Check for updates

### **OPEN ACCESS**

EDITED BY Tongling Shan, Chinese Academy of Agricultural Sciences, China

REVIEWED BY Jianmin Zhang, South China Agricultural University, China Di Ran, University of Illinois Chicago, United States

\*CORRESPONDENCE Bin Wu ⊠ wub@mail.hzau.edu.cn

RECEIVED 20 October 2023 ACCEPTED 15 January 2024 PUBLISHED 07 February 2024

#### CITATION

Wang L, Wu W, Zhao L, Zhu Z, Yao X, Fan J, Chen H, Song W, Huang X, Hua L, Qian P, Chen H, Peng Z and Wu B (2024) Fecal PCR survey and genome analysis of *Lawsonia intracellularis* in China. *Front. Vet. Sci.* 11:1324768. doi: 10.3389/fvets.2024.1324768

### COPYRIGHT

© 2024 Wang, Wu, Zhao, Zhu, Yao, Fan, Chen, Song, Huang, Hua, Qian, Chen, Peng and Wu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

## Fecal PCR survey and genome analysis of *Lawsonia intracellularis* in China

Lei Wang<sup>1,2,3</sup>, Wenqing Wu<sup>1,2</sup>, Lifeng Zhao<sup>1,2</sup>, Zhanwei Zhu<sup>1,2</sup>, Xinzhi Yao<sup>4</sup>, Jie Fan<sup>1,2</sup>, Hongjian Chen<sup>1,2</sup>, Wenbo Song<sup>1,2</sup>, Xi Huang<sup>1,2</sup>, Lin Hua<sup>1,2</sup>, Ping Qian<sup>1,2</sup>, Huanchun Chen<sup>1,2</sup>, Zhong Peng<sup>1,2,5</sup> and Bin Wu<sup>1,2</sup>\*

<sup>1</sup>State Key Laboratory of Agricultural Microbiology, College of Veterinary Medicine, Huazhong Agricultural University, Wuhan, China, <sup>2</sup>The Cooperative Innovation Center for Sustainable Pig Production, Huazhong Agricultural University, Wuhan, China, <sup>3</sup>Guangxi YangXiang Co., Ltd., Guigang, China, <sup>4</sup>College of Informatics, Hubei Key Laboratory of Agricultural Bioinformatics, Huazhong Agricultural University, Wuhan, China, <sup>5</sup>Hubei Hongshan Laboratory, Wuhan, China

Proliferative enteropathy caused by Lawsonia intracellularis is an important economic associated disease to pig industry, but the knowledge about the prevalence of L. intracellularis in pig farms in China is limited. In addition, there is no complete genome sequence available for *L. intracellularis* isolates from China. In this study, we developed a TaqMan qPCR for the screening of L. intracellularis by targeting the bacterial 16S rDNA gene. Laboratory evaluations revealed a good sensitivity and specificity on detecting L. intracellularis nucleic acid. Using this method, we investigated 891 fecal samples from apparently healthy pigs in 47 farms. The results demonstrated a screening positive rate of 37.3% (95% Cl, 34.1-40.5%) for the samples, and a farm screening positive rate of 93.6% (95% CI, 65.3–94.4%). The screening positive rate at herd level ranged from 6.67% (95% CI, 0.2-31.9%) to 40% (95% CI, 38-79.6%), while at animal level, the highest screening positive rate was found in 12-week-old pigs [85.7% (95% CI, 67.3-96.0%)]. Investigation of 705 diarrheal or bloody feces from symptomatic pigs revealed that the highest positive rate was found in replacement gilts which was 37.18% (95% CI, 45.1-89.5%). Secondly, we conducted the complete genome sequence of a L. intracellularis PPE-GX01-2022 from China through PacBio sequencing. The genome of PPE-GX01-2022 consisted of a chromosome of 1,439,110 bp in length and three plasmids of 193,063, 39,799, and 27,067 bp, respectively. This genome encoded 1,428 predicted proteins, 44 tRNAs, and 6 rRNAs. Sequence comparisons demonstrated that the genome sequence of PPE-GX01-2022 was highly homologous to those of two isolates from US, and these three isolates shared 1,378 core genes. The screening results suggest a high prevalence rate of L. intracellularis in Chinese pig farms. In addition, the genome sequence of the Chinese isolate was highly homologous to those of the field isolates from the US.

### KEYWORDS

 $\label{eq:lawsonia} \textit{Intracellularis}, \ \mathsf{TaqMan} \ \mathsf{qPCR}, \ \mathsf{screening}, \ \mathsf{pig} \ \mathsf{farms}, \ \mathsf{complete} \ \mathsf{genome} \\ \mathsf{sequence}$ 

## Background

The gram-negative microaerophilic obligate intracellular bacterium Lawsonia intracellularis can cause proliferative enteropathy (PE) in weaned and growing pigs <4 months, which is associated with decreased weight gain and low mortality; it can also lead to hemorrhagic enteritis (HE) in mature pigs older than 4 months, which is characterized by intestinal hemorrhages and sudden death (1). L. intracellularis is prevalent worldwide and has caused big economic losses. For example, it estimates that PE leads to an increase cost of 2-20 dollars of feed consumption in every diseased-pig and causes annual economic losses of \$20 million to the pig industry of the United States as well as £2-£4 million to the pig industry of the United Kingdom (2). However, many knowledge gaps on the biology, epidemiology, and pathogenesis of the bacterium remain to be bridged. For example, there are only eight whole genome sequences of L. intracellularis publicly available on NCBI GenBank database, and only two of them are complete genome sequences (3). Seven of these isolates are from pigs (three from Japan, one from South Korea, one from United Kingdom, two from United States) and the remaining one is recovered from a horse in United States. The genome sizes of the two complete sequences (N343, GenBank accession. CP004029; PHE/MN1-00, GenBank accession. AM180252) are 33.09 Mb, while those of the remaining six are 32.90 Mb, respectively. A lack of more genome sequences of L. intracellularis from different parts of the world limits the further understanding of the genetic evolutionary characteristics of the bacterium.

