either using the Robinson–Foulds (RF) metric or the modified matching cluster (MC) method which better accounts for sections of subtree congruence (Robinson and Foulds, 1981; Bogdanowicz and Giaro, 2013). These results indicate that the microbial communities of fig wasps in genus *Ceratosolen* are not formed randomly, but have a certain correlation with the host phylogeny.

Phylosymbiosis means the consistent relationship between hosts' phylogeny and their microbial community structures. Due to that the fig pollinators spend most of their lives in the enclosed environment provided by fig fruits, and their food source is provided steadily by the ovaries of female flowers, they are much less likely to be disturbed by external environment than those living freely in open environment. Under this circumstance, the closely related fig pollinators species within the same genus are speculated to have similar microbial community structure. However, our

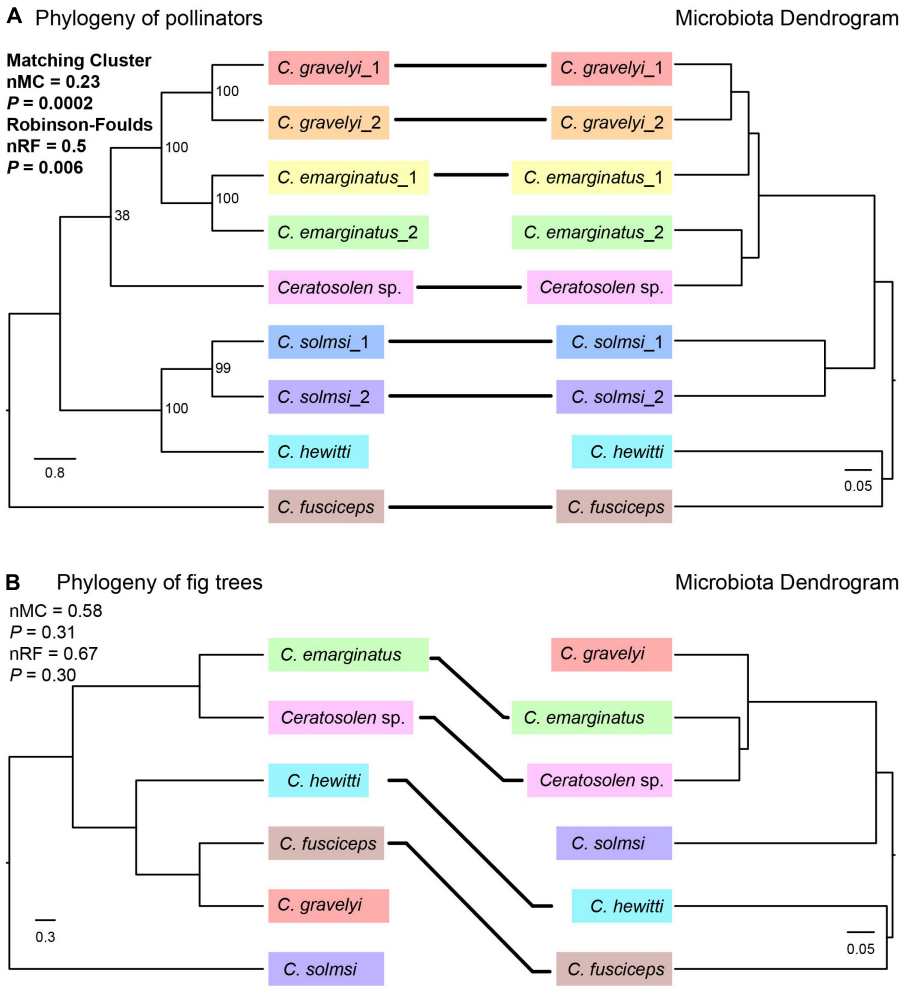

FIGURE 6 | Phylosymbiosis test between the phylogeny of pollinators **(A)** or fig trees **(B)** and microbiota dendrogram. The phylogeny of pollinators was based on CO1 and ITS2 sequences by ML method. The phylogeny of fig trees was based on ITS, ETS, and G3pdh genes by Bayesian method. For the convenience of comparison, the names of pollinators were used instead of the names of fig trees in the phylogenetic tree. Microbiota composition dendrogram was generated by the Bray–Curtis matrix. The nRF and nMC scores range from 0.0 (complete congruence) to 1.0 (complete incongruence). Solid lines connect species whose position is concordant between the phylogeny of pollinators or fig trees and microbiota dendrogram.

study shows that the microbial community structures are distinguishable among the *Ceratosolen* species. The differences in microbial compositions can reflect the phylogeny of the hosts, suggesting that host phylogeny plays a non-negligible role in shaping host-associated microbial communities. From another perspective, the microbial community may play a promoting/delaying role in host speciation. Hosts are better adapted to their native microbiota than to foreign ones (Brooks et al., 2016; Parker et al., 2018). For instance, transplanting interspecific microbial communities in *Peromyscus* deer mice significantly decreased their ability to digest food, and *Nasonia* wasps that received transplants of microbial communities from different wasp species had lower survival rate than those given their own microbiota (Brooks et al., 2016). In *Nasonia* parasitoid wasps, the microbiome shifts in hybrids, as a rare *Proteus* bacteria become dominant, so the larval hybrids then catastrophically succumb to bacterium-assisted lethality and reproductive isolation between the species (Cross et al., 2021). All these data showed that host characteristics could regulate microbial communities, and changes in the microbial communities were involved in host hybridization lethality, which reflected a mutually restrictive coevolutionary relationship between the hosts and associated microbiota. Therefore, the phylosymbiosis signals found in the *Ceratosolen* fig pollinators, on the one hand, indicate that different species of fig wasps have formed their own stable microbial communities during the process of systematic differentiation, and on the other hand, they also imply that the microbial community may play an important role in the speciation of closely related fig pollinators.

Although these fig pollinators we used have a close relationship within the same genus, and the symbiotic environment and food sources provided by various fig fruits are similar, the particular symbiosis environment provided

by fig fruits may have played an important role in shaping the different microbial communities of these closely related pollinators. The fig pollinators used in this study are associated to six different fig tree species. The fruits seem similar, but we cannot completely conclude that the fruits of different fig species provide the same source of nutrients for the fig pollinators living within them. Therefore, we cannot rule out the influence of diets on the bacterial community of these fig pollinators. There are few reports on bacterial communities associated with plant reproductive organs. Akduman et al. (2018) isolated and cultured bacteria of figs from *Ficus mauritiana* (La Re'union), *Ficus racemosa* (Vietnam), and *Ficus sycomorus* (South Africa), and isolated strains belonging to four bacterial phyla, Proteobacteria, Firmicutes, Bacteriodetes, and Actinobacteria. In our study, the bacterial communities of fig pollinators are mainly distributed in Proteobacteria, Firmicutes, and Bacteriodetes (**Figure 3A**), which is similar to the gut microbiota of other insects, such as *Nasonia* wasps, mosquitoes, and *Drosophila* flies (Brooks et al., 2016). Up to now, there are no studies that clearly explain the consistent relationship between figs and fig pollinators. However, there is a report that fungal communities in syconia and on pollinating wasps were similar (Martinson et al., 2012). Another study has shown that host phylogeny shapes the foliar endophytic fungal assemblages of *Ficus* (Liu et al., 2019). Fig pollinators only feed on fig fruits (Janzen, 1979) and carry the microbiota of their environment which is dictated by the evolution of the fig trees, but the fig trees do not show a phylosymbiotic pattern with the microbiota inside the pollinators (**Figure 6B**), thus do not fully dictate the evolution of the association wasp-microbiota. As have been shown in herbivorous rodents, although diet and geography had an impact on the structure of the natural microbiota, the effects of host phylogeny were stronger for both wild and captive animals (Weinstein et al., 2021), indicating that host genetic background is the most significant predictor of microbiota composition and stability than geography and diet in woodrats. At the same time, the enclosed living environment provided by fruits also naturally provides an opportunity for ecological space isolation for different fig pollinators to a large extent. In addition, sampling time or locations of different populations are diverse, which also provides an isolated environment in time or space. Therefore, the differences in nutritional conditions between populations and the temporal and spatial isolation may have played an important role in shaping the microbial community structures of these fig pollinators. The study of phylosymbiosis between fig pollinators and their associated microbiota would be more convincing when wasps are collected at the same site to avoid covariation, such as environmental (temperature, humidity, and so on), and under more strictly controlled conditions such as sex or nutritional variables that may introduce distortions in the analysis of the microbial community. However, the distinguishability of microbial communities caused by these factors should be a random separation pattern between populations, and the successful detection of phylosymbiosis signals indicates that the microbial community structures of fig pollinators may be mainly associated to the host phylogeny.

Our study provides the first thorough characterization of microbiota in fig wasps. Six of nine fig pollinator populations were infected with the intracellular symbiotic *Wolbachia*. *Wolbachia* as the dominant genera identified accounted for 61.96% of the whole bacterial community in the infected fig wasp populations, which was concordant with that *Wolbachia* was the dominant genus in Hymenoptera (Yun et al., 2014). Some *Wolbachia*-infected beetles and terrestrial isopods also showed high abundance of *Wolbachia* (Dittmer et al., 2014, 2016; Parker et al., 2020). The high abundance of *Wolbachia* in multiple hosts reminds us that endosymbionts are considerable factors shaping microbial diversity in the hosts. Despite *Wolbachia* have a widespread distribution in insects, little is known about how *Wolbachia* interacts directly with other bacteria within hosts. Our results suggest that *Wolbachia* can have a negative effect on the bacterial diversity of the fig pollinators. We find that the populations of *W+* group have a relatively high proportion of core OTUs, which is consistent with that the relative abundance of *Wolbachia* is significantly negatively correlated with the diversity indices across all populations. These results indicate that infection with *Wolbachia* can lead to a decrease in the bacterial diversity of fig wasps. In natural *Drosophila* populations, it was also showed that the microbiota composition varied significantly with the relative abundance of *Wolbachia* (Fromont et al., 2019). Other similar results have been found in the small brown planthopper that *Wolbachia* infection, for instance, appears to play a greater role in shaping the microbial community structure than abiotic factors, resulting in a sharp decline in the diversity and abundance of host bacterial taxa (Duan et al., 2020). These studies suggest that *Wolbachia* does have an impact on the bacterial diversity of insect hosts and the underlying mechanism deserves investigation.

In this study, *C. gravelyi*, *C. emarginatus*, and *C. solmsi* are discovered to have cryptic species based on distinct mitochondrial derived DNA barcode CO1, nuclear DNA marker ITS2 sequences, and indistinguishable morphological characteristics. We here thus explore the underlying reasons from the perspective of collection locations and microbiota community. First, *C. emarginatus*_1 and *C. emarginatus*_2 were collected from Guangxi and Hainan provinces, respectively, and both collection sites are about 400 km apart and separated by the Qiongzhou Strait. *C. solmsi*_1 and *C. solmsi*_2 were collected from Guangdong and Hainan provinces, respectively, and both collection sites are about 500 km apart and separated by the Qiongzhou Strait as well. Therefore, geographical isolation may be related to the differentiation between both pairs of cryptic species. Second, in the composition of microbiota, *C. solmsi*_1 is infected with *Wolbachia*, while *C. solmsi*_2 is not, and NMDS analysis can distinguish the microbial communities of *C. solmsi*_1 and *C. solmsi*_2 well. As intracellular symbiotic bacteria that widely infect insect species and can perform reproductive manipulation on the hosts (Hertig and Wolbach, 1924; Hilgenboecker et al., 2008; Werren et al., 2008), *Wolbachia* has been predicted to be the internal driving force for reproductive isolation (Bordenstein et al., 2001), which may also be the putative

reason for the differentiation of the cryptic species of the *C. solmsi*. Besides, in the species of *C. gravelyi*, although both populations of *C. gravelyi*_1 and *C. gravelyi*_2 were collected from the Xishuangbanna Tropical Botanical Garden (Yunnan Province, China), *C. gravelyi*_1 has more unique microbial strains in terms of the proportion of the core OTUs of the microbial community. As for bacterial abundance, *C. gravelyi*_1 and *C. gravelyi*_2 have significant difference (*t*-test, $P < 0.05$, the same below) only in the phylum of Bacteroidetes, while between *C. emarginatus*_1 and *C. emarginatus*_2, and between *C. solmsi*_1 and *C. solmsi*_2, there are significant differences in Proteobacteria and some other bacterial phyla. Among the core bacteria (**Supplementary Table 4**), we find significant differences in the abundance of the order of Pseudomonadales and the genus of *Acinetobacter* between each of the three pairs of cryptic species. Many bacteria belonging to *Acinetobacter* are nosocomial pathogens (Bergogne-Bérézin and Towner, 1996), but their physiological roles in insects are still unclear. Changes in the abundance of core microbiota may affect host fitness and thus contribute to their evolutionary divergence (Rosenberg et al., 2010). Therefore, besides the geographical isolation, the intracellular symbiotic *Wolbachia* or other microbial components may also play a role in the divergence of the cryptic species of the three fig pollinator species.

Interestingly, our high-throughput 16S rDNA microbial community profiles reveal that these fig pollinators may have been infected with a variety of intracellular symbiotic bacteria in their evolutionary history. In our survey of *Wolbachia* infection in the fig pollinators, the individual was considered to be uninfected if all three pairs of *Wolbachia*-specific primers failed to amplify *Wolbachia*-specific gene. In the final results, the OTUs classified as *Wolbachia* are also observed as a rare microbial community member in the *W*− populations, including *C. fusciceps* (2.9%), *C. hewitti* (3.6%), and *C. solmsi*_2 (0.2%). While in the six *W*+ populations, *Wolbachia* is detected at high abundance with relative abundance ranging from 20.3 to 96.4%. Due to that few OTUs belonging to *Wolbachia* were found in *W*− group, we speculate that these samples may contain footprints of historic *Wolbachia* infections that mostly have been lost, or these samples acquired a small amount of *Wolbachia* through horizontal transmission recently. In addition to *Wolbachia*, reads of some other intracellular symbiotic bacteria, such as *Buchnera*, *Spiroplasma*, *Arsenophonus*, *Blattabacterium*, *Rickettsiella*, *Serratia*, and *Candidatus Fritschea*, have also been detected in our samples, even though the relative sequence abundances of these bacteria are all lower than 0.01%. Excluding the possibility of contamination, these findings can provide evidences for the existence of multiple intracellular symbionts in the fig pollinators. Another very surprising finding is the phylum of Deinococcus-thermus, which has an abundance as high as 22.9 and 58.6% in *C. solmsi*_1 and *C. solmsi*_2, respectively, but the abundance in other populations is extremely low. Deinococcus-Thermus is considered as extremophiles bacteria (Griffiths and Gupta, 2007) and can be found in the gut of insects (Yun et al., 2014). It needs more work to explore why such a high proportion of Deinococcus-Thermus is present in the *C. solmsi* samples.

CONCLUSION

In this study, a phylosymbiosis signal was found between fig wasps of genus *Ceratosolen* and their microbial community, which help us to view the role of host phylogeny in shaping microbial community structure from a fundamental perspective. The presence of *Wolbachia* led to a decrease in host microbial diversity, and *Wolbachia* seemed to be the dominant species in the infected hosts. Our studies demonstrated that host phylogeny and the presence of endosymbiotic *Wolbachia* were the driving factors of bacterial community structure in *Ceratosolen* fig wasps. Collectively, the tripartite interactions of host, symbionts, and microbiota shape the dynamic stability of the holobiont ecosystem in fig wasps.

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://figshare.com/, doi: 10.6084/m9.figshare.16853656, doi: 10.6084/m9.figshare.17078801, and doi: 10.6084/m9.figshare.17075978.

AUTHOR CONTRIBUTIONS

JX and DH designed the project. JL and XW collected the fig wasp samples and completed data analyses. JL completed laboratory work. JL, JX, and XW wrote the manuscript. DH revised the manuscript. All authors have approved the final manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (Nos. 31970440, 31830084, and 32070466) and also supported by "the Fundamental Research Funds for the Central Universities", Nankai University (Nos. 96172158, 96173250, and 91822294).

ACKNOWLEDGMENTS

We are grateful to Wenquan Zhen, Yanzhou Zhang, and Gui Feng for collecting fig wasp samples. JL thanks Peiming Yuan especially for his moral support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.800190/full#supplementary-material

REFERENCES

Akduman, N., Rodelsperger, C., and Sommer, R. J. (2018). Culture-based analysis of *Pristionchus*-associated microbiota from beetles and figs for studying nematode-bacterial interactions. *PLoS One* 13:e0198018. doi: 10.1371/journal.pone.0198018

Berg, C. C. (1989). Classification and distribution of *Ficus*. *Experientia* 45, 605–611. doi: 10.1007/BF01975677

Bergogne-Bérézin, E., and Towner, K. J. (1996). *Acinetobacter* spp. as nosocomial pathogens: microbiological, clinical, and epidemiological features. *Clin. Microbiol. Rev.* 9, 148–165. doi: 10.1128/CMR.9.2.148

Bogdanowicz, D., and Giaro, K. (2013). On a matching distance between rooted phylogenetic trees. *Int. J. Appl. Math. Comput. Sci.* 23, 669–684. doi: 10.2478/amcs-2013-0050

Bogdanowicz, D., Giaro, K., and Wróbel, B. (2012). TreeCmp: comparison of trees in polynomial time. *Evol. Bioinform.* 8, 475–487. doi: 10.4137/ebo.s9657

Bokulich, N. A., Subramanian, S., Faith, J. J., Gevers, D., Gordon, J. I., Knight, R., et al. (2013). Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nat. Methods* 10, 57–59. doi: 10.1038/nmeth.2276

Bordenstein, S., O'hara, F., and Werren, J. (2001). Wolbachia-induced incompatibility precedes other hybrid incompatibilities in *Nasonia*. *Nature* 409, 707–710. doi: 10.1038/35055543

Brinker, P., Fontaine, M. C., Beukeboom, L. W., and Falcao Salles, J. (2019). Host, symbionts, and the microbiome: the missing tripartite interaction. *Trends Microbiol.* 27, 480–488. doi: 10.1016/j.tim.2019.02.002

Brooks, A. W., Kohl, K. D., Brucker, R. M., Van Opstal, E. J., and Bordenstein, S. R. (2016). Phylosymbiosis: relationships and functional effects of microbial communities across host evolutionary history. *PLoS Biol.* 14:e2000225. doi: 10.1371/journal.pbio.2000225

Brucker, R. M., and Bordenstein, S. R. (2012a). The roles of host evolutionary relationships (Genus: *Nasonia*) and development in structuring microbial communities. *Evolution* 66, 349–362. doi: 10.1111/j.1558-5646.2011.01454.x

Brucker, R. M., and Bordenstein, S. R. (2012b). Speciation by symbiosis. *Trends Ecol. Evol.* 27, 443–451. doi: 10.1016/j.tree.2012.03.011

Brucker, R. M., and Bordenstein, S. R. (2013). The hologenomic basis of speciation: gut bacteria cause hybrid lethality in the genus *Nasonia*. *Science* 341, 667–669. doi: 10.1126/science.1240659

Brust, R. A., Ballard, J. W. O., Driver, F., Hartley, D. M., Galway, N. J., and Curran, J. (1998). Molecular systematics, morphological analysis, and hybrid crossing identify a third taxon, *Aedes (Halaedes) wardangensis* sp.nov., of the *Aedes (Halaedes) australis* species-group (Diptera: Culicidae). *Rev. Can. Zool.* 76, 1236–1246. doi: 10.1139/cjz-76-7-1236

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

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

Chen, L., Cook, J. M., Xiao, H., Hu, H., Niu, L., and Huang, D. (2010). High incidences and similar patterns of *Wolbachia* infection in fig wasp communities from three different continents. *Insect Sci.* 17, 101–111. doi: 10.1111/j.1744-7917.2009.01291.x

Chiarello, M., Auguet, J. C., Bettarel, Y., Bouvier, C., Claverie, T., Graham, N. A. J., et al. (2018). Skin microbiome of coral reef fish is highly variable and driven by host phylogeny and diet. *Microbiome* 6:147. doi: 10.1186/s40168-018-0530-4

Chong, R. A., Park, H., and Moran, N. A. (2019). Genome evolution of the obligate endosymbiont *Buchnera aphidicola*. *Mol. Biol. Evol.* 36, 1481–1489. doi: 10.1093/molbev/msz082

Cisak, E., Wojcik-Fatla, A., Zajac, V., Sawczyn, A., Sroka, J., and Dutkiewicz, J. (2015). *Spiroplasma* - an emerging arthropod-borne pathogen? *Ann. Agric. Environ. Med.* 22, 589–593. doi: 10.5604/12321966.1185758

Corner, E. G. H. (1958). An introduction to the distribution of *Ficus*. *Reinwardtia* 4, 325–355.

Cross, K. L., Leigh, B. A., Hatmaker, E. A., Mikaelyan, A., Miller, A. K., and Bordenstein, S. R. (2021). Genomes of gut bacteria from *Nasonia* wasps shed light on phylosymbiosis and microbe-assisted hybrid breakdown. *mSystems* 6:e01342-20. doi: 10.1128/mSystems.01342-20

Cruaud, A., Jabbour-Zahab, R., Genson, G., Cruaud, C., Couloux, A., Kjellberg, F., et al. (2010). Laying the foundations for a new classification of Agaonidae (Hymenoptera: Chalcidoidea), a multilocus phylogenetic approach. *Cladistics* 26, 359–387. doi: 10.1111/j.1096-0031.2009.00291.x

Davis, R. E., Worley, J. F., Whitcomb, R. F., Ishijima, T., and Steere, R. L. (1972). Helical filaments produced by a mycoplasma-like organism associated with corn stunt disease. *Science* 176, 521–523. doi: 10.1126/science.176.4034.521

DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., et al. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* 72, 5069–5072. doi: 10.1128/AEM.03006-05

Dittmer, J., Beltran-Bech, S., Lesobre, J., Raimond, M., Johnson, M., and Bouchon, D. (2014). Host tissues as microhabitats for *Wolbachia* and quantitative insights into the bacterial community in terrestrial isopods. *Mol. Ecol.* 23, 2619–2635. doi: 10.1111/mec.12760

Dittmer, J., Lesobre, J., Moumen, B., and Bouchon, D. (2016). Host origin and tissue microhabitat shaping the microbiota of the terrestrial isopod *Armadillidium vulgare*. *FEMS Microbiol. Ecol.* 92:fiw063. doi: 10.1093/femsec/fiw063

Douglas, A. E. (2015). Multiorganismal insects: diversity and function of resident microorganisms. *Annu. Rev. Entomol.* 60, 17–34. doi: 10.1146/annurev-ento-010814-020822

Dray, S. E., and Dufour, A.-B. E. (2007). The ade4 package: implementing the duality diagram for ecologists. *J. Stat. Softw.* 22, 1–20.

Drummond, A. J., Suchard, M. A., Xie, D., and Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* 29, 1969–1973. doi: 10.1093/molbev/mss075

Duan, X., Sun, J., Wang, L., Shu, X., Guo, Y., Keiichiro, M., et al. (2020). Recent infection by *Wolbachia* alters microbial communities in wild *Laodelphax striatellus* populations. *Microbiome* 8:104. doi: 10.1186/s40168-020-00878-x

Dunaj, S. J., Bettencourt, B. R., Garb, J. E., and Brucker, R. M. (2020). Spider phylosymbiosis: divergence of widow spider species and their tissues' microbiomes. *BMC Evol. Biol.* 20:104. doi: 10.1186/s12862-020-01664-x

Edgar, R. C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* 10, 996–998. doi: 10.1038/nmeth.2604

Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27, 2194–2200. doi: 10.1093/bioinformatics/btr381

Ellison, S., Rovito, S., Parra-Olea, G., Vasquez-Almazan, C., Flechas, S. V., Bi, K., et al. (2019). The influence of habitat and phylogeny on the skin microbiome of amphibians in Guatemala and Mexico. *Microb. Ecol.* 78, 257–267. doi: 10.1007/s00248-018-1288-8

Ezard, T., Fujisawa, T., and Barraclough, T. (2015). *R Package Splits: Species' Limits by Threshold Statistics. Version 1.0-18/r45*. Available online at: http://R-Forge.R-project.org/projects/splits/ (accessed May 7, 2021).

Folmer, O., Black, M., Hoeh, W., Lutz, R., and Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotechnol.* 3, 294–299.

Fox, J., and Weisberg, S. (2019). *An R Companion to Applied Regression*. Thousand Oaks, CA: Sage.

Fromont, C., Adair, K. L., and Douglas, A. E. (2019). Correlation and causation between the microbiome, *Wolbachia* and host functional traits in natural populations of drosophilid flies. *Mol. Ecol.* 28, 1826–1841. doi: 10.1111/mec.15041

Goto, S., Anbutsu, H., and Fukatsu, T. (2006). Asymmetrical interactions between *Wolbachia* and *Spiroplasma* endosymbionts coexisting in the same insect host. *Appl. Environ. Microbiol.* 72, 4805–4810. doi: 10.1128/AEM.00416-06

Griffiths, E., and Gupta, R. S. (2007). Identification of signature proteins that are distinctive of the *Deinococcus-Thermus* phylum. *Int. Microbiol.* 10, 201–208. doi: 10.2436/20.1501.01.28

Groussin, M., Mazel, F., Sanders, J. G., Smillie, C. S., Lavergne, S., Thuiller, W., et al. (2017). Unraveling the processes shaping mammalian gut microbiomes over evolutionary time. *Nat. Commun.* 8:14319. doi: 10.1038/ncomms14319

Guegan, M., Zouache, K., Demichel, C., Minard, G., Tran Van, V., Potier, P., et al. (2018). The mosquito holobiont: fresh insight into mosquito-microbiota interactions. *Microbiome* 6:49. doi: 10.1186/s40168-018-0435-2

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood

phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321. doi: 10.1093/sysbio/syq010

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

Hertig, M., and Wolbach, S. B. (1924). Studies on Rickettsia-like micro-organisms in insects. *J. Med. Res.* 44, 329–374.

