Check for updates

OPEN ACCESS

EDITED BY Rodolfo García-Contreras, National Autonomous University of Mexico, Mexico

REVIEWED BY

Stephen Dela Ahator, UiT The Arctic University of Norway, Norway A W A Chathura Wikumpriya Gunasekara, Pukyong National University, Republic of Korea

*CORRESPONDENCE Fenfen Dong dongfenf@mail.ustc.edu.cn

RECEIVED 14 February 2025 ACCEPTED 20 March 2025 PUBLISHED 10 April 2025

CITATION

Zhang P, Wu X, Ji L, Yan W, Chen L and Dong F (2025) Comparative pan-genomic analysis reveals pathogenic mechanisms and genomic plasticity in *Vibrio parahaemolyticus* clinical and environmental isolates. *Front. Cell. Infect. Microbiol.* 15:1574627. doi: 10.3389/fcimb.2025.1574627

COPYRIGHT

© 2025 Zhang, Wu, Ji, Yan, Chen and Dong. 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.

Comparative pan-genomic analysis reveals pathogenic mechanisms and genomic plasticity in *Vibrio parahaemolyticus* clinical and environmental isolates

Peng Zhang, Xiaofang Wu, Lei Ji, Wei Yan, Liping Chen and Fenfen Dong*

Microbiology Laboratory, Huzhou Center for Disease Control and Prevention, Huzhou, Zhejiang, China

Introduction: *Vibrio parahaemolyticus* is a human pathogen capable of inducing bacterial gastroenteritis. Clinical strains of *V. parahaemolyticus* are considered pathogenic due to their possession of hemolysin and a type III secretion system (T3SS). Some environmental isolates are also acquiring corresponding virulence genes.

Methods: This study initially examines the infection characteristics of *V*. *parahaemolyticus*, and subsequently employs pan-genomic analysis to identify genes that exhibit significant differences in distribution between environmental and clinical isolates, thereby revealing their potential impact on virulence.

Results and discussion: The epidemiological analysis of clinical isolates suggests that infections of *V. parahaemolyticus* are more prevalent in warm seasons, with O4:KUT serotype presenting more severe symptoms. OrthoFinder analysis revealed that environmental isolates possess a higher number of core genes. PEPPAN and KEGG analysis revealed that the 10 genes exclusively found in clinical isolates were predominantly associated with virulence. Additionally, the functions of genes differentially distributed in the environment were significantly more diverse compared to those in clinical settings. Analysis of mobile genetic elements suggested that environmental isolates harbor more mobile genetic elements, implying a potential for an increased number of resistance genes. The pathogenic characteristics of the strains examined in this study, genomic diversity and variation in mobile genetic elements are highly significant for deepening our understanding of the pathogenic mechanisms of *V. parahaemolyticus* and for the development of strategies to prevent its infections.

KEYWORDS

Vibrio parahaemolyticus, pan-genome, mobile genetic element, pathogenic characteristic, genomic plasticity

1 Introduction

Vibrio parahaemolyticus, commonly found in estuarine and marine environments globally, is recognized as one of the most prevalent foodborne pathogens (Letchumanan et al., 2014). Consumption of raw or undercooked seafood can result in acute gastroenteritis, wound infections, and septicemia (Newton et al., 2012). While gastroenteritis may be self-limiting, infections can progress to sepsis, posing a life-threatening risk for individuals with a history of the disease (Letchumanan et al., 2014).

V. parahaemolyticus is classified through serotyping, which is determined by the combination of somatic (O) and capsular (K) antigens. Specifically, V. parahaemolyticus can be categorized into 13 O-antigen and 71 K-antigen types (Iguchi et al., 1995). In February 1996, a food poisoning outbreak caused by a novel O3: K6 serotype emerged in India, which rapidly spread and evolved into a pandemic clone (Okuda et al., 1997). Following the 1996 outbreak of the O3:K6 strain, V. parahaemolyticus has increasingly become one of the leading foodborne pathogens worldwide (Hara-Kudo et al., 2012). The O4:KUT serotype first appeared in Zhejiang Province in 2013 and quickly became the predominant serotype that year (Zhu et al., 2020). Since 2017, V. parahaemolyticus O4: KUT has remained the dominant serotype. However, in 2020, the O10:K4 serotype was detected for the first time in Huzhou City, eventually surpassing O3:K6 to become the new dominant serotype (Zhang et al., 2022). Similar trends have been observed in other regions, including Beijing (Huang et al., 2022b), Guangzhou (He et al., 2022), and Guangxi (Huang et al., 2022a). Notably, Thailand reported its first case of O10:K4 infection in 2021, isolated from an inpatient with acute diarrhea, confirming the emergence of the O10: K4 serotype in Southeast Asia (Okada et al., 2023).

V. parahaemolyticus is naturally prevalent in marine environments. However, recent observations indicate a rise in the incidence of V. parahaemolyticus in inland cities across China, potentially attributable to the contamination of freshwater fish with this bacterium (Lei et al., 2020). Over the past few years, V. parahaemolyticus has been isolated from various freshwater food sources, including crayfish, fish, shrimp, and sediments (Jiang et al., 2020; Chen et al., 2021). The majority of V. parahaemolyticus strains isolated from environmental sources are typically nonpathogenic, in contrast to clinical isolates, which often represent pathogenic clones. Pathogenic strains of V. parahaemolyticus are known to harbor virulence factors, including the genes encoding thermostable direct hemolysin (tdh) and tdh-related hemolysin (trh), whereas most environmental isolates lack these virulence genes (Martinez-Urtaza et al., 2008; Flores-Primo et al., 2014; Haley et al., 2014; Letchumanan et al., 2015). Our previous studies have demonstrated significant differences in the serotypes and virulence genes between clinical and environmental isolates of V. parahaemolyticus (Zhang et al., 2024).

The objective of this study is to explore the genomic characteristics of pathogenic *V. parahaemolyticus* isolates from both clinical and environmental sources, analyze the differences and relationships between them, and elucidate the underlying factors contributing to the pathogenicity of clinical isolates. We

collected 69 clinical isolates from patients and 59 isolates from environment. Employing pan-genomic approach, we identified distinctive features among genomes categorized by host origin. Our preliminary analysis has delineated genomic differences between clinical and environmental isolates, shedding light on the genetic basis of pathogenicity in clinical isolates.

2 Materials and methods

2.1 Bacterial isolate

The sequencing data have been deposited in the NCBI database under the BioProject accession number PRJNA1071824. Detailed information regarding the isolate has been presented in our previous article (Zhang et al., 2024). Metadata including gender, patient age, hospitalization status and main symptoms were collected. Mobile genetic elements, including plasmids and insertion sequences (IS), were identified using the online software MobileElementFinder version 1.0.3, which is accessible at https://cge.food.dtu.dk/services/ MobileElementFinder/ (Johansson et al., 2021). PhiSpy (Akhter et al., 2012) was used with its default parameters, employing GBK files (GenBank files) annotated by PROKKA (Seemann, 2014) as input to predict the presence of prophages. The histogram was generated using KaleidaGraph version 4.5.0. Multi-categorical alluvial diagrams were generated using RawGraphs 2.0 (Mauri et al., 2017). Virulence factors were predicted from the assembled genomes of each isolate using the Virulence Factor Database (VFDB) VFanalyzer tool, available at http://www.mgc.ac.cn/cgibin/VFs/v5/main.cgi?func=VFanalyzer. The heatmap was generated using the "pheatmap" package in R software.

2.2 Comparative genomics and phylogenetic analysis

The whole-genome DNA sequences from all samples were converted into protein sequences using gffread version 0.12.7 (Pertea and Pertea, 2020). Subsequently, these protein sequences were clustered with OrthoFinder version 2.5.5 (Emms and Kelly, 2019). The bar chart and associated plots were generated using KaleidaGraph version 4.5.0.

2.3 Pangenome analysis

Assembled contigs' FASTA files were initially annotated using PROKKA (Seemann, 2014). For the pangenome analysis, PEPPAN (Zhou et al., 2020) was employed with the default settings, utilizing the GFF3 files annotated by PROKKA as inputs for the pipeline. Subsequently, the PEPPAN_parser script was applied to generate the final output. We comparatively analyzed each identified gene from clinical and environmental sources. Statistical analysis was performed using the chi-square test or Fisher's exact test. Given the extensive number of genes, we focused our analysis on those with a

10.3389/fcimb.2025.1574627

p-value less than 0.01, indicating significant differences in distribution. To clarify the distribution patterns, we use the following notation: C-Yes to indicate the presence of a gene in clinical isolates, C-No to denote its absence in clinical isolates, E-Yes to represent its presence in environmental isolates, and E-No to signify its absence in environmental isolates. Those genes exhibiting significant differences in distribution underwent COG (Cluster of Orthologous Groups) annotation using the eggNOG 4.5 server (Huerta-Cepas et al., 2016). Pathway analysis of enriched COG classes was explored using the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database (Kanehisa and Goto, 2000). The R packages "dotplot" and "clusterProfile" were used to analyze and visualize the KEGG analysis. The R packages "ggplot " and "clusterProfile" were used to analyze and visualize the Gene Ontology (GO) analysis.

2.4 Phylogenetic analysis

We constructed a phylogenetic tree using methods described in previous literature (Zhang et al., 2024). A phylogenetic tree was constructed based on 69 single nucleotide polymorphisms (SNPs) identified among clinical isolates. Variant calling was performed against the V. parahaemolyticus RIMD 2210633 (GCF_000196095.1) genome using the Snippy pipeline (https://github.com/tseemann/ snippy). Core SNP alignment was generated using snippy-core v.4.6.0, and recombinant regions were filtered out from the alignment using Gubbins v.3.0.0 (Croucher et al., 2015). SNPs were subsequently extracted from the alignment using SNP-sites (Page et al., 2016). The core SNP alignment was used to infer a phylogenetic tree with IQ-TREE, employing the maximum likelihood method and the TVMe+ASC nucleotide substitution model (Kalyaanamoorthy et al., 2017; Minh et al., 2020). Finally, the tree was visualized using FigTree v.1.4.4. We employed the snpdists v0.8.2 software to convert FASTA format files into a matrix of SNP distances

2.5 Statistical analysis

Statistical processing was performed using either the chi-square test or Fisher's exact test in R software. The test selection depended on sample size and occurrence frequency. Chi-square test was used for samples \geq 40 with frequencies \geq 5. Fisher's exact test was applied for smaller samples or frequencies \leq 5. The Odds Ratio (OR) was used to assess the distribution patterns of each gene between clinical and environmental isolates.