Serological methods particular detection of antibodies against L. intracellularis through enzyme linked immunosorbent assay (ELISA) are commonly applied for monitoring the prevalence of L. intracellularis infection in clinical investigations (1, 4). For example, a study investigated the seroprevalence of L. intracellularis in pigs at different ages in 134 farm sites in 1998 and 43 farm sites in 2008 in Great Britain and the Republic of Ireland (5). Both two surveys in this study revealed higher than 90% seropositive pigs at 20-23 weeks of age on British farms (97.8% in 1998; 93.1% in 2008) and Irish farms (97% in 1998; 92.9% in 2008) (5). In another serological investigation conducted on farms in France and Spain, postweaning pigs on 29 French farms (33 total) and 20 Spanish farms (29 total) had a pattern of infection characterized by seroconversion in the grower period, representing a respective farm-seropositivity of 87.88 and 68.97% (6). In addition to serological methods, several molecular methods have been also developed for detecting L. intracellularis DNA in tissue specimens or fecal samples, and real-time quantitative PCR (qPCR) is the mostly-notable one (1, 4). A couple of previous studies reported the development of qPCR to detect L. intracellularis DNA based on different target genes, including ubiE (7), the 16S ribosomal RNA (rRNA) gene (8, 9), aspA (10, 11), and the 16S ribosomal DNA (rDNA) gene (12, 13). While these methods demonstrate good diagnostic specificity and sensitivity under laboratory validation, there are not used in clinical investigations as field detection is beyond these studies. However, there are still studies involve in the application of qPCR for clinical investigation. For example, a study investigated 144 herds in five European countries (Germany, Denmark, Spain, the Netherlands and the United Kingdom) by examination of 6,450 fecal samples using commercial qPCR kits ("Kylt<sup>®</sup> PIA [*Lawsonia intracellularis*]" and "Kylt<sup>®</sup> Quantitative standard for *Lawsonia intracellularis*", AniCon Labor GmbH, Hoeltinghausen, Germany) and demonstrated a herd prevalence of 90.3% (79.2–100.0%) among the tested herds (14). All the above findings suggest a high prevalence of *L. intracellularis* on pig farms around the world.

China is the largest pig rearing and pork producing country in the world. However, current data on the prevalence of L. intracellularis on Chinese pig farms is still limited. A recent study investigated 3,586 serum samples in eight major pig-producing provinces during 2019-2020 and revealed that 2,837 (79.1%, 95% CI: 77.7-80.4%) were seropositive for L. intracellularis (15). In an earlier study, examination of 1,060 serum samples from 14 commercial farms through China revealed a 57% (95% CI: 50-64%) seroprevalence of L. intracellularis (16). These sero-epidemiological studies suggest a high prevalence of L. intracellularis in Chinese pig farms. However, it is still lack of much data on the prevalence of L. intracellularis from the perspective of etiology. In this study, we developed a TaqMan qPCR to detect L. intracellularis DNA from porcine fecal samples. The objective is to investigate the prevalence of L. intracellularis in several Chinese pig farms by detecting bacterial DNA using qPCR. In addition, we also aim to generate a complete genome sequence of a Chinese L. intracellularis isolate from clinical samples to help understand the genomic characteristics of L. intracellularis prevalent in China.

### Results

## Development of a TaqMan qPCR for the detection of *L. intracellularis*

A TaqMan qPCR for detecting L. intracellularis was developed based on the 16S rRNA gene (GenBank accession no. U30147), following the protocol described previously (13). Optimization of the amplification conditions revealed that an optimal reaction occurred when the concentrations of the forward and reverse primers set as 0.5 µM while the concentration of the probe set as  $0.25\,\mu\text{M}$ . The limit of the qPCR method detecting the standard plasmid was 36 copies/µl (Table 1). Construction of the standard curves demonstrated that there was a strong linear correlation  $(R^2 = 0.9999; y = -3.547$ lgx+40.520) between the Ct values and the corresponding copy numbers of the 16S rDNA gene of L. intracellularis. Evaluation of specificity showed that the qPCR method only gave amplification curve to L. intracellularis (Supplementary Figure 1). The coefficient of variation (CV) values of the method detecting different concentrations of the standard plasmids were lower than 1.5 (Table 2). There was no difference ( $\chi^2$ = 1, p > 0.05) on the true positive rate of clinical samples detected by the qPCR method developed in this study and the qPCR method reported previously (17) (Table 3). The detection limit of the qPCR

Abbreviations: ANI, Average nucleotide identity; ARG, Antimicrobial resistant gene; ELISA, Enzyme linked immunosorbent assay; HE, Hemorrhagic enteritis; PE, Proliferative enteropathy; qPCR, Real-time quantitative PCR; VFGs, Virulence factor genes.

TABLE 1	Limits of the	TaqMan	qPCR	method	detecting	the	standard
plasmid.							

Dilution folds	Plasmid concentrations (ng/µl)	DNA copies (copies/µl)	CT values
$10^{-2}$	412.2	$6.73  imes 10^9$	14.103
$10^{-3}$	$412.2 \times 10^{-2}$	$6.73 \times 10^{8}$	17.343
$10^{-4}$	$412.2 \times 10^{-3}$	$6.73 \times 10^{7}$	20.943
10 <sup>-5</sup>	$412.2\times10^{-4}$	$6.73 \times 10^{6}$	24.517
10 <sup>-6</sup>	$412.2 \times 10^{-5}$	$6.73 \times 10^{5}$	28.243
10 <sup>-7</sup>	$412.2 \times 10^{-6}$	$6.73  imes 10^4$	31.977
10 <sup>-8</sup>	$412.2 \times 10^{-7}$	$6.73 \times 10^{3}$	34.962
10 <sup>-9</sup>	$412.2 \times 10^{-8}$	$6.73  imes 10^2$	No detection result
$10^{-10}$	$412.2 \times 10^{-9}$	6.73 × 10	No detection result
Nucleotide-free water	0	0	No detection result

TABLE 2 Repeatability of the TaqMan qPCR method detecting the standard plasmid.