Hilgenboecker, K., Hammerstein, P., Schlattmann, P., Telschow, A., and Werren, J. H. (2008). How many species are infected with *Wolbachia*?-A statistical analysis of current data. *FEMS Microbiol. Lett.* 281, 215–220. doi: 10.1111/j.1574-6968.2008.01110.x

Hughes, G. L., Dodson, B. L., Johnson, R. M., Murdock, C. C., Tsujimoto, H., Suzuki, Y., et al. (2014). Native microbiome impedes vertical transmission of *Wolbachia* in *Anopheles* mosquitoes. *Proc. Natl. Acad. Sci. U.S.A.* 111, 12498–12503. doi: 10.1073/pnas.1408888111

Huot, C., Clerissi, C., Gourbal, B., Galinier, R., Duval, D., and Toulza, E. (2019). Schistosomiasis vector snails and their microbiota display a phylosymbiosis pattern. *Front. Microbiol.* 10:3092. doi: 10.3389/fmicb.2019.03092

Janzen, D. H. (1979). How to be a fig. *Annu. Rev. Ecol. Syst.* 10, 13–51. doi: 10.1146/annurev.es.10.110179.000305

Jeyaprakash, A., and Hoy, M. A. (2000). Long PCR improves *Wolbachia* DNA amplification: *wsp* sequences found in 76% of sixty-three arthropod species. *Insect Mol. Biol.* 9, 393–405. doi: 10.1046/j.1365-2583.2000.00203.x

Jones, M., Ghoorah, A., and Blaxter, M. (2011). jMOTU and Taxonerator: turning DNA Barcode sequences into annotated operational taxonomic units. *PLoS One* 6:e19259. doi: 10.1371/journal.pone.0019259

Kartzinel, T. R., Hsing, J. C., Musili, P. M., Brown, B. R. P., and Pringle, R. M. (2019). Covariation of diet and gut microbiome in African megafauna. *Proc. Natl. Acad. Sci. U.S.A.* 116, 23588–23593. doi: 10.1073/pnas.1905666116

Kassambara, A. (2020). ggpubr: 'ggplot2' Based Publication Ready Plots. Available online at: https://CRAN.R-project.org/package=ggpubr (accessed June 26, 2021).

Katoh, K., Rozewicki, J., and Yamada, K. D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform.* 20, 1160–1166. doi: 10.1093/bib/bbx108

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16, 111–120. doi: 10.1007/BF01731581

Kjer, K. M., Baldridge, G. D., and Fallon, A. M. (1994). Mosquito large subunit ribosomal RNA: simultaneous alignment of primary and secondary structure. *Biochim. Biophys. Acta* 1217, 147–155. doi: 10.1016/0167-4781(94)90028-0

Kueneman, J. G., Parfrey, L. W., Woodhams, D. C., Archer, H. M., Knight, R., and Mckenzie, V. J. (2014). The amphibian skin-associated microbiome across species, space and life history stages. *Mol. Ecol.* 23, 1238–1250. doi: 10.1111/mec.12510

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Lagier, J. C., Dubourg, G., Million, M., Cadoret, F., Bilen, M., Fenollar, F., et al. (2018). Culturing the human microbiota and culturomics. *Nat. Rev. Microbiol.* 16, 540–550. doi: 10.1038/s41579-018-0041-0

Lanfear, R., Calcott, B., Ho, S. Y., and Guindon, S. (2012). Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Mol. Biol. Evol.* 29, 1695–1701. doi: 10.1093/molbev/mss020

Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T., and Calcott, B. (2017). PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol. Biol. Evol.* 34, 772–773. doi: 10.1093/molbev/msw260

Letunic, I., and Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* 49, W293–W296. doi: 10.1093/nar/gkab301

Lim, S. J., and Bordenstein, S. R. (2020). An introduction to phylosymbiosis. *Proc. Biol. Sci.* 287: 20192900. doi: 10.1098/rspb.2019.2900

Liu, J., Zhao, J., Wang, G., and Chen, J. (2019). Host identity and phylogeny shape the foliar endophytic fungal assemblages of *Ficus*. *Ecol. Evol.* 9, 10472–10482. doi: 10.1002/ece3.5568

Lopez-Madrigal, S., Latorre, A., Moya, A., and Gil, R. (2015). The link between independent acquisition of intracellular gamma-endosymbionts and concerted evolution in *Tremblaya princeps*. *Front. Microbiol.* 6:642. doi: 10.3389/fmicb.2015.00642

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

Martinson, E. O., Herre, E. A., Machado, C. A., and Arnold, A. E. (2012). Culture-free survey reveals diverse and distinctive fungal communities associated with developing figs (*Ficus* spp.) in Panama. *Microb. Ecol.* 64, 1073–1084. doi: 10.1007/s00248-012-0079-x

Nguyen, L.-T., Schmidt, H. A., Von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300

O'Brien, P. A., Webster, N. S., Miller, D. J., and Bourne, D. G. (2019). Host-microbe coevolution: applying evidence from model systems to complex marine invertebrate holobionts. *mBio* 10:e02241-18. doi: 10.1128/mBio.02241-18

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., et al. (2020). *vegan*: *Community Ecology Package*. Available online at: https://CRAN.R-project.org/package=vegan (accessed October 16, 2020).

O'Neill, S. L., Giordano, R., Colbert, A. M., Karr, T. L., and Robertson, H. M. (1992). 16S rRNA phylogenetic analysis of the bacterial endosymbionts associated with cytoplasmic incompatibility in insects. *Proc. Natl. Acad. Sci. U.S.A.* 89, 2699–2702. doi: 10.1073/pnas.89.7.2699

Parker, E. S., Dury, G. J., and Moczek, A. P. (2018). Transgenerational developmental effects of species-specific, maternally transmitted microbiota in *Onthophagus* dung beetles. *Ecol. Entomol.* 44, 274–282. doi: 10.1111/een.12703

Parker, E. S., Newton, I. L. G., and Moczek, A. P. (2020). (My microbiome) would walk 10,000 miles: maintenance and turnover of microbial communities in introduced dung beetles. *Microb. Ecol.* 80, 435–446. doi: 10.1007/s00248-020-01514-9

Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., et al. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Syst. Biol.* 55, 595–609. doi: 10.1080/10635150600852011

Price, M. N., Dehal, P. S., and Arkin, A. P. (2009). FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* 26, 1641–1650. doi: 10.1093/molbev/msp077

Puillandre, N., Lambert, A., Brouillet, S., and Achaz, G. (2012). ABGD, automatic barcode gap discovery for primary species delimitation. *Mol. Ecol.* 21, 1864–1877. doi: 10.1111/j.1365-294X.2011.05239.x

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

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Robinson, D. F., and Foulds, L. R. (1981). Comparison of phylogenetic trees. *Math. Biosci.* 53, 131–147. doi: 10.1016/0025-5564(81)90043-2

Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D. L., Darling, A., Hohna, S., et al. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539–542. doi: 10.1093/sysbio/sys029

Rosenberg, E., Koren, O., Reshef, L., Efrony, R., and Zilber-Rosenberg, I. (2007). The role of microorganisms in coral health, disease and evolution. *Nat. Rev. Microbiol.* 5, 355–362. doi: 10.1038/nrmicro1635

Rosenberg, E., Sharon, G., Atad, I., and Zilber-Rosenberg, I. (2010). The evolution of animals and plants via symbiosis with microorganisms. *Environ. Microbiol. Rep.* 2, 500–506. doi: 10.1111/j.1758-2229.2010.00177.x

Salcedo-Porras, N., Umana-Diaz, C., Bitencourt, R. O. B., and Lowenberger, C. (2020). The role of bacterial symbionts in triatomines: an evolutionary perspective. *Microorganisms* 8:1438. doi: 10.3390/microorganisms8091438

Santos-Garcia, D., Mestre-Rincon, N., Ouvrard, D., Zchori-Fein, E., and Morin, S. (2020). *Portiera* gets wild: genome instability provides insights into the evolution of both whiteflies and their endosymbionts. *Genome Biol. Evol.* 12, 2107–2124. doi: 10.1093/gbe/evaa216

Tamura, K., and Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* 10, 512–526. doi: 10.1093/oxfordjournals.molbev.a040023

Thao, M. L., Moran, N. A., Abbot, P., Brennan, E. B., Burckhardt, D. H., and Baumann, P. (2000). Cospeciation of psyllids and their primary prokaryotic endosymbionts. *Appl. Environ. Microbiol.* 66, 2898–2905. doi: 10.1128/aem.66.7.2898-2905.2000

Tinker, K. A., and Ottesen, E. A. (2020). Phylosymbiosis across deeply diverging lineages in omnivorous cockroaches. *Appl. Environ. Microbiol.* 86:e02513-19. doi: 10.1128/aem.02513-19

Weinstein, S. B., Martinez-Mota, R., Stapleton, T. E., Klure, D. M., Greenhalgh, R., Orr, T. J., et al. (2021). Microbiome stability and structure is governed by host phylogeny over diet and geography in woodrats (*Neotoma* spp.). *Proc. Natl. Acad. Sci. U.S.A.* 118:e2108787118. doi: 10.1073/pnas.2108787118

Werren, J. H., Baldo, L., and Clark, M. E. (2008). *Wolbachia*: master manipulators of invertebrate biology. *Nat. Rev. Microbiol.* 6, 741–751. doi: 10.1038/nrmicro1969

Yang, K., Xie, K., Zhu, Y. X., Huo, S. M., Hoffmann, A., and Hong, X. Y. (2020). Wolbachia dominate *Spiroplasma* in the co-infected spider mite *Tetranychus truncatus*. *Insect. Mol. Biol.* 29, 19–37. doi: 10.1111/imb.12607

Yun, J.-H., Roh, S. W., Whon, T. W., Jung, M.-J., Kim, M.-S., Park, D.-S., et al. (2014). Insect gut bacterial diversity determined by environmental habitat, diet, developmental stage, and phylogeny of host. *Appl. Environ. Microbiol.* 80, 5254–5264. doi: 10.1128/AEM.01226-14

Zhang, D., Gao, F., Jakovlić, I., Zou, H., Zhang, J., Li, W. X., et al. (2020). PhyloSuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol. Ecol. Resour.* 20, 348–355. doi: 10.1111/1755-0998.13096

Zhang, J., Kapli, P., Pavlidis, P., and Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* 29, 2869–2876. doi: 10.1093/bioinformatics/btt499

Zhou, W., Rousset, F., and O'neil, S. (1998). Phylogeny and PCR-based classification of *Wolbachia* strains using *wsp* gene sequences. *Proc. Biol. Sci.* 265, 509–515. doi: 10.1098/rspb.1998.0324

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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 16 March 2022
doi: 10.3389/fgene.2022.758304

Experimental Evolution Reveals Redox State Modulates Mycobacterial Pathogenicity

Zheng Jiang[1,2†], *Zengfang Zhuang*[1,2†] *and Kaixia Mi*[1,2*]

[1]*CAS Key Laboratory of Pathogenic Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China,* [2]*Savaid Medical School, University of Chinese Academy of Sciences, Beijing, China*

Understanding how *Mycobacterium tuberculosis* has evolved into a professional pathogen is helpful in studying its pathogenesis and for designing vaccines. We investigated how the evolutionary adaptation of *M. smegmatis* mc^{2}51 to an important clinical stressor H_2O_2 allows bacteria to undergo coordinated genetic mutations, resulting in increased pathogenicity. Whole-genome sequencing identified a mutation site in the *fur* gene, which caused increased expression of *katG*. Using a Wayne dormancy model, mc^{2}51 showed a growth advantage over its parental strain mc^{2}155 in recovering from dormancy under anaerobic conditions. Meanwhile, the high level of KatG in mc^{2}51 was accompanied by a low level of ATP, which meant that mc^{2}51 is at a low respiratory level. Additionally, the redox-related protein Rv1996 showed different phenotypes in different specific redox states in *M. smegmatis* mc^{2}155 and mc^{2}51, *M. bovis* BCG, and *M. tuberculosis* mc^{2}7000. In conclusion, our study shows that the same gene presents different phenotypes under different physiological conditions. This may partly explain why *M. smegmatis* and *M. tuberculosis* have similar virulence factors and signaling transduction systems such as two-component systems and sigma factors, but due to the different redox states in the corresponding bacteria, *M. smegmatis* is a nonpathogen, while *M. tuberculosis* is a pathogen. As mc^{2}51 overcomes its shortcomings of rapid removal, it can potentially be developed as a vaccine vector.

Keywords: mycobacterial pathogenicity, KatG, Fur, H_2O_2 resistant, TB

OPEN ACCESS

Edited by:
Wei Huang,
Johns Hopkins University,
United States

Reviewed by:
Liang Wang,
Xuzhou Medical University, China
Jeffrey Morris,
University of Alabama at Birmingham,
United States

***Correspondence:**
Kaixia Mi
mik@im.ac.cn

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Genetics

Received: *13 August 2021*
Accepted: *10 February 2022*
Published: *16 March 2022*

Citation:
Jiang Z, Zhuang Z and Mi K (2022) Experimental Evolution Reveals Redox State Modulates Mycobacterial Pathogenicity. Front. Genet. 13:758304. doi: 10.3389/fgene.2022.758304

INTRODUCTION

Tuberculosis (TB) is caused by the pathogen *Mycobacterium tuberculosis* and remains a public health threat, resulting in 1.4 million deaths in 2020 (WHO, 2021). The prevalence of multidrug-resistant *M. tuberculosis* and the rising cases of co-infection with HIV increase this health concern. The World Health Organization (WHO) has estimated that a quarter of the world's population is infected with *M. tuberculosis* (WHO, 2021). This latent state may be extended as long as the life of the infected host, but unfortunately, the reactive rate is approximately 5–10% of infected individuals (Flynn and Chan, 2001). TB is treated with chemotherapy, and the latent state of mycobacteria prolongs the time of treatment, which is one of the causes for the development of mycobacterial resistance.

As one of the world's most successful human pathogens, *M. tuberculosis* has evolved elegant strategies to escape the immune defensive system of the host. For example, D'Arcy et al. observed that *M. tuberculosis*-containing phagosomes do not fuse with the lysosome inside infected macrophages (Brown et al., 1969; Armstrong and Hart, 1971). Several studies have indicated that *M. tuberculosis*

being an intracellular pathogen is partially due to its ability to survive and persist in macrophages, in hostile environments with oxidative stress, in low pH, and under starvation and other stresses (Cohen et al., 2018; Nauseef, 2019). When *M. tuberculosis* infects macrophages, mycobacteria must overcome exogenous reactive oxygen species (ROS), a classical innate defense mechanism against infection (Pieters, 2008). In addition, during latency, *M. tuberculosis* continues to be exposed to oxidative stress, and thus, the accumulation of mutations caused by oxidative DNA damage is predicted as a potential risk for resistance to antibiotics (Ford et al., 2011). Clinical investigation has shown that cells from TB patients produce less ROS than those of healthy individuals (Jaswal et al., 1992; Kumar et al., 1995). In addition, ROS can also be associated with the treatment of TB (Piccaro et al., 2014; Hu et al., 2021). To avoid ROS attack, an evolved detoxified system is essential for *M. tuberculosis* survival, persistence, and subsequent reactivation (Piccaro et al., 2014).

As a mycobacterial model of *M. tuberculosis*, *M. smegmatis* has contributed to understanding the functions of mycobacteria (Aldridge et al., 2012; Kieser et al., 2015; Gray et al., 2016). The essential genes in *M. tuberculosis* have a corresponding ortholog gene in *M. smegmatis* (Dragset et al., 2019; Judd et al., 2021). *M. smegmatis* mc^{2}155 and *M. tuberculosis* H37Rv share 2547 mutually orthologous genes (Dragset et al., 2019). At least under certain conditions, *M. tuberculosis* and *M. smegmatis* have similar growth mechanisms. The assumed virulence factors identified in *M. tuberculosis* such as PhoPR and DosR/S/T (Gonzalo-Asensio et al., 2014; Mehra et al., 2015) are also present in *M. smegmatis*. However, compared to virulent *M. tuberculosis*, *M. smegmatis* is a nonpathogenic mycobacterium. *M. tuberculosis* can persist in the infected host, while the host can quickly remove *M. smegmatis* (Anes et al., 2003; Anes et al., 2006). Considering that *M. tuberculosis* with high resistance to hydrogen peroxide (H_2O_2) persists in the lungs, while *M. smegmatis* with lower resistance to H_2O_2 exists in the soil, we hypothesized that the redox state of *M. tuberculosis* and *M. smegmatis* may adapt to the corresponding redox environment. The differences between the two bacteria are due to their corresponding degree of resistance to H_2O_2. To test this hypothesis, we selected a series of H_2O_2-resistant mutant strains using a clinically important stressor H_2O_2 and identified the highly H_2O_2-resistant mycobacterial strain mc^{2}51 (Li et al., 2014b). Compared to the wild-type *M. smegmatis* mc^{2}155 strain, the minimum inhibitory concentration (MIC) of H_2O_2 was more than 80-fold higher, while the MIC of mc^{2}51 to H_2O_2 was 3.125 mM, similar to that of *M. bovis* BCG (0.625 mM) (Li et al., 2014a) and *M. tuberculosis* (0.625 mM), while that of mc^{2}155 was 0.039 mM (Li et al., 2014b). mc^{2}51 exhibited a slow growth rate similar to *M. tuberculosis*.

In this study, we first showed that H_2O_2-resistant mc^{2}51 had a growth advantage both in mice and macrophages compared with wild-type mc^{2}155, which indicated that the higher resistance to H_2O_2 of mycobacteria is related to higher virulence. Similar to *M. tuberculosis* that can survive under hypoxia in the lungs and resuscitate under appropriate conditions, mc^{2}51 presented a growth advantage to recover from dormancy using the Wayne dormancy model. Furthermore, we showed that Fur mutant carrying the A28V point mutation dysregulated *katG* levels, which was the main cause for resistance to H_2O_2 and low levels of ATP. Additionally, the redox-related protein Rv1996, responsible for regulating gene expression, exhibited different phenotypes associated with isoniazid susceptibility in *M. smegmatis* mc^{2}155 and mc^{2}51, *M. bovis* BCG, and *M. tuberculosis* mc^{2}7000. Thus, the same protein presents different phenotypes under different physiological conditions. Our results suggest that the difference in the corresponding redox status causes the difference in pathogenicity between *M. tuberculosis* and *M. smegmatis*.

RESULTS

The *Mycobacterium tuberculosis*-Like *M. smegmatis* Mutant Strain mc^{2}51 Displayed Improved Virulence

Previous studies showed that *M. smegmatis* with the gain of H_2O_2 resistance, named mc^{2}51, improves growth fitness in mycobacteria under stress (**Table 1**) (Li et al., 2014a; Li et al., 2014b). Prior to selecting H_2O_2-adapted mycobacterial mutants, we first measured the MIC of wild-type mc^{2}155 to H_2O_2, which was 0.039 mM. We used 0.0293 mM H_2O_2 for the initial screening, which was lower than the MIC of mc^{2}155 to H_2O_2 (**Table 1**). The whole process was performed as follows: cultures were started from glycerol-frozen stocks and grown to log phase (OD_{600} of 0.6–0.8). Then, the culture was diluted 1,000 times and grown under 0.0293 mM H_2O_2 until the OD_{600} reached logarithmic (log) phase. Cultures were then further diluted 1:1,000, and an additional 0.0293 mM of H_2O_2 was added to the culture. This process was repeated until the H_2O_2 concentrations reached 0.4395 mM. In further rounds of culture, H_2O_2 was added in steps of 0.0879 mM instead of 0.0293 mM until a concentration of 1.5 mM was reached (**Supplementary Figure S1**; **Supplementary Table S1**). The actual MIC of H_2O_2 of the mutant strain, named mc^{2}51, selected at 1.5 mM, was 3.125 mM. In a previous study, the mutated mycobacterial strain, mc^{2}51, evolved into an *M. tuberculosis*-like strain that presented slow growth and improved growth fitness under stress (Li et al., 2014b), which H_2O_2 resistance in mycobacteria linked to virulence. Previous studies including ours have associated isoniazid (INH) with H_2O_2 resistance (Timmins and Deretic, 2006; Hu et al., 2021). The MIC of isoniazid (INH) for mc^{2}51 was dramatically reduced to ~1% of the MIC for mc^{2}155 (**Table 1**). To test our hypothesis that the H_2O_2-resistant strain mc^{2}51 would be more persistent in the host than its parental strain mc^{2}155, we performed non-invasive intranasal infections, which can induce respiratory mucosal immune responses and is a promising way of vaccination for respiratory infection diseases. As shown in **Figure 1A**, intranasal infection of C57BL/6 mice with ~1 × 10^7/50 μl and bacterial colony-forming units (CFUs) were counted 1 day after

TABLE 1 | List of bacteria in this study.

Strains	Genotype or relevant characteristics	Source or references	H_2O_2 (mM)	INH (μg/ml)
mc^2155	*Mycobacterium smegmatis*	Snapper et al. (1990)	0.039	10
$mc^2$51	Highly H_2O_2-resistant *Mycobacterium tuberculosis*-like *Mycobacterium smegmatis*	Li et al. (2014a); Li et al. (2014b)	3.125	0.1
Δ*katG*	*Mycobacterium smegmatis* Δ*katG::hyg*^{R}	This study	N/A	N/A
pMV261-*katG*/mc^2155	*Mycobacterium smegmatis* harboring pMV261- *katG*	This study	N/A	N/A
mFur/mc^2155	*Mycobacterium smegmatis mFur* at the T28A site	This study	0.64	N/A
pMV261/mc^2155	*Mycobacterium smegmatis* harboring pMV261 Kan^R	This study	N/A	2.5
pMV261-*rv1996*/mc^2155	*Mycobacterium smegmatis* harboring pMV261-*rv1996*, Kan^R	This study	N/A	10
pMV261/$mc^2$51	Highly H_2O_2-resistant *Mycobacterium smegmatis* harboring pMV261 Kan^R	This study	N/A	0.1
pMV261-*rv1996*/$mc^2$51	Highly H_2O_2-resistant *Mycobacterium smegmatis* harboring pMV261-*rv1996*, Kan^R	This study	N/A	0.1
pMV261/BCG	*Mycobacterium bovis* BCG Pasteur harboring pMV261, *Kan*^{R}	This study	N/A	0.05
pMV261-*rv1996*/BCG	*Mycobacterium bovis* BCG Pasteur harboring pMV261-*rv1996*, *Kan*^{R}	This study	N/A	0.025
pMV261/$mc^2$7000	*Mycobacterium tuberculosis ΔpanCD harboring pMV261, Kan*^{R}	This study	N/A	0.05
pMV261-*rv1996*/$mc^2$7000	*Mycobacterium tuberculosis ΔpanCD harboring pMV261-rv1996, Kan*^{R}	This study	N/A	0.05

N/A: not available.

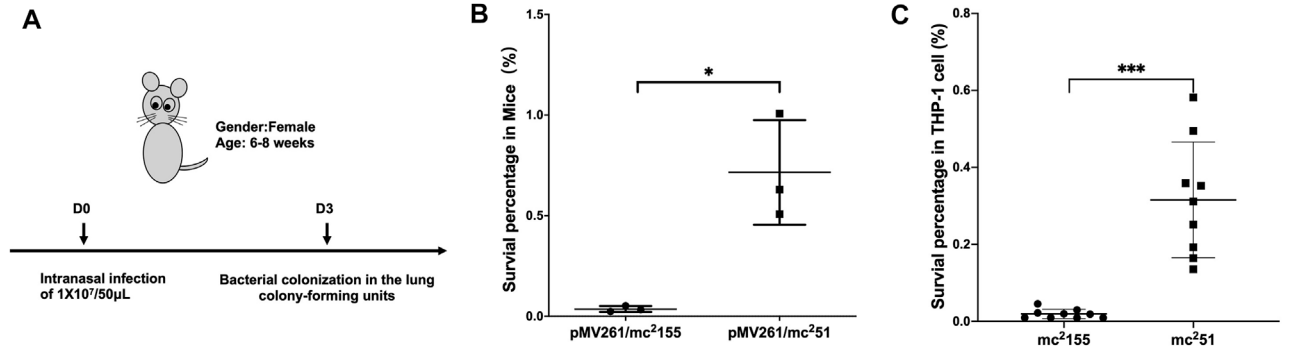

FIGURE 1 | *M. smegmatis* $mc^2$51 displayed improved virulence in mice and THP-1 cells. **(A)** The schematic diagram of the mice infection. C57BL/6 mice were intranasally infected as described in *Materials and Methods* with ~1 × 10^7/50 μl of *M. smegmatis* $mc^2$155 or $mc^2$51. At the 3rd day after infection, the whole lung homogenates were plated to determine bacterial numbers. **(B)** The H_2O_2-resistant mutant strain $mc^2$51 has a significantly higher bacterial load in the infected lung. The lung burdens 3 days after infection with $mc^2$155 (black circles) or $mc^2$51 (black squares) were measured, and the survival percentage was calculated as described in *Materials and Methods*. Each group consisted of three mice. *$p < 0.05$. **(C)** The H_2O_2-resistant mutant strain $mc^2$51 has a survival advantage over wild-type $mc^2$155 in the macrophage-like cell line THP-1. THP-1 cells were infected with $mc^2$155 (black circles) or $mc^2$51 (black squares) at a multiplicity of infection of 1 h after infection, THP-1 cell lysates were collected, and the intracellular bacilli were measured, and the survival percentage was calculated as described in *Materials and Methods*. Each group consisted of nine repeats. ***$p < 0.001$.

infection. Compared to $mc^2$155 with the percentage of survival in the infected lung of 0.03653 ± 0.01462% (n = 3), the $mc^2$51 strain with that of 0.7153 ± 0.2597% (n = 3) had a significantly higher survival percentage (p = 0.0451) (**Figure 1B**). After infection with live mycobacteria, the bacilli enter alveolar macrophages (AMΦs) and persist within them. To exclude the possibility that the higher CFUs were caused by the larger number of infected bacilli being captured by AMΦs, macrophage-killing assays were explored using the THP-1 cell line. The survival percentage of mutant $mc^2$51 is 0.3158 ± 0.1502 (n = 9), while the survival of wild-type $mc^2$155 was 0.01921 ± 0.01223 (n = 6) (p = 0.0003) (**Figure 1C**). These results indicated that $mc^2$51 exhibited enhanced virulence.