3 Results

3.1 Patient characteristics

To study the characteristics of patients infected with V. parahaemolyticus, we collected metadata from 2020 to 2023 for

patients with diarrhea. The data included gender, age, hospitalization status, main symptoms, and onset time, and was gathered from a total of 9,413 patients. Among these patients analyzed, the infection rates were 3.07% (151/4,911) in male patients and 3.53% (159/4,502) in female patients. Statistical analysis revealed no significant gender-based difference in V. parahaemolyticus infection rates (p = 0.24). The age group most affected was 18-40 years old, followed by those aged 41-65. The majority of patients did not require hospitalization; however, 6% of the patients needed hospital care. Hospitalized patients were found across all age groups. The primary symptoms reported were nausea, vomiting, abdominal pain, and diarrhea. In severe cases, patients also experienced fever and dehydration (Figure 1A). We categorize symptoms such as dehydration, fatigue, and fever as systemic and more severe, while considering digestive system symptoms like abdominal pain and diarrhea to be milder in comparison. We compared the clinical characterizations among patients infected with O10:K4, O3:K6, and O4:KUT serotypes and discovered that the phenotype associated with O4:KUT was significantly more severe than those associated with O10:K4 and O3:K6 (Figure 1B). The months with the highest incidence of illness were May to October (Figure 1A).

3.2 Evolution of clinical isolates and patient symptoms analysis

To further investigate the pathogenicity of clinical isolates, we utilized the sequenced isolates for additional analysis. Initially, we constructed an evolutionary tree for the clinical isolates and observed that a cluster of O4:KUT demonstrated increased phylogenetic distance compared to other isolates (Figure 2; Supplementary Table S1). This O4:KUT cluster was further divided into two distinct branches (Figure 2). While there was minor variation in the number of SNPs within each O4:KUT branch, significant differences were observed between branches. Additionally, substantial SNP differences exist between O4:KUT and other serotypes (Supplementary Table S1). To gain a clearer understanding of the patient's clinical symptoms, we have categorized them into four severity levels: level 1 for abdominal pain and diarrhea, level 2 for abdominal pain, diarrhea, and vomiting, level 3 for abdominal pain, diarrhea, vomiting, and fever, and level 4 for the additional presence of thirst and fatigue. We found that the majority of patients with severe symptoms were infected with the O4:KUT serotype. The disease infection was not associated with gender (Figure 2). To investigate the underlying reasons for the more severe clinical symptoms associated with the O4:KUT serotype compared to O3:K6 and O10:K4, we predicted the virulence genes in these clinical isolates. For clarity, we focused only on virulence genes that exhibited differential distribution patterns across the isolates, excluding those with uniform distribution. Our analysis revealed differences in the distribution of six virulence genes, tcpA, wbfU, wbfY and wecA were predominantly found in the O4:KUT serotype (Figure 2).



3.3 Orthologs identification

Next, we intended to investigate the genomic differences between clinical and environmental isolates. To identify orthologous groups within the coding genes of both clinical and environmental isolates, we submitted 128 samples to OrthoFinder to identify orthologous groups. Our objective was to ascertain the distribution of gene families among these isolates. As shown in Figure 3A, the majority of non-core gene families are not broadly distributed. Furthermore, we investigated the differences in the distribution of these orthogroups between clinical and environmental isolates. OrthoFinder identified 13,003 orthogroups in both clinical and environmental isolates. We determined that 12.84% (1,670) of these orthologous groups were exclusively present in all clinical isolates, while 29.81% (3,876) were exclusively present in all environmental isolates. Furthermore, we observed that 65.55% (8,523) of the orthologous groups were present in at least two clinical isolates, while 53.91% (7,010) were found in at least two environmental isolates. Additionally, 21.61% (2,810) of the orthologous groups were unique to all clinical isolates, and 16.28% (2,117) were unique to all environmental isolates. In our analysis of the 128 isolates, we identified core orthogroups comprising 11.92% (1,550) of the total, unique orthogroups accounting for 35.39% (4,602), and accessory orthogroups representing 52.69% (6,851) (Figures 3B-D).

3.4 Pan-genomic analysis

Following our analysis, we sought to explore the differences between clinical and environmental isolates at the individual gene level. For this pan-genomic analysis, we employed the software

PEPPAN. The PEPPAN output revealed a total of 15,783 genes among all isolates. The strict core gene set, present in 100% of the sequences, comprised 3,712 genes. Additionally, 83 genes were identified as core genes present in 99%-100% of the sequences. The soft core gene category, shared by 95%-99% of the isolates, included 102 genes. The accessory gene set, also known as the shell genes, which are shared by 15% to less than 95% of the strains, contained 875 genes. Conversely, the cloud genes, found in 0% to less than 15% of the strains, numbered 11,011 in total (Supplementary Figure S1). A total of 1,677 genes showed statistically significant differences in their distribution patterns between clinical and environmental isolates (Supplementary Table S2). Figure 4 displays the distribution of 259 genes with distinct names. Among the 259 genes identified, 120 were found predominantly in clinical isolates, while 139 were primarily detected in environmental isolates (Supplementary Table S3). We found that 10 genes were exclusively distributed in clinical isolates, and these genes were identified as atpB, sctC, yscU, ISVsa5, ISVch8, ssaV, yopJ, spaP, pdeL and tdh2. Among the 1677 genes examined, only 464 genes corresponded to the KO (KEGG Orthology) numbers. Of these, 226 genes showed a distribution bias toward clinical isolates, while 238 genes were more prevalent in environmental isolates (Supplementary Table S4). The KEGG results indicate that clinical isolates exhibit differentially enriched pathways, including transcription factors and bacterial secretion systems, compared to environmental isolates. In contrast, environmental isolates show distinct pathway enrichment, glycosyltransferases and lipopolysaccharide biosynthesis proteins, relative to clinical isolates (Figure 5A, B). We further analyzed the differentially distributed genes by mapping them to the COG database. Among these, only 40 genes could be mapped to COG categories, with 18 genes showing a bias toward clinical isolates and 22 genes exhibiting a preference for environmental isolates