Dilutes	s Repeats		$\begin{array}{l} \text{Mean} \pm \\ \text{standard} \\ \text{deviation} \end{array}$	Coefficient of variation (CV) values	
	1	2	3		
$6.73  imes 10^8$	18.63	19.02	18.51	$18.72\pm0.267$	1.42
$6.73  imes 10^7$	21.85	21.85	21.85	$21.89\pm0.075$	0.34
$6.73  imes 10^{6}$	24.75	24.75	24.66	$24.69\pm0.049$	0.19

method was also similar to that of the previously reported method (36 vs. 30 copies/ $\mu$ l).

## Fecal PCR investigation of *L. intracellularis* in pig farms in China

To understand the prevalence of L. intracellularis on Chinese pig farms, we collected 891 fecal samples from apparently healthy pigs in 47 farms for bacterial DNA screening. The results revealed that 37.3% (95% CI: 34.1-40.5%) of the samples were positive for L. intracellularis DNA, with a farm positive rate of 93.6% (95% CI: 65.3-94.4%). High positive screening rates were observed in pigs aged between 72- and 155-day-old (Figures 1A, B). In particular, the highest positive investigation rate (92.31%) was found in pigs at 95 days of age (Figures 1A, B). We next investigated 705 diarrheal or bloody feces from different pig herds (boars, sows, gilts, weaners, finisher pigs and fattening pigs) delivered from 125 pig farms (Supplementary Table 2). The results demonstrated the positive screening rates among different herds ranging from 8.97 to 37.18% (Figure 1C). Regarding diarrheal or bloody samples from pigs at different ages, highest positive screening rates were found in pigs at 87-, 151-, 162-, 165-, 196-, and 226 days of age (Figure 1D).

TABLE 3 Comparison of the TaqMan qPCR method and a previously reported qPCR (China AQSIQ) on detecting clinical samples.

China AQSIQ	TaqMai	n qPCR	Total	Statistical analysis	
	Numbers of positive samples	Numbers of negative samples			
Numbers of positive samples	18	0	18		
Numbers of negative samples	0	50	50	$\chi^2 = 1, p > 0.05$	
Total	18	50	68		

## Generation of the first complete genome sequence of *L. intracellularis* from China

The complete genome sequence of a L. intracellularis isolate, designated PPE-GX01-2022, was generated using PacBio sequencing. The genome of this isolate consisted of a chromosome of 1,439,110 bp in length and three plasmids of 193,063 bp (GX01-2022-P1), 39,799 bp (GX01-2022-P2), and 27,067 bp (GX01-2022-P3), respectively (Figure 2). Gene annotation revealed that the complete genome sequence of PPE-GX01-2022 encoded 1,428 predicted proteins, as well as 44 tRNAs, 6 rRNAs (5S rRNA, 2; 16S rRNA, 2; 23S rRNA, 2), and 5 other non-coding RNAs (ncRNAs; Table 4). No CRISPR elements or prophage regions were identified in the bacterial genome. Prediction of antimicrobial resistance genes (ARGs) demonstrated that PPE-GX01-2022 contained 27 genes conferring resistance to macrolides (seven genes), peptide antibiotics (five genes), fluoroquinolones (three genes), glycopeptide antibiotics (two genes), isoniazid (two genes), aminoglycosides (one gene), elfamycins (one gene), fosfomycin (one gene), fusidic acids (one gene), monobactams (one gene), mupirocin (one gene), pyrazinamide (one gene), and sulfonamides (one gene; Supplementary Table 3). However, prediction of virulence factor genes (VFGs) only identified two VFGs in the genome of PPE-GX01-2022 (VFG046458 associated with bacterial adherence and invasion; VFG002891 similar to bacterial O-antigen; Supplementary Table 4). Notably, none of these ARGs or VFGs were situated on the plasmids.

Phylogenetic analysis using the whole genome sequence data revealed that PPE-GX01-2022 was closely related to *L. intracellularis* isolates from the other regions in the world (Figure 3A). Since the currently available commercial PPE attenuated vaccine (Enterisol<sup>®</sup> Ileitis; Boehringer Ingelheim) was developed based on a US clinical isolate, we compared the complete genome sequence of PPE-GX01-2022 against those of the two US clinical isolates N343 (GenBank accession no. NC\_020127) and PHE/MN1-00 (GenBank accession no. NC\_008011), respectively (Figures 3B, C). The results demonstrated that the genome sequence of PPE-GX01-2022 shared a 99.93% average nucleotide identity to those of N343 and PHE/MN1-00 (Table 4). Interestingly,



the genome sequences of the three plasmids carried by PPE-GX01-2022 were also highly homologous (average nucleotide identity > 99.9%) to those of the plasmids carried by N343 and PHE/MN1-00, respectively. Pan-genome analysis revealed that *L. intracellularis* isolates PPE-GX01-2022, N343, and PHE/MN1-00 shared a total of 1,378 core genes as well as 27, 1, and 3 isolate-specific genes, respectively (Figure 3D; Supplementary Table 5). Most of the specific genes carried by PPE-GX01-2022 encoded hypothetic proteins (16 genes), and the remaining 11 genes encoded proteins participating in translation, ribosomal structure and biogenesis (three genes), carbohydrate transport and metabolism (two genes),

inorganic ion transport and metabolism (two genes), energy production and conversion (one gene), replication, recombination and repair (one gene), cell wall/membrane/envelope biogenesis (one gene), and function unknown (one gene), according to the clusters of orthologous groups of proteins (COG) functional analysis (Supplementary Table 5).

### Discussion

*L. intracellularis* is an obligately intracellular bacterium and is very difficult to be routinely cultured *in vitro*; therefore, the



confirmation of L. intracellularis infection mainly relied on several alternative methods instead of bacterial culture (1). Amongst using laboratory-based PCR assays confirming L. intracellularis DNA in fecal or intestinal mucosal samples is one of the widely used method (4). In this study, we developed a TaqMan qPCR for detecting L. intracellularis by targeting the bacterial 16S rDNA gene. It is worthy of note that the 16S rDNA gene is a target commonly used for the development of DNA-based diagnostic methods of L. intracellularis infection in many studies (13, 18-20). A laboratory evaluation revealed that the method developed in this study displayed good specificity on detecting L. intracellularis, and in particular, the qPCR results for detecting clinical samples were identical with those detected by a previously published qPCR method (17). These findings suggest that the qPCR method we developed in this study could be applied as an effective alternative method for L. intracellularis detection in clinical investigations.