Mutant *fur* Altered the Intracellular Redox State and Was Conductive to Latency and Resuscitation *via* Modulation of KatG Levels

M. tuberculosis has evolved to survive in hypoxic conditions. Over more than 100 years of research, *M. tuberculosis* has been confirmed to be an obligate aerobic bacterium that cannot replicate under hypoxic conditions. However, *M. tuberculosis* has incredible survivability in long-term anaerobic environments.

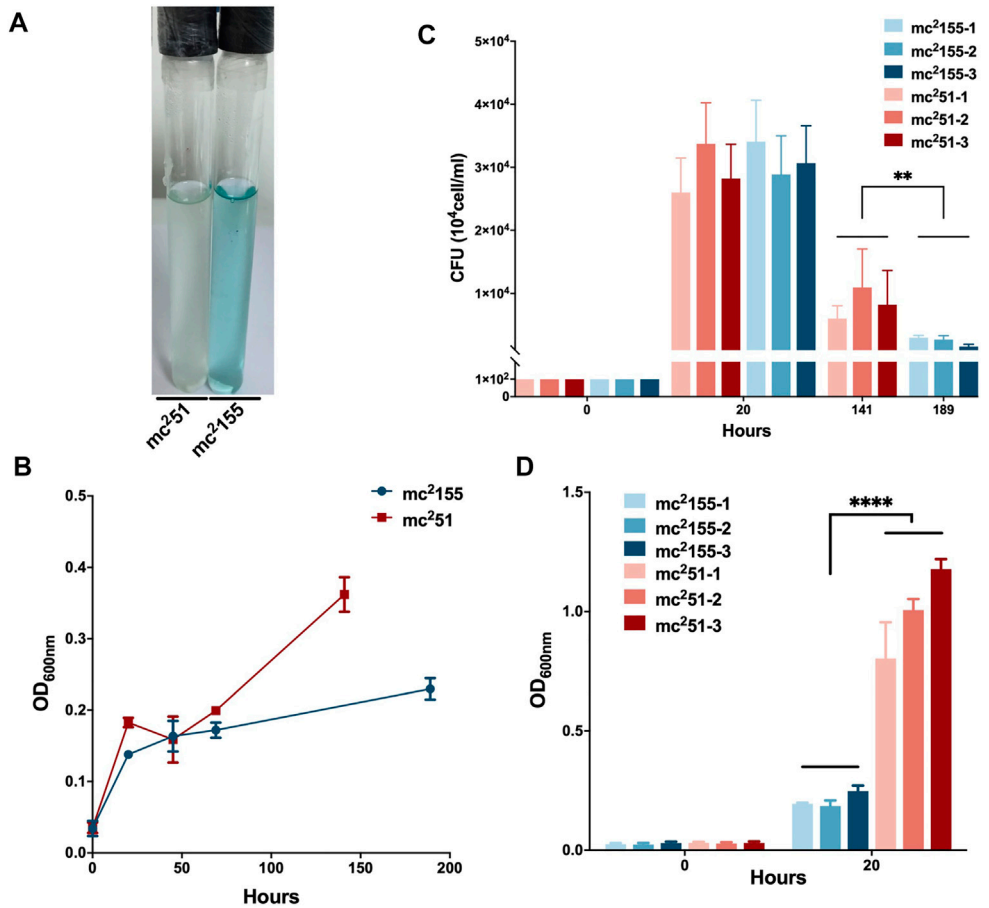

FIGURE 2 | Survival in hypoxic environments and resuscitation of *M. smegmatis* mc^2 155 and mc^{2}51. **(A)** The oxygen tension indicator methylene blue of mc^{2}51 culture (left) changes to colorless on day 6, while the indicator of mc^{2}155 culture (right) stays blue in the Wayne dormancy model. The cultures were initially inoculated in anaerobic tubes at an OD_{600} of 0.01 with a headspace ratio of 0.5. The cultures were stirred at 120 rpm. Methylene blue (1.5 mg/L) was used as an oxygen tension indicator. Methylene blue changes from blue to colorless under reducing conditions. Data present results of three biological replicates. Growth rates of strains mc^{2}155 (navy blue) and mc^{2}51 (brown) were measured by measuring the OD_{600} **(B)** and by determination of CFUs **(C)** after plating on 7H10. Data are presented as the mean ± standard deviation of three independent replicates. $^{**}p < 0.01$. **(D)** The H_2O_2-resistant mutant strain mc^{2}51 shows a growth advantage of recovering from dormancy in an anaerobic state over mc^{2}155. After the indicator methylene blue in the culture of the Wayne dormancy model became colorless, cultures of mc^{2}155 (navy blue) or mc^{2}51 (brown) were collected and reinoculated into 7H9 combined with the heart infusion medium at an OD_{600} of 0.01 and aerobic shaken at 200 rpm. The CFUs were determined 20 h after inoculation. Data are presented as the mean ± standard deviation of three independent replicates. $^{****}p < 0.0001$.

Evidence suggests that *M. tuberculosis* has the ability to reduce the respiratory system to low levels and maintain vitality (Loebel et al., 1933). Due to respiratory depression, ATP is maintained at low levels, which guarantees minimal metabolic activity to ensure membrane integrity under hypoxic conditions. We first compared the survival of mc^{2}51 and mc^{2}155 strains under hypoxic conditions. We established a Wayne dormancy model (Wayne and Hayes, 1996), consisting of mycobacterial strain cultures grown in the 7H9 medium to an OD_{600} of 1.0, which were then transferred to anaerobic tubes containing 1×10^6 cells/ml with a headspace ratio of the culture system of 0.5. Methylene blue (1.5 mg/L) was added as an indicator of oxygen status. The OD_{600} and CFUs at different indicated times and the color transition time were measured. The indicator of mc^{2}51 had become colorless on the 6th day (141 h); on the contrary, the indicator of mc^{2}155 was still in blue at this time (**Figure 2A**). For mc^{2}155, the indicator became colorless in about 8 days (189 h), indicating that it has entered an anaerobic state. The mc^{2}51 entered the anaerobic state faster in the later stages, and the number of viable bacteria in the anaerobic state was significantly higher than mc^{2}155 (**Figures 2B,C**).

For the activation experiment, bacteria cultured in the anaerobic conditions were collected and diluted to 1×10^6 cells/ml in 7H9 and the brain–heart infusion medium and were then grown under aerobic conditions. The three independent mc^{2}51 clones were set, and each clone set up three replicates, all of which were grown better than the three independent clones of mc^{2}155 (**Figure 2D**). Thus, mc^{2}51 had the growth advantage of recovering from the dormancy of the anaerobic state over mc^{2}155. We also measured the intracellular ATP levels of mc^{2}51 and mc^{2}155. As expected,

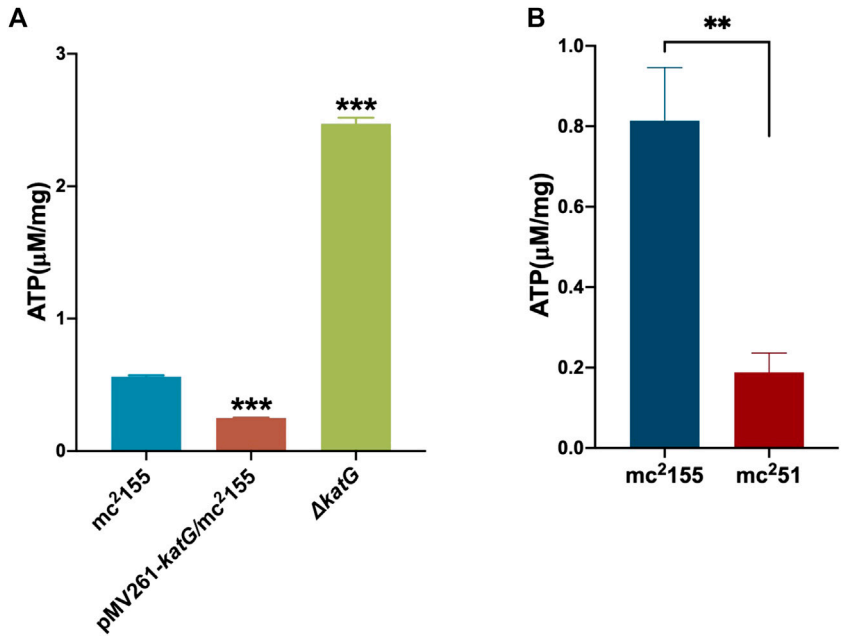

FIGURE 3 | ATP content in mc^{2}51 is similar to that in pMV261-*katG*/mc^{2}155. **(A)** The detection of ATP content in mc^{2}155, Δ*katG*, and pMV261-*katG*/mc^{2}155. The cultures of mc^{2}155 (navy blue), pMV261-*katG*/mc^{2}155 (brown), and Δ*katG* (olive drab) at an OD_{600} of 0.8 were collected. ATP levels were measured in relative light unit (RLU) using a Cytation 3 Cell Imaging Multi-Mode Reader. The corresponding ATP concentrations were calculated according to the ATP standard curve and further converted to μM/mg protein. Data are presented as the mean ± standard deviation of three independent replicates. ***$p < 0.001$. **(B)** The intracellular ATP content of mc^{2}51 and mc^{2}155. The cultures of mc^{2}155 (navy blue) or mc^{2}51 (brown) at an OD_{600} of 0.8 were collected. ATP levels were measured in relative light unit (RLU) using a Cytation 3 Cell Imaging Multi-Mode Reader. The corresponding ATP concentrations were calculated according to the ATP standard curve and further converted to μM/mg protein. Data are presented as the mean ± standard deviation of three independent replicates. **$p < 0.01$.

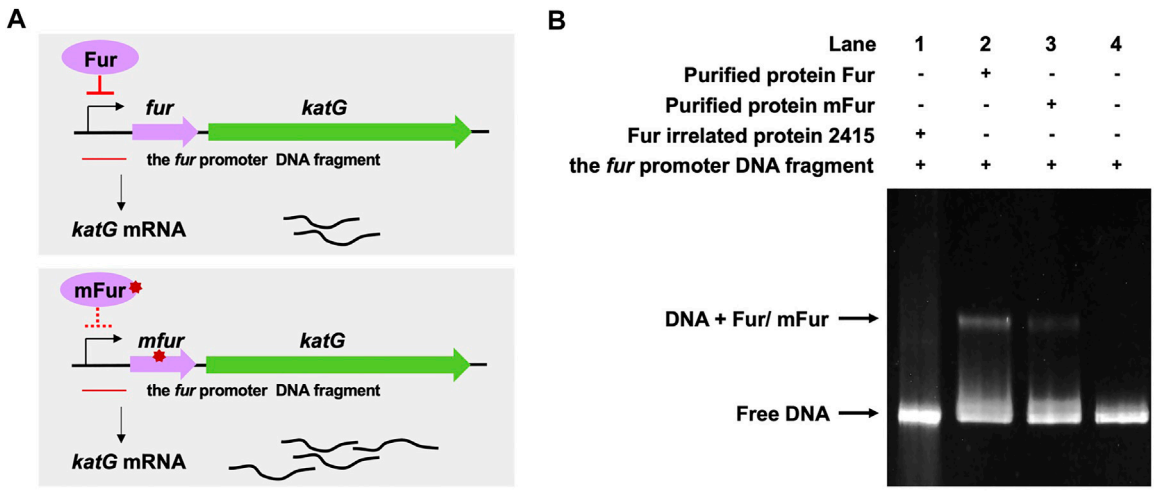

FIGURE 4 | A28V Fur mutant protein decreased DNA binding to the *fur* promoter. **(A)** Genetic organization of the *fur-katG* and the schematic diagram of Fur negative regulation of *katG* (upper panel). Genetic organization of the *mfur-katG* and the schematic diagram of mFur resulted in derepression of *katG* with increasing *katG* mRNA (bottom panel). The red line indicates the *fur* promoter DNA fragment. **(B)** Electrophoretic mobility shift assays (EMSAs) of the binding of Fur/mFur protein to the *fur* promoter DNA fragment. Purified MSMEG_2415 protein (2415), irrelated to Fur, expressed in *E. coli*, was run in the first lane of a 4–20% polyacrylamide gel. Gel shift caused by Fur (lane 2) and mFur (lane 3) is shown. The image shown is representative of at least three experiments.

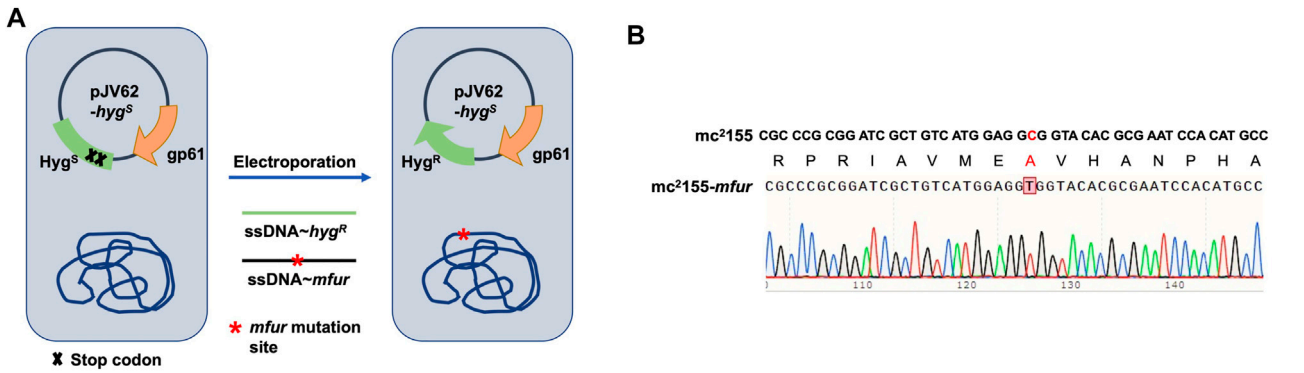

FIGURE 5 | Generation of a point mutation furA28V in fur on the chromosome in *M. smegmatis*. **(A)** Strategies for recombineering. The site-directed mutagenesis of fur was obtained using Phage Che9c gp61-mediated recombination. Plasmids pJV62-*hyg*^{S} express Che9c gene product gp61, to facilitate single-stranded DNA carrying the point mutation recombination. The clones carrying the mutated site were selected on 7H9 with Hyg^R and Kan^R. **(B)** The sequencing result of a point mutation in fur. Sanger sequencing showed the nuclear acid point mutation site c/t, resulting in the protein switch from A to V at site 28th.

the abundance of ATP in mc^{2}51 (0.1885 ± 0.0481 μM/mg) was lower than that of mc^{2}155 (0.8138 ± 0.1324 μM/mg) (**Figure 3B**).

We previously performed whole-genome sequencing to compare differences in mc^{2}155 and mc^{2}51 at the genome level. Whole-genome sequencing revealed that there were 29 single-nucleotide polymorphisms (SNPs) in mc^{2}51. All 29 SNPs were cloned, and each was transformed into mc^{2}155 strains, and only the *fur* (*msmeg_3460*), located upstream of *msmeg_3461*, encoding KatG (**Figure 4A**), could restore the resistance phenotype of H_2O_2 (Li et al., 2014b). The *fur*-encoded protein Fur negatively regulated *katG* expression (Pym et al., 2001). The A28V Fur mutation (mFur) in mc^{2}51 may also affect the expression of *katG*. To verify this hypothesis, we examined the binding of mFur to the target DNA (the promoter region of the *fur*) using electrophoretic mobility shift assays (EMSAs). Compared to wild-type Fur protein, EMSA showed that mFur decreased DNA binding (**Figure 4B**), which resulted in the *katG* transcription dysregulation by mFur. In addition, the RNA of mc^{2}155 and mc^{2}51 was extracted and quantified. The results showed that compared to the mc^{2}155 strain, the expression of the catalase–peroxidase (KatG) encoding gene *katG* of the mc^{2}51 strain was significantly upregulated to ~61.82-fold that of the wild-type strain. Taken together, mFur increases the KatG protein level in mc^{2}51. KatG is a dual enzyme for catalase and peroxidase, which hydrolyzes ROS (Ng et al., 2004). Thus, the mc^{2}51 may maintain ATP at lower levels through KatG, compared to mc^{2}155 levels; that is, the abundance of KatG may affect the mycobacterial redox state and, thus, change the susceptibility to H_2O_2. We then constructed the Δ*katG* (mc^{2}155 with knockout *katG*) and pMV261-*katG*/mc^{2}155 (mc^{2}155 with overexpression of *katG*) strains, and their respective ATP content was tested. As shown in **Figure 3A**, the KatG level negatively correlated with the ATP level. Furthermore, we constructed the specific site mutant of the *fur* gene (*mfur*) in wild-type mc^{2}155 causing an amino acid change of A28V of Fur (**Figures 5A,B**) by using recombination protein gp61 from Che9c mycobacteriophage (van Kessel et al., 2008), to construct mc^{2}155-*mfur* (**Table 1**). As we expected, the Fur mutation at A28V induced high resistance to H_2O_2 with the MIC of H_2O_2 in mc^{2}155-*mfur* being 0.64 mM. We showed that the point mutation of *fur* dysregulation of *katG* expression is a major factor leading to the phenotype H_2O_2 resistance.

The Same Protein Performs Different Functions in Different Redox States

As a successful human pathogen, *M. tuberculosis* has unique respiration properties. *M. tuberculosis* excretes alkaline supernatants, which is in contrast to other strains that excrete acidic supernatants (Merrill, 1930). The difference between secreted compounds with different acid–base properties suggested that *M. tuberculosis* has a distinctive redox state. As shown in **Table 2**, the comparative genomic analysis shows that PhoPR and DosR/S/T, identified as virulence factors of *M. tuberculosis*, are present in *M. smegmatis*. The signaling transduction systems such as the two-component systems and the sigma factors of *M. tuberculosis* are homologous in *M. smegmatis* (**Table 3**). Different phenotypes might be due to different redox states. Thus, we considered that the same redox-regulated related protein might perform different functions in different redox states. To test this hypothesis, we examined the biological function of a universal stress protein Rv1996 that increases the expression of KatG, in various mycobacterial strains (Hu et al., 2015). Both previous studies and our studies have linked isoniazid action with redox states, and H_2O_2 resistance is negatively correlated with INH susceptibility in mycobacteria (Bhaskar et al., 2014; Hu et al., 2015; Vilcheze et al., 2017). We used INH as a chemical probe for monitoring mycobacterial redox states and measured the MICs of INH to the corresponding mycobacterial strains (**Figure 6**). As predicted, the MICs of INH differed across the tested mycobacterial strains: the MIC of INH in pMV261-*rv1996*/mc^{2}7000 was equal to that of pMV261/mc^{2}7000; the MIC of INH in pMV261-*rv1996*/BCG was lower than that in pMV261/BCG; the MIC of INH in pMV261-*rv1996*/mc^{2}51 was equal to that of pMV261/mc^{2}51; and in mc^{2}155, the opposite results were observed with the MIC of

TABLE 2 | Conservation of *M. tuberculosis* H37Rv TCSS in *M. smegmatis*.

Gene	*M. tuberculosis* rv#	Mtb	Msm	References
RegX3-SenX3	Rv0491-Rv0490	+	+	James et al. (2012); Parish et al. (2003b); Rifat and Karakousis (2014)
HK1-HK2-TcrA	Rv0600c-Rv0601c-Rv0602c	*	–	Shrivastava and Das (2007)
PhoP-PhoR	Rv0757-Rv0758	+	+	Walters et al. (2006)
NarL-NarS	Rv0844c-Rv0845c	+	+	Schnell et al. (2008)
PrrA-PrrB	Rv0903c-Rv0902c	+	+	Arora et al. (2021); Nowak et al. (2006)
MprA-MprB	Rv0981-Rv0982	+	+	He and Zahrt (2005); Sureka et al. (2007)
KdpD-KdpE	Rv1028c-Rv1027c	+	+	Parish et al. (2003a); Steyn et al. (2003)
TrcR-TrcS	Rv1032c-Rv1034c	+	+	Haydel et al. (2002)
MtrA-MtrB	Rv3245c-Rv3247c	+	+	Fol et al. (2006); Li et al. (2010); Plocinska et al. (2012)
TcrX-TcrY	Rv3765c-Rv3764c	+	+	Bhattacharya et al. (2010)
PdtaS-PdtaR	Rv3220c-Rv1626	+	+	Boshoff et al. (2004); Shrivastava and Das (2007)

Msm, Mycobacterium smegmatis*; Mtb,* Mycobacterium tuberculosis *CDC1551. +Genes encoding the sensor kinase and the response regulator are present and genetically linked. –Genes encoding both the sensor kinase and the response regulator are absent. *Genes encoding the two sensor kinases have been fused, and the gene encoding this fused sensor kinase is genetically linked to the response regulator.*

TABLE 3 | Sigma factor genes in mycobacteria.

Sigma	*M. tuberculosis* rv#	Mtb	Msm	References
SigA(σ^{A})	Rv2703	+	+	Gomez et al. (1998)
SigB (σ^{B})	Rv2710	+	+	Lee et al. (2008b); Mukherjee and Chatterji (2005)
SigC (σ^{C})	Rv2069	+	–	Sun et al. (2004)
SigD(σ^{D})	Rv3414c	+	+	Calamita et al. (2005); Raman et al. (2004)
SigE(σ^{E})	Rv1221	+	+	Song et al. (2008)
SigF (σ^{F})	Rv3286c	+	+	Rodrigue et al. (2007)
SigG (σ^{G})	Rv0182c	+	+	Lee et al. (2008a)
SigH (σ^{H})	Rv3223c	+	+	Song et al. (2008)
SigI (σ^{I})	Rv1189	+	–	Homerova et al. (2008)
SigJ (σ^{J})	Rv3328c	+	+	Homerova et al. (2008)
SigK (σ^{K})	Rv0445c	+	–	Veyrier and Behr, (2008)
SigL (σ^{L})	Rv0735	+	+	Hahn et al. (2005)
SigM(σ^{M})	Rv3911	+	+	Agarwal et al. (2007); Raman et al. (2006)

+, presence of the gene; –, absence of the gene.

INH in pMV261-*rv1996*/mc^{2}155 being lower than that of INH in pMV261/mc^{2}155 (**Figure 6**).

DISCUSSION

Understanding how *M. tuberculosis* evolved into a professional pathogen is of benefit to the study of its pathogenesis and design of vaccines. The combination of experimental evolution and whole-genome sequencing provides a powerful method for identifying adaptive mutations and elucidating the specific genotype–phenotype relationship (Elena and Lenski, 2003; Lenski, 2017). Historically, the most successful example of continuous selective cultures is *M. bovis BCG*, the only anti-TB vaccine, which was attenuated after 13 years of continuous *in vitro* passages of *M. bovis BCG*. We previously used a similar adaptive evolution strategy to select H_2O_2-resistant *M. smegmatis* strains by using a clinically key stressor H_2O_2 (Li et al., 2014b). Preliminary results showed that the mc^{2}51 strain was highly resistant to H_2O_2 and had greater susceptibility to INH, compared to mc^{2}155. The mc^{2}51 phenotype showed an *M. tuberculosis*-like *M. smegmatis* phenotype. Altogether, the mutant *M. smegmatis* mc^{2}51 exhibited higher virulence.

The whole-genome sequencing showed the presence of gene mutations in *fur,* and the mutant Fur resulted in *katG* levels (**Figure 4B**). In the Wayne dormancy model, mc^{2}51 shows a growth advantage of recovering from dormancy under anaerobic conditions over mc^{2}155 (**Figure 2B**). In parallel, a high level of *katG* in mc^{2}51 is accompanied by lower ATP levels, which implied mc^{2}51 exhibited at a lower level of respiration (**Figure 3B**). Moreover, we showed that a redox-related protein Rv1996 exhibits a different phenotype under different specific redox states in *M. smegmatis* mc^{2}155 and mc^{2}51, *M. bovis* BCG, and *M. tuberculosis* mc^{2}7000 (**Figure 6**). This study indicated that the same genotype presents different phenotypes under different physiological conditions. We at least partially explain why *M. smegmatis* and *M. tuberculosis* have similar virulent factors, including a two-component system and sigma factors (**Tables 2**, **3**), but *M. smegmatis* is a nonpathogen and *M. tuberculosis* is a pathogen.

M. tuberculosis is a successful human pathogen. It is considered to be derived from the environment (Gutierrez

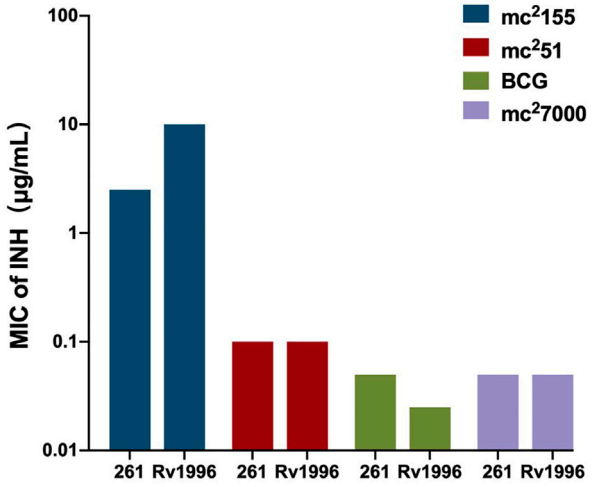

FIGURE 6 | Comparison of isoniazid susceptibility in different mycobacterial strains harboring pMV261-*rv1996* and the corresponding control strain harboring pMV261.The minimum inhibitory concentrations (MICs) of isoniazid were determined by inoculating each bacterial strain in 7H9 containing serially isoniazid (INH). The values of MIC were recorded. 261 present pMV261. Rv1996 present pMV261-*rv1996*. The image represents the results of three independent repeats.

et al., 2005; Wolfe et al., 2007) and has adapted to the immune environment of the human body through long-term evolution. Its successfully established infection is partially attributed to its survival capacity and persistence in macrophages (Podinovskaia et al., 2013). To defend against mycobacterial infection, the host produces ROS, as an important innate defense mechanism. Consequently, *M. tuberculosis* has evolved a hierarchy and unique antioxidant function and maintains a low level of respiration, manifested by slow growth and persistence in the host. In contrast, *M. smegmatis* is present in the soil, which is a totally different environment from the host (Zhang and Furman, 2021). In **Table 2**, we show that *M. tuberculosis* and *M. smegmatis* have similar genotypes; however, they show different phenotypes, in terms of INH susceptibility, H_2O_2 resistance, and virulence. We believe that this striking difference is due to H_2O_2 resistance. The selected resistance to H_2O_2 of mc^{2}51 shows improved virulence in both the macrophage-killing assay and in an animal model (**Figure 1**). In fact, several studies have shown that abiotic stress can improve the virulence phenotype of bacterial pathogens (Sundberg et al., 2014; Li et al., 2021). Our study also supports the sit-and-wait hypothesis (Wang et al., 2017), that is, bacterial environmental abiotic stress and virulence evolution. In addition, this study also suggests that we can use mc^{2}51 as a model strain, replacing mc^{2}155, to study the regulation of redox homeostasis of *M. tuberculosis*.