S0102 S0102 S0104 OKUUT make O S0104 OKUUT make F S0104 OKUUT make F S0104 OKUUT make F OKUUT make F F S0107 OKUUT make F OKUUT make F F S0108 OKUUT make F OKUUT make F F S0108 OKUUT make F F F F F S0108 OKUUT make F </th <th></th> <th>Reference</th> <th>Main Symptom and Signs</th> <th>s Serotype</th> <th>Gender</th> <th>Age</th> <th>Virulence genes</th>		Reference	Main Symptom and Signs	s Serotype	Gender	Age	Virulence genes
S0124 0.00 CAUUT famale 9 S0124 CAUUT make 9 3 S0124 CAUUT make 9 3 S0181 CAUUT make 8 4 4 S0181 CAUUT famale 8 4 4 S0181 CAUUT famale 8 4 4 4 S0184 CAUUT famale 3 4 <	Г	S0182	4.00	O4:KUT	male	60	
S0179 4.00 0.4/UT male 73 S0187 6.00 0.4/UT male 73 8 8 S0187 6.00 0.4/UT male 73 8 8 8 S0180 0.0 0.4/UT male 73 8 8 8 8 S0180 0.4/UT male 33 8		S0124	1.00	O4:KUT	female	56	
S0164 6.0 OKUT make 6.0 0.4 0.0		S0179	4.00	O4:KUT	male	57	
S0181 50181 50 CAKUT male 65 64 64 65 S0181 S0 CAKUT ferale 70 64 <		S0184	4.00	O4:KUT	male	73	
30181 5019 0.4KUT female 48 48 30189 2.00 0.4KUT female 23 48 48 30189 2.00 0.4KUT female 23 48 48 30180 2.00 0.4KUT female 23 48 48 30180 2.00 0.4KUT female 23 48 48 30181 200 0.4KUT female 43 48 48 S2354 200 0.4KUT female 40 48 48 S2353 200 0.4KUT female 40 48 48 S2353 130 0.20 6.6 female 40 48 48 48 S2354 130 0.36 female 14 48 48 48 48 48 48 48 48 48 48 48 48 48 48 48 48 48 48		S0187	4.00	O4:KUT	male	65	
Solles 2200 CARUT female 20 CARUT female 20 SS164 C200 CARUT female 22 CARUT female 22 SS164 C200 CARUT female 22 CARUT female 23 CARUT female 20 CARUT female 21 CARUT female 21 CARUT female 21 CARUT female 21 CARUT female		S0181	4.00	O4:KUT	female	48	
30.10 2.00 0,4,001 Hemaio 20 0,4,001 Hemaio 20 25244 200 0,4,001 Hemaio 30 0,4,001 Hemaio 40 50 50 0,4,001 Hemaio 40 10 0,4,001 Hemaio 50 50 0,4,001 Hemaio 50 50 0,4,001 Hemaio 50 50 0,3,00 Hemaio 50 10 0,3,60 Hemaio 31 0,3,00 Hemaio 50 50 0,3,00 Hemaio 50 10		S0189	4.00	O4:KUT	female	76	
Construction Construction<		S2646	2.00	04:KUT	temale	25	
S2942 300 0.440.77 rase 23 10		S2544	4.00	O4:KUT	temale male	32	
Special Special <t< td=""><th></th><th>S2642</th><td>3.00</td><td>04.KUT</td><td>male</td><td>23</td><td></td></t<>		S2642	3.00	04.KUT	male	23	
S5235 2.00 0.4XUT Finade 43 S2394 4.00 0.4XUT Finade 64 S2393 4.00 0.4XUT Finade 63 S01813 4.00 0.4XUT Finade 53 64 S1596 1.00 0.4X0 Finade 53 64 64 S1596 1.00 0.4X0 Finade 64 64 64 S0131 1.00 0.4X0 Finade 53 64 64 S0134 1.00 0.4X0 Finade 53 64 64 S0134 1.00 0.4X0 Finade 53 64 64 S0261 2.00 0.4X0 finade 53 64 64 S0338 3.00 0.4X0 finade 53 64 64 S1475 1.00 0.0X4 finade 53 64 64 S1475 1.00 0.0X4 finade 53 64 64 S1475 1.00 0.0X4 finade 53		S2661	4.00	O4:KUT	male	43	
S2339 400 OHKUT fermale 60 80 80 80 S2339 400 OHKUT fermale 55 80 80 80 S0131 2.00 OKK0 fermale 34 80 90 S0131 2.00 OKK0 fermale 34 90		S2535	2.00	O4:KUT	female	45	
S2399 400 OKKUT female 66 S2393 400 OKKUT female 55 S0131 200 OSK6 fmale 54 S1596 100 OSK6 fmale 14 60 S0131 100 OSK6 fmale 14 60 60 S0134 100 OSK6 fmale 51 60 60 60 S0134 100 OSK6 fmale 53 60		S2394	4.00	O4:KUT	female	40	
1 S0183 4.00 O.KIUT female 35 50183 2.00 O.SK6 female 55 50195 100 O.SK6 female 56 50194 100 O.SK6 female 56 50194 100 O.SK6 female 56 51958 3.00 O.SK6 male 21 56 52641 2.00 O.SK6 female 53 66 50261 2.00 O.SK6 female 53 66 502621 2.00 O.SK6 female 53 66 502621 2.00 O.SK6 female 53 66 50185 4.00 O.SK6 female 54 66 50465		S2399	4.00	O4:KUT	female	59	
SU1813 2.00 OK MUT Fendle 54 SU131 2.00 OX MS fendle 34 SU131 2.00 OX MS fendle 54 SU134 1100 OX MS fendle 60 SU2401 100 OX MS fendle 53 SU2441 2.00 OX MS fendle 53 SU2441 2.00 OX MS fendle 53 SU2641 2.00 OX MS fendle 53 SU2641 2.00 OX MS fendle 53 SU2623 2.00 OX MS fendle 54 64 SU265 1.00 O1044 male 76 64 SU265 1.00 O1044 male 56 64 64 64 SU265 1.00 O1044 male 56 64		S2393	4.00	O4:KUT	female	35	
S0131 2.00 03K6 fmaile 34 S1266 1.00 03K6 fmaile 64 S0134 1.00 03K6 fmaile 21 6 S1566 1.00 03K6 maile 21 6 6 S1544 2.00 03K6 maile 65 6		S0183	4.00	O4:KUT	female	55	
0.1.000 1.000 0.38 bit maile 60 0.0140 1.000 0.38 bit maile 60 0.0134 1.000 0.38 bit maile 60 0.0140 0.200 0.38 bit maile 51 0.02401 2.000 0.38 bit maile 53 4 4 4 0.02401 2.000 0.38 bit female 53 4 <th></th> <th>S1596</th> <td>2.00</td> <td>03:K6</td> <td>female</td> <td>34</td> <td></td>		S1596	2.00	03:K6	female	34	
S0134 100 O3.68 minite 21 S1589 400 O3.68 male 21 S2243 300 O3.68 male 51 S2241 200 O3.68 female 53 S2240 200 O3.68 female 53 S22622 200 O3.68 female 53 S0185 S2222 200 O3.68 female 53 S0185 S0185 GMAR female 25 50 S0185 S0185 O10.64 fmale 26 50 S1475 100 O10.64 fmale 28 50 S0437 100 O10.64 fmale 28 51 S1404 200 O10.64 fmale 28 51 S0467 200 O10.64 fmale 28 50 S0467 200 O10.64 female 30 61 S0467 200		S0240	1.00	03:K6	male formal-	41	
S1589 100 0.368 male 32 S2243 300 0.368 male 52 S2241 2.00 0.368 female 33 S2241 2.00 0.368 female 34 S2240 2.00 0.368 female 46 S2222 2.00 0.368 female 46 S2222 2.00 0.368 female 35 S0185 4.00 0.368 female 35 S0185 4.00 0.368 female 35 S0185 4.00 0.01644 male 25 S0185 4.00 0.01644 male 34 S1517 4.00 01044 male 37 S0433 1.00 01044 male 32 4 4 S0447 2.00 01044 female 32 4 4 4 S0467 2.00 01044 female 32 4 4 4 4 4 4 4 4 4		S0134	1.00	03:K0	remale male	09 21	
S2243 3.8 0.348 male 55 S2240 2.00 0.348 fermale 63 S0261 2.00 0.348 fermale 63 S22622 2.00 0.348 fermale 53 S22623 2.00 0.348 fermale 53 S0185 2.00 0.348 fermale 53 S0185 3.00 0.348 fermale 54 S1475 1.00 0.1044 male 26 S1523 4.00 0.1044 male 28 S1474 1.00 0.1044 male 28 S1523 4.00 0.1044 fermale 38 S0433 1.00 0.1044 fermale 38 S0467 2.00 0.1044 fermale 51 S0468 2.00 0.1044 fermale 51 S0463 1.00 0.1044 fermale 51 S0463 1.00 0.1044 fermale 51 S0466 1.00 0.1044 fe		S1589	4.00	O3:K6	male	32	
S2641 200 03.68 female 63 S0261 200 03.68 female 64 S2223 200 03.68 female 63 S2222 200 03.68 female 22 S0388 300 03.68 female 22 S0388 400 03.68 female 22 S0185 400 03.68 female 25 S0185 400 03.68 female 26 S1475 100 01644 female 27 S1473 100 01644 female 37 S1523 400 01644 female 38 S1542 200 01644 female 38 S0467 200 01644 female 31 S0467 200 01644 female 32 S0468 200 01644 female 32 S0469 200 01644 female 32 S0469 200 01644 female 32		S2543	3.00	O3:K6	male	55	
S02640 2.00 03K6 female 63 S02621 2.00 03K6 female 53 S02622 2.00 03K6 female 53 S0338 30360 03K6 female 22 S0338 30300 03K6 female 22 S0338 30300 03K6 female 22 S0338 30300 03K6 female 24 S1475 100 010K4 male 26 S1473 100 010K4 male 26 S1473 100 010K4 male 26 S1442 100 010K4 female 51 S0433 100 010K4 female 51 S0467 2.00 010K4 female 51 S0467 2.00 010K4 female 63 S0468 2.00 010K4 female 63 S0469 2.00 010K4 female 63 S0468 1000 010K4 female 63 <th>-</th> <th>S2641</th> <td>2.00</td> <td>O3:K6</td> <td>female</td> <td>53</td> <td></td>	-	S2641	2.00	O3:K6	female	53	
S0261 2.00 O3K6 male 46 S2622 2.00 O3K6 female 21 S0383 300 O3K6 female 22 S0185 2000 O3K6 female 22 S0185 2000 O3K6 female 25 S2005 1000 O10K4 male 35 S1424 100 O10K4 male 34 S1424 100 O10K4 male 37 S1475 0100 O10K4 male 38 S2442 200 O10K4 male 31 S1504 2.00 O10K4 female 33 S0470 1000 O10K4 female 31 S0467 2.00 O10K4 female 46 S0467 2.00 O10K4 female 42 S0463 2.00 O10K4 female 42 S0463 2.00 O10K4 female 42 S0463 2.00 O10K4 male 51		S2640	2.00	03:K6	female	63	
S2623 2.00 0.3k6 female 23 S0338 3.00 0.3k6 female 25 S0185 4.00 0.3k6 female 75 S1475 1.00 0.1k4 male 36 S1533 4.00 0.1k4 male 38 S0433 1.00 0.1k4 male 38 S2542 4.00 0.1k4 male 38 S0467 2.00 0.1k4 female 63 S0467 0.00 0.1k4 female 63 S0468 2.00 0.1k4 female 42 S0468 2.00 0.1k4 female 42 S0468 2.00 0.1k4 female 42 S0468 1.00 0.1k4 female 42		S0261	2.00	O3:K6	male	46	
S2822 2.00 O3:K6 female 22 S01338 S0100 O3:K6 female 75 S11475 1.00 O10:K4 female 35 S1424 1.00 O10:K4 female 36 S1424 1.00 O10:K4 female 36 S1424 1.00 O10:K4 female 36 S1523 G.00 O10:K4 female 37 S0433 1.00 O10:K4 female 38 S2542 G.00 O10:K4 female 63 S0470 1.00 O10:K4 female 63 S0467 2.00 O10:K4 female 63 S0467 2.00 O10:K4 female 63 S0468 2.00 O10:K4 female 63 S0468 2.00 O10:K4 female 63 S0463 1.00 O10:K4 female 63 S0463 1.00 O10:K4 female 63 S0391 1.00 O10:K4 <td< td=""><th></th><th>S2623</th><td>2.00</td><td>O3:K6</td><td>female</td><td>53</td><td></td></td<>		S2623	2.00	O3:K6	female	53	
S0185 3.00 03.06 male 25 S1475 1.00 O10.44 fmale 35 S1475 1.00 O10.44 fmale 35 S1475 1.00 O10.44 fmale 35 S1424 1.00 O10.44 fmale 37 S1523 0.00 O10.44 male 37 S0433 1.00 O10.44 fmale 38 S1524 0.00 O10.44 fmale 38 S1647 2.00 O10.44 fmale 68 S1504 2.00 O10.44 female 63 S0467 1.00 O10.44 female 63 S0468 2.00 O10.44 female 60 S0468 2.00 O10.44 female 61 0 S0468 2.00 O10.44 female 62 0 0 S0463 1.00 O10.44 female 63 0 0 0 S0463 1.00 O10.44 female 63		S2622	2.00	O3:K6	female	22	
S1475 0.03 K4 male 75 S2605 1.00 016 K4 male 95 S1475 1.00 016 K4 male 96 S1424 1.00 016 K4 male 98 S1523 4.00 016 K4 male 24 9 S1517 4.00 016 K4 male 38 9 9 9 S0433 1.00 016 K4 male 38 9		50338	3.00	O3:K6	male	25	
S2605 100 010K4 female 34 S1424 100 010K4 male 34 S1523 30 010K4 male 34 S1517 400 010K4 male 37 S0433 100 010K4 male 38 S0470 100 010K4 female 63 S0467 200 010K4 female 63 S0468 200 010K4 female 63 S0465 100 010K4 female 64 S0465 100 010K4 female 65 S0465 100 010K4 female 54 44 S0465 100 010K4 female 53<		S1475	4.00	O3:K6	female	75	
S1424 00 010K4 male 34 S1523 400 010K4 male 34 S1523 400 010K4 male 34 S0433 100 010K4 male 36 S0433 100 010K4 male 38 S1504 200 010K4 female 63 S0467 200 010K4 female 33 S1473 300 010K4 female 63 S0468 2.00 010K4 female 63 S0463 100 010K4 female 63 S0453 2.00 010K4 female 63 S0466 100 010K4 female 63 S0391 100 010K4 female 53 S0391 100 010K4 female 53		S2605	1.00	010:K4	famale	30 60	
S1523 400 O10.K4 male 28 S1517 400 O10.K4 male 37 S2542 400 O10.K4 male 38 S1517 100 O10.K4 male 38 S2542 400 O10.K4 female 63 S0470 100 O10.K4 female 63 S0467 2.00 O10.K4 female 63 S0467 2.00 O10.K4 female 63 S0468 2.00 O10.K4 female 63 S0468 2.00 O10.K4 female 63 S0463 100 O10.K4 female 63 S0463 100 O10.K4 female 63 S0466 100 O10.K4 female 64 S0453 1.00 O10.K4 female 65 S0454 100 O10.K4 male 53 S0331 100 O10.K4 female 65 S2722 100 O10.K4 female		S1424	1.00	010:K4	male	34	