In this study, we investigated two groups of samples. The first group included fecal samples randomly collected from healthy pigs at different ages in 47 Chinese pig farms, and the investigation revealed a high farm positive rate of L. intracellularis [93.6% (95% CI: 65.3-94.4%)]. While few studies have reported the prevalence of L. intracellularis based on the detection of bacterial DNA on pig farms in China, two seroprevalence studies using ELISA assays detecting specific antibodies against L. intracellularis have demonstrated a high bacterial seropositivity (over 70%) (15, 16). From another article published in Chinese, the farm positive rate of L. intracellularis on 21 pig farms in China was as high as 100% (21). The above findings suggest that the prevalence of L. intracellularis infections might be common on Chinese pig farms, and more attentions should be paid on the prevention and control of the disease. This might be very important as the Chinese government issued a policy (Ministry of Agriculture and Rural Affairs Announcement No. 194, 07-10-2019) banning the addition of antibiotics to feed to promote animal growth on July 1st, 2020, for the purpose of combating the global rapid increasing of antimicrobial resistance. It is noteworthy that the infection

Strains	PPE-GX01- 2022	N343	PHE/MN1-00
GenBank accession numbers		NC_020127	NC_008011
Total genome sizes (bp)	1,699,039	1,719,192	1,719,014
GC Content (%)	32.93	33.09	33.09
Number of plasmids	3	3	3
Size of chromosome (bp)	1,439,110	1,457,568	1,457,619
Size of plasmid 1 (bp)	193,063	194,613	194,553
Size of plasmid 2 (bp)	39,799	39,878	39,794
Size of plasmid 3 (bp)	27,067	27,133	27,048
CDSs	1,428	1,432	1,432
tRNAs	44	44	45
rRNAs	6	6	6
Other non-coding RNAs	5	4	4
CRISPR elements	0	0	0
Prophages	0	0	0
Average nucleotic	de identity		
PPE-GX01-2022	100%	99.93%	99.93%
N343	99.93%	100%	99.99%
PHE/MN1-00	99.93%	99.99%	100%

TABLE 4 General features of the complete genome sequences of *L. intracellularis* isolates from swine.

rate of *L. intracellularis* at farm level in several other regions in the word is also high. For example, a previous study investigated the prevalence of *L. intracellularis* infection in 65 Korean herds through indirect immunofluorescence antibody technique and found an infection rate of 100% on pig farms (22). A more recent epidemiological study revealed a positive rate of 90.3% (79.2–100.0%) at farm level in six European countries (Germany, Denmark, Spain, the Netherlands and the United Kingdom) by detecting *L. intracellularis* DNA from fecal samples using qPCR (14).

We also investigated another group of samples in this study. This group of samples included diarrheal and bloody feces from different pig herds (boars, sows, gilts, weaners, finisher pigs and fattening pigs), and these samples were delivered by 125 farms for *L. intracellularis* diagnosis. Investigation on this group of samples demonstrated an average positive rate of 23.83%. This value is also similar to positive rate of a recent epidemiological study performed in six European countries by detecting *L. intracellularis* DNA from fecal samples from diarrheal pigs using qPCR [26.2% (15.9–41.5%)] (14). The above findings indicate that diarrhea caused by *L. intracellularis* should not be ignored. In addition, our screening of fecal samples from

both healthy pigs and/or diarrheal pigs demonstrated a higher positive rate in old pigs than in young pigs. This result agrees well with the report that *L. intracellularis* affects more frequently post-weaned pigs between 6 and 20 weeks of age (1, 23). It is noteworthy that several published studies investigating the seroprevalence of *L. intracellularis* antibody among swine herds in both China and other countries have demonstrated a higher seropositivity in old pigs than in young pigs (16, 24). These findings suggest that old pigs might be at higher risk compared to young pigs.

However, this study also has several limitations. The strict biosecurity actions implemented in pig farms for the purpose of African Swine Fever control increases the difficulty of including more farms in more regions for investigation. Therefore, the survey conducted in this study is a snapshot survey and the data given by this study may not reflect the overall prevalence of *L. intracellularis* on pig farms across China. In addition, a combining use of PCR tests and serological investigation may help to make more solid conclusions. However, collecting blood samples is also not easy at this period. Despite of these limitations, this study still provides insights on the prevalence of *L. intracellularis* on Chinese pig farms, and in the next we intend to conduct a more comprehensive investigation on more farms located on more geographical regions.

As of October 7th, 2022, there are only eight assembled genome sequences of L. intracellularis in NCBI Genome database (3), and none of them is from China. The lack of such information may limit the understanding of genomic characteristics of L. intracellularis prevalent in pigs in China. Among the eight assembled genome sequences publicly available, two genome sequences were complete [isolate N343 from a sow (25) and isolate PHE/MN1-00 from a pig; both of them were isolated in Minnesota, United States (US)]. In this study, we generated the first full genome sequence of L. intracellularis from China. Interestingly, the overall genomic characteristics of the Chinese isolate, including the numbers of plasmids, the sizes of the chromosome and the plasmids, the average GC content, and the numbers of predicted genes, were similar to those of the two US isolates. In addition, the genome sequence of the Chinese isolate was highly identical to those of the two US isolates. These findings suggest that L. intracellularis prevalent in China may genetically be related to those prevalent in the US. A possible reason to explain this result might be the frequent import of breeding pigs of China from the United States and Europe, and the pig genetics, bacterial populations and other pig husbandry features of China may therefore largely the same as elsewhere in the global pig industry (26). Indeed, a recent study has demonstrated that L. intracellularis isolates associated with disease of pigs in different geographical regions in the world belong to a genetically monomorphic clonal lineage (27).

### Conclusion

In summary, we developed a TaqMan qPCR for the detection of *L. intracellularis* by targeting the bacterial 16S rDNA gene. Using the TaqMan qPCR developed herein, we investigated the prevalence of *L. intracellularis* in pigs belonging to different ages on 47 farms in China. The investigation results indicated a high prevalence rate of *L. intracellularis* on Chinese pig farms. We also generated the



first complete genome sequence of *L. intracellularis* from China, and sequence comparisons showed that the genome sequence of the Chinese isolate was highly homologous to those of the field isolates from the United States.