We previously sequenced the whole genome of mc^{2}51 strain (Li et al., 2014a) and identified 29 SNPs, compared to mc^{2}51. Confirmed with our previous study (Li et al., 2014b), we found that only the *fur* gene can partially complement the resistant phenotype. This suggested that the *fur* mutation facilitated elevated H_2O_2 resistance, although it was not entirely responsible for the high resistance observed in mc^{2}51. We also mutated the *fur* in wild-type mc^{2}155 by genome editing and produced a phenotype similar to mc^{2}51, which is highly resistant to H_2O_2 (**Table 1**). Large-scale whole-genome sequencing studies on the evolutionary history of tuberculosis also show that key tract mutations at the transcription site will have a critical impact on the particular phenotype (Gagneux, 2018). For example, the change of PhoP in BCG allows infection with bovine pathogenic bacteria capable of infecting humans (Gonzalo-Asensio et al., 2014; Broset et al., 2015). This study reminds us that when designing vaccines, greater attention should be paid to regulators, which may be more efficient targets. *M. smegmatis* is an effective vaccine for TB and HIV (Sweeney et al., 2011; Kim et al., 2017). The disadvantage of *M. smegmatis* as a vaccine vector is its transient infection and difficulty to establish a persistent infection and produce adaptive immunity. The *M. tuberculosis*-like mutant *M. smegmatis* mc^{2}51 may be developed as a vaccine vector.

By comparing the survival of mc^{2}155 and mc^{2}51 in the Wayne dormancy animal model, we also found that low respiratory levels are beneficial for survival under anaerobic conditions and resurrection (**Figure 2**). In the future, we plan to use these strains to compare physiological indicators such as NADH/NAD$^+$, NADPH/NADP$^+$, and ATP, to further understand the mechanisms underlying *M. tuberculosis* resurrection. This study provides insight into H_2O_2-resistant mechanisms in mycobacteria and has important implications for linking mycobacteria redox capacity and persistence infection in mice.

MATERIALS AND METHODS

Strains and Growth Conditions

The H_2O_2-resistant *Mycobacterium smegmatis* strain mc^{2}51 was screened in the Mi lab (Li et al., 2014b). *Mycobacterium tuberculosis* Δ*panCD* (named mc^{2}7000) (Sambandamurthy et al., 2002) was kindly gifted by J Deng. The *M. smegmatis* wild-type mc^{2}155, mutant strain mc^{2}51, *M. bovis* BCG Pasteur, and *M. tuberculosis* mc^{2}7000 were cultured in Middlebrook 7H9 (Becton Diskinson Sparks, MD, United States) supplemented with ADS (10% albumin, dextrose, and saline), 0.05% Tween 80 (Sigma, St. Louis, MO, United States), and 0.5% glycerol (Beijing Modern Eastern Fine Chemical Co., Ltd., Beijing, China) for liquid culture and Middlebrook 7H10 (Becton Diskinson Sparks, MD, United States) supplemented with ADS for bacterial colony culture. The colony-forming units (CFUs) of mycobacterial strains were determined by plating serial dilutions of cultures on Middlebrook 7H10 agar plates and incubating at 37°C in an atmosphere of 5% CO_2 for the indicated time. For mc^{2}7000 culture, panthothenate (24 mg/L) was added. When required, kanamycin (25 mg/L, Amresco, United States) and hygromycin (50 mg/L, Sigma, United States) were added. All bacterial strains used in this study are listed in **Table 1**.

Determination of MIC to Isoniazid and H_2O_2

The susceptibility of isoniazid (INH) or H_2O_2 of mycobacteria was evaluated using the modified broth microdilution method

(Franzblau et al., 1998). In brief, INH or H_2O_2 was serial diluted using the 7H9 medium. The diluted fold was 1.25- or 2-fold, when required. Then, 40 μl of diluted INH or H_2O_2 was mixed with 40 μl of mycobacterial suspension with 1×10^7 cells/ml in each well of 96-well microtiter plates and then incubated at 37°C for the indicated days. As an indicator, 0.02% resazurin was added to individual samples, and the color switches from blue to pink were recorded after 4 h. All the experiments were performed in triplicate. The abundance of the cultures was measured using a microplate reader (FLUOstar OPTIMA, BMG Labtech). A difference of two serial dilutions or more indicated a significant difference in the INH or H_2O_2 susceptibility of bacterial strains.

Mice Infection

Female pathogen-free C57BL/6 mice (aged 6–8 weeks) were purchased from Vital River (Beijing, China). For the intranasal infection of *M. smegmatis*, mice were anesthetized by intraperitoneal injection of pentobarbital sodium (60 mg/kg), and ~10^7 CFU/50 μl PBS of $mc^2$51 or $mc^2$155 was introduced dropwise through the nostril of each mouse. The bacterial burden throughout the infection was monitored by collecting whole lung tissue at the indicated times after the mice were euthanized, and serial dilutions were then plated on 7H10 agar plates. The dose of infection was confirmed on day one after infection by plating whole lung homogenates from three mice on 7H10 agar. The percentage of survival was calculated as (CFUs after infection/CFUs before infection) × 100%.

Macrophage-Killing Assay

Human-derived cell line THP-1 (ATCC TIB-202) was cultured in RPMI medium with 10% fetal bovine serum (FBS, GIBCO, United States). THP-1 cells were activated with 100 ng/ml phorbo-12-myristate-13-acetate (PMA, Sigma, United States) overnight. Infection was carried out at a multiplicity of infection (MOI) of 10 for 1 h at 37°C and 5% CO_2 atmosphere. The infected THP-1 cells were washed with RPMI 3 times and then chased for 1 h. The cells with intracellular bacilli were then washed and lysed in sterile cold PBST (PBS with 0.05% Tween 20). Lysates were then vortexed, diluted, and plated on 7H10 agar plates as previously described (Chan et al., 1992). The percentage of survival was calculated as (CFUs after infection/CFUs before infection) × 100%.

Wayne Dormancy Model and Dormancy Exit

Mycobacterial strains were cultured under hypoxic conditions as described by Wayne and Hayes (Wayne and Hayes, 1996). In brief, cultures were initiated at an OD_{600} of ~0.01 (1×10^6) and incubated in anaerobic tubes with sealed caps. The headspace ratio of the cultures was 0.5. The cultures were stirred using an 8 mm Teflon stir bar (Fisher Scientific, United States) at 200 rpm. Methylene blue (1.5 mg/L) was used as an oxygen tension indicator. It changes in color from blue to colorless under low oxygen tension. The color transition time was recorded. All experiments were performed in triplicate. Growth was monitored by measuring the OD_{600} and by determination of CFUs after plating on 7H10.

The indicator methylene blue in the culture became colorless, indicating that the bacteria entered anaerobic conditions. Then, after collecting the bacteria in the anaerobic tube, they were washed with the culture medium or PBS three times and resuspended in a culture medium (7H9 and brain–heart infusion medium), the concentration was adjusted to the same amount of OD_{600} (OD_{600} of ~1.0) and diluted for the CFU count, 1:100 or 1:50 into a fresh culture medium was shaken at 37°C, to monitor the status of the bacteria, and the OD_{600} was measured. Three independent mycobacterial strain clones were set, and each clone set up three replicates.

Measurement of Intracellular ATP

The ATP Assay Kit was purchased from Beyotime Biotechnology (Beijing, China). The intracellular ATP assay was performed following the protocol provided by the manufacturer. In brief, the sample measurements were prepared as follows: cultures of indicated mycobacteria were obtained to an OD_{600} of 0.8. Bacteria were collected by low-temperature centrifugation at the maximum speed, and the pellet was washed with precooled PBS buffer 3 times. A 300 μl volume of ATP detection lysate and 0.5 ml volume of glass beads were added for cell lysis. The lysate obtained was centrifuged at low temperature for 5 min, and the supernatant was placed on ice for later use. The preparation of the standard solution of gradient concentration ATP was performed as follows: the ATP standard solution was serially diluted into 7 concentrations of 10, 3.333, 1.111, 0.37, 0.1234, 0.04115, and 0.01371 μM and stored on ice for later use. The preparation of the working solution for the detection of ATP was performed as follows: an appropriate amount of ATP detection reagent was prepared according to the number of samples, and then, a 90% final volume of ATP detection reagent diluent was added. The prepared working fluid was placed on ice for further use. The determination of the ATP level was performed as follows: 1) the prepared ATP detection working solution was dispensed into 1.5 ml centrifuge tubes, 100 μl per tube, and was allowed to incubate at room temperature for 5 min to allow full reaction of the ATP in the centrifuge tube; 2) during the test, 20 μl of each sample (standard or total protein sample) was added to a 1.5 ml centrifuge tube containing 100 μl of ATP detection working solution and was mixed quickly with a pipette and incubated for 2 s to complete the reaction before using a Cytation 3 Cell Imaging Multi-Mode Reader to determine the relative light unit (RLU); 3) a standard curve was constructed to measure and determine the concentration of the sample by converting the RLU into an ATP concentration; and 4) in order to eliminate the error caused by the difference in the amount of protein during sample preparation, the BCA protein concentration determination kit produced by Beyotime Biotechnology (Beijing, China) was used to determine the protein concentration in the sample. The ATP concentration was converted to μM/mg protein.

Electrophoretic Mobility Shift Assay

The coding regions of *fur* and *mfur* were amplified from $mc^2$155 and $mc^2$51 genomic DNA and cloned into the *Escherichia coli* expression vector pET23b (+) (Novagen, Madison, WI, United States) in-frame fused with a C-terminal His_6-tag sequence to construct the plasmids pET23b-*fur* and pET23b-*mfur*. The final constructs were transformed into BL21 (DE3) for expression, and recombinant Fur/mFur proteins were purified using Ni-NTA

agarose (Qiagen, California, United States). The proteins were induced by the addition of 1 mM IPTG at 16°C for 12 h. Protein purification was performed as described previously (Li et al., 2014c). The protocols of the recombinant protein purification are available on request. The recombinant protein MSMEG_2415 was purified as described previously (Li et al., 2014c) and used as a negative control for EMSA, while MSMEG_2415 is irrelated to Fur. The DNA fragment containing the promoter region of *fur* for gel shift experiments was amplified by PCR with specific primers (forward: 5′-CGTTGGAAAACAACCATTGCAAG-3′, reverse: 5′-CATCCGCAGTTGGGCTTCGAAC-3′). Binding reaction mixtures in 20 μl of binding buffer (20 mM Tris HCl pH 8.0, 1 mM dithiothreitol (DTT), 50 mM KCl, and 5 mM $MgCl_2$) containing 0.15 pmol of the DNA fragment were incubated with purified Fur/mFur protein (0.5 nmol) for 30 min at 30°C. Reaction mixtures were loaded on a 4–20% polyacrylamide gel containing 0.5 × TBE. Gels were run at 70 V at 4°C for 3 h. The gel was stained with Good-view and photographed for the image.

Generation of the *katG* Knockout and KatG Overexpression Strains

The knockout *katG* strain was constructed using mycobacteriophage-based specialized transduction (Bardarov et al., 2002; Li et al., 2014c). The upstream and downstream sequences of *katG* were amplified from *M. smegmatis* genome DNA. The knockout vector was constructed using phAE159 (Hsu and Jacobs, unpublished data). The mycobacteriophage used for knockout was obtained using MaxPlax packaging extract (Epicentre Biotechnologies, Madison, WI, United States), and a *katG* knockout strain was obtained by phage transduction, named Δ*katG*. The KatG overexpression strain was constructed using pMV261 to yield pMV261-*katG*, and the constructed plasmid was electroporated into mc^2155, yielding pMV261-*katG*/mc^2155. The detailed information for construction of all the mycobacterial strains and primers for plasmid construction is available on request.

Generation of *Fur* Point Mutation on the Chromosome in *M. smegmatis*

The single-strand (ss) DNA oligonucleotides used for recombineering were ordered from Genewiz (Suzhou, China). The site-directed mutagenesis of *fur* was obtained using Phage Che9c gp61-mediated recombination (van Kessel et al., 2008). The detailed information on primers for construction of the *fur* mycobacterial strain (named mc^2155-*mfur*) is available on request. The coding region containing *fur* point mutation in the genome (encoding mFur) was amplified and sequenced by Genewiz (Suzhou, China).

Generation of the *rv1996* Overexpression Mycobacterial Strains

The *rv1996* gene was amplified and constructed and cloned into pMV261 to yield pMV261-*rv1996*. The constructed pMV261-*rv1996* plasmid was transformed into mycobacterial strains, *M. smegmatis* mc^2155 and mc^251, *M. bovis* BCG Pasteur, and *M. tuberculosis* mc^27000, and the corresponding strains, named pMV261-*rv1996*/mc^2155, pMV261-*rv1996*/mc^251, pMV261-*rv1996*/BCG, and pMV261-*rv1996*/mc^27000. The empty vector pMV261 was transformed into the corresponding mycobacterial strains, named pMV261/mc^2155, pMV261/mc^251, pMV261/BCG, and pMV261/mc^27000.

Statistical Analysis

Each experiment was carried out at least twice with three–nine mice or samples per group. The CFUs and OD_{600} were analyzed using an unpaired *t*-test (Version 8.0 for Windows GraphPad Software). The ATP content was analyzed using ANOVA tests (Version 8.0 for Windows GraphPad Software). $^{****}p < 0.0001$, $^{***}p < 0.001$, $^{**}p < 0.01$, and $^{*}p < 0.05$.

Animal Ethics

This study was performed in strict accordance with the recommendations of the Ethics Committee established in the Guide for the Care and Use of Laboratory Animals of the Institute of Microbiology, Chinese Academy of Sciences (IMCAS). The protocol was approved by the Committee on the Ethics of Animal Experiments of the IMCAS. The mice were bred under specific pathogen-free conditions at the IMCAS laboratory animal facility. All animal experiments were conducted under isoflurane anesthesia, and all efforts were made to minimize suffering.

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 following: BioProject: PRJNA233977; BioSample: SAMN02951866, https://www.ncbi.nlm.nih.gov/nuccore/JAJD00000000.

AUTHOR CONTRIBUTIONS

KM conceived and designed the experiments; ZJ and ZZ performed the experiments; KM wrote the manuscript; and KM, ZJ, and ZZ revised the manuscript. All authors have read and agreed to the published version of the manuscript.

FUNDING

This work was supported by grants from the Ministry of Science and Technology of China (2018YFC1603900 and 2017YFA0505901 to KM), National Natural Science Foundation of China (31970136 and 32170181 to KM), and International Joint Research Project of the Institute of Medical Science, University of Tokyo (Extension-2019-K3006 to KM).

ACKNOWLEDGMENTS

The authors thank J. Deng for providing the mycobacterial strain mc^{2}7000. They also thank Tong Yin for her help in preparing the experimental materials.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.758304/full#supplementary-material

Supplementary Figure S1 | Evolutionary selection of H_2O_2-resistant mutations in *M. smegmatis* Cultures were started from glycerol-frozen stocks and grown to log phase (OD600 of 0.6 - 0.8). Then the cultures were diluted 1:1000 into 5 ml of 7H9 media containing 10% ADS. Hydrogen peroxide (H_2O_2) was then added to a concentration of 0.0293 mM and cultures were grown until the OD600 reached log phase. Cultures were then further diluted 1:1000 and an additional 0.0293 mM of H2O2 was added to the culture. This process was repeated until the H_2O_2 concentration reached 0.4395 mM. In further rounds of culture, $H_2O_2$was added in steps of 0.0879 mM, instead of 0.0293 mM until an H2O2 concentration of 1.5 mM was reached. To ensure that the H_2O_2-resistant phenotype was caused by a chromosomal mutation, selected cultures were sub-cultured for 10 passages and then streaked on plates to obtain single colonies. The distinctive single colonies were then inoculated in liquid culture and actual MIC of H_2O_2 was determined.

REFERENCES

Agarwal, N., Woolwine, S. C., Tyagi, S., and Bishai, W. R. (2007). Characterization of the *Mycobacterium tuberculosis* Sigma Factor SigM by Assessment of Virulence and Identification of SigM-dependent Genes. *Infect. Immun.* 75, 452–461. doi:10.1128/IAI.01395-06

Aldridge, B. B., Fernandez-Suarez, M., Heller, D., Ambravaneswaran, V., Irimia, D., Toner, M., et al. (2012). Asymmetry and Aging of Mycobacterial Cells lead to Variable Growth and Antibiotic Susceptibility. *Science* 335, 100–104. doi:10.1126/science.1216166

Anes, E., Kühnel, M. P., Bos, E., Moniz-Pereira, J., Habermann, A., and Griffiths, G. (2003). Selected Lipids Activate Phagosome Actin Assembly and Maturation Resulting in Killing of Pathogenic Mycobacteria. *Nat. Cel Biol* 5, 793–802. doi:10.1038/ncb1036

Anes, E., Peyron, P., Staali, L., Jordao, L., Gutierrez, M. G., Kress, H., et al. (2006). Dynamic Life and Death Interactions between Mycobacterium Smegmatis and J774 Macrophages. *Cell Microbiol* 8, 939–960. doi:10.1111/j.1462-5822.2005.00675.x

Armstrong, J. A., and Hart, P. D. A. (1971). Response of Cultured Macrophages to *Mycobacterium tuberculosis*, with Observations on Fusion of Lysosomes with Phagosomes. *J. Exp. Med.* 134, 713–740. doi:10.1084/jem.134.3.713

Arora, G., Bothra, A., Prosser, G., Arora, K., and Sajid, A. (2021). Role of post-translational Modifications in the Acquisition of Drug Resistance in *Mycobacterium tuberculosis*. *FEBS J.* 288, 3375–3393. doi:10.1111/febs.15582

Bardarov, S., Bardarov, S., Pavelka, M. S., Sambandamurthy, V., Larsen, M., Tufariello, J., et al. (2002). Specialized Transduction: an Efficient Method for Generating Marked and Unmarked Targeted Gene Disruptions in *Mycobacterium tuberculosis*, M. Bovis BCG and M. Smegmatis. *Microbiology (Reading)* 148, 3007–3017. doi:10.1099/00221287-148-10-3007

Bhattacharya, M., Biswas, A., and Das, A. K. (2010). Interaction Analysis of TcrX/Y Two Component System from *Mycobacterium tuberculosis*. *Biochimie* 92, 263–272. doi:10.1016/j.biochi.2009.11.009

Boshoff, H. I. M., Myers, T. G., Copp, B. R., McNeil, M. R., Wilson, M. A., and Barry, C. E., 3rd (2004). The Transcriptional Responses of *Mycobacterium tuberculosis* to Inhibitors of Metabolism. *J. Biol. Chem.* 279, 40174–40184. doi:10.1074/jbc.M406796200

Broset, E., Martín, C., and Gonzalo-Asensio, J. (2015). Evolutionary Landscape of the *Mycobacterium tuberculosis* Complex from the Viewpoint of PhoPR: Implications for Virulence Regulation and Application to Vaccine Development. *mBio* 6, e01289–01215. doi:10.1128/mBio.01289-15

Brown, C. A., Draper, P., and Hart, P. D. A. (1969). Mycobacteria and Lysosomes: a Paradox. *Nature* 221, 658–660. doi:10.1038/221658a0

Calamita, H., Ko, C., Tyagi, S., Yoshimatsu, T., Morrison, N. E., and Bishai, W. R. (2005). The *Mycobacterium tuberculosis* SigD Sigma Factor Controls the Expression of Ribosome-Associated Gene Products in Stationary Phase and Is Required for Full Virulence. *Cel Microbiol* 7, 233–244. doi:10.1111/j.1462-5822.2004.00454.x

Chan, J., Xing, Y., Magliozzo, R. S., and Bloom, B. R. (1992). Killing of Virulent *Mycobacterium tuberculosis* by Reactive Nitrogen Intermediates Produced by Activated Murine Macrophages. *J. Exp. Med.* 175, 1111–1122. doi:10.1084/jem.175.4.1111

Cohen, S. B., Gern, B. H., Delahaye, J. L., Adams, K. N., Plumlee, C. R., Winkler, J. K., et al. (2018). Alveolar Macrophages Provide an Early *Mycobacterium tuberculosis* Niche and Initiate Dissemination. *Cell Host & Microbe* 24, 439–446. doi:10.1016/j.chom.2018.08.001

Dragset, M. S., Ioerger, T. R., Zhang, Y. J., Mærk, M., Ginbot, Z., Sacchettini, J. C., et al. (2019). Genome-wide Phenotypic Profiling Identifies and Categorizes Genes Required for Mycobacterial Low Iron Fitness. *Sci. Rep.* 9, 11394. doi:10.1038/s41598-019-47905-y

Elena, S. F., and Lenski, R. E. (2003). Evolution Experiments with Microorganisms: The Dynamics and Genetic Bases of Adaptation. *Nat. Rev. Genet.* 4, 457–469. doi:10.1038/nrg1088

Flynn, J. L., and Chan, J. (2001). Immunology of Tuberculosis. *Annu. Rev. Immunol.* 19, 93–129. doi:10.1146/annurev.immunol.19.1.93

Fol, M., Chauhan, A., Nair, N. K., Maloney, E., Moomey, M., Jagannath, C., et al. (2006). Modulation of *Mycobacterium tuberculosis* Proliferation by MtrA, an Essential Two-Component Response Regulator. *Mol. Microbiol.* 60, 643–657. doi:10.1111/j.1365-2958.2006.05137.x

Ford, C. B., Lin, P. L., Chase, M. R., Shah, R. R., Iartchouk, O., Galagan, J., et al. (2011). Use of Whole Genome Sequencing to Estimate the Mutation Rate of *Mycobacterium tuberculosis* during Latent Infection. *Nat. Genet.* 43, 482, 486. doi:10.1038/ng.811

Franzblau, S. G., Witzig, R. S., McLaughlin, J. C., Torres, P., Madico, G., Hernandez, A., et al. (1998). Rapid, Low-Technology MIC Determination with Clinical *Mycobacterium tuberculosis* Isolates by Using the Microplate Alamar Blue Assay. *J. Clin. Microbiol.* 36, 362–366. doi:10.1128/Jcm.36.2.362-366.1998

Gagneux, S. (2018). Ecology and Evolution of *Mycobacterium tuberculosis*. *Nat. Rev. Microbiol.* 16, 202–213. doi:10.1038/nrmicro.2018.8

Gomez, M., Doukhan, L., Nair, G., and Smith, I. (1998). sigAis an Essential Gene inMycobacterium Smegmatis. *Mol. Microbiol.* 29, 617–628. doi:10.1046/j.1365-2958.1998.00960.x

Gonzalo-Asensio, J., Malaga, W., Pawlik, A., Astarie-Dequeker, C., Passemar, C., Moreau, F., et al. (2014). Evolutionary History of Tuberculosis Shaped by Conserved Mutations in the PhoPR Virulence Regulator. *Proc. Natl. Acad. Sci.* 111, 11491–11496. doi:10.1073/pnas.1406693111

Gray, T. A., Clark, R. R., Boucher, N., Lapierre, P., Smith, C., and Derbyshire, K. M. (2016). Intercellular Communication and Conjugation Are Mediated by ESX Secretion Systems in Mycobacteria. *Science* 354, 347–350. doi:10.1126/science.aag0828

Gutierrez, M. C., Brisse, S., Brosch, R., Fabre, M., Omaïs, B., Marmiesse, M., et al. (2005). Ancient Origin and Gene Mosaicism of the Progenitor of *Mycobacterium tuberculosis*. *Plos Pathog.* 1, e5. doi:10.1371/journal.ppat.0010005

Hahn, M.-Y., Raman, S., Anaya, M., and Husson, R. N. (2005). The *Mycobacterium tuberculosis* Extracytoplasmic-Function Sigma Factor SigL Regulates Polyketide Synthases and Secreted or Membrane Proteins and Is Required for Virulence. *J. Bacteriol.* 187, 7062–7071. doi:10.1128/JB.187.20.7062-7071.2005

Haydel, S. E., Benjamin, W. H., Dunlap, N. E., and Clark-Curtiss, J. E. (2002). Expression, Autoregulation, and DNA Binding Properties of the *Mycobacterium tuberculosis* TrcR Response Regulator. *J. Bacteriol.* 184, 2192–2203. doi:10.1128/Jb.184.8.2192-2203.2002

He, H., and Zahrt, T. C. (2005). Identification and Characterization of a Regulatory Sequence Recognized by *Mycobacterium tuberculosis* Persistence Regulator MprA. *J. Bacteriol.* 187, 202–212. doi:10.1128/JB.187.1.202-212.2005

Homerova, D., Halgasova, L., and Kormanec, J. (2008). Cascade of Extracytoplasmic Function Sigma Factors inMycobacterium Tuberculosis: Identification of a Ïf J-dependent Promoter Upstream ofsigI. *FEMS Microbiol. Lett.* 280, 120–126. doi:10.1111/j.1574-6968.2007.01054.x

Hu, X., Zhou, X., Yin, T., Chen, K., Hu, Y., Zhu, B., et al. (2021). The Mycobacterial DNA Methyltransferase HsdM Decreases Intrinsic Isoniazid Susceptibility. *Antibiotics* 10 (11), 1323. doi:10.3390/antibiotics10111323