S1517 4.00 O10.K4 male 37 S0433 1.00 O10.K4 male 38 S1504 2.00 O10.K4 female 66 S0470 1.00 O10.K4 female 63 S0477 2.00 O10.K4 female 33 S0467 2.00 O10.K4 male 51 S0468 2.00 O10.K4 male 61 S0469 2.00 O10.K4 male 63 S0468 2.00 O10.K4 female 63 S0469 2.00 O10.K4 female 64 S0463 1.00 O10.K4 male 2 S0463 2.00 O10.K4 female 66 S0463 1.00 O10.K4 male 2 S0466 1.00 O10.K4 male 67 S0391 1.00 O10.K4 male 53 S0391 1.00 O10.K4 male 53 S0364 1.00 O10.K4 male 53		S1523	4.00	O10:K4	male	26	
S0433 1.00 O10.K4 male 38 S2542 4.00 O10.K4 male 58 S0470 1.00 O10.K4 male 63 S0470 1.00 O10.K4 male 12 S0477 0.00 O10.K4 male 12 S0467 2.00 O10.K4 female 60 S0468 2.00 O10.K4 female 61 S0468 2.00 O10.K4 female 63 S0468 2.00 O10.K4 female 64 S0468 2.00 O10.K4 female 64 S0463 1.00 O10.K4 female 64 S0453 2.00 O10.K4 female 67 S0466 1.00 O10.K4 male 2 0 S0466 1.00 O10.K4 male 51 0 S1474 3.00 O10.K4 male 53 0 0 S2722 1.00 O10.K4 female 53 0 0 <th></th> <th>S1517</th> <td>4.00</td> <td>O10:K4</td> <td>male</td> <td>37</td> <td></td>		S1517	4.00	O10:K4	male	37	
S2542 400 O10:K4 female 56 S1504 200 O10:K4 female 56 S0470 100 O10:K4 female 51 0 0 S0467 2.00 O10:K4 female 33 0 0 0 S0468 2.00 O10:K4 female 60 0 <td< td=""><th></th><th>S0433</th><td>1.00</td><td>O10:K4</td><td>male</td><td>38</td><td></td></td<>		S0433	1.00	O10:K4	male	38	
S1004 2.00 O10:K4 female 63 S0470 1.00 O10:K4 male 51 60 S1473 3.00 O10:K4 male 51 60 S0467 2.00 O10:K4 male 51 60 S0468 2.00 O10:K4 female 63 60 60 S0463 1.00 O10:K4 female 63 60 60 60 S0463 1.00 O10:K4 female 63 60 <t< td=""><th></th><th>S2542</th><td>4.00</td><td>O10:K4</td><td>female</td><td>58</td><td></td></t<>		S2542	4.00	O10:K4	female	58	
S0470 1.00 O10:K4 male 12 S0467 2.00 O10:K4 female 33 S0468 2.00 O10:K4 female 63 S0469 2.00 O10:K4 female 63 S0469 2.00 O10:K4 female 63 S0463 1.00 O10:K4 female 63 S0463 1.00 O10:K4 female 64 S0463 1.00 O10:K4 male 29 0 0 S0466 1.00 O10:K4 male 67 0 0 0 S0431 1.00 O10:K4 male 66 0		S1504	2.00	O10:K4	female	63	
S0437 2.00 O10 K4 temale 33 S0468 2.00 O10 K4 temale 63 S0468 2.00 O10 K4 temale 63 S1441 2.00 O10 K4 temale 63 S0463 1.00 O10 K4 temale 63 S0463 1.00 O10 K4 temale 22 0 0 S0466 1.00 O10 K4 temale 22 0 0 0 S0466 1.00 O10 K4 male 57 0 0 0 S0466 1.00 O10 K4 male 67 0 0 0 S1474 3.00 O10 K4 temale 53 0 0 0 S2491 2.00 O10 K4 temale 53 0 0 0 0 S2722 1.00 O10 K4 temale 53 0 0 0 S0360 1.00 O10 K4 temale 53 0 0 0 S0360 1.00		50470	1.00	010:K4	male	12	
S0465 2.00 010.K4 female 61 S0469 2.00 010.K4 female 63 S0463 1.00 010.K4 female 63 S0463 1.00 010.K4 female 22 1.0 1.0 S0463 1.00 010.K4 female 22 1.0 1.0 1.0 S0466 1.00 010.K4 male 29 1.0		S1473	2.00	010:K4	temale	33	
S0469 2.00 O10:K4 female 63 S1484 2.00 O10:K4 female 36 S0463 1.00 O10:K4 male 20 0 S0463 1.00 O10:K4 male 20 0 0 S0453 2.00 O10:K4 male 20 0 0 S0391 1.00 O10:K4 male 67 0 0 S0391 1.00 O10:K4 female 66 0 0 0 S1474 3.00 O10:K4 female 50 0		S0468	2.00	010:K4	female	51	
S1484 2.00 010:K4 female 36 S0463 1.00 010:K4 female 42 S0463 2.00 010:K4 male 29 S0466 1.00 010:K4 male 2 S0466 1.00 010:K4 male 2 1 1 S1474 3.00 010:K4 male 66 1 1 1 S1474 3.00 010:K4 female 66 1		S0469	2.00	010:K4	female	63	
S0463 1.00 010.K4 female 42 S0463 2.00 010.K4 male 29 S0466 1.00 010.K4 male 29 S1474 3.00 010.K4 male 67 S0391 1.00 010.K4 male 66 S2724 1.00 010.K4 female 53 S2724 1.00 010.K4 female 58 S2722 1.00 010.K4 female 53 S2722 1.00 010.K4 female 53 S2722 1.00 010.K4 female 53 S3360 1.00 010.K4 female 53 S0341 4.00 010.K4 female 43 S0360 1.00 010.K4 male 45 S22659 4.00 010.K4 male 45 S2398 4.00 04.KUT female 63 S2398 4.00 04.KUT female 63 S2581 2.00 010.K4 female		S1484	2.00	O10:K4	female	36	
S0453 2.00 010:K4 male 29 0 0 S0466 1.00 010:K4 male 2 0 0 0 S1474 3.00 010:K4 male 67 0		S0463	1.00	010:К4	female	42	
S0466 1.00 010:K4 male 2 0 0 S0391 1.00 010:K4 male 67 0 0 S0391 1.00 010:K4 female 66 0 0 0 S2491 2.00 010:K4 female 53 0 0 0 S2724 1.00 010:K4 female 53 0 0 0 S2722 1.00 010:K4 female 53 0 0 0 S2660 4.00 010:K4 female 53 0 0 0 S0364 1.00 010:K4 male 44 0		S0453	2.00	O10:K4	male	29	
S14/4 3.00 O10:K4 male 67 S0391 1.00 O10:K4 female 66 S2491 2.00 O10:K4 female 50 S2724 1.00 O10:K4 female 58 S2724 1.00 O10:K4 female 58 S2722 1.00 O10:K4 female 58 S2660 4.00 O10:K4 female 53 S0341 4.00 O10:K4 female 53 S0364 1.00 O10:K4 male 44 S0360 1.00 O10:K4 male 45 S0360 1.00 O10:K4 male 35 S1472 3.00 O10:K4 male 45 S2398 4.00 O4:KUT female 63 S2391 1.00 O10:K4 female 63 S2395 1.00 O10:K4 female 65 S2581 1.00 O10:K4 female 52 S2581 1.00 O10:K4 female <th></th> <th>S0466</th> <td>1.00</td> <td>O10:K4</td> <td>male</td> <td>2</td> <td></td>		S0466	1.00	O10:K4	male	2	
S0391 1.00 O10:K4 female 66 S2491 2.00 O10:K4 female 50 S2724 1.00 O10:K4 male 53 S2722 1.00 O10:K4 male 53 S2724 1.00 O10:K4 female 58 S2722 1.00 O10:K4 female 58 S1598 3.00 O10:K4 female 53 S0341 4.00 O10:K4 female 43 S0364 1.00 O10:K4 male 45 S2659 4.00 O10:K4 male 45 S2398 4.00 O4:KUT female 63 S2291 1.00 O10:K4 female 65 S2619 1.00 O10:K4 female 53 S2581 2.00 O10:K4 female		514/4	3.00	O10:K4	male	67	
32491 2.00 010.K4 female 50 S2724 1.00 010.K4 male 53 S2722 1.00 010.K4 male 53 S2660 4.00 010.K4 female 53 S1598 3.00 010.K4 female 53 S0341 4.00 010.K4 female 44 S0364 1.00 010.K4 male 44 S0364 1.00 010.K4 male 45 S2559 4.00 010.K4 male 35 S1472 3.00 010.K4 female 63 S2398 4.00 04.KUT female 63 S2398 4.00 04.KUT female 66 S22619 1.00 010.K4 female 69 S2521 1.00 010.K4 female 53 S2588 1.00 010.K4 female 53 S2589 2.00 010.K4 female 52 S2589 2.00 010.K4 female <th></th> <th>50391</th> <td>1.00</td> <td>O10:K4</td> <td>female</td> <td>66</td> <td></td>		50391	1.00	O10:K4	female	66	
S2722 1.00 O10:K4 male 53 S2660 3.00 O10:K4 female 53 S1598 3.00 O10:K4 male 53 S0341 4.00 O10:K4 female 53 S0364 1.00 O10:K4 male 43 S0364 1.00 O10:K4 male 44 S0360 1.00 O10:K4 male 45 S2659 4.00 O10:K4 male 63 S2659 4.00 O10:K4 male 63 S22659 0.00 O10:K4 female 63 S2398 4.00 O4:KUT female 63 S22619 1.00 O10:K4 female 69 S2585 2.00 O10:K4 female 69 S2585 2.00 O10:K4 female 52 S2581 1.00 O10:K4 female 52 S2589 1.00 O10:K4 female 52 S2589 2.00 O10:K4 male		S2491	2.00	010:K4	female	50	
100.44 icitize 56 100.44 icitize 56 100.44 icitize 56 100.44 icitize 53 100.44 icitize 54 100.44 icitize 54 100.44 icitize 55 100.44 icitize 53		S2722	1.00	010:K4	maie	03 59	
S1598 3.00 O10:K4 female 5.3 S0341 4.00 O10:K4 female 5.3 S0364 1.00 O10:K4 female 5.3 S0360 1.00 O10:K4 male 44 S0360 1.00 O10:K4 male 45 S1472 3.00 O10:K4 male 63 S1472 3.00 O10:K4 female 63 S2398 4.00 O4:KUT female 65 S2395 4.00 O4:KUT female 65 S2585 2.00 O10:K4 female 69 S2585 2.00 O10:K4 female 53 S2588 1.00 O10:K4 female 52 S2588 1.00 O10:K4 female 52 S2589 2.00 O10:K4 fem		S2660	4.00	010:K4	male	29	
S0341 4.00 O10:K4 female 43 S0364 1.00 O10:K4 male 44 S0360 1.00 O10:K4 male 45 S2659 4.00 O10:K4 male 45 S1472 3.00 O10:K4 female 63 S2398 4.00 O4:KUT female 65 S2395 4.00 O4:KUT female 65 S2585 S2619 1.00 O10:K4 female 69 S2585 2.00 O10:K4 female 53 0 0 S2582 1.00 O10:K4 female 53 0 0 0 S2588 1.00 O10:K4 female 53 0 <		S1598	3.00	O10:K4	female	53	
S0364 1.00 010.K4 male 44 S0360 1.00 010.K4 male 45 S2659 3.00 010.K4 male 35 S1472 3.00 010.K4 female 63 S2398 4.00 04:KUT female 63 S2295 0.00 010.K4 female 69 S22619 1.00 010.K4 female 69 S2585 2.00 010.K4 female 69 S2585 2.00 010.K4 female 53 S2588 1.00 010:K4 female 53 S2589 2.00 010:K4 female 52 S2589 2.00 010:K4 female 52 S2589 2.00 010:K4 female 52 S2580 1.00 010:K4 female 52 S2589 2.00 010:K4 male 21 S2584 2.00 010:K4 female 33 S2584 2.00 010:K4 female<		S0341	4.00	O10:K4	female	43	
S0360 1.00 010.K4 male 45 S2659 4.00 010.K4 male 35 S1472 3.00 010.K4 female 63 S2398 4.00 04.KUT female 63 S2398 4.00 04.KUT female 63 S2398 4.00 04.KUT female 65 S2298 1.00 010.K4 male 75 S2585 2.00 010.K4 female 52 S2588 1.00 010.K4 female 52 S2588 1.00 010.K4 female 52 S2589 2.00 010.K4 male 21 S2584 2.00 010.K4 male 39 S2584 2.00 010.K4 mot known<		S0364	1.00	O10:K4	male	44	
S2659 4.00 O10.K4 male 35 0 0 S1472 3.00 O10.K4 female 63 0 0 S2398 4.00 O10.K4 female 63 0 0 0 S2398 4.00 O4.KUT female 65 0 0 0 0 S2585 2.00 O10.K4 female 69 0 0 0 0 S2585 2.00 O10.K4 female 63 0		S0360	1.00	O10:K4	male	45	
Si1472 3.00 O10:K4 female 63 0 0 63 Si2398 0.00 O4:KUT female 33 0 0 0 Si2395 0.00 O4:KUT female 65 0 0 0 Si2395 1.00 O10:K4 female 69 0 0 0 Si2585 2.00 O10:K4 female 69 0 0 0 Si2585 1.00 O10:K4 female 53 0 </td <th></th> <th>S2659</th> <td>4.00</td> <td>O10:K4</td> <td>male</td> <td>35</td> <td></td>		S2659	4.00	O10:K4	male	35	
S2395 4.00 O4:KUT female 33 Image: S2395 Image: S		S1472	3.00	O10:K4	female	63	
32.395 31.00 O4.KUT female 65 S2619 1.00 O4.KUT female 69 S2585 2.00 O10.K4 female 69 S2582 1.00 O10.K4 female 75 S2588 1.00 O10.K4 female 52 S2588 1.00 O10.K4 female 52 S2591 2.00 O10.K4 female 52 S2589 2.00 O10.K4 female 52 S2590 1.00 O10.K4 female 39 S2584 2.00 O10.K4 female 39 S2584 2.00 O10.K4 female 39 S2584 2.00 O10.K4 female 39		S2398	4.00	O4:KUT	female	33	
32019 1.00 010:K4 female 69 S2585 2.00 010:K4 not known not known S2592 1.00 010:K4 female 75 S2588 1.00 010:K4 female 53 S2591 2.00 010:K4 female 52 S2589 2.00 010:K4 female 52 S2590 1.00 010:K4 female 39 S2584 2.00 010:K4 not known not known		52395	4.00	O4:KUT	female	65	
2.00 O10:K4 not known not known not known S2592 1.00 O10:K4 female 75 0 0 0 S2588 1.00 O10:K4 female 53 2 0 0 0 S2589 2.00 O10:K4 female 52 0 0 0 S2589 2.00 O10:K4 female 21 0		S2585	1.00	010:K4	female	69	
I.00 Ol0:K4 remain 75 S2588 1.00 Ol0:K4 female 53 S2591 2.00 Ol0:K4 female 52 S2589 2.00 Ol0:K4 female 52 S2589 2.00 Ol0:K4 male 21 1 S2590 1.00 Ol0:K4 female 39 S2584 2.00 Ol0:K4 not known female 1		S2592	2.00	010:K4	not known	not known	
S2591 2.00 O10.K4 female 52 S2589 2.00 O10.K4 female 52 S2590 1.00 O10.K4 female 52 S2590 1.00 O10.K4 female 39 S2584 2.00 O10.K4 not known not known S2584 2.00 O10.K4 female 39		S2588	1.00	010:K4	remale female	/0 53	
S2589 2.00 O10:K4 male 21 0 0 S2590 1.00 O10:K4 female 39 0		S2591	2.00	010:K4	female	52	
S2590 1.00 O10:K4 female 39 S2584 2.00 O10:K4 not known not known		S2589	2.00	O10:K4	male	21	
S2584 <u>2.00</u> O10:K4 not known not known i bar a bar 		S2590	1.00	010:К4	female	39	
1 (cp <i>pli</i> /1 (cp	L	S2584	2.00	010:К4	not known	not known	
							wc: wbi wbi pilA