## **Methods**

### Primer design, standard plasmid construction, and TaqMan qPCR development

To develop a TaqMan qPCR for the detection of *L. intracellularis*, we designed a pair of primers (LI-F: 5'-CG<u>GGATCCCCGATCTAAGAGGATAATC-3'</u> [*Bam*HI]; LI-R: 5'-CG<u>GAATTC</u>GATCCAAAGACCTTCATC-3' [*Eco*RI]) based on the bacterial 16S rRNA gene (GenBank accession no. U30147). Genomic DNA of *L. intracellularis* was extracted from a PPE commercial live vaccine (Enterisol<sup>®</sup> Ileitis; Boehringer Ingelheim, St. Joseph, Missouri) using a commercial bacterial DNA preparation kit (TIANGEN, Beijing, China). The target sequence (138 bp) of LI-F/LI-R was amplified from the genomic DNA of *L. intracellularis* through PCR performed in a 25-µl reaction volume containing 1-µl of the template DNA, 1.5-µl of the forward/reverse primer (10 pmol/ul), 12.5-µl of 2×Phanta Flash Master Mix (Vazyme, Nanjing, China), and 8.5-µl nucleotide free water. PCR running conditions were 95°C for 5 min, followed by 35 cycles of 95°C for 30s, 57°C for 35 s, and 72°C for 35s, with a final extension at 72°C for 10 min. PCR product with correct size (138 bp) was confirmed by Sanger sequencing (Supplementary Text 1). Afterwards, the PCR product and the pcDNA3.1 vector were double digested by *Bam*HI/*Eco*RI (TAKARA, Tokyo, Japan), and they were ligated using a DNA ligase (TAKARA, Tokyo, Japan). Finally, the standard plasmid pcDNA-LI (6.73 × 10<sup>10</sup> copies/µl) was generated by transforming the connected product into *E. coli* DH5 $\alpha$  (TAKARA, Tokyo, Japan).

To optimize the amplification conditions for qPCR, the reaction was performed in a 25- $\mu$ l volume, which contains template DNA 1- $\mu$ l (2.69 × 10<sup>9</sup> copies/ $\mu$ l), AceQ<sup>®</sup> Uniwersal U+ Probe Master Mix (Vazyme, Nanjing, China) 12.5- $\mu$ l, each of the forward and reverse primers (0.1, 0.2, 0.25, 0.5, 0.75, or 1.0  $\mu$ M), the TaqMan probe (FAM-CACACTGGAACTGGAACACG-TAMRA;

0.05, 0.1, 0.2, 0.25, or 0.5  $\mu$ M), and nuclease-free water up to 25- $\mu$ l. PCR assay was performed on a CFX96 Touch Real-Time PCR Detection System (Bio-Rad, Hercules, CA) with the following conditions: 95°C for 10 min, followed by 40 cycles of 95°C for 15 s, annealing at different temperatures 60°C for 60 s. Fluorescence was recorded at 50°C for 2 min.

To construct the standard curve, a series of 10-fold dilutions of pcDNA-LI (6.73  $\times$  10<sup>9</sup>-6.73  $\times$  10<sup>5</sup> copies/µl) were used as the template to perform the qPCR assay, which was performed in a 25- $\mu$ l volume containing 2- $\mu$ l of the template DNA, AceQ<sup>®</sup> Uniwersal U<sup>+</sup> Probe Master Mix (Vazyme, Nanjing, China) 12.5-µl, each of the forward and reverse primers  $0.5 \,\mu$ M (final concentration), the TaqMan probe 0.25 µM (final concentration), and nucleasefree water up to 25-µl. PCR was performed on an CFX96 Touch Real-Time PCR Detection System (Bio-Rad, Hercules, CA) with the following conditions: 95°C for 10 min, followed by 40 cycles of 95°C for 15 s, annealing at different temperatures 60°C for 60 s. Fluorescence was recorded at 50°C for 2 min. Standard curves were generated based on the cycle threshold (Ct) values and the copy numbers (lg values) of the template DNA. Coefficients of determination  $(R^2)$  were calculated using GraphPad Prism v. 8.0.1 (https://www.graphpad.com/scientific-software/prism/). The experiment was also performed three times independently to verify the repeatability. The detection limit was calculated following the formula  $Ct = -3.547 \times \lg DNA$  copies + 40.520 ( $R^2 = 0.9999$ ).

To evaluate the sensitivity, a series of 10-fold dilutions of pcDNA-LI (6.73 × 10<sup>5</sup>-6.73 × 10 copies/µl) were used as the template to perform the qPCR assay. To evaluate the specificity, genomic DNAs of *Escherichia coli* (EC), *Salmonella* (Salm), *Glaesserella parasuis* (GPS), *Brachyspira hyodysenteriae* (Bh), *Streptococcus suis* (Ss), Pseudorabies virus (PRV), Japanese encephalitis virus (JEV), as well as DNAs extracted from intestinal tissues of healthy pigs were used as the templates to perform the qPCR assays. Genomic DNA of *L. intracellularis* and nuclease-free water were used as positive and negative controls, respectively. The efficacy of the qPCR detecting *L. intracellularis* was also compared with that of a qPCR published previously (17).

### Investigation of clinical samples

To investigate the prevalence of L. intracellularis in pigs in China, we randomly collected 891 fecal samples from apparently healthy pigs at different ages in 47 farms [seven sow farms (size ranged from 500 sows per farm to 5,000 sows per farm) and 38 fattening farms (size ranged from 600 pigs per farm to 5,000 pigs per farm); Supplementary Table 1] in five provinces located in south China (Guangxi, Guangdong), northeast China (Liaoning), central China (Hubei), and east China (Shanghai). The DNA of L. intracellularis in these samples was investigated using the TaqMan qPCR developed in this study. In addition, a total of 705 diarrheal or bloody feces from symptomatic pigs delivered from 125 Chinese farms (Supplementary Table 2) were also investigated. Total DNA was extracted from the samples using a DNA/RNA Extraction Kit (Vazyme, China) and was used as the template DNA for the PCR, which was performed in a 25-µl volume containing 2-µl of the template DNA, AceQ<sup>®</sup> Uniwersal U<sup>+</sup> Probe Master Mix (Vazyme, China) 12.5- $\mu$ l, each of the forward and reverse primers 0.5  $\mu$ M (final concentration), the TaqMan probe 0.25  $\mu$ M (final concentration), and nuclease-free water up to 25- $\mu$ l. PCR was performed on an CFX96 Touch Real-Time PCR Detection System (Bio-Rad, Hercules, CA) with the following conditions: 95°C for 10 min, followed by 40 cycles of 95°C for 15 s, annealing at different temperatures 60°C for 60 s. Fluorescence was recorded at 50°C for 2 min.