Hu, X., Li, X., Huang, L., Chan, J., Chen, Y., Deng, H., et al. (2015). Quantitative Proteomics Reveals Novel Insights into Isoniazid Susceptibility in Mycobacteria Mediated by a Universal Stress Protein. *J. Proteome Res.* 14, 1445–1454. doi:10.1021/pr5011058

James, J. N., Hasan, Z.-u., Ioerger, T. R., Brown, A. C., Personne, Y., Carroll, P., et al. (2012). Deletion of SenX3-RegX3, a Key Two-Component Regulatory System of Mycobacterium Smegmatis, Results in Growth Defects under Phosphate-Limiting Conditions. *Microbiology (Reading)* 158, 2724–2731. doi:10.1099/mic.0.060319-0

Jaswal, S., Dhand, R., Sethi, A. K., Kohli, K. K., and Ganguly, N. K. (1992). Oxidative Metabolic Status of Blood Monocytes and Alveolar Macrophages in the Spectrum of Human Pulmonary Tuberculosis. *Scand. J. Clin. Lab. Invest.* 52, 119–128. doi:10.3109/00365519209088775

Judd, J. A., Canestrari, J., Clark, R., Joseph, A., Lapierre, P., Lasek-Nesselquist, E., et al. (2021). A Mycobacterial Systems Resource for the Research Community. *mBio* 12. doi:10.1128/mBio.02401-20

Kieser, K. J., Boutte, C. C., Kester, J. C., Baer, C. E., Barczak, A. K., Meniche, X., et al. (2015). Phosphorylation of the Peptidoglycan Synthase PonA1 Governs the Rate of Polar Elongation in Mycobacteria. *Plos Pathog.* 11, e1005010. doi:10.1371/journal.ppat.1005010

Kim, B.-J., Gong, J.-R., Kim, G.-N., Kim, B.-R., Lee, S.-Y., Kook, Y.-H., et al. (2017). Recombinant Mycobacterium Smegmatis with a pMyong2 Vector Expressing Human Immunodeficiency Virus Type I Gag Can Induce Enhanced Virus-specific Immune Responses. *Sci. Rep.* 7, 44776. doi:10.1038/srep44776

Kumar, V., Jindal, S. K., and Ganguly, N. K. (1995). Release of Reactive Oxygen and Nitrogen Intermediates from Monocytes of Patients with Pulmonary Tuberculosis. *Scand. J. Clin. Lab. Invest.* 55, 163–169. doi:10.3109/00365519509089609

Lee, J.-H., Geiman, D. E., and Bishai, W. R. (2008a). Role of Stress Response Sigma Factor SigG in *Mycobacterium tuberculosis*. *J. Bacteriol.* 190, 1128–1133. doi:10.1128/JB.00511-07

Lee, J.-H., Karakousis, P. C., and Bishai, W. R. (2008b). Roles of SigB and SigF in the *Mycobacterium tuberculosis* Sigma Factor Network. *J. Bacteriol.* 190, 699–707. doi:10.1128/JB.01273-07

Lenski, R. E. (2017). Experimental Evolution and the Dynamics of Adaptation and Genome Evolution in Microbial Populations. *Isme J.* 11, 2181–2194. doi:10.1038/ismej.2017.69

Li, F., Xiong, X.-S., Yang, Y.-Y., Wang, J.-J., Wang, M.-M., Tang, J.-W., et al. (2021). Effects of NaCl Concentrations on Growth Patterns, Phenotypes Associated with Virulence, and Energy Metabolism in *Escherichia coli* BW25113. *Front. Microbiol.* 12, 705326. doi:10.3389/fmicb.2021.705326

Li, X., Liu, F., Hu, Y., and Mi, K. (2014a). Draft Genome Sequence of Mc 2 51, a Highly Hydrogen Peroxide-Resistant Mycobacterium Smegmatis Mutant Strain. *Genome Announc* 2. doi:10.1128/genomeA.00092-14

Li, X., Tao, J., Han, J., Hu, X., Chen, Y., Deng, H., et al. (2014b). The Gain of Hydrogen Peroxide Resistance Benefits Growth Fitness in Mycobacteria under Stress. *Protein Cell* 5, 182–185. doi:10.1007/s13238-014-0024-5

Li, X., Tao, J., Hu, X., Chan, J., Xiao, J., and Mi, K. (2014c). A Bacterial Hemerythrin-like Protein MsmHr Inhibits the SigF-dependent Hydrogen Peroxide Response in Mycobacteria. *Front. Microbiol.* 5, 800. doi:10.3389/fmicb.2014.00800

Li, Y., Zeng, J., Zhang, H., and He, Z.-G. (2010). The Characterization of Conserved Binding Motifs and Potential Target Genes for *M. tuberculosis* MtrAB Reveals a Link between the Two-Component System and the Drug Resistance of M. Smegmatis. *BMC Microbiol.* 10, 242. doi:10.1186/1471-2180-10-242

Loebel, R. O., Shorr, E., and Richardson, H. B. (1933). The Influence of Adverse Conditions upon the Respiratory Metabolism and Growth of Human Tubercle Bacilli. *J. Bacteriol.* 26, 167–200. doi:10.1128/Jb.26.2.167-200.1933

Mehra, S., Foreman, T. W., Didier, P. J., Ahsan, M. H., Hudock, T. A., Kissee, R., et al. (2015). The DosR Regulon Modulates Adaptive Immunity and Is Essential forMycobacterium tuberculosisPersistence. *Am. J. Respir. Crit. Care Med.* 191, 1185–1196. doi:10.1164/rccm.201408-1502OC

Merrill, M. H. (1930). Carbohydrate Metabolism of Organisms of the Genus Mycobacterium. *J. Bacteriol.* 20, 235–286. doi:10.1128/jb.20.4.235-286.1930

Mukherjee, R., and Chatterji, D. (2005). Evaluation of the Role of Sigma B in Mycobacterium Smegmatis. *Biochem. Biophysical Res. Commun.* 338, 964–972. doi:10.1016/j.bbrc.2005.10.038

Nauseef, W. M. (2019). The Phagocyte NOX2 NADPH Oxidase in Microbial Killing and Cell Signaling. *Curr. Opin. Immunol.* 60, 130–140. doi:10.1016/j.coi.2019.05.006

Ng, V. H., Cox, J. S., Sousa, A. O., MacMicking, J. D., and McKinney, J. D. (2004). Role of KatG Catalase-Peroxidase in Mycobacterial Pathogenesis: Countering the Phagocyte Oxidative Burst. *Mol. Microbiol.* 52, 1291–1302. doi:10.1111/j.1365-2958.2004.04078.x

Nowak, E., Panjikar, S., Konarev, P., Svergun, D. I., and Tucker, P. A. (2006). The Structural Basis of Signal Transduction for the Response Regulator PrrA from *Mycobacterium tuberculosis*. *J. Biol. Chem.* 281, 9659–9666. doi:10.1074/jbc.M512004200

Parish, T., Smith, D. A., Kendall, S., Casali, N., Bancroft, G. J., and Stoker, N. G. (2003a). Deletion of Two-Component Regulatory Systems Increases the Virulence of *Mycobacterium tuberculosis*. *Infect. Immun.* 71, 1134–1140. doi:10.1128/IAI.71.3.1134-1140.2003

Parish, T., Smith, D. A., Roberts, G., Betts, J., and Stoker, N. G. (2003b). The senX3-regX3 Two-Component Regulatory System of *Mycobacterium tuberculosis* Is Required for Virulence. *Microbiology (Reading)* 149, 1423–1435. doi:10.1099/mic.0.26245-0

Piccaro, G., Pietraforte, D., Giannoni, F., Mustazzolu, A., and Fattorini, L. (2014). Rifampin Induces Hydroxyl Radical Formation in *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.* 58, 7527–7533. doi:10.1128/AAC.03169-14

Pieters, J. (2008). *Mycobacterium tuberculosis* and the Macrophage: Maintaining a Balance. *Cell Host & Microbe* 3, 399–407. doi:10.1016/j.chom.2008.05.006

Plocinska, R., Purushotham, G., Sarva, K., Vadrevu, I. S., Pandeeti, E. V. P., Arora, N., et al. (2012). Septal Localization of the *Mycobacterium tuberculosis* MtrB Sensor Kinase Promotes MtrA Regulon Expression. *J. Biol. Chem.* 287, 23887–23899. doi:10.1074/jbc.M112.346544

Podinovskaia, M., Lee, W., Caldwell, S., and Russell, D. G. (2013). Infection of Macrophages withMycobacterium Tuberculosisinduces Global Modifications to Phagosomal Function. *Cel Microbiol* 15, 843–859. doi:10.1111/cmi.12092

Pym, A. S., Domenech, P., Honore, N., Song, J., Deretic, V., and Cole, S. T. (2001). Regulation of Catalase-Peroxidase (KatG) Expression, Isoniazid Sensitivity and Virulence by furA of *Mycobacterium tuberculosis*. *Mol. Microbiol.* 40, 879–889. doi:10.1046/j.1365-2958.2001.02427.x

Raman, S., Hazra, R., Dascher, C. C., and Husson, R. N. (2004). Transcription Regulation by the *Mycobacterium tuberculosis* Alternative Sigma Factor SigD and its Role in Virulence. *J. Bacteriol.* 186, 6605–6616. doi:10.1128/Jb.186.19.6605-6616.2004

Raman, S., Puyang, X., Cheng, T.-Y., Young, D. C., Moody, D. B., and Husson, R. N. (2006). *Mycobacterium tuberculosis* SigM Positively Regulates Esx Secreted Protein and Nonribosomal Peptide Synthetase Genes and Down Regulates Virulence-Associated Surface Lipid Synthesis. *J. Bacteriol.* 188, 8460–8468. doi:10.1128/JB.01212-06

Rifat, D., and Karakousis, P. C. (2014). Differential Regulation of the Two-Component Regulatory System senX3-regX3 in *Mycobacterium tuberculosis*. *Microbiology (Reading)* 160, 1125–1133. doi:10.1099/mic.0.077180-0

Rodrigue, S., Brodeur, J., Jacques, P.-E., Gervais, A. L., Brzezinski, R., and Gaudreau, L. (2007). Identification of Mycobacterial σ Factor Binding Sites by Chromatin Immunoprecipitation Assays. *J. Bacteriol.* 189, 1505–1513. doi:10.1128/JB.01371-06

Sambandamurthy, V. K., Wang, X., Chen, B., Russell, R. G., Derrick, S., Collins, F. M., et al. (2002). A Pantothenate Auxotroph of *Mycobacterium tuberculosis* Is Highly Attenuated and Protects Mice against Tuberculosis. *Nat. Med.* 8, 1171–1174. doi:10.1038/nm765

Schnell, R., Ågren, D., and Schneider, G. (2008). 1.9 Å Structure of the Signal Receiver Domain of the Putative Response Regulator NarL fromMycobacterium Tuberculosis. *Acta Cryst. Sect F* 64, 1096–1100. doi:10.1107/S1744309108035203

Shrivastava, R., and Das, A. K. (2007). Temperature and Urea Induced Conformational Changes of the Histidine Kinases from *Mycobacterium tuberculosis*. *Int. J. Biol. Macromolecules* 41, 154–161. doi:10.1016/j.ijbiomac.2007.01.010

Snapper, S. B., Melton, R. E., Mustafa, S., Kieser, T., and Jr, W. R. J., Jr. (1990). Isolation and Characterization of Efficient Plasmid Transformation Mutants of Mycobacterium Smegmatis. *Mol. Microbiol.* 4, 1911–1919. doi:10.1111/j.1365-2958.1990.tb02040.x

Song, T., Song, S.-E., Raman, S., Anaya, M., and Husson, R. N. (2008). Critical Role of a Single Position in the −35 Element for Promoter Recognition by *Mycobacterium tuberculosis* SigE and SigH. *J. Bacteriol.* 190, 2227–2230. doi:10.1128/Jb.01642-07

Steyn, A. J. C., Joseph, J., and Bloom, B. R. (2003). Interaction of the Sensor Module of *Mycobacterium tuberculosis* H37Rv KdpD with Members of the Lpr Family. *Mol. Microbiol.* 47, 1075–1089. doi:10.1046/j.1365-2958.2003.03356.x

Sun, R., Converse, P. J., Ko, C., Tyagi, S., Morrison, N. E., and Bishai, W. R. (2004). *Mycobacterium tuberculosis* ECF Sigma Factor sigC Is Required for Lethality in Mice and for the Conditional Expression of a Defined Gene Set. *Mol. Microbiol.* 52, 25–38. doi:10.1111/j.1365-2958.2003.03958.x

Sundberg, L.-R., Kunttu, H. M. T., and Valtonen, E. (2014). Starvation Can Diversify the Population Structure and Virulence Strategies of an Environmentally Transmitting Fish Pathogen. *BMC Microbiol.* 14, 67. doi:10.1186/1471-2180-14-67

Sureka, K., Dey, S., Datta, P., Singh, A. K., Dasgupta, A., Rodrigue, S., et al. (2007). Polyphosphate Kinase Is Involved in Stress-Induced mprAB-sigE-Rel Signalling in Mycobacteria. *Mol. Microbiol.* 65, 261–276. doi:10.1111/j.1365-2958.2007.05814.x

Sweeney, K. A., Dao, D. N., Goldberg, M. F., Hsu, T., Venkataswamy, M. M., Henao-Tamayo, M., et al. (2011). A Recombinant Mycobacterium Smegmatis Induces Potent Bactericidal Immunity against *Mycobacterium tuberculosis*. *Nat. Med.* 17, 1261–1268. doi:10.1038/nm.2420

Timmins, G. S., and Deretic, V. (2006). Mechanisms of Action of Isoniazid. *Mol. Microbiol.* 62, 1220–1227. doi:10.1111/j.1365-2958.2006.05467.x

van Kessel, J. C., Marinelli, L. J., and Hatfull, G. F. (2008). Recombineering Mycobacteria and Their Phages. *Nat. Rev. Microbiol.* 6, 851–857. doi:10.1038/nrmicro2014

Veyrier, F., Saïd-Salim, B., and Behr, M. A. (2008). Evolution of the Mycobacterial SigK Regulon. *J. Bacteriol.* 190, 1891–1899. doi:10.1128/JB.01452-07

Walters, S. B., Dubnau, E., Kolesnikova, I., Laval, F., Daffe, M., and Smith, I. (2006). The *Mycobacterium tuberculosis* PhoPR Two-Component System Regulates Genes Essential for Virulence and Complex Lipid Biosynthesis. *Mol. Microbiol.* 60, 312–330. doi:10.1111/j.1365-2958.2006.05102.x

Wang, L., Liu, Z., Dai, S., Yan, J., and Wise, M. J. (2017). The Sit-And-Wait Hypothesis in Bacterial Pathogens: A Theoretical Study of Durability and Virulence. *Front. Microbiol.* 8, 2167. doi:10.3389/fmicb.2017.02167

Wayne, L. G., and Hayes, L. G. (1996). An *In Vitro* Model for Sequential Study of Shiftdown of *Mycobacterium tuberculosis* through Two Stages of Nonreplicating Persistence. *Infect. Immun.* 64, 2062–2069. doi:10.1128/Iai.64.6.2062-2069.1996

Who (2021). *Global Tuberculosis Report 2021*. Geneva, Switzland: WHO.

Wolfe, N. D., Dunavan, C. P., and Diamond, J. (2007). Origins of Major Human Infectious Diseases. *Nature* 447, 279–283. doi:10.1038/nature05775

Zhang, Z., and Furman, A. (2021). Soil Redox Dynamics under Dynamic Hydrologic Regimes - A Review. *Sci. Total Environ.* 763, 143026. doi:10.1016/j.scitotenv.2020.143026

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.

frontiers | Frontiers in Microbiology

ORIGINAL RESEARCH
published: 06 April 2022
doi: 10.3389/fmicb.2022.817494

The Role of the Two-Component QseBC Signaling System in Biofilm Formation and Virulence of Hypervirulent *Klebsiella pneumoniae* ATCC43816

Jingnan Lv[1†], Jie Zhu[1†], Ting Wang[2†], Xiaofang Xie[1], Tao Wang[1], Zhichen Zhu[1], Liang Chen[3,4], Fengyun Zhong[5] and Hong Du[1]**

[1] Department of Clinical Laboratory, The Second Affiliated Hospital of Soochow University, Suzhou, China, [2] Department of Clinical Laboratory, Suzhou Science and Technology Town Hospital, Suzhou, China, [3] Hackensack Meridian Health Center for Discovery and Innovation, Nutley, NJ, United States, [4] Department of Medical Sciences, Hackensack Meridian School of Medicine, Nutley, NJ, United States, [5] Department of General Surgery, The Second Affiliated Hospital of Soochow University, Suzhou, China

OPEN ACCESS

Edited by:
Wei Huang,
Johns Hopkins University,
United States

Reviewed by:
Yiquan Zhang,
Jiangnan University, China
Bijay Khajanchi,
United States Food and Drug Administration, United States

***Correspondence:**
Hong Du
hong_du@126.com
Fengyun Zhong
zfysz@126.com

† These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *18 November 2021*
Accepted: *26 January 2022*
Published: *06 April 2022*

Citation:
Lv J, Zhu J, Wang T, Xie X, Wang T, Zhu Z, Chen L, Zhong F and Du H (2022) The Role of the Two-Component QseBC Signaling System in Biofilm Formation and Virulence of Hypervirulent Klebsiella pneumoniae ATCC43816. Front. Microbiol. 13:817494. doi: 10.3389/fmicb.2022.817494

Hypervirulent *Klebsiella pneumoniae* (hvKP) is an evolving infectious pathogen associated with high mortality. The convergence of hypervirulence and multidrug resistance further challenges the clinical treatment options for *K. pneumoniae* infections. The QseBC two-component system (TCS) is a component of quorum-sensing regulatory cascade and functions as a global regulator of biofilm growth, bacterial motility, and virulence in *Escherichia coli*. However, the functional mechanisms of QseBC in hvKP have not been reported, and we aim to examine the role of QseBC in regulating virulence in hvKP strain ATCC43816. The CRISPR-Cas9 system was used to construct *qseB*, *qseC*, and *qseBC* knockout in ATCC43816. No significant alterations in the growth and antibiotic susceptibility were detected between wild-type and mutants. The deletion of *qseC* led to an increase of biofilm formation, resistance to serum killing, and high mortality in the *G. mellonella* model. RNAseq differential gene expression analysis exhibited that gene-associated biofilm formation (*glgC, glgP, glgA, gcvA, bcsA, ydaM, paaF, ptsG*), bacterial type VI secretion system (*virB4, virB6, virB10, vgrG, hcp*), and biosynthesis of siderophore (*entC, entD, entE*) were significantly upregulated in comparison with the wild-type control. In addition, *qseB*, *ygiW* (encode OB-family protein), and AraC family transcriptional regulator IT767_23090 genes showed highest expressions in the absence of QseC, which might be related to increased virulence. The study provided new insights into the functional importance of QseBC in regulating the virulence of hvKP.

Keywords: hypervirulent, *Klebsiella pneumoniae*, QseBC, two-component system, biofilm formation

INTRODUCTION

Hypervirulent *Klebsiella pneumoniae* (hvKP) is a gram-negative opportunistic pathogen that has become a worldwide concern due to increasing cases of life-threatening infections in healthy individuals (Harada et al., 2019; Russo and Marr, 2019; Zafar et al., 2019). HvKP is such an invasive strain in part because of its virulence factors that protect it from immunological responses to survive. These virulence factors include capsule, siderophore, fimbriae, biofilm, but the mechanisms of hypervirulence are not fully defined (Choby et al., 2020; Zhu et al., 2021).

Two-component systems (TCS) are dominant bacterial signal-transduction systems (Groisman, 2016). Quorum sensing is a ubiquitous chemical communication process that bacteria use to response to environmental cues (Abisado et al., 2018). The QseBC two-component system is associated with quorum sensing and functions as a regulator of virulence in a wide range of species (Weigel and Demuth, 2016). QseC, a membrane-bound sensor kinase, senses signals (epinephrine/norepinephrine/autoinducer-3) and phosphorylates the response regulator QseB to control the virulence of enterohemorrhagic *Escherichia coli* (EHEC) (Sperandio et al., 2002; Clarke et al., 2006). Activation of QseBC results in enhanced motility and upregulation of flagellar genes in *Salmonella Typhimurium* (Bearson and Bearson, 2008). The *qseC* mutant is attenuated for virulence in rabbits (EHEC) and piglets (*S. Typhimurium*) (Merighi et al., 2009). Likewise, this function of QseBC had been reported in rare species, such as *Edwardsiella tarda*, *Aggregatibacter actinomycetemcomitans*, *Aeromonas hydrophila*. Disruption of *qseB* and *qseC* exhibited impaired flagellar motilities as well as *in vivo* competitive abilities of *Edwardsiella tarda* (Wang et al., 2011). In the oral pathogen *Aggregatibacter actinomycetemcomitans*, activation of QseBC by autoinducer-2 (AI-2) resulted in increased biofilm growth (Novak et al., 2010). Moreover, Khajanchi et al. (2012) found that the decrease in virulence of *Aeromonas hydrophila*Δ*qseB* were associated with reduced production of protease and cytotoxic enterotoxin. However, the regulon and function of QseBC in hvKP is currently unknown. In order to explore the role of QseBC in, *qseB* and *qseC* single and double knock-out mutants of the hvKP strain ATCC43816 were constructed and studied.

MATERIALS AND METHODS

Bacterial Strains, Plasmids, Primers and Growth Conditions

Strains, plasmids and primers used in this study are listed in **Supplementary Table 1**. The ST439 K2 hypervirulent *K. pneumoniae* ATCC43816 was obtained from the American Type Culture Collection (ATCC). *E. coli* DH5α and *K. pneumoniae* strains were grown aerobically in lysogeny broth (LB) medium (1% tryptone, 0.5% yeast, 1%NaCl). Strains with temperature-sensitive plasmid pCasKP and L-arabinose-inducible expression plasmid pBAD24 were grown at 30°C and in the presence of 0.2% arabinose, respectively. Selective antibiotics were added at the following concentrations: rifampicin 100 μg/ml, apramycin 30 μg/ml, and ampicillin 100 μg/ml.

Construction of the *qseB, qseC, qseBC* Mutants

The hvKP ATCC43816 *qseB*, *qseC*, and *qseBC* mutants were constructed using the CRISPR/Cas9 mediated genome-editing system (Wang et al., 2018). The plasmid pCasKP was electroporated into ATCC43816, followed by selection on agar containing 30 μg/ml apramycin and PCR confirmation. The 20-nt base-pairing region (N20) of an sgRNA was designed through the online web server.[1] Primers HindIII-QseB/QseC-N20-For and XbaI-gRNA-Rev were used to amplify the sgRNA fragment from plasmid pSGKP-Rif. The PCR product was subsequently ligated into HindIII and XbaI digested pSGKP-Rif to construct the final pSGKP-Rif with targeted sgRNA (pSGKP-QseB/QseC-N20). We utilized the linear homologous DNA fragment as the repair templates. The linear double-stranded DNA (dsDNA) and the pSGKP-QseB/QseC-N20 plasmid were co-transformed into the L-arabinose-induced pCasKP-positive ATCC43816 to generate the *qseB*, *qseC*, and *qseBC* deletions. The cultures were plated on LB agar plates containing 30 μg/ml apramycin and 100 μg/ml rifampicin, and the mutants were confirmed in selected colonies by both PCR and sequencing.

Complementation of the *qseC* Mutant

The *qseC* gene was amplified from the chromosomal DNA of ATCC43816 using primers QseC-C-For and QseC-C-Rev. The PCR products were then assembled into the multiple cloning site (MCS) of vector pBAD24 using the ClonExpress® Ultra One Step Cloning Kit (Vazyme Biotech Co., Ltd., China). The recombinant plasmid pBAD24-*qseC* was transformed into *E. coli* DH5α, and then primers pBAD24-For and pBAD24-Rev were used to confirm positive colonies on LB agar with 100 μg/ml ampicillin. The *qseC* mutant was transformed with the pBAD24-*qseC* plasmid and named as Δ*qseC*-pC*qseC*. ATCC43816 and Δ*qseC* carrying the empty vector pBAD24 were used as control.

Bacterial Growth Curves

Growth curves of ATCC43816, Δ*qseB*, Δ*qseC*, Δ*qseBC*, ATCC43816-pBAD24, Δ*qseC*-pBAD24, and Δ*qseC*-pC*qseC* were determined by subculturing in LB medium and growth. Briefly, overnight cultures of ATCC43816 and mutants were diluted 1:100 in 20 ml of fresh LB broth and grown at 37°C with shaking at 250 rpm. The cell density was detected per hour by OD_{600} measurements. The experiment was conducted with three independent cultures.

Antimicrobial Susceptibility Testing

Antimicrobial susceptibility testing was performed using a BD Phoenix™ 100 Automated Microbiology System. Results were interpreted according to the Clinical and Laboratory Standards Institute (CLSI M100, 30th edition) breakpoints.

[1] http://crispr.dfci.harvard.edu/SSC/

Biofilm Formation Assay

Biofilm formation was estimated by crystal violet (CV) assay as previously described (Khajanchi et al., 2012; Fu et al., 2018). ATCC43816, Δ*qseB*, Δ*qseC*, Δ*qseBC*, ATCC43816-pBAD24, Δ*qseC*-pBAD24, and Δ*qseC*-pC*qseC* were grown overnight in LB broth and then sub-cultured 1:100 into fresh LB, and grown for 6 h; 20 μl of culture was transferred to a 96-cell polystyrene microtiter plate containing 180 μl brain heart infusion (BHI). Afterward, the plate was incubated at 37°C for 36 h. The medium was then decanted and planktonic cells were washed off gently with sterile water. Adhered biofilms were stained for 30 min at room temperature with 250 μl 1% CV (Beyotime, China), and the purple area was dissolved with 33% glacial acetic acid. Finally, biofilm quantification was measured by using microtiter plate reader (TECAN Infinite M200 Pro NanoQuant) at a wavelength of 595 nm. Absorbance data were obtained from three replicate experiments. In addition, we performed the above procedure by using plastic centrifuge tubes to visualize results of biofilm formation.