FIGURE 2

Phylogenetic tree of clinical isolates. A phylogenetic tree depicting the genetic relationships among 69 clinical isolates, adjacent columns on the right detail serotypes, symptom profiles, gender, age of patients and virulence genes for each isolate. patients were categorized into four severity levels based on clinical manifestations: level 1 for abdominal pain and diarrhea, level 2 for abdominal pain, diarrhea, and vomiting, level 3 for abdominal pain, diarrhea, vomiting, and fever, and level 4 for the additional presence of thirst and fatigue. In the heatmap, the presence of virulence genes is indicated by red, while their absence is represented by blue.

Tree scale: 0.01

(Supplementary Table S5). However, no significant differences were observed in the COG functional annotations of these genes between clinical and environmental isolates (Figures 5C, D).

3.5 Mobile genetic element analysis

The pan-genomic results indicate significant genetic differences between clinical isolates and environmental isolates. Consequently, we characterized the genome plasticity by evaluating mobile gene elements (MGEs). Many MGEs also carry accessory genes. The benefits of accessory gene carriage can be seen in the success of integrons, elements first identified on MGEs that appear adapted for the acquisition, assembly, and expression of accessory genes (Partridge et al., 2009; Lacotte et al., 2017). In this study, we identified insertion sequences and prophages within the isolates. Clinical isolates contained one, three, or four insertion sequences, with over half carrying four such sequences (Figure 6A). Environmental isolates harbored between 0 to 6 insertion sequences (Figure 6B). Specifically, four insertion sequences—namely ISVvu6, ISVpa1, ISVpa2, and ISVal1—were identified in clinical isolates. In contrast, environmental isolates exhibited 11 different insertion sequences, with ISVvu5 being the most frequently detected (Figure 6C).

Regarding prophages, 17% of clinical isolates were found to carry the pp1 prophage, whereas approximately 19% of environmental isolates contained pp1. Additionally, 2% of the isolates carried both pp1 and the pp2 prophage (Figure 6D).



Overall orthogroup statistics and comparison. (A) Bar plot showing the number of orthologous clusters. (B–D) Summary of core, unique, and accessory orthogroups in the genomes of clinical (B), environmental (C) and all isolates (D). The numbers indicate the total counts and percentages of each orthogroup category.



FIGURE 4

Heatmap displaying the pan-genome genes across 128 strains. Strain labels were located at the bottom, and pan-genome gene clusters were indicated on the left. The presence of genes was indicated by red color, while absence was denoted by blue.