# Next-genetration sequencing and bioinformatical analysis

To generate the complete genome sequence of L. intracellularis, the ileal sample collected from a pig at 120 days of age died from L. intracellularis infection on a farm in Guangxi Province in China was selected for DNA extraction using a DNA/RNA Extraction Kit (Vazyme, China). After quantification using the TBS-380 fluorometer (Turner BioSystems Inc., Sunnyvale, CA), DNA with high quality (OD260/280 = 1.8-2.0, >6 ug) was used for sequencing library preparation. The complete genome sequence of L. intracellularis PPE-GX01-2022-Li was sequenced using a combination of PacBio RS and Illumina sequencing platforms. For Illumina sequencing, DNA libraries constructed using an Illumina TruSeq<sup>TM</sup> Nano DNA Sample Prep Kit following the manufacturer's recommendations were sequenced on an Illumina NovaSeq 6000 platform. We also generated 20 kb insert whole genome shotgun libraries which were sequenced on a Pacific Biosciences RS instrument using standard methods. Briefly, an aliquot of 8 µg DNA was spun in a Covaris g-TUBE (Covaris, MA) at 6,000 rpm for 60s using an Eppendorf 5424 centrifuge (Eppendorf, NY). DNA fragments were then purified, end-repaired and ligated with SMRTbell sequencing adapters following manufacturer's recommendations (Pacific Biosciences, CA). Resulting sequencing libraries were purified three times using 0.45 × volumes of Agencourt AMPure XPbeads (Beckman Coulter Genomics, MA) following the manufacture's recommendations.

After sequencing, a total of 58,848.9 Mb raw data was generated by the Illumina strategy. The raw paired end reads were trimmed and quality controlled by Trimmomatic (http://www.usadellab.org/ cms/uploads/supplementary/Trimmomatic) (28). Following this step, a total of 53,967.8 Mb clean data (Q30: 96.72%) was obtained, and this data was used to evaluate the complexity of the genome and correct the PacBio long reads. PacBio sequencing yielded 17,450,119,530 bp reads (N<sub>50</sub>, 8,267 bp; N<sub>90</sub>, 6,864 bp) for the samples. Next, the sequence data was assembled using ABySS 2.0 (http://www.bcgsc.ca/platform/bioinfo/software/abyss) (29) with multiple-Kmer parameters, and the PacBio corrected long reads were assembled using Canu (https://github.com/ marbl/canu) (30). Subsequently, GapCloser (https://sourceforge. net/projects/soapdenovo2/files/GapCloser/) (31) was applied to fill up the remaining local inner gaps and correct the single base polymorphism for the final assembly results. The complete circle of the genome was drawn with Circos v0.64 (http://circos.ca/) (32).

The complete genome sequence of PPE-GX01-2022 was annotated using RAST Server (33). Gene functions were also annotated against the COG database (http://www.ncbi.nlm.nih.

gov/COG) (34) by blastp module. Antimicrobial resistance genes (ARGs) and virulence factor genes (VFGs) were determined by searching the genome sequence against the CARD database (https://card.mcmaster.ca/) (35) and VFDB database (http:// www.mgc.ac.cn/VFs/) (36), respectively. Prophages and CRISPR elements were determined by PHASTER (37) and CRISPRfinder (38), respectively. Average nucleotide identities (ANI) between two complete genome sequences of PPE-GX01-2022 and the US isolates N343 (GenBank accession no. NC\_020127) and PHE/MN1-00 (GenBank accession no. NC 008011) were calculated using the ANI calculator (http://enve-omics.ce.gatech. edu/ani/) (39). Sequence comparisons were also performed and visualized using EasyFig 2.2.5 (40). Orthologous genes shared by the genomes of PPE-GX01-2022, N343 and PHE/MN1-00, as well as isolate-specific genes possessed by the three isolates were determined using OrthoMCL v.2.0.3 (41) with default parameters. Phylogenetic analysis was performed as previously described (27). To generate a maximum likelihood (ML) tree, the whole genome sequences of Escherichia coli strain E2865 (GenBank accession no. AP018808), Salmonella enterica strain FDAARGOS 878 (GenBank accession no. CP065719), L. intracellularis isolates Fu/JPN (GenBank accession no. QNHO01000000), Ni/JPN (GenBank accession no. QNHN01000000), Ib2/JPN (GenBank QNHM01000000), CBNU010 accession no. (GenBank accession no. JAIPUV01000000), LR189 (GenBank accession PRDD01000000), E40504 (GenBank accession no. no. MTPJ01000000), N343, PHE/MN1-00, and PPE-GX01-2022 were aligned using the MAFFT software (version 7.037) (42) with default parameters. Afterwards, the fasta file obtained from the alignment was input for constructing the ML tree using RAxML (version 8) (43) with the GTR/GAMMA model.

### Statistical analysis

Statistical analysis was performed using "Chi-square test" in the STADA software. P < 0.05 was set as result significant.

### Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://www.ncbi.nlm. nih.gov/genbank/, CP107054; https://www.ncbi.nlm.nih.gov/genbank/, CP107055; https://www.ncbi.nlm.nih.gov/genbank/, CP107056; https://www.ncbi.nlm.nih.gov/genbank/, CP107057.

### **Ethics statement**

The animal study was approved by Institutional Ethics Committees (IECs) of Huazhong Agricultural University (approval number: HZAUSW-2022-0022). The study was conducted in accordance with the local legislation and institutional requirements.