Serum Killing Assay

Overnight cultures of ATCC43816, Δ*qseB*, Δ*qseC*, Δ*qseBC*, ATCC43816-pBAD24, Δ*qseC*-pBAD24, and Δ*qseC*-pC*qseC* grown in LB were sub-cultured 1:100 and grown to late exponential phase; 250 μl bacteria (10^5 CFU/ml) were incubated with 750 μl serum for 2 h. Samples were serially diluted and plated onto LB agar, and counted colonies were incubated overnight.

Galleria mellonella Killing Assay

G. mellonella larvae were purchased from Tianjin Huiyude Biotech Company, Tianjin, China. Each group included 40 larvae, and bacterial inoculum (10^5 CFU/larva) was administered as 10 μl injection into the left proleg. Saline inoculated larvae were included as negative control. Larvae were incubated at 37°C, in the dark, and scored every 24 h for survival or death (no response to stimulation recorded as dead). The killing assays were performed in triplicate using three different batches of *G. mellonella* larvae.

RNA Sequencing and Differential Expression Analysis

The RNA-sequencing was performed by Shanghai Majorbio Bio-pharm Technology Co., Ltd. Total RNA was extracted from the late-exponentially growing ATCC43816, Δ*qseB*, Δ*qseC*, and Δ*qseBC* using TRIzol® Reagent (Invitrogen) followed by genomic DNA removal using DNase (Takara). RNAseq libraries were prepared according TruSeq™ RNA sample preparation Kit from Illumina (San Diego, CA, United States). The ribosomal RNA (rRNA) was removed using Ribo-Zero Magnetic Kit (epicenter). Paired-end RNA sequencing was performed with the Illumina HiSeq × TEN (2 × 150 bp read length).

The data were analyzed with the free online platform of Majorbio Cloud Platform.[2] Raw sequencing reads were trimmed and filtered for high quality reads before being mapped to ATCC43816 genomes (Accession: PRJNA675363) using Bowtie2.[3] RSEM[4] was used to quantify expression levels of genes and transcripts. Differentially expressed genes (DEGs) were identified using DESeq2 packages[5]. The differential expression was considered as statistically significant if $p < 0.05$ and $|\log2(\text{fold change})| > 2$. KEGG and GO enrichment was conducted using KOBAS 2.0[6] and Goatools.[7]

Total RNA Isolation, cDNA Generation and Real-Time PCR Processing

Overnight culture of ATCC43816, Δ*qseB*, Δ*qseC*, Δ*qseBC*, ATCC43816-pBAD24, Δ*qseC*-pBAD24, and Δ*qseC*-pC*qseC* were diluted 1:100 in LB media and grown for 6 h to late exponential phase. Cells were collected and treated with Trizol (Sigma) to extract total RNA, followed by DNase treatment as above. The transcript levels of genes were tested by quantitative reverse-transcription PCR (qRT-PCR). HiScript III RT SuperMix kit and Taq Pro Universal SYBR qPCR Master Mix kit (Vazyme Biotech Co., Ltd., China) was used according to the manufacturer's instructions. Differences in gene expression were normalized with the expression of 16S rRNA and calculated by the $2^{-\Delta\Delta Ct}$ method. All of the qRT-PCR assays were repeated at least three times.

Statistical Analyses

The Student *t*-test was used to detect the differences between groups in the experiments. Survival curves were assessed by log-rank rest. $p < 0.05$ was considered statistically significant.

Data Availability

The raw sequencing files in this study are available at the NCBI Sequence Read Archive (SRA) under BioProject PRJNA784457.

RESULTS

Deletion of *qseB, qseC,* and *qseBC* Did Not Affect the Growth and Antimicrobial Susceptibility

To investigate the roles of QseBC in hvKP, Δ*qseB*, Δ*qseC*, and Δ*qseBC* were constructed by CRISPR/Cas9-mediated genome editing. The correct mutants were confirmed by PCR and Sanger sequencing. The colony morphologies of the three mutants on LB agar plates were similar to that of the wild-type. The growth curves of three mutants in LB broth over a period of 24 h were also similar to that of the wild-type strain (**Figure 1**). Antimicrobial susceptibility testing showed that ATCC43816, Δ*qseB*, Δ*qseC*, and Δ*qseBC* were susceptible to all tested antibacterial drugs. There was no

[2]www.majorbio.com

[3]http://bowtie-bio.sourceforge.net/bowtie2/index.shtml

[4]http://deweylab.github.io/RSEM/

[5]https://genomebiology.biomedcentral.com/articles/10.1186/s13059-014-0550-8

[6]http://kobas.cbi.pku.edu.cn

[7]https://github.com/tanghaibao/GOatools

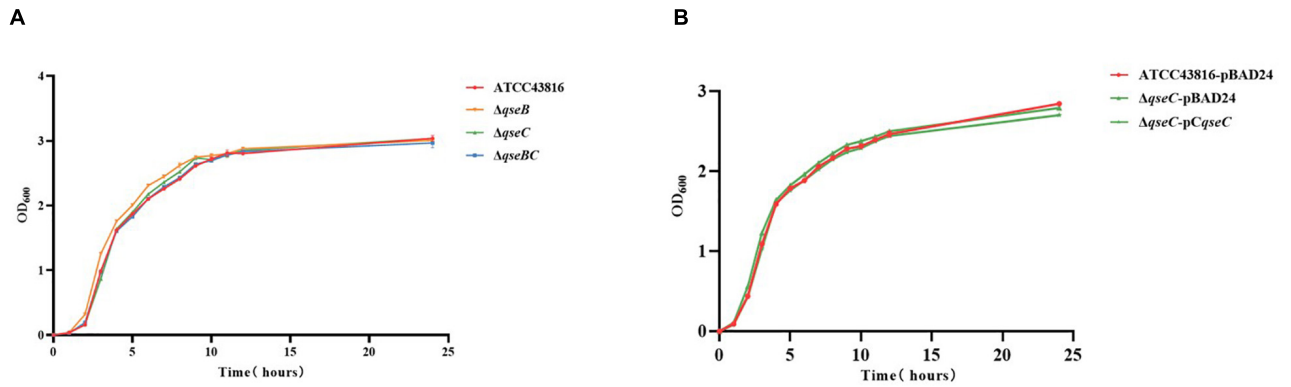

FIGURE 1 | (A) The growth rates of ATCC43816, Δ*qseB*, Δ*qseC* and Δ*qseBC* were similar by measuring the optical density (OD_{600}) at per hour over a period of 24 h. **(B)** Growth curves of ATCC43816-pBAD24, Δ*qseC*-pBAD24 and the complement strain Δ*qseC*-pCqseC.

significant difference of the minimum inhibitory concentration (MIC) values in ATCC43816, Δ*qseB*, Δ*qseC*, and Δ*qseBC* against 18 antibiotics examined (**Supplementary Table 2**). These results indicated that deletion of *qseB*, *qseC*, and *qseBC* did not affect growth and antimicrobial susceptibility of hvKP ATCC43816.

Deletion of *qseC* Increased Biofilm Formation

That QseBC regulated biofilm formation had been reported in *E. coli* (Sharma and Casey, 2014; Gou et al., 2019). Here, the CV staining assay was applied to examine whether QseBC affected biofilm formation in hvKP ATCC43816. As shown in **Figure 2**, Δ*qseC* formed a significantly increased, solid-surface-associated biofilm in the polystyrene plate, with about a twofold increase in the CV staining when compared with the biofilm formed by ATCC43816. Compensation of *qseC* partially decreased the biofilm formation (**Figures 3B,C**). Since growth was not altered after *qseC*-deletion (**Figure 1**), the biofilm changes were not a result of growth differences. However, Δ*qseB* and Δ*qseBC* produced a similar level of biofilms to that from the wild-type strain. These data suggested to us that QseC negatively controlled biofilm formation in hvKP ATCC43816.

Deletion of *qseC* Increased Serum Resistance of ATCC43816

Serum killing assay was used to study serum resistance in the wild-type and mutants *in vitro*. Viable counts of both wild-type strain and mutants were > 100% throughout the time period, suggesting that they all had a high level of defense against serum bactericidal activity (**Figure 3A**). A significant difference was identified between ATCC43816 and Δ*qseC*. After 2 h of incubation, the mean survival count of Δ*qseC* showed 3.02×10^5 CFU/ml and exhibited ∼3 times higher than that of ATCC43816 and complemented strain Δ*qseC*-pC*qseC* (**Figure 3B**), while serum-resistant levels of Δ*qseB* and Δ*qseBC* had no significant difference with that of ATCC43816 in serum killing assay.

Deletion of *qseC* Increased *in vivo* Virulence of ATCC43816

The *Galleria mellonella* infection model has been suggested to be a valid method for evaluating *K. pneumoniae* pathogenicity (Insua et al., 2013). To examine the influence of QseBC on *in vivo* virulence of hvKP ATCC43816, the *G. mellonella* killing assay was then performed. The *qseB* and *qseBC* mutants remained as viable as ATCC43816 in experimentally infected *G. mellonella*. After 72 h of infection, the survival rate of Δ*qseC* was about less than 20%. By contrast, the survival rate of ATCC43816, *qseB*, and *qseBC* mutants was about 40%, which was significantly higher than that of the Δ*qseC* (**Figure 4A**). As shown in **Figure 4B**, the survival rate was significantly increased in Δ*qseC*-pC*qseC* relative to Δ*qseC*. These data demonstrated that Δ*qseC* had relatively higher *in vivo* virulence in comparison to the wild-type strain.

Profiling Gene Expression of Δ*qseB*, Δ*qseC*, and Δ*qseBC*

To explore possible transcriptomic contributions of QseBC to virulence in hvKP ATCC43816, we examined the gene expression profiles of the late-exponentially growing ATCC43816, Δ*qseB*, Δ*qseC*, and Δ*qseBC*. We identified differentially expressed genes (DEGs) by comparing the RNA-sequencing data with the control strain ATCC43816. DEGs were identified by adjusted $p < 0.05$ and $|\log2(\text{fold change})| > 2$. Compared to the wild-type strain, the number of DEGs in Δ*qseB*, Δ*qseC*, and Δ*qseBC* groups was 58 (49 upregulated and 9 downregulated genes), 216 (186 upregulated and 30 downregulated genes), and 193 (171 upregulated and 22 downregulated genes), respectively (**Supplementary Tables 3–5**). In our analysis, we detected 12 DEGs that were shared by Δ*qseB* and Δ*qseC*, 13 DEGs shared by Δ*qseB* and Δ*qseBC*, 54 DEGs shared by Δ*qseC* and Δ*qseBC*, and 21 DEGs shared by all three mutants (**Figure 5A**). Notably, among the 129 DEGs unique to Δ*qseC*, IT767_23090 (AraC family transcriptional regulator family) and

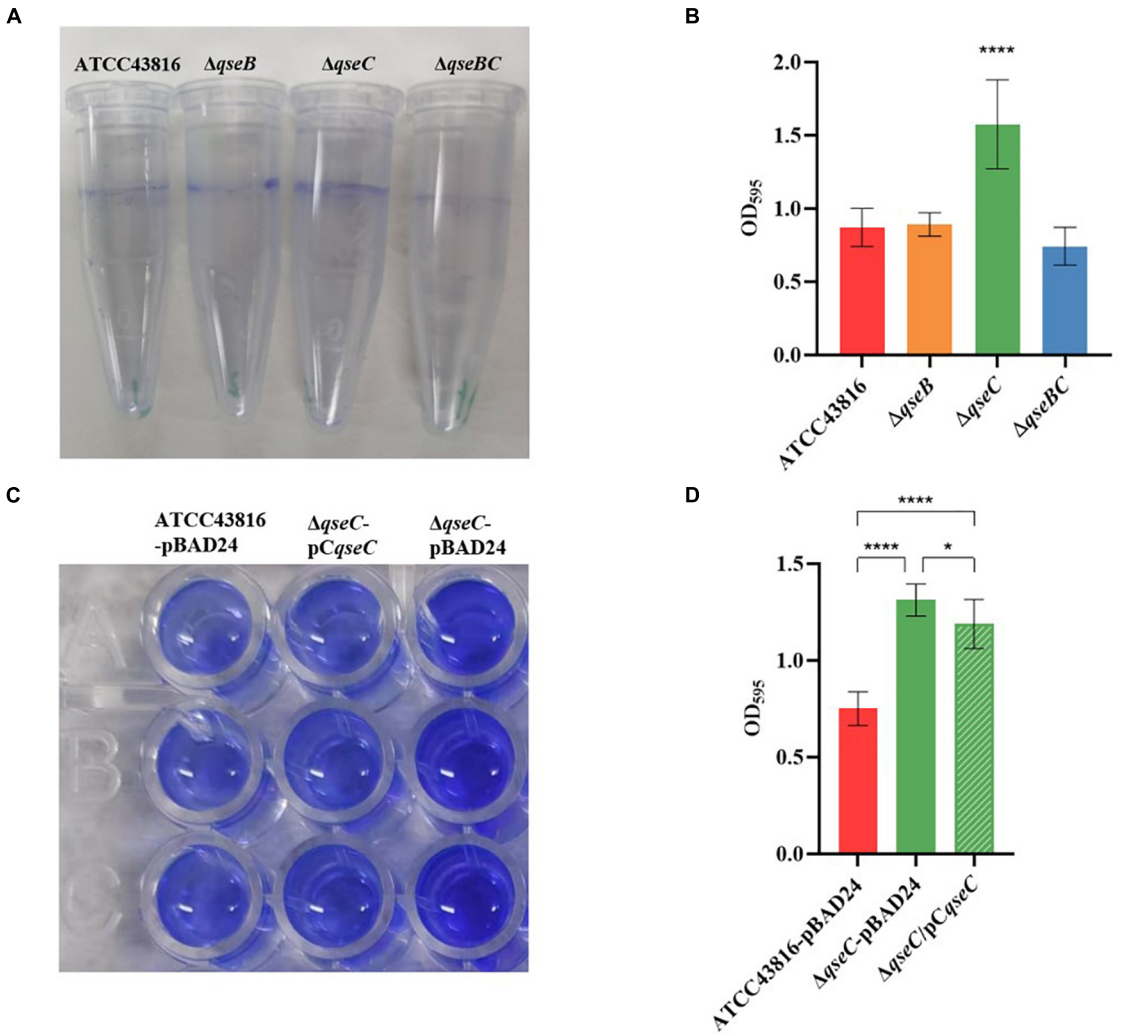

FIGURE 2 | (A) Detection of biofilm formed by ATCC43816, Δ*qseB*, Δ*qseC* and Δ*qseBC*. Image of biofilm stained with 1% CV and adhered on plastic centrifuge tubes. **(B)** Biofilm were stained with 1% CV and washed with sterile water. The extracted color was dissolved with 33% glacial acetic acid and measured at OD_{595}. Asterisks (****) represent statistically significant differences in biofilm formation between ATCC43816 and its *qseC*mutant ($P < 0.0001$, *t* test). **(C)** Image of biofilm formed by ATCC43816-pBAD24, Δ*qseC*-pBAD24 and the complement strain Δ*qseC*-pC*qseC*. Biofilm were stained with 1% CV, washed with sterile water and then dissolved with 33% glacial acetic acid. **(D)** The extracted color was measured at OD_{595} to show biofilm production. Asterisks (****) represent statistically significant differences in biofilm formation between ATCC43816-pBAD24 and Δ*qseC*-pBAD24/Δ*qseC*-pC*qseC* ($P < 0.0001$, *t* test). The biofilm formation of Δ*qseC*-pC*qseC* was significantly lower compared with that of Δ*qseC*-pBAD24 ($P = 0.0366$, *t* test).

ygiW showed the highest expression levels (up to 300-fold increase). Compared with ATCC43816, Δ*qseC* apparently had a significantly high-expression of the *qseB* gene. The gene expressions of selected genes were confirmed by qRT-PCR and the qRT-PCR showed similar results as the RNA-seq analysis (**Figures 6**, **7A**). The transcript levels of *qseB*, *ygiW*, and *araC* were restored in Δ*qseC*-pC*qseC* (**Figure 7B**).

The 216 DEGs of Δ*qseC* group were performed by Gene Ontology (GO) enrichment analysis. Twenty significantly enriched GO terms from three categories (biological process, cellular component, and molecular function) were shown in **Figure 5B**. Most DEGs were enriched in two GO terms of cellular component, integral component of membrane and plasma membrane. In addition, DGEs were also subject to KEGG pathway enrichment analysis. Cellular community, signal transduction, membrane transport, energy metabolism, carbohydrate metabolism, and amino acid metabolism were significantly enriched in KEGG pathways (**Figure 5C**). Among the cellular community pathway, genes associated biofilm formation (*glgC, glgP, glgA, gcvA, bcsA, ydaM, paaF, ptsG*) in Δ*qseC* were significantly up-regulated as comparison with the control, which is consistent with the increased biofilm production described above. Moreover, several up-regulated DEGs in pathway involved in virulence, including the bacterial secretion system (*virB4, virB6, virB10, vgrG, hcp*) and biosynthesis of the siderophore group (*entC, entD, entE*), were identified. The

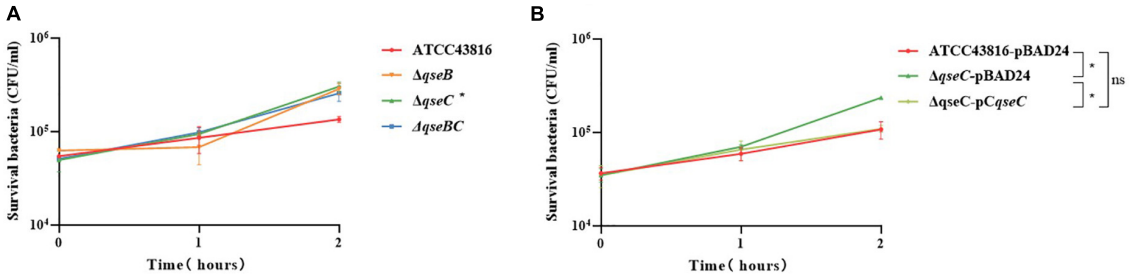

FIGURE 3 | (A) Effect of human serum on ATCC43816, Δ*qseB*, Δ*qseC* and Δ*qseBC*. The bacteria were mixed with serum from healthy human volunteers for 2 hours. Δ*qseC* had significantly higher resistance when compared to control ATCC43816 ($P = 0.0438$, two tailed unpaired *t* test). **(B)** Effect of human serum on ATCC43816-pBAD24, Δ*qseC*-pBAD24 and the complement strain Δ*qseC*-pC*qseC*. Δ*qseC*-pBAD24 had significantly higher resistance when compared to control ATCC43816-pBAD24 ($P = 0.0327$, two tailed unpaired *t* test).

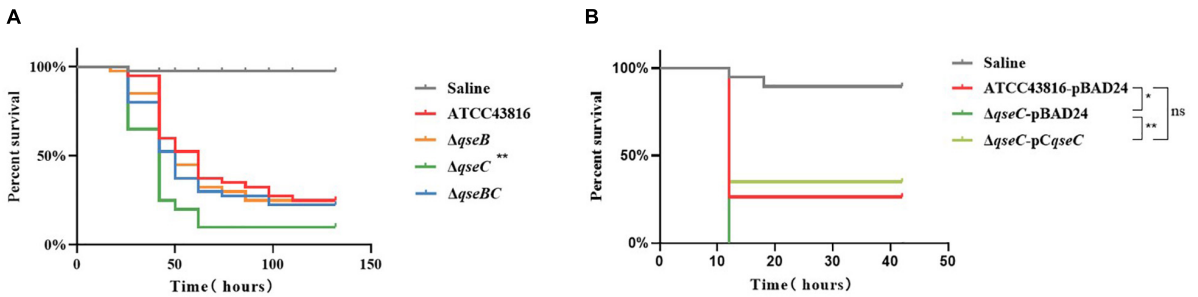

FIGURE 4 | (A) Larvae survival of ATCC43816, Δ*qseB*, Δ*qseC* and Δ*qseBC*. *Galleria mellonella* were injected with 105 CFU wild-type and mutants (40 larvae per group). The survival of *G. mellonella* infected with Δ*qseC* was significantly lower compared with that of the wild-type group ($P = 0.0012$, Log-rank test). **(B)** Larvae survival of ATCC43816-pBAD24, Δ*qseC*-pBAD24 and the complement strain Δ*qseC*-pC*qseC*. The survival of *G. mellonella* infected with Δ*qseC*-pC*qseC* was significantly higher than that of Δ*qseC*-pBAD24 ($P = 0.0040$, Log-rank test).

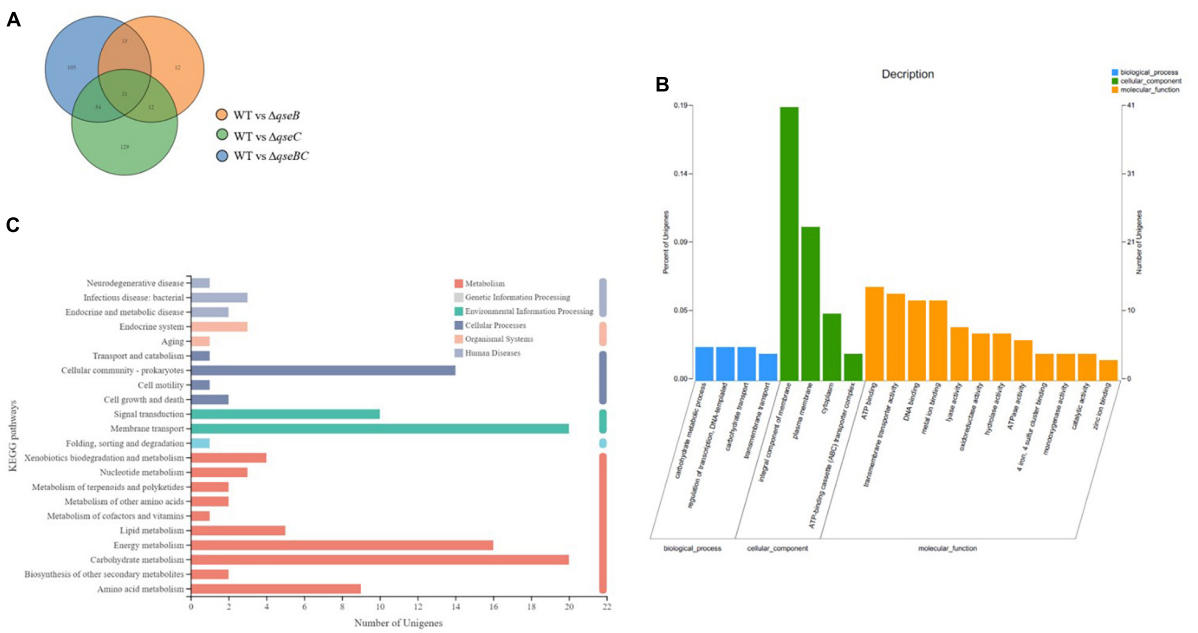

FIGURE 5 | Analysis of total RNA sequencing. **(A)** The Venn diagram shows the overlapped DEGs numbers across Δ*qseB*, Δ*qseC* and Δ*qseBC*. **(B)** Significantly enriched GO terms of DGEs in the *qseC*mutant. **(C)** Significantly enriched KEGG pathways of DEGs in the *qseC*mutant.

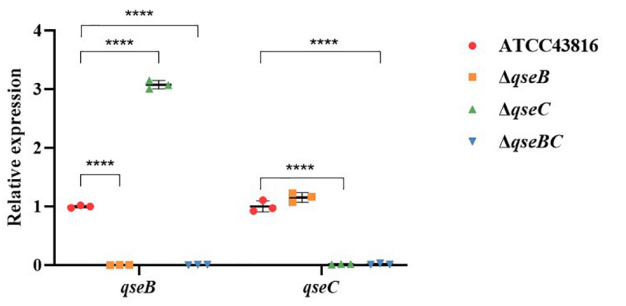

FIGURE 6 | Expression of genes *qseB*, *qseC* in ATCC43816, Δ*qseB*, Δ*qseC* and Δ*qseBC*. Cultures of the wild-type and mutants were grown in LB medium to late-exponential phase, and expressions of genes *qseC*, *qseB* were measured by qRT-PCR and normalized to the expression in the wild-type strain. Data were mean ± standard deviation of three replicates. Significant expression differences between the wild-type andmutants were determined using two-way ANOVA, $^{****}P < 0.0001$.

upregulation of these genes may be associated with the increased virulence observed in the *in vivo* and *in vitro* models, which deserve further studies.

DISCUSSION

In this study, we investigated the effect of QseBC on the pathogenic bacteria virulence and biofilm formation in hvKp strain ATCC43816. The results showed that *qseC*-deletion increased serum resistance in the serum killing assay, but Δ*qseB* and Δ*qseBC* showed no significant difference compared with the wild-type strain. The results suggested that Δ*qseC* could increase serum resistance and possibly enhances the virulence of hvKP. In order to further test this hypothesis, we conducted the *in vivo G. mellonella* killing assay. The results showed that there was a significantly higher virulence of the *qseC* mutant strain when compared to the control ATCC43816. The results of biofilm formation testing also showed that there was a significantly increased biofilm production in Δ*qseC*, while the levels of biofilm production in Δ*qseB* and Δ*qseBC* mutants were similar to those from the wild-type strain. These data suggested that the QseC negatively controlled biofilm formation in hvKP ATCC43816, and Δ*qseC*, but not Δ*qseB* and Δ*qseBC*, had higher serum resistance and *in vivo* virulence in the *G. mellonella* model. Several previous studies have demonstrated that QseC has a regulatory function on the motility, biofilm formation, and virulence in some other pathogens. For example, Yang et al. (2014) found that motility was diminished and biofilm formation was decreased in an *E. coli qseC* mutant. However, our study found that QseC is negatively correlated with the virulence in ATCC43816. Our result appears to be contradicted with the results for *E. coli*, and an explanation is the TCS QseBC system may behave differently among different species.