4 Discussion

This study found that the age group with the highest infection rate of V. *parahaemolyticus* was between 18 to 40 years old (Figure 1), which was consistent with previous research (Wang

et al., 2021). Individuals over the age of 66 had a lower risk of infection, aligning with previous research findings (Gong et al., 2018; Wang et al., 2021). The diversity in age distribution may reflect natural changes in host immunity (Simon et al., 2015) and alterations in dietary habits associated with aging. Additionally, the





clinical and environmental isolates. (**D**) Proportional distribution of distinct prophage between clinical and environmental isolates.

primary symptoms reported by patients treated in Huzhou First People's Hospital include abdominal pain, diarrhea, nausea, and vomiting. In severe cases, systemic symptoms such as fever and thirst occur, which is consistent with other research findings (Qi et al., 2016). However, the incidence of fever and dehydration is relatively low (Figure 1). Furthermore, we found that most patients infected with the O4:KUT serotype exhibit more severe clinical manifestations (Figures 1, 2). TCP, composed of *tcpA*, is a toxin-

coregulated pilus that mediates the intestinal colonization of V. cholerae (Taylor et al., 1987). wbfU was identified in clinical, shrimp-pathogenic strains but were absent in non-pathogens (Xue et al., 2023). The presence of tcpA and wbfU may enhance the virulence of O4:KUT serotype, although the underlying molecular mechanisms require further investigation. Additionally, the regulatory roles of wbfY and wecA in bacterial pathogenicity remain to be elucidated. This serotype first appeared in Zhejiang province in 2013 and was the most dominant that year (Zhu et al., 2020). According to our long-term monitoring results, *V. parahaemolyticus* O4:KUT has been a local dominant serotype since 2017 (Zhang et al., 2022). Our previous research has identified at least five ST types within the O4:KUT serotype (Zhang et al., 2024). Nearly all O4:KUT isolates in this study are associated with more severe symptoms (Figure 2). In 2021, the positive strains of *V. parahaemolyticus* in Huzhou were almost exclusively O10:K4, leading to a relative decrease in other serotypes, including O4:KUT. Moreover, the number of infections from May to October each year between 2020 and 2023 was significantly higher than that from November to April (Figure 1). This pattern aligns with previous research (Lu et al., 2020), which suggests that warmer seasons are more conducive to contracting *V. parahaemolyticus*.

Using the Orthofinder software, we detected a total of 13,003 orthogroups and found that environmental isolates had a higher proportion of orthogroups belonging to the core orthogroups than clinical isolates. A smaller proportion belonged to unique orthogroups (Figure 3). Bacteria frequently encounter numerous environmental abiotic stresses (heat, cold, osmotic, salt, oxidation, pH, and radiation) and biotic stresses (antimicrobial compounds and microbial toxins) in their natural life cycle (Beales, 2004), and a greater number of core orthogroups may be necessary for survival under such complex environmental pressures. The PEPPAN software analysis revealed 15,783 genes among these 128 isolates, which exceeds the pan-genome size reported in previous studies (Meparambu Prabhakaran et al., 2022). This increase is attributed to the higher number of V. parahaemolyticus isolates included in this study, suggesting that the pan-genomic gene count rises with an increase in the quantity of isolates. However, the difference in the number of core genes, 3,712, compared to previous studies is relatively small, indicating that the core genome is relatively stable. On average, V. parahaemolyticus possesses 4,630 genes (Meparambu Prabhakaran et al., 2022), with core genes constituting over 80% of this total. We found atpB, sctC, yscU, ISVsa5, ISVch8, ssaV, yopJ, spaP, pdeL and tdh2 were exclusively distributed in clinical isolates (Figure 4). YscU, SctC, SsaV, YopJ are secreted proteins or effector proteins associated with the type III secretion system (Login and Wolf-Watz, 2015; Ma and Ma, 2016; Yu et al., 2018), while SpaP is involved in bacterial adhesion (Ma et al., 1991). PdeL is a cyclic di-GMP phosphodiesterase. The bacterial second messenger, c-di-GMP, regulates the formation, movement, cell cycle progression, development, and virulence of bacterial biofilms (Hengge, 2016). Thermostable direct hemolysin (TDH) is a proteinaceous toxin that is considered a major virulence factor associated with the ability of V. parahaemolyticus strains to cause foodborne gastroenteritis (Shimohata and Takahashi, 2010; Letchumanan et al., 2014). The tdh2 gene exhibits 97.2% homology with *tdh1* and has been identified as primarily responsible for the phenotypic expression of hemolytic activity (Bhowmik et al., 2014). Therefore, most of these genes, which are exclusively found in clinical isolates, are associated with virulence.

KEGG pathway enrichment analyses have shown differences between clinical isolates and environmental isolates. Both clinical and environmental isolates share secretion system, transporters, prokaryotic defense system, bacterial motility proteins and biofilm formation (Figure 5). The protein secretion system is crucial for bacterial growth and is involved in a variety of processes. Almost all bacteria possess some secretion systems that secrete various substrates (Green and Mecsas, 2016). The KEGG annotation results indicate that clinical and environmental isolates protect themselves through prokaryotic defense systems. These prokaryotic defense systems can be classified into two broad groups, differing in their modes of action. The first group consists of defense systems that operate on the principle of self-non-self discrimination, with DNA typically being the target of discriminatory recognition; these mechanisms can be considered as forms of prokaryotic immunity (Makarova et al., 2013). Additionally, within this group is the DNA phosphorothioation system (known as the DND system), which marks DNA through phosphorothioation and degrades unmodified DNA (He et al., 2007; Liang et al., 2007; Xu et al., 2010). The ability of Vibrio spp. to adapt to and survive within eukaryotic hosts, as well as to endure varying aquatic environmental conditions, is largely attributed to their capacity to form biofilms (Yildiz and Visick, 2009; Lutz et al., 2013). A biofilm is a structured community of bacterial cells embedded within a self-produced matrix of extracellular polymeric substances (EPS). This matrix facilitates nutrient acquisition and provides a protective barrier against the penetration of antimicrobial agents (Hathroubi et al., 2017). Previous studies showed that the flagellar motility plays an important role in the formation of biofilm in the life cycle of vibrios (Echazarreta and Klose, 2019). Lipopolysaccharide biosynthesis is one pathway that is enriched in environmental isolates. Lipopolysaccharide (LPS) is a type of glycolipid found in the outer leaflet of the outer membrane (OM) of Gram-negative bacteria (Muhlradt and Golecki, 1975; Kamio and Nikaido, 1976). The presence of LPS on the cell surface enhances the barrier function of the OM, which renders many antibiotics ineffective that are typically used to treat infections caused by Gram-positive pathogens (Nikaido, 2003). Numerous transcription factors orchestrate the spatiotemporal regulation of virulence factors in Vibrio vulnificus, integrating diverse environmental cues such as nutrient availability, bacterial cell density, and the presence of antimicrobial agents (Choi and Choi, 2022). Previous study suggested that transcription factor CytR plays an important role in V. cholerae pathogenesis (Das et al., 2020). Therefore, the regulatory pathways mediated by transcription factors in clinical isolates, as opposed to environmental isolates, may play a critical role in determining their pathogenicity. Furthermore, a significant number of genes remain unmatched to KO numbers, and further research is needed to elucidate the enriched pathways associated with these genes.

Previous studies have shown that among these 128 isolates, environmental isolates carry a more diverse array of resistance genes compared to clinical isolates (Zhang et al., 2024). The majority of AMR-encoding genes in *V. parahaemolyticus* are primarily acquired through horizontal gene transfer (HGT) (Pazhani et al., 2021). This study indicates that a greater variety of mobile genetic elements can be

detected in environmental isolates, and a higher proportion of these isolates harbor prophages (Figure 6). Insertion sequence (IS) elements are a class of small mobile genetic elements that are ubiquitously distributed across the genomes of most bacterial species. Accumulating evidence indicates that IS elements play a significant role in mediating insertion mutations, facilitating genome rearrangements, promoting the dissemination of antibiotic resistance genes and virulence factors both within and between species, and modulating the expression of neighboring genes (Mahillon and Chandler, 1998). Previous studies have demonstrated that ISVpa2 exhibits the ability to insert at multiple genomic sites. Furthermore, ISVpa2 has been shown to induce genetic rearrangements, including insertional inactivation of target genes and adjacent deletions. Additionally, ISVpa2 can be horizontally transferred among species within the genus Vibrio (Kamruzzaman and Nishibuchi, 2008). The transfer of MGEs between different bacteria is facilitated by HGT. As a major force in microbial evolution, HGT enables the acquisition of novel functions on a large scale and allows for rapid adaptation to new niches (Wiedenbeck and Cohan, 2011). Previous research has shown that HGT can significantly drive adaptation in food systems and other environments managed by humans (Rossi et al., 2014; Akanni et al., 2015). Previous studies have revealed that ISVal1 facilitates the HGT of prophages, plasmids, and genomic islands among V. parahaemolyticus populations in response to adverse environmental conditions and selective pressures. Additionally, ISVal1 has been shown to mediate the formation of MGEs associated with antibiotic resistance (Fu et al., 2021). ISVsa3 has been identified as being strongly associated with MDR in Salmonella (Lewis et al., 2023). ISVsa5 has been shown to play a crucial role in modulating the expression of *bla*_{CTX-M} in *E. coli* (Black et al., 2024). Further investigation is required to elucidate the roles and mechanisms of other mobile genetic elements (MGEs). Therefore, we hypothesize that the elevated concentration of mobile genetic elements (MGEs) in environmental isolates contributes to enhanced environmental adaptation.

In summary, statistical analysis of patients treated at the First People's Hospital of our city over recent years has revealed that the primary population affected by *V. parahaemolyticus* is aged 18-40, with a higher infection rate observed during warmer seasons. Analysis of infection characteristics revealed that patients infected with the O4: KUT serotype exhibited more severe symptoms. Although clinical isolates had a smaller number of serotypes, their core orthogroups still accounted for a smaller proportion compared to environmental isolates. Pan-genomic analysis results indicated that there was a significant difference in gene expression between clinical and environmental isolates, which may be attributed to the higher presence of mobile genetic elements (MGEs) in environmental isolates.

Data availability statement

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

Ethics statement

The manuscript presents research on animals that do not require ethical approval for their study.

Author contributions

PZ: Funding acquisition, Project administration, Writing – review & editing. XW: Investigation, Methodology, Writing – original draft. LJ: Investigation, Methodology, Writing – original draft. WY: Investigation, Methodology, Writing – original draft. LC: Investigation, Methodology, Writing – original draft. FD: Funding acquisition, Investigation, Methodology, Project administration, Writing – original draft.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by the Huzhou Science and Technology Project (grant number 2023GYB07), the Medical Science and Technology Projects of Zhejiang Province (grant number 2024KY1658), and the Natural Science Foundation of Zhejiang Province (grant number LTGY23C010001).