### Author contributions

LW: Data curation, Funding acquisition, Investigation, Methodology, Writing – original draft. WW: Investigation, Methodology, Writing – review & editing. LZ: Methodology, Writing – review & editing. ZZ: Methodology, Writing – review & editing. XY: Methodology, Software, Writing – review & editing. JF: Methodology, Writing – review & editing. HoC: Methodology, Writing – review & editing. WS: Methodology, Writing – review & editing. XH: Methodology, Writing – review & editing. LH: Methodology, Writing – review & editing. PQ: Supervision, Writing – review & editing. HuC: Supervision, Writing – review & editing. ZP: Funding acquisition, Project administration, Supervision, Writing – original draft, Writing – review & editing. BW: Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review & editing.

### Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported in part by National Center of Technology Innovation for Pigs (No. NCTIP-XD/B11), Yingzi Tech and Huazhong Agricultural University Intelligent Research Institute of Food Health (No. IRIFH202209), Hubei Provincial Key Research and Development Program of China (No. 2021BBA085), and the Major Innovation Program of Yangxiang Company (No. 20191230). The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

### Conflict of interest

LW is an employee of Guangxi YangXiang Co., Ltd.

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

## Publisher's note

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

### Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fvets.2024. 1324768/full#supplementary-material

#### SUPPLEMENTARY FIGURE 1

Evaluation of the specificity of the qPCR method developed in this study by generating amplification curves of the method on detecting different pathogens.

#### SUPPLEMENTARY TABLE 1

Numbers of fecal samples collected from healthy pigs at different ages in 47 Chinese farms.

#### SUPPLEMENTARY TABLE 2

Numbers of fecal samples from diarrheal pigs at different ages in 125 Chinese farms.

### References

1. Vannucci FA, Gebhart CJ, McOrist S. Proliferative enteropathy. *Dis Swine*. (2019) 58:898–911. doi: 10.1002/9781119350927.ch58

2. McOrist S, Smith SH, Green LE. Estimate of direct financial losses due to porcine proliferative enteropathy. *Vet Rec.* (1997) 140:579–81. doi: 10.1136/vr.140.22.579

3. NCBI. Genome Assembly and Annotation Report of Lawsonia intracellularis. (2022). Available online at: https://www.ncbi.nlm.nih.gov/genome/browse/#!/ prokaryotes/1101/ (accessed October 7, 2022).

4. Campillo M, Smith SH, Gally DL, Opriessnig T. Review of methods for the detection of *Lawsonia intracellularis* infection in pigs. *J Vet Diagn Invest.* (2021) 33:621–31. doi: 10.1177/10406387211003551

5. Hands I, McOrist S, Blunt R, Lawrence K. Current infection patterns of porcine proliferative enteropathy in Great Britain and the Republic of Ireland. *Vet Rec.* (2010) 167:343–4. doi: 10.1136/vr.c3763

6. Chouet S, Prieto C, Mieli L, Veenhuizen MF, McOrist S. Patterns of exposure to *Lawsonia intracellularis* infection on European pig farms. *Vet Rec.* (2003) 152:14–7. doi: 10.1136/vr.152.1.14

7. Nathues H, Holthaus K, Grosse Beilage E. Quantification of *Lawsonia intracellularis* in porcine faeces by real-time PCR. *J Appl Microbiol.* (2009) 107:2009–16. doi: 10.1111/j.1365-2672.2009.04389.x

8. Richter B, Ladinig A, Nedorost N, Weissenböck H. A TaqMan quantitative polymerase chain reaction assay for the detection of *Lawsonia intracellularis* in fecal and tissue samples from pigs. *J Vet Diagn Invest.* (2010) 22:70–3. doi: 10.1177/104063871002200112

9. Ståhl M, Kokotovic B, Hjulsager CK, Breum S, Angen Ø. The use of quantitative PCR for identification and quantification of *Brachyspira pilosicoli*, *Lawsonia intracellularis* and *Escherichia coli* fimbrial types F4 and F18 in pig feces. *Vet Microbiol.* (2011) 151:307–14. doi: 10.1016/j.vetmic.2011.03.013

10. Willems H, Reiner G. A multiplex real-time PCR for the simultaneous detection and quantitation of *Brachyspira hyodysenteriae*, *Brachyspira pilosicoli* and *Lawsonia intracellularis* in pig faeces. *Berl Munch Tierarztl Wochenschr.* (2010) 123:205–9.

11. Wattanaphansak S, Gebhart CJ, Anderson JM, Singer RS. Development of a polymerase chain reaction assay for quantification of *Lawsonia intracellularis*. *J Vet Diagn Invest.* (2010) 22:598–602. doi: 10.1177/104063871002200416

12. Zmudzki J, Szczotka A, Podgórska K, Nowak A, Grzesiak A, Dors A, et al. Application of real-time PCR for detection of *Lawsonia intracellularis* and *Brachyspira hyodysenteriae* in fecal samples from pigs. *Pol J Vet Sci.* (2012) 15:267–73. doi: 10.2478/v10181-011-0144-2

13. Lindecrona RH, Jensen TK, Andersen PH, Møller K. Application of a 5' nuclease assay for detection of *Lawsonia intracellularis* in fecal samples from pigs. *J Clin Microbiol.* (2002) 40:984–7. doi: 10.1128/JCM.40.3.984-987.2002

14. Arnold M, Crienen A, Swam H, von Berg S, Jolie R, Nathues H. Prevalence of *Lawsonia intracellularis* in pig herds in different European countries. *Porcine Health Manag.* (2019) 5:31. doi: 10.1186/s40813-019-0137-6

15. Xiao N, Li J, Li M, Zhou H, Lin H, Fan H. Isolation and *in vitro* cultivation of *Lawsonia intracellularis* from China. *Vet Microbiol.* (2022) 270:109457. doi: 10.1016/j.vetmic.2022.109457

16. Wu Z, Ling Y, Tian D, Pan Q, Heegaard PM, He C. Seroprevalence of *Lawsonia intracellularis* antibodies in intensive pig farms in China. *BMC Vet Res.* (2014) 10:100. doi: 10.1186/1746-6148-10-100

17. China AQSIQ. *Quarantine Protocol for Real-Time PCR of Lawsona intracellularis* (*SN/T 3488-2013*). Chengdu: Sichuan Entry-Exit Examination and Quarantine Bureau (2013).