In order to further explore the mechanism how QseC negatively regulate the virulence, transcriptional profiling was implemented. The RNA-seq results suggested that in Δ*qseC*, the gene associated with biofilm formation (*glgC, glgP, glgA, gcvA, bcsA, ydaM, paaF, ptsG*), bacterial type VI secretion system (*virB4, virB6, virB10, vgrG, hcp*), and biosynthesis of siderophore (*entC, entD, entE*) were significantly up-regulated in comparison to the control. It is worth noting that deletion of *qseC* resulted in the highest overexpression of *ygiW* and *qseB*. In the oral pathogen *Aggregatibacter actinomycetemcomitans*, the *qseBC* genes are co-expressed

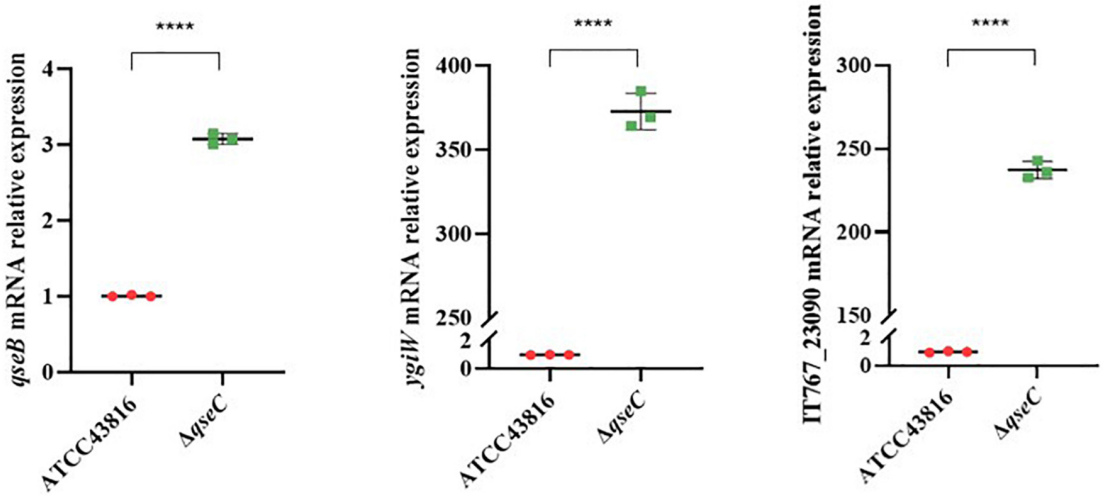

FIGURE 7 | (A) Expression of genes *qseB*, *ygiW*, IT767_23090 in ATCC43816 and Δ*qseC*. Cultures of ATCC43816 and Δ*qseC* were grown in LB medium to late-exponential phase, and expressions of genes *qseB*, *ygiW*, IT767_23090 were measured by qRT-PCR and normalized to the expression in the wild-type strain. Data were mean ± standard deviation of three replicates. Significant expression differences between the wild-type and *qseC*mutant were determined using *t* test. $^{****}P < 0.0001$. **(B)** Expression of genes *qseB*, *ygiW*, IT767_23090 in ATCC43816-pBAD24, Δ*qseC*-pBAD24 and the complement strain Δ*qseC*-p*CqseC*. Data were mean ± standard deviation of three replicates. Significant expression differences between the wild-type and *qseC*mutant were determined using *t* test. $^{*}P < 0.05$, $^{***}P < 0.001$, $^{****}P < 0.0001$.

with *ygiW*, and the transcription of *ygiW-qseBC* operon is directed by the promoter in upstream region of *ygiW* that contains QseB binding sites (Juárez-Rodríguez et al., 2013, 2014). However, analysis of the ATCC43816 genome (accession no. CP009208) showed that *ygiW* was transcribed from the opposite strand as *qseBC* (Weigel and Demuth, 2016), which also suggested that the gene expression regulation for *qseBC* may be different in *A. actinomycetemcomitans* and *K. pneumoniae*. In the absence of the sensor protein QseC, QseB regulator cannot receive the signal of automatic phosphorylation. Therefore, the expression of *qseB* is supposed to be downregulated due to the lack of autophosphorylation signals from QseC. However, in our study, the expression of *qseB* was significantly upregulated by approximately four times, indicating that QseB could accept autophosphorylation signals from other sources. For example, Guckes et al. (2013) reported that in uropathogenic *E. coli*, polymyxin resistance (Pmr) sensor kinase PmrB could active QseB in the absence of QseC. Interestingly, Δ*qseB* and Δ*qseBC* have similar levels of biofilm formation, serum resistance, and *G. mellonella* killing effects to those from the wild-type strain. We therefore hypothesized that in the presence of both QseB and QseC, QseC works as sensor and receives signal, and transfers the phosphate groups to the intracellular response regulator QseB. The phosphorylated QseB directly regulates the expression of iron carrier-related genes, and other virulence-related genes, leading to the changes of bacterial virulence and biofilm formation. In the absence of QseC, QseB could compensatorily receive phosphate groups from other sensor kinase, while the compensatory pathway may be inhibited in the normal condition. When *qseB* was knocked out (both in Δ*qseB* and Δ*qseBC*), the *qseB* regulatory pathway was interrupted, and Δ*qseB* and Δ*qseBC* behaved similarly in the *in vivo* and *in vitro* models.

We also analyzed the differential gene expressions between Δ*qseB*, Δ*qseC*, Δ*qseBC*. IT767_23090, belonging to AraC family transcriptional regulator family, was upregulated by 305-fold. In *Bacillus subtilis*, the AraC family transcriptional regulatory protein PrkC is used to phosphorylate proteins in the metabolic pathway (Libby et al., 2015). In *Vibrio cholerae*, AraC family transcribed regulated protein ToxT promotes virulence by activating cholera toxin and coordinately regulating the expression of fimbriae (Thomson et al., 2015). The AraC family transcribed regulatory protein RarA leads to multiple drug resistance by regulating the expression of active efflux aggregation in *K. pneumoniae* (Jiménez-Castellanos et al., 2016). The study by Chen et al. (2015) showed that the AraC family transcription regulator protein RSP increases the expression of *Staphylococcus aureus* virulence by upregulating the accessory gene regulator (Agr) controlled toxins phenol-soluble modules (PSMs) and alpha-toxin. As such, we suspected that Δ*qseC* may in part work through upregulating the AraC family transcriptional regulator IT767_23090 to confer to higher virulence in hvKP ATCC43816. However, further studies are needed to determine the role of IT767_23090 in the increased virulence in Δ*qseC*.

In summary, our data showed that Δ*qseC* leads to the high expression of *qseB, ygiW,* IT767_23090, and genes associated biofilm formation, bacterial type VI secretion system, and biosynthesis of siderophore, which might contribute to the increased of virulence and biofilm formation. This research provides new insights into the role of QseBC in biofilm formation and virulence of hvKP ATCC43816.

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

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

FUNDING

This study was supported by the grants from the Natural Science Fund of China (82002204), the Science Foundation of Jiangsu Province Health Department (ZDB2020014), the Science Foundation of Suzhou Health Department (LCZX202106), and the Medical and Health Science and Technology Plan Project of Suzhou National New and Hi-Tech Industrial Development Zone (2019Q013).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.817494/full#supplementary-material

REFERENCES

Abisado, R. G., Benomar, S., Klaus, J. R., Dandekar, A. A., and Chandler, J. R. (2018). Bacterial quorum sensing and microbial community interactions. *mBio* 9:e02331-17.

Bearson, B. L., and Bearson, S. M. (2008). The role of the QseC quorum-sensing sensor kinase in colonization and norepinephrine-enhanced motility of *Salmonella enterica* serovar Typhimurium. *Microb. Pathog.* 44, 271–278. doi: 10.1016/j.micpath.2007.10.001

Chen, Z., Wang, Y., Tian, L., Zhu, X., Li, L., Zhang, B., et al. (2015). First report in China of *Enterobacteriaceae* clinical isolates coharboring blaNDM-1 and blaIMP-4 drug resistance genes. *Microb. Drug Resist.* 21, 167–170. doi: 10.1089/mdr.2014.0087

Choby, J. E., Howard-Anderson, J., and Weiss, D. S. (2020). Hypervirulent *Klebsiella pneumoniae* - clinical and molecular perspectives. *J. Intern. Med.* 287, 283–300.

Clarke, M. B., Hughes, D. T., Zhu, C., Boedeker, E. C., and Sperandio, V. (2006). The QseC sensor kinase: a bacterial

adrenergic receptor. *Proc. Natl. Acad. Sci. U. S. A.* 103, 10420–10425.

Fu, L., Huang, M., Zhang, X., Yang, X., Liu, Y., Zhang, L., et al. (2018). Frequency of virulence factors in high biofilm formation bla(KPC-2) producing *Klebsiella pneumoniae* strains from hospitals. *Microb. Pathog.* 116, 168–172. doi: 10.1016/j.micpath.2018.01.030

Gou, Y., Liu, W., Wang, J. J., Tan, L., Hong, B., Guo, L., et al. (2019). CRISPR-Cas9 knockout of qseB induced asynchrony between motility and biofilm formation in *Escherichia coli*. *Can. J. Microbiol.* 65, 691–702. doi: 10.1139/cjm-2019-0100

Groisman, E. A. (2016). Feedback control of two-component regulatory systems. *Annu. Rev. Microbiol.* 70, 103–124. doi: 10.1146/annurev-micro-102215-095331

Guckes, K. R., Kostakioti, M., Breland, E. J., Gu, A. P., Shaffer, C. L., Martinez, C. R. III, et al. (2013). Strong cross-system interactions drive the activation of the QseB response regulator in the absence of its cognate sensor. *Proc. Natl. Acad. Sci. U. S. A.* 110, 16592–16597. doi: 10.1073/pnas.1315320110

Hao, M., He, Y., Zhang, H., Liao, X. P., Liu, Y. H., Sun, J., et al. (2020). crispr-cas9-mediated carbapenemase gene and plasmid curing in carbapenem-resistant *Enterobacteriaceae*. *Antimicrob. Agents Chemother.* 64:e00843-20. doi: 10.1128/AAC.00843-20

Harada, S., Aoki, K., Yamamoto, S., Ishii, Y., Sekiya, N., Kurai, H., et al. (2019). Clinical and molecular characteristics of *Klebsiella pneumoniae* isolates causing bloodstream infections in Japan: occurrence of hypervirulent infections in health care. *J. Clin. Microbiol.* 57:e01206-19. doi: 10.1128/JCM.01206-19

Insua, J. L., Llobet, E., Moranta, D., Pérez-Gutiérrez, C., Tomás, A., Garmendia, J., et al. (2013). Modeling *Klebsiella pneumoniae* pathogenesis by infection of the wax moth Galleria mellonella. *Infect. Immun.* 81, 3552–3565. doi: 10.1128/IAI.00391-13

Jiménez-Castellanos, J. C., Wan Ahmad Kamil, W. N., Cheung, C. H., Tobin, M. S., Brown, J., Isaac, S. G., et al. (2016). Comparative effects of overproducing the AraC-type transcriptional regulators MarA, SoxS, RarA and RamA on antimicrobial drug susceptibility in *Klebsiella pneumoniae*. *J. Antimicrob. Chemother.* 71, 1820–1825. doi: 10.1093/jac/dkw088

Juárez-Rodríguez, M. D., Torres-Escobar, A., and Demuth, D. R. (2013). ygiW and qseBC are co-expressed in Aggregatibacter actinomycetemcomitans and regulate biofilm growth. *Microbiology* 159, 989–1001. doi: 10.1099/mic.0.066183-0

Juárez-Rodríguez, M. D., Torres-Escobar, A., and Demuth, D. R. (2014). Transcriptional regulation of the Aggregatibacter actinomycetemcomitans ygiW-qseBC operon by QseB and integration host factor proteins. *Microbiology* 160, 2583–2594. doi: 10.1099/mic.0.083501-0

Khajanchi, B. K., Kozlova, E. V., Sha, J., Popov, V. L., and Chopra, A. K. (2012). The two-component QseBC signalling system regulates in vitro and in vivo virulence of Aeromonas hydrophila. *Microbiology* 158, 259–271. doi: 10.1099/mic.0.051805-0

Libby, E. A., Goss, L. A., and Dworkin, J. (2015). The Eukaryotic-like Ser/Thr Kinase PrkC regulates the essential WalRK two-component system in *Bacillus subtilis*. *PLoS Genet.* 11:e1005275. doi: 10.1371/journal.pgen.1005275

Merighi, M., Septer, A. N., Carroll-Portillo, A., Bhatiya, A., Porwollik, S., Mcclelland, M., et al. (2009). Genome-wide analysis of the PreA/PreB (QseB/QseC) regulon of *Salmonella enterica* serovar Typhimurium. *BMC Microbiol.* 9:42. doi: 10.1186/1471-2180-9-42

Novak, E. A., Shao, H., Daep, C. A., and Demuth, D. R. (2010). Autoinducer-2 and QseC control biofilm formation and in vivo virulence of Aggregatibacter actinomycetemcomitans. *Infect Immun* 78, 2919–2926. doi: 10.1128/IAI.01376-09

Russo, T. A., and Marr, C. M. (2019). Hypervirulent *Klebsiella pneumoniae*. *Clin. Microbiol. Rev.* 32:e00001-19.

Sharma, V. K., and Casey, T. A. (2014). *Escherichia coli* O157:H7 lacking the qseBC-encoded quorum-sensing system outcompetes the parental strain in colonization of cattle intestines. *Appl. Environ. Microbiol.* 80, 1882–1892. doi: 10.1128/AEM.03198-13

Sperandio, V., Torres, A. G., and Kaper, J. B. (2002). Quorum sensing *Escherichia coli* regulators B and C (QseBC): a novel two-component regulatory system involved in the regulation of flagella and motility by quorum sensing in *E. coli*. *Mol. Microbiol.* 43, 809–821. doi: 10.1046/j.1365-2958.2002.02803.x

Thomson, J. J., Plecha, S. C., and Withey, J. H. (2015). A small unstructured region in *Vibrio cholerae* ToxT mediates the response to positive and negative effectors and ToxT proteolysis. *J. Bacteriol.* 197, 654–668. doi: 10.1128/JB.02068-14

Wang, X., Wang, Q., Yang, M., Xiao, J., Liu, Q., Wu, H., et al. (2011). QseBC controls flagellar motility, fimbrial hemagglutination and intracellular virulence in fish pathogen Edwardsiella tarda. *Fish Shellfish Immunol.* 30, 944–953. doi: 10.1016/j.fsi.2011.01.019

Wang, Y., Wang, S., Chen, W., Song, L., Zhang, Y., Shen, Z., et al. (2018). CRISPR-Cas9 and CRISPR-assisted cytidine deaminase enable precise and efficient genome editing in *Klebsiella pneumoniae*. *Appl. Environ. Microbiol.* 84:e01834-18. doi: 10.1128/AEM.01834-18

Weigel, W. A., and Demuth, D. R. (2016). QseBC, a two-component bacterial adrenergic receptor and global regulator of virulence in *Enterobacteriaceae* and Pasteurellaceae. *Mol. Oral Microbiol.* 31, 379–397. doi: 10.1111/omi.12138

Yang, K., Meng, J., Huang, Y. C., Ye, L. H., Li, G. J., Huang, J., et al. (2014). The role of the QseC quorum-sensing sensor kinase in epinephrine-enhanced motility and biofilm formation by *Escherichia coli*. *Cell Biochem. Biophys.* 70, 391–398. doi: 10.1007/s12013-014-9924-5

Zafar, S., Hanif, S., Akhtar, H., and Faryal, R. (2019). Emergence of hypervirulent K. pneumoniae causing complicated UTI in kidney stone patients. *Microb. Pathog.* 135:103647. doi: 10.1016/j.micpath.2019.103647

Zhu, J., Wang, T., Chen, L., and Du, H. (2021). Virulence factors in hypervirulent *Klebsiella pneumoniae*. *Front. Microbiol.* 12:642484. doi: 10.3389/fmicb.2021.642484

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.

frontiers | Frontiers in Microbiology

ORIGINAL RESEARCH
published: 19 May 2022
doi: 10.3389/fmicb.2022.848906

Incomplete Concordance Between Host Phylogeny and Gut Microbial Community in Tibetan Wetland Birds

Tingbei Bo[1,2†], Gang Song[3†], Shiyu Tang[3,4], Mengru Zhang[3,4], Zhiwei Ma[3,5], Hongrui Lv[3,4], Yun Wu[3,4], Dezhi Zhang[3], Le Yang[6], Dehua Wang[1,2,7] and Fumin Lei[3,4*]*

[1] State Key Laboratory of Integrated Management of Pest Insects and Rodents, Institute of Zoology, Chinese Academy of Sciences, Beijing, China, [2] CAS Center for Excellence in Biotic Interactions, University of Chinese Academy of Sciences, Beijing, China, [3] Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing, China, [4] University of Chinese Academy of Sciences, Beijing, China, [5] School of Ecology and Environment, Anhui Normal University, Wuhu, China, [6] Tibet Plateau Institute of Biology, Lhasa, China, [7] School of Life Sciences, Shandong University, Qingdao, China

OPEN ACCESS

Edited by:
Rodrigo Pulgar Tejo,
University of Chile, Chile

Reviewed by:
Binghua Sun,
Anhui University, China
Tongtong Li,
Zhejiang University of Technology, China

***Correspondence:**
Dehua Wang
dehuawang@sdu.edu.cn
Fumin Lei
leifm@ioz.ac.cn

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: 05 January 2022
Accepted: 08 April 2022
Published: 19 May 2022

Citation:
Bo T, Song G, Tang S, Zhang M, Ma Z, Lv H, Wu Y, Zhang D, Yang L, Wang D and Lei F (2022) Incomplete Concordance Between Host Phylogeny and Gut Microbial Community in Tibetan Wetland Birds. Front. Microbiol. 13:848906. doi: 10.3389/fmicb.2022.848906

Gut microbial communities of animals play key roles in host evolution, while the relationship between gut microbiota and host evolution in Tibetan birds remains unknown. Herein, we sequenced the gut microbiota of 67 wild birds of seven species dwelling in the Tibetan wetlands. We found an obvious species-specific structure of gut microbiota among these plateau birds whose habitats were overlapped. Different from plateau mammals, there was no strict synergy between the hierarchical tree of gut microbial community and species phylogeny. In brown-headed gulls (*Larus brunnicephalus*) as an example, the structure of gut microbiota differed in different habitats, and the relative abundance of bacteria, such as *Lactobacillus*, *Streptococcus, Paracoccus, Lachnospiraceae,* and *Vibrio*, significantly correlated with altitude. Finally, we found various pathogenic bacteria in the birds of these plateau wetlands, and the interspecific differences were related to their diet and living environments.

Keywords: birds, gut microbiota, Tibetan Plateau, phylogeny, adaptation

INTRODUCTION

Previous studies have revealed that gut microbiota plays crucial roles in many physiological processes of vertebrates, including intestinal morphology, digestive functions, and immune health (Backhed et al., 2005). McFall-Ngai et al. (2013) proposed that animals and microbes could be co-evolving together. For example, in African ape species, the dissimilarity between gut microbial communities is correlated with the evolutionary distance between the host species (Moeller et al., 2014). Studies on eight small mammal species distributed in the Qinghai-Tibet Plateau (QTP) and Inner Mongolia grassland show that the gut microbial composition is associated with the host phylogeny (Li et al., 2017).

Birds are a group of warm-blooded vertebrates that show strong adaptability to the environmental changes and are characterized by complex life history and diversified feeding habits and migration patterns (Kohl, 2012). Previous work has revealed that the taxonomic composition of a bird's microbiome differs with the host species (Waite and Taylor, 2014; Hird et al., 2015), and a high variation has also been observed within the species (Hird et al., 2014). One study on woodlarks (*Lullula arborea*) and skylarks (*Alauda arvensis*) shows that niche environment plays a stronger

role in shaping the composition of bird gut microbiota than phylogeny (Veelen et al., 2017). The Tibetan Plateau is well known for its extreme environmental conditions, such as extreme cold, strong UV light, and low oxygen levels, which pose a challenge to the survival and reproduction of birds. Although there are intensive studies on the evolutionary history and adaptation of birds in QTP (Qu et al., 2013, 2020, 2021; Zhu et al., 2018), little is known about the gut microbial ecology of such high-plateau birds and the relationship between host phylogeny and gut microbiota.

In the present study, we selected seven species of birds that are commonly encountered in the plateau wetlands of north Tibet during the breeding season. Bar-headed goose (*Anser indicus*), ruddy shelduck (*Tadorna ferruginea*), and brown-headed gull (*Larus brunnicephalus*) are commonly found in large areas of Northwest China, Mongolia, Kazakhstan, and Russia (Liu and Chen, 2021), while they are the dominant species in plenty of water bodies located in north Tibet during the summer season. The common redshank (*Tringa totanus*) and lesser sand plover (*Charadrius mongolus*) are two waders of smaller body size often found in shallow water bodies, such as lake coasts and riversides. The hill pigeon (*Columba rupestris*) inhabits relatively drier environments and rests on rock reefs, whereas the plain-backed snowfinch (*Pyrgilauda blanfordi*) dwells on highland meadows and lakeside grasslands. These species belong to four avian orders (Anseriformes, Charaderiiformes, Columbiformes, and Passeriformes), representing the various and typical ecological groups of birds inhabiting high-altitude wetlands of north Tibet. Therefore, we studied the gut microbiota of these target species to evaluate the characteristics of the gut microbiota and assess the relationship between host phylogeny and gut microbiota. In this context, we propose to get clues from the gut microbiome that facilitate the adaptation of wild birds to high altitudes.

MATERIALS AND METHODS

Sample Collection

As most of the birds migrate to distant places to escape from extremely harsh conditions during the winters, we carried out field trips and sampling work from 20 July to 15 August, 2021, the time corresponding to the late breeding season of birds in the north Tibet. We found that all these species except for plain-backed snowfinch perched and always moved in flocks. We observed the bird flock for more than 30 min, identified the species, and counted their number by viewing through 8 × 42 binoculars and a 25 × 60 spotting scope. Then, fresh fecal samples were collected from the spots where the flocks rested. We captured the plain-backed snowfinch using mist nets. The trapped birds were removed into clean bird bags independently for 2 h. Then the fecal samples were collected, and the birds were released after the morphology measurements using a clipper. The samples were placed in 2-ml sterilized storage tubes and stored in liquid nitrogen immediately. After the field trip, the samples were immersed in dried ice during transportation and later stored in a lab freezer at −80°C.

16S rRNA Gene Sequencing Analysis

The DNA was extracted from the fecal samples by using the QIAamp DNA stool Mini Kit from Qiagen (Germany), according to the manufacturer's instructions. The 16S rRNA gene, comprising V3 and V4 regions, was amplified by PCR using composite-specific bacterial primers (338F 5′-ACTCCTACGGG AGGCAGCA-3′; 806R 5′-GGACTACHVGGGTWTCTAAT-3′). High-throughput pyrosequencing of the PCR products was performed on an Illumina MiSeq platform. The raw paired-end reads from the original DNA fragments were merged using FLASH32 and assigned to each sample according to the unique barcodes. For the bioinformatics analysis, high-quality reads were performed, and all the effective reads from each sample were clustered into operational taxonomic units (OTUs) based on 97% sequence similarity according to UCLUST33. For alpha diversity analysis, we rarified the OTUs to several metrics, including the Shannon index. For beta diversity analysis, principal coordinate analysis (PCoA) was performed using Bray–Curtis and weighted UniFrac distance matrices. The method of unweighted pair group method with arithmetic mean (UPGMA) was used to establish the phylogenetic trees for bacteria and judge the differences between the samples. The LDA effect size (LEfSe) analysis was performed for the quantitative analysis of biomarkers within each group. Briefly, LEfSe analysis, LDA threshold of > 3, used the non-parametric factorial Kruskal–Wallis (KW) sum rank test.

Host Phylogeny Extraction and Statistical Analysis

TimeTree[1] was used to build a phylogenetic tree with a timeline for these seven species of birds (Hedges et al., 2006; Kumar et al., 2017). TimeTree is a public knowledge base that provides information on the evolutionary timescale of life. Data from thousands of published studies are assembled into a searchable tree of life scaled to time. All prinicpal coordinate analysis (PCoA) were based on Bray–Curtis and weighted UniFrac distances using evenly sampled OTU abundances. Significance for PCoA (β-diversity) analyses was checked using PERMANOVA. The Bray–Curtis dissimilarity metrics were compared with the distance matrix of the host branch length using Mantel's test to test the correlation between the host phylogeny and gut microbiota. The microbial tree was built by UPGMA based on all the samples collected from seven species of birds. The correlation between microbiome and environmental factors was analyzed by Spearman correlation. Differences between groups were statistically analyzed using KW test with a level of $P < 0.05$ ($^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$).

RESULTS

We sequenced the V3 and V4 regions of bacterial 16S rRNA genes of 67 fecal samples from seven species of birds: bar-headed goose (*Anser indicus*), brown-headed gull (*Larus brunnicephalus*), ruddy shelduck (*Tadorna ferruginea*), lesser sand plover (*Charadrius mongolus*), common redshank

[1]http://www.timetree.org

(*Tringa totanus*), hill pigeon (*Columba rupestris*), and plain-backed snowfinch (*Pyrgilauda blanfordi*). The sampling work was carried out at five localities of the Tibet Autonomous Region in summer of 2021, and the regions from west to east were as follows: GJX, GZXDC, NMX, SLC, BGX, NMC, and LKZX (**Supplementary Tables 1, 2** and **Figure 1A**). In PCoA analysis, the gut microbial assemblages were clustered by Bray–Curtis and weighted UniFrac distance matrices (PERMANOVA, $P < 0.05$, **Figure 1B** and **Supplementary Figure 1A**). Microbial composition among groups overlapped across most host species at the phylum level, but differences were observed among Firmicutes, Bacteroidetes, Cyanobacteria, and Proteobacteria (**Figure 1C**). For instance, Proteobacteria were abundant in bar-headed goose and ruddy shelduck, Bacteroidetes were abundant in lesser sand plover, and Cyanobacteria were abundant in common redshank (**Figure 1C**). At the genus level, the microbiota of brown-headed gull was mainly dominated by *Breznakia*, *Cetobacterium, Vibrio, Alkalibacterium,* and *Halomonas*. *Rickettsiella* was higher in common redshank, *Lactobacillus* and *Enterococcus* were higher in hill pigeons, and *Acinetobacter* and *Exiguobacterium* were predominant in bar-headed goose (**Figure 1D**). The results of PICRUSt revealed higher stability in functions, in general, with respect to the variability seen at the taxonomic level. The LEfSe analysis showed statistically significant differences in the lipid metabolism, energy metabolism, metabolism of co-factors, and vitamins at level 2 in the KEGG hierarchy (**Supplementary Figure 1B**). The Mantel test between the distance metrics of host genes and Bray–Curtis dissimilarity showed that there was no significant overall correlation between host phylogeny and gut microbiota ($r = 0.1018$, $P = 0.2$). Our data showed incomplete concordance between host phylogeny and gut microbial communities in the Tibetan wetland birds (**Figures 1E,F**). At the genus level, we also observed several common potentially pathogenic bacteria, including *Erysipelatoclostridium, Escherichia-Shigella, Helicobacter, Nodosilinea_PCC-7104, Staphylococcus,* and *Vibrio*. *Erysipelatoclostridium* accounted for a relatively high proportion of bacteria, particularly in bar-headed gooses. *Helicobacter* was found to be abundant in hill pigeons (**Figure 1G**).