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.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

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/fcimb.2025.1574627/ full#supplementary-material

10.3389/fcimb.2025.1574627

References

Akanni, W. A., Siu-Ting, K., Creevey, C. J., McInerney, J. O., Wilkinson, M., Foster, P. G., et al. (2015). Horizontal gene flow from Eubacteria to Archaebacteria and what it means for our understanding of eukaryogenesis. *Philos. Trans. R Soc. Lond B Biol. Sci.* 370, 20140337. doi: 10.1098/rstb.2014.0337

Akhter, S., Aziz, R. K., and Edwards, R. A. (2012). PhiSpy: a novel algorithm for finding prophages in bacterial genomes that combines similarity- and composition-based strategies. *Nucleic Acids Res.* 40, e126. doi: 10.1093/nar/gks406

Beales, N. (2004). Adaptation of microorganisms to cold temperatures, weak acid preservatives, low pH, and osmotic stress: A review. *Compr. Rev. Food Sci. Food Saf.* 3, 1–20. doi: 10.1111/j.1541-4337.2004.tb00057.x

Bhowmik, S. K., Pazhani, G. P., and Ramamurthy, T. (2014). Phylogenetic and in silico functional analyses of thermostable-direct hemolysin and tdh-related encoding genes in Vibrio parahaemolyticus and other Gram-negative bacteria. *BioMed. Res. Int.* 2014, 576528. doi: 10.1155/2014/576528

Black, C. A., Benavides, R., Bandy, S. M., Dallas, S. D., Gawrys, G., So, W., et al. (2024). Diverse Role of bla(CTX-M) and Porins in Mediating Ertapenem Resistance among Carbapenem-Resistant Enterobacterales. *Antibiotics (Basel)* 13, 185. doi: 10.3390/antibiotics13020185

Chen, H., Dong, S., Yan, Y., Zhan, L., Zhang, J., Chen, J., et al. (2021). Prevalence and population analysis of vibrio parahaemolyticus isolated from freshwater fish in Zhejiang province, China. *Foodborne Pathog. Dis.* 18, 139–146. doi: 10.1089/fod.2020.2798

Choi, G., and Choi, S. H. (2022). Complex regulatory networks of virulence factors in Vibrio vulnificus. *Trends Microbiol.* 30, 1205–1216. doi: 10.1016/j.tim.2022.05.009

Croucher, N. J., Page, A. J., Connor, T. R., Delaney, A. J., Keane, J. A., Bentley, S. D., et al. (2015). Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res.* 43, e15. doi: 10.1093/nar/gku1196

Das, S., Chourashi, R., Mukherjee, P., Gope, A., Koley, H., Dutta, M., et al. (2020). Multifunctional transcription factor CytR of Vibrio cholerae is important for pathogenesis. *Microbiol. (Reading)* 166, 1136–1148. doi: 10.1099/mic.0.000949

Echazarreta, M. A., and Klose, K. E. (2019). Vibrio flagellar synthesis. Front. Cell Infect. Microbiol. 9, 131. doi: 10.3389/fcimb.2019.00131

Emms, D. M., and Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* 20, 238. doi: 10.1186/s13059-019-1832-y

Flores-Primo, A., Pardio-Sedas, V., Lizarraga-Partida, L., Lopez-Hernandez, K., Uscanga-Serrano, R., and Flores-Hernandez, R. (2014). Seasonal abundance of total and pathogenic Vibrio parahaemolyticus isolated from American oysters harvested in the Mandinga Lagoon System, Veracruz, Mexico: implications for food safety. *J. Food Prot* 77, 1069–1077. doi: 10.4315/0362-028X.JFP-13-482

Fu, S., Yang, Q., Wang, Q., Pang, B., Lan, R., Wei, D., et al. (2021). Continuous genomic surveillance monitored the *in vivo* evolutionary trajectories of vibrio parahaemolyticus and identified a new virulent genotype. *mSystems* 6, e01254-20. doi: 10.1128/mSystems.01254-20

Gong, X. H., Wu, H. Y., Li, J., Xiao, W. J., Zhang, X., Chen, M., et al. (2018). Epidemiology, aetiology and seasonality of infectious diarrhoea in adult outpatients through active surveillance in Shanghai, China, 2012-2016: a cross-sectional study. *BMJ Open* 8, e019699. doi: 10.1136/bmjopen-2017-019699

Green, E. R., and Mecsas, J. (2016). Bacterial secretion systems: an overview. *Microbiol. Spectr.* 4, 10.1128/microbiolspec.VMBF-0012-2015. doi: 10.1128/ microbiolspec.VMBF-0012-2015

Haley, B. J., Kokashvili, T., Tskshvediani, A., Janelidze, N., Mitaishvili, N., Grim, C. J., et al. (2014). Molecular diversity and predictability of Vibrio parahaemolyticus along the Georgian coastal zone of the Black Sea. *Front. Microbiol.* 5, 45. doi: 10.3389/ fmicb.2014.00045

Hara-Kudo, Y., Saito, S., Ohtsuka, K., Yamasaki, S., Yahiro, S., Nishio, T., et al. (2012). Characteristics of a sharp decrease in Vibrio parahaemolyticus infections and seafood contamination in Japan. *Int. J. Food Microbiol.* 157, 95–101. doi: 10.1016/j.ijfoodmicro.2012.04.019

Hathroubi, S., Mekni, M. A., Domenico, P., Nguyen, D., and Jacques, M. (2017). Biofilms: microbial shelters against antibiotics. *Microb. Drug Resist.* 23, 147–156. doi: 10.1089/mdr.2016.0087

He, P., Zhang, J., Zhou, Y., Hou, S., Tao, X., Wang, A., et al. (2022). Genomic characteristics of a notable emerging serotype O10:K4 of Vibrio parahaemolyticus from food-borne cluster events in Guangzhou, China. *J. Infect.* 85, 702–769. doi: 10.1016/j.jinf.2022.09.012

He, X., Ou, H. Y., Yu, Q., Zhou, X., Wu, J., Liang, J., et al. (2007). Analysis of a genomic island housing genes for DNA S-modification system in Streptomyces lividans 66 and its counterparts in other distantly related bacteria. *Mol. Microbiol.* 65, 1034–1048. doi: 10.1111/j.1365-2958.2007.05846.x

Hengge, R. (2016). Trigger phosphodiesterases as a novel class of c-di-GMP effector proteins. *Philos. Trans. R Soc. Lond B Biol. Sci.* 371, 20150498. doi: 10.1098/rstb.2015.0498

Huang, Y., Du, Y., Wang, H., Tan, D., Su, A., Li, X., et al. (2022a). New variant of vibrio parahaemolyticus, sequence type 3, serotype O10:K4, China, 2020. *Emerg. Infect. Dis.* 28, 1261–1264. doi: 10.3201/eid2806.211871

Huang, Y., Lyu, B., Zhang, X., Tian, Y., Lin, C., Shen, L., et al. (2022b). Vibrio parahaemolyticus O10:K4: an emergent serotype with pandemic virulence traits as predominant clone detected by whole-Genome sequence analysis - Beijing municipality, China, 2021. *China CDC Wkly* 4, 471–477. doi: 10.46234/ccdcw2022.106

Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M. C., et al. (2016). eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res.* 44, D286–D293. doi: 10.1093/nar/gkv1248

Iguchi, T., Kondo, S., and Hisatsune, K. (1995). Vibrio parahaemolyticus O serotypes from O1 to O13 all produce R-type lipopolysaccharide: SDS-PAGE and compositional sugar analysis. *FEMS Microbiol. Lett.* 130, 287–292. doi: 10.1111/j.1574-6968.1995.tb07733.x

Jiang, H., Yu, T., Yang, Y., Yu, S., Wu, J., Lin, R., et al. (2020). Co-occurrence of Antibiotic and Heavy Metal Resistance and Sequence Type Diversity of Vibrio parahaemolyticus Isolated From Penaeus vannamei at Freshwater Farms, Seawater Farms, and Markets in Zhejiang Province, China. *Front. Microbiol.* 11, 1294. doi: 10.3389/fmicb.2020.01294

Johansson, M. H. K., Bortolaia, V., Tansirichaiya, S., Aarestrup, F. M., Roberts, A. P., and Petersen, T. N. (2021). Detection of mobile genetic elements associated with antibiotic resistance in Salmonella enterica using a newly developed web tool: MobileElementFinder. *J. Antimicrob. Chemother.* 76, 101–109. doi: 10.1093/jac/ dkaa390

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

Kamio, Y., and Nikaido, H. (1976). Outer membrane of Salmonella typhimurium: accessibility of phospholipid head groups to phospholipase c and cyanogen bromide activated dextran in the external medium. *Biochemistry* 15, 2561–2570. doi: 10.1021/bi00657a012

Kamruzzaman, M., and Nishibuchi, M. (2008). Detection and characterization of a functional insertion sequence, ISVpa2, in Vibrio parahaemolyticus. *Gene* 409, 92–99. doi: 10.1016/j.gene.2007.11.012

Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30. doi: 10.1093/nar/28.1.27

Lacotte, Y., Ploy, M. C., and Raherison, S. (2017). Class 1 integrons are low-cost structures in Escherichia coli. ISME J. 11, 1535–1544. doi: 10.1038/ismej.2017.38

Lei, T., Jiang, F., He, M., Zhang, J., Zeng, H., Chen, M., et al. (2020). Prevalence, virulence, antimicrobial resistance, and molecular characterization of fluoroquinolone resistance of Vibrio parahaemolyticus from different types of food samples in China. *Int. J. Food Microbiol.* 317, 108461. doi: 10.1016/j.ijfoodmicro.2019.108461

Letchumanan, V., Chan, K. G., and Lee, L. H. (2014). Vibrio parahaemolyticus: a review on the pathogenesis, prevalence, and advance molecular identification techniques. *Front. Microbiol.* 5, 705. doi: 10.3389/fmicb.2014.00705

Letchumanan, V., Yin, W. F., Lee, L. H., and Chan, K. G. (2015). Prevalence and antimicrobial susceptibility of Vibrio parahaemolyticus isolated from retail shrimps in Malaysia. *Front. Microbiol.* 6, 33. doi: 10.3389/fmicb.2015.00033

Lewis, G. L., Fenton, R. J., Moriyama, E. N., Loy, J. D., and Moxley, R. A. (2023). Association of ISVsa3 with Multidrug Resistance in Salmonella enterica Isolates from Cattle (Bos taurus). *Microorganisms* 11, 631. doi: 10.3390/microorganisms11030631

Liang, J., Wang, Z., He, X., Li, J., Zhou, X., and Deng, Z. (2007). DNA modification by sulfur: analysis of the sequence recognition specificity surrounding the modification sites. *Nucleic Acids Res.* 35, 2944–2954. doi: 10.1093/nar/gkm176

Login, F. H., and Wolf-Watz, H. (2015). YscU/flhB of yersinia pseudotuberculosis harbors a C-terminal type III secretion signal. J. Biol. Chem. 290, 26282–26291. doi: 10.1074/jbc.M114.633677

Lu, Y., Yang, L., Meng, J., Zhao, Y., Song, Y., Zhu, Y., et al. (2020). Microevolution of vibrio parahaemolyticus isolated from clinical, acute hepatopancreatic necrosis disease infecting shrimps, and aquatic production in China. *Microbes Environ.* 35, ME19095. doi: 10.1264/jsme2.ME19095

Lutz, C., Erken, M., Noorian, P., Sun, S., and McDougald, D. (2013). Environmental reservoirs and mechanisms of persistence of Vibrio cholerae. *Front. Microbiol.* 4, 375. doi: 10.3389/fmicb.2013.00375

Ma, J. K., Kelly, C. G., Munro, G., Whiley, R. A., and Lehner, T. (1991). Conservation of the gene encoding streptococcal antigen I/II in oral streptococcci. *Infect. Immun.* 59, 2686–2694. doi: 10.1128/iai.59.8.2686-2694.1991

Ma, K. W., and Ma, W. (2016). YopJ family effectors promote bacterial infection through a unique acetyltransferase activity. *Microbiol. Mol. Biol. Rev.* 80, 1011–1027. doi: 10.1128/MMBR.00032-16

Mahillon, J., and Chandler, M. (1998). Insertion sequences. *Microbiol. Mol. Biol. Rev.* 62, 725–774. doi: 10.1128/MMBR.62.3.725-774.1998

Makarova, K. S., Wolf, Y. I., and Koonin, E. V. (2013). Comparative genomics of defense systems in archaea and bacteria. *Nucleic Acids Res.* 41, 4360–4377. doi: 10.1093/nar/gkt157

Martinez-Urtaza, J., Lozano-Leon, A., Varela-Pet, J., Trinanes, J., Pazos, Y., and Garcia-Martin, O. (2008). Environmental determinants of the occurrence and distribution of Vibrio parahaemolyticus in the rias of Galicia, Spain. *Appl. Environ. Microbiol.* 74, 265–274. doi: 10.1128/AEM.01307-07

Mauri, M., Elli, T., Caviglia, G., Uboldi, G., and Azzi, M. (2017). RAWGraphs: A Visualisation Platform to Create Open Outputs. In *Proceedings of the 12th Biannual Conference on Italian SIGCHI Chapter (CHItaly '17)*. New York, NY, USA: Association for Computing Machinery, 1–5. doi: 10.1145/3125571.3125585

Meparambu Prabhakaran, D., Patel, H. R., Sivakumar Krishnankutty Chandrika, S., and Thomas, S. (2022). Genomic attributes differ between Vibrio parahaemolyticus environmental and clinical isolates including pathotypes. *Environ. Microbiol. Rep.* 14, 365–375. doi: 10.1111/1758-2229.13000

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., et al. (2020). Corrigendum to: IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37, 2461. doi: 10.1093/molbev/msaa131

Muhlradt, P. F., and Golecki, J. R. (1975). Asymmetrical distribution and artifactual reorientation of lipopolysaccharide in the outer membrane bilayer of Salmonella typhimurium. *Eur. J. Biochem.* 51, 343–352. doi: 10.1111/j.1432-1033.1975.tb03934.x

Newton, A., Kendall, M., Vugia, D. J., Henao, O. L., and Mahon, B. E. (2012). Increasing rates of vibriosis in the United States, 1996-2010: review of surveillance data from 2 systems. *Clin. Infect. Dis.* 54 Suppl 5, S391–S395. doi: 10.1093/cid/cis243

Nikaido, H. (2003). Molecular basis of bacterial outer membrane permeability revisited. *Microbiol. Mol. Biol. Rev.* 67, 593–656. doi: 10.1128/MMBR.67.4.593-656.2003

Okada, K., Roobthaisong, A., Hearn, S. M., Okada, P. A., Doung-Ngern, P., Wongboot, W., et al. (2023). Emergence of Vibrio parahaemolyticus serotype O10: K4 in Thailand. *Microbiol. Immunol.* 67, 201–203. doi: 10.1111/1348-0421.13055

Okuda, J., Ishibashi, M., Hayakawa, E., Nishino, T., Takeda, Y., Mukhopadhyay, A. K., et al. (1997). Emergence of a unique O3:K6 clone of Vibrio parahaemolyticus in Calcutta, India, and isolation of strains from the same clonal group from Southeast Asian travelers arriving in Japan. *J. Clin. Microbiol.* 35, 3150–3155. doi: 10.1128/jcm.35.12.3150-3155.1997

Page, A. J., Taylor, B., Delaney, A. J., Soares, J., Seemann, T., Keane, J. A., et al. (2016). SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb. Genom* 2, e000056. doi: 10.1099/mgen.0.000056

Partridge, S. R., Tsafnat, G., and Coiera, E. (2009). and Iredell, J. R. Gene cassettes and cassette arrays in mobile resistance integrons. *FEMS Microbiol. Rev.* 33, 757–784. doi: 10.1111/j.1574-6976.2009.00175.x

Pazhani, G. P., Chowdhury, G., and Ramamurthy, T. (2021). Adaptations of vibrio parahaemolyticus to stress during environmental survival, host colonization, and infection. *Front. Microbiol.* 12, 737299. doi: 10.3389/fmicb.2021.737299

Pertea, G., and Pertea, M. (2020). GFF utilities: gffRead and gffCompare. *F1000Res* 9, ISCB Comm J-304. doi: 10.12688/f1000research

Qi, X. L., Wang, H. X., Bu, S. R., Xu, X. G., Wu, X. Y., and Lin, D. F. (2016). Incidence rates and clinical Symptoms of Salmonella, Vibrio parahaemolyticus, and Shigella infections in China, 1998-2013. J. Infect. Dev. Ctries 10, 127–133. doi: 10.3855/jidc.6835

Rossi, F., Rizzotti, L., Felis, G. E., and Torriani, S. (2014). Horizontal gene transfer among microorganisms in food: current knowledge and future perspectives. *Food Microbiol.* 42, 232–243. doi: 10.1016/j.fm.2014.04.004

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153

Shimohata, T., and Takahashi, A. (2010). Diarrhea induced by infection of Vibrio parahaemolyticus. J. Med. Invest. 57, 179-182. doi: 10.2152/jmi.57.179

Simon, A. K., Hollander, G. A., and McMichael, A. (2015). Evolution of the immune system in humans from infancy to old age. *Proc. Biol. Sci.* 282, 20143085. doi: 10.1098/ rspb.2014.3085

Taylor, R. K., Miller, V. L., Furlong, D. B., and Mekalanos, J. J. (1987). Use of phoA gene fusions to identify a pilus colonization factor coordinately regulated with cholera toxin. *Proc. Natl. Acad. Sci. U.S.A.* 84, 2833–2837. doi: 10.1073/pnas.84.9.2833

Wang, L. P., Zhou, S. X., Wang, X., Lu, Q. B., Shi, L. S., Ren, X., et al. (2021). Etiological, epidemiological, and clinical features of acute diarrhea in China. *Nat. Commun.* 12, 2464. doi: 10.1038/s41467-021-22551-z

Wiedenbeck, J., and Cohan, F. M. (2011). Origins of bacterial diversity through horizontal genetic transfer and adaptation to new ecological niches. *FEMS Microbiol. Rev.* 35, 957–976. doi: 10.1111/j.1574-6976.2011.00292.x

Xu, T., Yao, F., Zhou, X., Deng, Z., and You, D. (2010). A novel host-specific restriction system associated with DNA backbone S-modification in Salmonella. *Nucleic Acids Res.* 38, 7133–7141. doi: 10.1093/nar/gkq610

Xue, M., Gao, Q., Yan, R., Liu, L., Wang, L., Wen, B., et al. (2023). Comparative genomic analysis of shrimp-Pathogenic vibrio parahaemolyticus LC and intraspecific strains with emphasis on virulent factors of mobile genetic elements. *Microorganisms* 11, 2752. doi: 10.3390/microorganisms11112752

Yildiz, F. H., and Visick, K. L. (2009). Vibrio biofilms: so much the same yet so different. *Trends Microbiol.* 17, 109–118. doi: 10.1016/j.tim.2008.12.004

Yu, X. J., Grabe, G. J., Liu, M., Mota, L. J., and Holden, D. W. (2018). SsaV interacts with ssaL to control the translocon-to-Effector switch in the salmonella SPI-2 type three secretion system. *mBio* 9, e01149-18. doi: 10.1128/mBio.01149-18

Zhang, P., Wu, X., Ji, L., Yan, W., Chen, L., Lu, Z., et al. (2024). Prevalence and virulence of Vibrio parahaemolyticus isolated from clinical and environmental samples in Huzhou, China. *BMC Genomics* 25, 1187. doi: 10.1186/s12864-024-11106-3

Zhang, P., Wu, X., Yuan, R., Yan, W., Xu, D., Ji, L., et al. (2022). Emergence and predominance of a new serotype of Vibrio parahaemolyticus in Huzhou, China. *Int. J. Infect. Dis.* 122, 93–98. doi: 10.1016/j.ijid.2022.05.023

Zhou, Z., Charlesworth, J., and Achtman, M. (2020). Accurate reconstruction of bacterial pan- and core genomes with PEPPAN. *Genome Res.* 30, 1667–1679. doi: 10.1101/gr.260828.120

Zhu, Q., Chen, X., Liu, Y., Wang, R., Chen, J., and Chen, Y. (2020). Virulence, antimicrobial susceptibility, molecular and epidemiological characteristics of a new serotype of vibrio parahaemolyticus from diarrhea patients. *Front. Microbiol.* 11, 2025. doi: 10.3389/fmicb.2020.02025