18. La T, Collins AM, Phillips ND, Oksa A, Hampson DJ. Development of a multiplex-PCR for rapid detection of the enteric pathogens *Lawsonia intracellularis*,

SUPPLEMENTARY TABLE 3

Antimicrobial resistance genes possessed by *L. intracellularis* PPE-GX01-2022.

SUPPLEMENTARY TABLE 4 Virulence factor genes possessed by *L. intracellularis* PPE-GX01-2022.

SUPPLEMENTARY TABLE 5 Strain-specific genes possessed by *L. intracellularis* PPE-GX01-2022, PHE/MN1-00, or N343.

SUPPLEMENTARY TEXT 1 Nucleotide sequence amplified from the genomic DNA of *L. intracellularis* with using the primers designed in this study.

Brachyspira hyodysenteriae, and Brachyspira pilosicoli in porcine faeces. Lett Appl Microbiol. (2006) 42:284–8. doi: 10.1111/j.1472-765X.2005.01831.x

19. Li Y, Wang J, Wang J, Liu L, Zhang R, Shi R, et al. A real-time loop-mediated isothermal amplification method for rapid detection of *Lawsonia intracellularis* in porcine fecal samples. *J Microbiol Methods.* (2018) 151:62–5. doi: 10.1016/j.mimet.2018.06.002

20. Pławińska J, Jakubowski T, Rzewuska M, Binek M. Occurrence of *Lawsonia intracellularis* and *Brachyspira* spp. infection in swine suffering from diarrhea. *Pol J Vet Sci.* (2004) 7:203–6.

21. Qu XY, Yao HC, Zhang X, Huang T, Li XH. Epidemiological survey of porcine ileitis on intensive swine farms in East China. *Chin J Prev Vet Med.* (2012) 34:873–7. doi: 10.3969/j.issn.1008-0589.2012.11.06

22. Lee SW, Kim TJ, Park SY, Song CS, Chang HK, Yeh JK, et al. Prevalence of porcine proliferative enteropathy and its control with tylosin in Korea. *J Vet Sci.* (2001) 2:209–12. doi: 10.4142/jvs.2001.2.3.209

23. Pascua C, Costinara L, Merneaa I, Tătara D, Hermana V. Prevalence of *Lawsonia intracellularis* Infections in Pig Herds from the Western Romania. Agri Agri Sci Proc. (2015) 6:378-81. doi: 10.1016/j.aaspro.2015.08.100

24. Che TN. Seroprevalence of Lawsonia intralence of Lawsonia intracellularis Antibody Among Swine Acellularis Antibody Among Swine Herds in Thailand, December 2021-March 2022. (Theses and Dissertations), Chulalongkorn University, Bangkok, Thailand (2022).

25. Sait M, Aitchison K, Wheelhouse N, Wilson K, Lainson FA, Longbottom D, et al. Genome sequence of *Lawsonia intracellularis* strain N343, isolated from a sow with hemorrhagic proliferative enteropathy. *Genome Announc.* (2013) 1:13. doi: 10.1128/genomeA.00027-13

26. Megens H-J, Crooijmans RP, Cristobal MS, Hui X, Li N, Groenen MA. Biodiversity of pig breeds from China and Europe estimated from pooled DNA samples: differences in microsatellite variation between two areas of domestication. *Genet Select Evol.* (2008) 40:1–26. doi: 10.1186/1297-9686-40-1-103

27. Bengtsson RJ, Wee BA, Yebra G, Bacigalupe R, Watson E, Guedes RMC, et al. Metagenomic sequencing of clinical samples reveals a single widespread clone of *Lawsonia intracellularis* responsible for porcine proliferative enteropathy. *Microb Genom.* (2020) 6:358. doi: 10.1099/mgen.0.000358

28. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics*. (2014) 30:2114–20. doi: 10.1093/bioinformatics/btu170

29. Jackman SD, Vandervalk BP, Mohamadi H, Chu J, Yeo S, Hammond SA, et al. ABySS 20: resource-efficient assembly of large genomes using a Bloom filter. *Genome Res.* (2017) 27:768–77. doi: 10.1101/gr.214346.116

30. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. CANU: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* (2017) 27:722–36. doi: 10.1101/gr.215087.116

31. Xu M, Guo L, Gu S, Wang O, Zhang R, Peters BA, et al. TGS-GapCloser: a fast and accurate gap closer for large genomes with low coverage of error-prone long reads. *Gigascience*. (2020) 9:giaa094. doi: 10.1093/gigascience/giaa094

32. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, et al. Circos: an information aesthetic for comparative genomics. *Genome Res.* (2009) 19:1639–45. doi: 10.1101/gr.092759.109

33. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics*. (2008) 9:75. doi: 10.1186/1471-2164-9-75

34. Tatusov RL, Galperin MY, Natale DA, Koonin EV. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* (2000) 28:33–6. doi: 10.1093/nar/28.1.33

35. Alcock BP, Huynh W, Chalil R, Smith KW, Raphenya AR, Wlodarski MA, et al. CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Res.* (2023) 51:D690–9. doi: 10.1093/nar/gkac920

36. Chen L, Yang J, Yu J, Yao Z, Sun L, Shen Y, et al. VFDB: a reference database for bacterial virulence factors. *Nucleic Acids Res.* (2005) 33:D325-8. doi: 10.1093/nar/gki008

37. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, et al. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* (2016) 44:W16-21. doi: 10.1093/nar/gkw387

38. Grissa I, Vergnaud G, Pourcel C. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* (2007) 35:W52–7. doi: 10.1093/nar/gkm360

39. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM, et al. hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol.* (2007) 57:81–91. doi: 10.1099/ijs.0.64483-0

40. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. *Bioinformatics*. (2011) 27:1009–10. doi: 10.1093/bioinformatics/btr039

41. Li L, Stoeckert CJ Jr, Roos DS. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* (2003) 13:2178–89. doi: 10.1101/gr.1224503

42. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* (2002) 30:3059–66. doi: 10.1093/nar/gkf436

43. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogeneis. *Bioinformatics.* (2014) 30:1312–3. doi: 10.1093/bioinformatics/btu033