To further explore the impact of habitat differences, we compared the gut microbiota diversity of brown-headed gulls inhabiting five geographical regions. Alpha diversity was higher in the SLC and NMC (**Figure 2A**). In PCoA analysis, most of the gut microbial assemblages clustered by locality (PERMANOVA, $P = 0.001$, **Figure 2B** and **Supplementary Figure 1C**). There were differences in the proportion of Firmicutes, Cyanobacteria, and Proteobacteria among the five localities (**Supplementary Figure 1D**). UPGMA clustering showed that microbial clustering was not entirely consistent within the localities, which may be attributed to individual dispersal due to migration (**Supplementary Figure 1E**). Microbial composition was closely related to habitat, for example, Bacteroidetes and Proteobacteria were positively correlated with altitude, while Firmicutes were negatively correlated with altitude (**Figure 2D**). *Lactobacillus, Streptococcus, Paracoccus,* and *Lachnospiraceae* populations were negatively correlated with altitude, while *Vibrio* was positively correlated with altitude (**Figure 2C** and **Supplementary Table 3**). To explore the impact of bird species on microbial diversity, we compared the four water species (bar-headed goose, brown-headed gull, ruddy shelduck, and common redshank) distributed in the same habitat, BGX. The results of PCoA showed that although the birds lived in the same area, significant differences in the gut microbiota were observed among the species (PERMANOVA, $P < 0.001$, **Figure 2E** and **Supplementary Figure 1F**). The microbial composition and proportion were also different at the genus level (**Supplementary Figure 1G**).

DISCUSSION

This is the first study to investigate the composition and diversity of the gut microbiota in the wild birds dominating high-plateau regions and determine the relationship between host phylogeny and their gut microbiota. Our results indicate that the impact of environmental factors on the gut microbiota of plateau birds would be the main factor rather than phylogeny. Therefore, we should carefully interpret the information about gut microbiota while inferring the evolution of plateau birds.

Features of Gut Microbiota of Plateau Birds

The results of our study indicate that the gut microbiota of the plateau birds is dominated by Firmicutes, Cyanobacteria, and Proteobacteria. The characteristic feature of Firmicutes is the synthesis of short-chain fatty acids (SCFAs), which are found to be higher in other plateau endemic animals, like plateau yaks (*Bos mutus*) (Zhang et al., 2016) and plateau pika (*Ochotona curzoniae*) (Li et al., 2016). Higher SCFA content can satisfy the physiological or energy demands of the host in cold and hypoxic high-altitude environments. Although living in the same extreme environments of the Qinghai-Tibet Plateau, the seven species of birds show differences in the composition of the microbiota. The different levels of gut microbial diversity between the host species may be partly explained by the differences in diet quality. For example, bar-headed goose and ruddy shelduck mainly feed on invertebrates and plants, while the hill pigeon and plain-backed snowfinch feed on seeds and insects. Although bar-headed goose and ruddy shelduck have similar gut microbiota and diet, the synthetic pathways related to functional adaptation might be different. For instance, bar-headed geese have more gut bacteria related to amino acid metabolism and lipid metabolism, while the proportion of bacteria involved in energy, co-factor, and vitamin metabolism is higher in ruddy shelduck. This finding reflects the species specificity of gut microbiota and ecological niche differentiation between the two species. The common redshank and lesser sand plover are two small-sized waders often found together in shallow water bodies. Their habitats and feeding habits are similar, but the structure of gut microbiota is not similar. Bacteroidetes are abundant in lesser sand plover, while Cyanobacteria are abundant in common redshank. We also noticed that different birds harbor different types of potentially pathogenic microbiota. We observed that the feces of bar-headed goose contains a large proportion of *Erysipelatoclostridium*, which may be due to the pollution of the water in which they live or contact with the

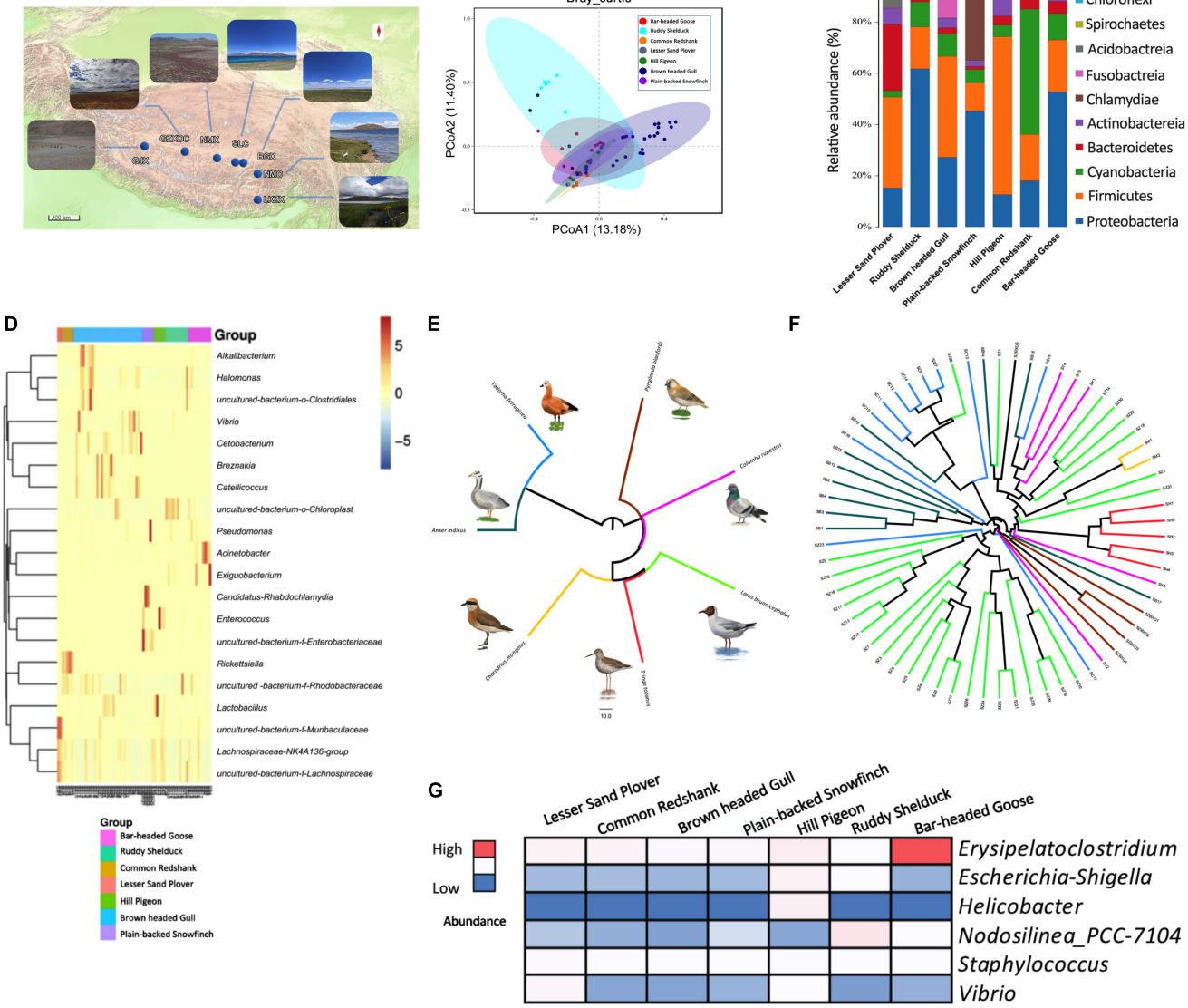

FIGURE 1 | Gut microbial diversity of birds and phylogenetic tree based on variations in complete mitochondrial gene sequences and composition of the gut microbiome. **(A)** Sampling location and photo record of habitats. **(B)** PCoA plots of Bray–Curtis distance for 67 birds. **(C)** Result of gut microbiota in seven species of birds at the phylum level. **(D)** Heat map of the microbial composition of seven species of birds at the genus level. **(E)** Host phylogeny based on mitochondrial genes. **(F)** Gut microbiome phylogeny based on Bray–Curtis dissimilarity. Branches of the same color represent the same species. The closer the samples, the shorter the branch length, indicating that the species composition of the two samples is more similar. **(G)** Differences in the abundance of pathogens among different species.

pathogenic bacteria in water. Hill pigeons usually live closely with humans, so they carry more number of *Escherichia-Shigella* and *Helicobacter* species in their feces. Therefore, harboring and transmission of potentially pathogenic bacteria by the wild birds of the plateau is a topic that is worthy of further study in the future.

Relationship Between Host Species Phylogeny and Their Gut Microbiota

Previous studies on mammals show that various internal and external factors influence the composition and diversity of the gut microbiota, including host phylogeny, gut morphology, diet, physiological status, geography, and body weight (Ley et al., 2006, 2008; Santacruz et al., 2010; Linnenbrink et al., 2013). Therefore, co-evolution between host and microbial lineages has played a key role in the adaptation of mammals to their diverse lifestyles. Most mammalian microbiomes are strongly correlated with host phylogeny (Ley et al., 2008). Findings in other vertebrates also showed that closely related lineages harbor more similar gut microbiota than distantly related lineages (Ochman et al., 2010; Sanders et al., 2014). As for plateau animals, for example, Pika species, the host relationship affects

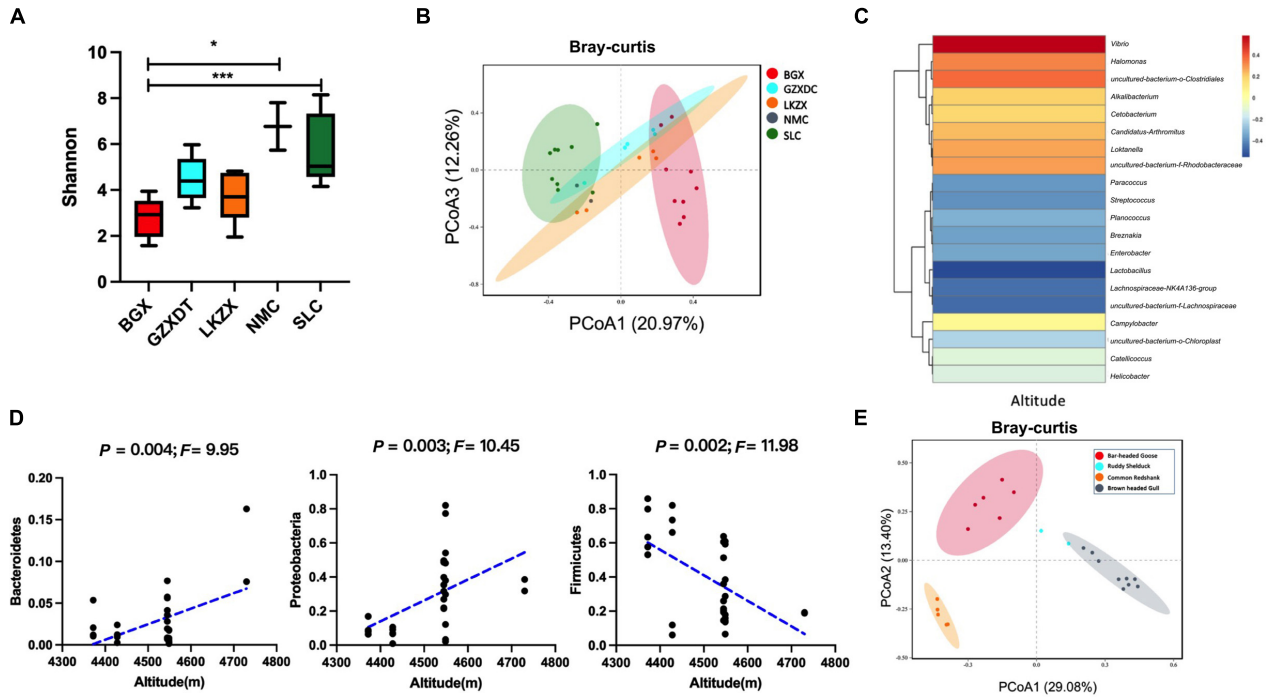

FIGURE 2 | Gut microbial diversity of brown-headed gulls from five habitats and gut microbial diversity of four species of birds in BGX. **(A)** The Shannon index. **(B)** PCoA plots of Bray–Curtis distance. **(C)** Correlation between three phyla and altitude (Spearman). **(D)** Heat map of correlation between microbial composition and altitude at the genus level. **(E)** PCoA plots of Bray–Curtis distance of four species in BGX. Data indicate mean ± SEM values. $^{*}P < 0.05$ and $^{***}P < 0.001$.

the structure of the gut microbiota apparently, showing a significant overall correlation between host phylogeny and gut microbiota (Li et al., 2017). However, in birds, only a few studies have found a correlation between the gut microbiome and host phylogeny. For example, studies on four waterbird species in Israel, great cormorant (*Phalacrocorax carbo*), little egret (*Egretta garzetta*) black-crowned night heron (*Nycticorax nycticorax*), and black-headed gull (*Larus ridibundus*), suggest a correlation between the host phylogeny and their intestinal microbial community hierarchical tree, thus displaying phylosymbiosis (Laviad-Shitrit et al., 2019). A study including 491 species of birds showed a weak correlation between host factors and microbiome composition, while most of their intestinal bacteria exhibited no host specificity (Song et al., 2020). Our results show that the phylogeny of birds does not influence their intestinal microbial community hierarchical tree in Tibet wetlands, which is inconsistent with the existing research on plateau mammals and may reflect a unique feature of high-plateau birds.

Altitude Affects the Composition of Gut Microbiota

Previous research has shown that the avian microbiomes have relatively low stability (Song et al., 2020; Bodawatta et al., 2021) and high malleability to environmental and dietary changes (Bodawatta et al., 2021). Similarly, our results show that brown-headed gulls from different localities have different intestinal microbial structures and that the gut microbiome of these birds inhabiting the same habitat is more similar. The study on woodlarks and skylarks has also shown that sharing an ecological niche among hosts (either species or individuals) leads to the convergence of their microbiota (Veelen et al., 2017). Through correlation analysis, we found that altitude is the main factor that affects the proportion of key bacteria. With an increase in altitude, the proportion of Firmicutes, including *Lactobacillus* and *Lachnospiraceae*, in brown-headed gulls decreases. *Lactobacillus* play major roles in the SCFAs (Borda-Molina et al., 2018), food digestion, and energy conversion. Lachnospiraceae include obligate anaerobic bacteria that affect the host health by producing short-chain fatty acids, participating in bile acid metabolism, and promoting colonization resistance to intestinal pathogens (Sorbara et al., 2020; Xu et al., 2021). Therefore, our results imply that low-altitude brown-headed gulls may rely on the relative abundance of *Lactobacillus* and *Streptococcus* to adapt to the complex food and pathogen-rich environment. At higher altitudes, the proportion of Bacteroidetes and Proteobacteria is increased, including *Vibrio* and *Paracoccus*, which may be attributed to the weakening of immune function, suggesting that the gut microbiome may be involved in regulating the immune functions of high-altitude birds (Qu et al., 2013). However, microbial clustering is not completely consistent with geographical separation, which suggests that there might be some individual migrants and shared microbiome across different localities in brown-headed gulls.

CONCLUSION

Although the composition of microbiomes is different among species, our results show that the gut microbial community is not in line with the phylogenetic relationship in Tibet wetland birds. Environmental factors, such as altitude, are the key factors that affect the gut microbiota of plateau birds. Our results provide a new idea for exploring the relationship between gut microbiota and the evolution and adaptability of high-plateau animals, and suggest that one should be cautious while taking into account the gut microbiota to infer the evolution pattern in plateau birds.

DATA AVAILABILITY STATEMENT

Raw sequence data are deposited in the NCBI Sequence Read Archive under accession BioProject PRJNA787368.

AUTHOR CONTRIBUTIONS

ZM, HL, YW, DZ, and LY: field work and experiment. TB, GS, ST, MZ, and DW: writing. DW and FL: supervision and project management. TB, GS, and FL: funding acquisition. TB and GS: data analysis. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the Second Tibetan Plateau Scientific Expedition and Research program (No. 2019QZKK0304) and the National Natural Scientific Foundation of China (No. 32070434). The experimental and sequencing work was supported by the Central Guidance on Local Science and Technology Development Fund of Tibet Autonomous Region (XZ202201YD0015C) and the National Science and Technology Major Project (No. 2018ZX10101004). We also acknowledge the support of the Postdoctoral Research Foundation of China (2021M693158) and Young Elite Scientists Sponsorship Program by CAST (2021QNRC001) and ISZS to TB.

ACKNOWLEDGMENTS

We thank Prof. Huan Li of Lanzhou University for his suggestions and help in writing the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.848906/full#supplementary-material

REFERENCES

Backhed, F., Ley, R. E., Sonnenburg, J. L., Peterson, D. A., and Gordon, J. I. (2005). Host-bacterial mutualism in the human intestine. *Science* 307, 1915–1920. doi: 10.1126/science.1104816

Bodawatta, K. H., Hird, S. M., Grond, K., Poulsen, M., and Jønsson, K. A. (2021). Avian gut microbiomes taking flight. *Trends Microbiol.* 30, 268–280. doi: 10.1016/j.tim.2021.07.003

Borda-Molina, D., Seifert, J., and Camarinha-Silva, A. (2018). Current perspectives of the chicken gastrointestinal tract and its microbiome. *Comput. Struct. Biotechnol.* 16, 131–139. doi: 10.1016/j.csbj.2018.03.002

Hedges, S. B., Dudley, J., and Kumar, S. (2006). TimeTree: a public knowledge-base of divergence times among organisms. *Bioinformatics* 22, 2971–2972. doi: 10.1093/bioinformatics/btl505

Hird, S. M., Carstens, B. C., Cardiff, S. W., Dittmann, D. L., and Brumfield, R. T. (2014). Sampling locality is more detectable than taxonomy or ecology in the gut microbiota of the brood- parasitic Brown-headed Cowbird (*Molothrus ater*). *PeerJ.* 2:321.

Hird, S. M., Sánchez, C., Carstens, B. C., and Brumfield, R. T. (2015). Comparative gut microbiota of 59 neotropical bird species. *Front. Microbiol.* 6:1403. doi: 10.3389/fmicb.2015.01403

Kohl, K. D. (2012). Diversity and function of the avian gut microbiota. *J. Comp. Physiol. B.* 182, 591–602. doi: 10.1007/s00360-012-0645-z

Kumar, S., Stecher, G., Suleski, M., and Hedges, S. B. (2017). TimeTree: a resource for timelines, timetrees, and divergence times. *Mol. Biol. Evol.* 34, 1812–1819. doi: 10.1093/molbev/msx116

Laviad-Shitrit, S., Izhaki, I., Lalzar, M., and Halpern, M. (2019). Comparative analysis of intestine microbiota of four wild waterbird species. *Front Microbiol.* 10:1911. doi: 10.3389/fmicb.2019.01911

Ley, R. E., Hamady, M., Lozupone, C., Turnbaugh, P. J., Ramey, R. R., Bircher, J. S., et al. (2008). Evolution of mammals and their gut microbes. *Science* 320, 1647–1651.

Ley, R. E., Peterson, D. A., and Gordon, J. I. (2006). Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell* 124, 837–848. doi: 10.1016/j.cell.2006.02.017

Li, H., Li, T., Beasley, D. E., Heděnec, P., Xiao, Z., Zhang, S., et al. (2016). Diet diversity is associated with beta but not alpha diversity of pika gut microbiota. *Front. Microbiol.* 7:758. doi: 10.3389/fmicb.2016.01169

Li, H., Qu, J. P., Li, T. T., Yao, M. J., Li, J. Y., and Li, X. Z. (2017). Gut microbiota may predict host divergence time during Glires evolution. *FEMS Microbiol. Ecol.* 93:009. doi: 10.1093/femsec/fix009

Linnenbrink, M., Wang, J., Hardouin, E. A., Künzel, S., Metzler, D., and Baines, J. F. (2013). The role of biogeogra- phy in shaping diversity of the intestinal microbiota in house mice. *Mol. Ecol.* 22, 1904–1916. doi: 10.1111/mec.12206

Liu, Y., and Chen, S. H. (2021). *The CNG Field Guide to the Birds of China*. Changsha: Hunan Science & Technology Press.

McFall-Ngai, M., Hadfield, M. G., Bosch, T. C., Carey, H. V., Domazet-Lošo, T., Douglas, A. E., et al. (2013). Animals in a bacterial world, a new imperative for the life sciences. *PNAS* 110, 3229–3236. doi: 10.1073/pnas.1218525110

Moeller, A. H., Li, Y., Mpoudi Ngole, E., Ahuka-Mundeke, S., Lonsdorf, E. V., Pusey, A. E., et al. (2014). Rapid changes in the gut microbiome during human evolution. *PNAS* 111, 431–435. doi: 10.1073/pnas.1419136111

Ochman, H., Worobey, M., Kuo, C. H., Ndjango, J. B., Peeters, M., Hahn, B. H., et al. (2010). Evolutionary relationships of wild hominids recapitulated by gut microbial communities. *PLoS Biol.* 8:e1000546. doi: 10.1371/journal.pbio.1000546

Qu, Y. H., Zhao, H. W., Han, N. J., Zhou, G. Y., Song, G., Gao, B., et al. (2013). Ground tit genome reveals avian adaptation to living at high altitudes in the Tibetan plateau. *Nat. Commun.* 4:2071. doi: 10.1038/ncomms3071

Qu, Y., Chen, C., Chen, X., Hao, Y., She, H., Wang, M., et al. (2021). The evolution of ancestral and species-specific adaptations in snowfinches at the Qinghai-Tibet Plateau. *Proc. Natl. Acad. Sci. U.S.A.* 118:e2012398118. doi: 10.1073/pnas.2012398118

Qu, Y., Chen, C., Xiong, Y., She, H., Zhang, Y. E., Cheng, Y., et al. (2020). Rapid phenotypic evolution with shallow genomic differentiation during early stages of high elevation adaptation in Eurasian tree sparrows. *Natl. Sci. Rev.* 7, 113–127. doi: 10.1093/nsr/nwz138

Sanders, J. G., Powell, S., Kronauer, D. J., Vasconcelos, H. L., Frederickson, M. E., and Pierce, N. E. (2014). Stability and phylogenetic correlation in gut microbiota: lessons from ants and apes. *Mol. Ecol.* 23, 1268–1283. doi: 10.1111/mec.12611

Santacruz, A., Collado, M. C., Garcia-Valdes, L., Segura, M. T., Martín-Lagos, J. A., Anjos, T., et al. (2010). Gut microbiota composition is associated with body weight, weight gain and biochemical parameters in pregnant women. *Br. J. Nutr.* 104, 83–92. doi: 10.1017/S0007114510000176

Song, S. J., Sanders, J. G., Delsuc, F., Metcalf, J., Amato, K., Taylor, M. W., et al. (2020). Comparative analyses of vertebrate gut microbiomes reveal convergence between birds and bats. *mBio* 11, 2901–2919. doi: 10.1128/mBio.02901-19

Sorbara, M. T., Littmann, E. R., Fontana, E., Moody, T. U., Kohout, C. E., Gjonbalaj, M., et al. (2020). Functional and genomic variation between human-derived isolates of lachnospiraceae reveals inter-and intra-species diversity. *Cell Host Microbe* 28, 134.e4–146.e4. doi: 10.1016/j.chom.2020.05.005

Veelen, H. P., Salles, J. F., and Tieleman, B. I. (2017). Multi-level comparisons of cloacal, skin, feather and nest-associated microbiota suggest considerable influence of horizontal acquisition on the microbiota assembly of sympatric woodlarks and skylarks. *Microbiome* 5:156. doi: 10.1186/s40168-017-0371-6

Waite, D. W., and Taylor, M. W. (2014). Characterizing the avian gut microbiota: membership, driving influences, and potential function. *Front. Microbiol.* 5:223. doi: 10.3389/fmicb.2014.00223

Xu, J., Zhang, J. N., Sun, B. H., Liu, Q., Ma, J., Zhang, Q., et al. (2021). The role of genotype and diet in shaping gut microbiome in a genetic Vitamin A deficient mouse model. *J. Genet. Genomics* 49, 155–164. doi: 10.1016/j.jgg.2021.08.015

Zhang, Z. G., Xu, D. M., Wang, L., Long, R. J., Zhao, F. Q., and Shi, P. (2016). Convergent evolution of rumen microbiomes in high-altitude mammals. *Curr. Biol.* 26, 1873–1879. doi: 10.1016/j.cub.2016.05.012

Zhu, X., Guan, Y., Signore, A. V., Natarajan, C., DuBay, S. G., Cheng, Y., et al. (2018). Divergent and parallel routes of biochemical adaptation in high-altitude passerine birds from the Qinghai-Tibet Plateau. *Proc. Natl. Acad. Sci. U.S.A.* 115, 1865–1870. doi: 10.1073/pnas.1720487115

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.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership